Optimal Designs for Multivariate Linear Models

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

von D.H.S. Moudar Soumaya
(akademischer Grad, Vorname, Name)

geb. am 01. Januar 1979 in Lattakia-Syrien

genehmigt durch die Fakultät für Mathematik
der Otto-von-Guericke-Universität Magdeburg

Gutachter: Prof. Dr. Rainer Schwabe
(akademischer Grad, Vorname, Name)

Prof. Dr. Wolfgang Bischoff
(akademischer Grad, Vorname, Name)

eingereicht am: 26.06.2013

verteidigt am: 29.11.2013

Acknowledgment

The foundation stone of this work was laid from my advisor Prof. Rainer Schwabe, as he had suggested the idea and supported me by building the remainder stones with rich discussions, suggestions and kind encouragements, so i would like to thank him cordially for that and for his help and solidarity by overcoming the different difficulties in Germany.
I am thankful to the team of the institute of mathematical stochastics at the Otto-von-Guericke University of Magdeburg for their kindness, encouragement, and warm environment by working and supporting by overcoming any problems and for the nice meetings with cake. I would like to thank Prof.Norbert Gaffke for the creative ideas by reduction the long calculation of the inverse for the partitioned matrices, and for his help and learning me the theoretical background of mathematical statistics, and Fritjof Freise for his friendly help by correcting my German texts, and Kerstin Altenkirch for his solidarity and helping me, Prof. Waltraud Kahle for teaching me the mathematical statistics, Survival Analysis and computer intensive processes, Prof. Gerd Christoph for teaching me the stochastic processes.
I could not have done this work without the scholarship from Al-Baath university in my home of Syria, so i would like to thank my inner advisor Prof. Dr. Haytham Farah and the secretary of the statistical department for supporting, helping and encouraging me. I would like to thank Dr. Kamal Jony and the staff of the mathematic and physic department at Tishreen university in Lattakia, who have taught and motivated me.
I would like to thank my friends and relatives, who have motivated and helped me.
I am very grateful to my father and two mothers, who have struggled for a better future for my siblings and me, motivated and gave us hope.
I am grateful to my wife Ruba, who has been my best supporter in Germany even though i have had less time to spend it with her and she has been patient with me, has financed me in the last months, has motivated me, and has donated me my best years with a lot of love and gave me hope by the birth our child Ali, who has simplified all problems with his smile and nice simple speeches.

Abstract

In this work we investigate the problem of characterizing optimal experimental designs for a wide class of multivariate statistical models. In particular, we consider SUR models of seemingly unrelated regression, which describe models for multivariate observations, where the components do not seem to have anything in common at a first sight. However these observations are indirectly related to each other by a correlation structure between the components. These SUR models were originally introduced in economic applications, but may nowadays also be used in other practical problems for example in the biosciences, when several processes like pharmacokinetics and pharmacodynamics may be measured at different time points at the same individuals. For the control variables (e.g. measurement times) of the observations the settings of the single components (processes) may vary across the components within each individual. The aim of our investigation is now to determine the best settings for each component within each individual, if the correlation structure between the components is incorporated in the statistical analysis.
In contrast to a primary guess experimental conditions according to a MANOVA structure of multivariate regression, for which the settings coincide for all components, are optimal only in the case of uncorrelated observations. In all other cases under mild regularity conditions product type designs turn out to be better, which contain all possible combination of those settings, which are optimal in the corresponding univariate models.
For a more detailed characterization we first introduce the basics of optimal design theory for univariate linear models and specify the SUR models considered. It is then shown that product type designs are optimal with regard to the D-criterion of minimizing the determinant of the variance covariance matrix or to the IMSE-criterion of minimizing the mean expected quadratic deviation of the estimated response function, as long as the marginal models contain an intercept. This result will be extended to more general SUR models containing nested multiplicative or additive structures.
Additionally different variants are proposed for a G-criterion of minimizing the maximal expected deviation of the estimated response function and are investigated for MANOVA and SUR models. Also the efficiency of the OLS estimator is compared with the optimal Gauß-Markov estimator, if product type designs are used.
Finally, a consideration of univariate marginal models without intercepts yields the unexpected result that product type designs retain their optimality only for small to moderate correlations, whereas they may lose their optimality in the presence of larger correlations.with depended components and individuals to the corresponding problem for univariate models with correlated observations.

Zusammenfassung

In dieser Arbeit untersuchen wir das Problem der Charakterisierung optimaler Versuchspläne für eine weite Klasse von multivariaten statistischen Beobachtungsmodellen. Speziell betrachten wir Modelle für scheinbar unverbundene Regression, d.h. für multivariate Beobachtungen, die auf den ersten Blick nichts mit einander zu tun haben, aber durch in denen die Beobachtungen durch eine Korrelationsstruktur miteinander verbunden sind. Diese sogenannten SUR-Modelle wurden ursprünglich in ökonomischen Anwendungen eingeführt, können aber auch in anderen praktischen Problemen z.B. in den Biowissenschaften eingesetzt werden, wenn mehrere Prozesse wie Pharmakokinetik und Pharmakodynamik zu unterschiedlichen Zeitpunkten an denselben Individuen beobachtet werden können. Die Kontrollvariablen (z.B. Messzeitpunkte) für die Beobachtungen der verschiedenen Komponenten (Prozesse) können dabei innerhalb eines Individuums von Komponente zu Komponente variieren. Ziel ist es nun, die besten Einstellungen für die einzelnen Komponenten zu bestimmen, wobei bei der Analyse der Beobachtungen die Korrelation zwischen den Beobachtungen ausgenutzt werden soll.

Entgegen einer ersten Vermutung sind Versuchseinstellungen mit einer MANOVA-Struktur der multivariaten Regression, bei denen die Einstellungen für alle Komponenten übereinstimmen, nur bei Unkorreliertheit der Komponenten optimal. In allen anderen Fällen erweisen sich unter Regularitätsvoraussetzungen an die Modelle sogenannte produktartige Versuchseinstellungen als besser, bei denen alle Versuchseinstellungen von Designs, die optimal in den zugehörigen univariaten Marginalmodellen sind, miteinander kombiniert werden.

Um dies genauer zu charakterisieren, werden zuerst die Grundlagen der Theorie optimaler Versuchsplanung für univariate lineare Modelle vorgestellt und die betrachteten multivariaten SUR-Modelle spezifiziert. Es wird dann generell gezeigt, dass bezüglich üblicher Kriterien wie des D-Kriteriums zur Minimierung der Determinante der Varianz-Kovarianzmatrix oder des IMSE-Kriteriums zur Minimierung der gemittelten erwarteten quadratischen Abweichungen der geschätzten Wirkungsfunktionen produktartige Pläne optimal sind, sofern die univariaten marginalen Modelle Achsenabschnitte enthalten. Dies wird auf SUR-Modellen mit verschiedenen geschachtelten, multiplikativen oder additiven Strukturen verallgemeinert.

Weiterhin werden einerseits verschiedene G-Optimalitätskriterien zur Minimierung der maximalen erwarteten Abweichung der geschätzten Wirkungsfunktionen vorgeschlagen und beim Einsatz von MANOVA- und SUR-Modellen untersucht und andererseits die Effizienz des OLS-Schätzers gegenüber dem optimalen Gauß-Markov-Schätzer bei der Verwendung von produktartigen Plänen untersucht.

Bei der Betrachtung von marginalen Modellen ohne Achsenabschnitt ergibt sich schließlich das unerwartete Ergebnis, dass produktartige Pläne nur für moderate Korrelationen optimal sind und diese Optimalität für starke Korrelationen verlieren können.

Contents

1 Introduction

When a phenomena or process is described by a statistical model, which is assumed to be correct, then the optimal design theory can play a roll by improving the performance of the experiment The work of *(Smith (1918)*) was the entrance to this area of research , where the maximal variance for the prediction resp. the G- or global optimal design criterion had been discussed, then followed by many works about optimal designs in the univariate case.
The optimal design theory in the multivariate case was brought to light extensively by the work of *(Fedorov (1972)*), where the multivariate version of the equivalence theorems for D- and linear criteria were derived. After that, there are many works in the multivariate case, such as multivariate linear models, their components are hierarchically nested and are considered by the work of *(Krafft and Schaefer (1992)*), and resulted in the usage of a method, which was developed by *(Dette (1990)*), that the D-optimal designs for these kind of models is the joint D-optimal design for the corresponding marginal univariate models.
The MANOVA-models or homogeneous multivariate linear models are considered in the work of *(Chang (1994)*) and the result was, that the reduction of the D-optimal design problem for these type of models on the corresponding univariate problem is possible, so the D-optimal designs for one of the homogeneous univariate linear models is the D-optimal for the MANOVA-model.
An Extension to both works *(Krafft and Schaefer (1992)*) and *(Chang (1994)*) was the work of *(Kurotschka and Schwabe (1996)*), where the Result of *(Chang (1994)*) is extended for the A-, c- and E-optimality criteria and the result of *(Krafft and Schaefer (1992)*) is proven by another technique and broaded for multivariate linear models with heterogeneous marginal components under the condition of the block diagonal form of the information matrices.
Bivariate linear models with multi-factor marginal models by Kronecker product form for the variance covariance matrix of the error variables are discussed in the work of *(Schwabe (1996)*) and it is shown, that the reduction of the bivariate problem on its corresponding univariate problems by the optimality of the product type designs is possible.
A special bivariate linear model with heterogeneous marginal components, where there are some joint parameters by both components is considered in the work of *(Changa et al. (2001)*) and it is shown by some examples, that the D-optimality for some designs by use of the equivalence theorem is restricted to intervals for the correlation term.
Dealing with the optimal design problem in the multivariate case is explored through algorithms or optimization methods as the positive definite programming, see for example the works of *(Wijesinha and Khuri (1987)*) and *(Atashgah and Seifi (2009)*).
SUR models with different marginal structures are considered in this work and it is proven based on the equivalence theorems for D- and linear criteria, that the product type designs are D- and linear optimal if the weight matrix is block diagonal(4.1), for SUR models with intercepts by the marginals with one-factor marginal components, multiplicative marginal components by different nesting cases for general form for the information matrices, and for additive marginal components by different nesting cases as the hierarchical form of *(Krafft and Schaefer (1992)*), for SUR models without intercepts by the marginal components for block diagonal form of the information matrices. Where a practical

example for SUR models without intercepts by the marginals, which have block diagonal or non block diagonal form can be respectively the multivariate chemical and spring balance regression models. These results hold for general structure for the variance covariance matrix of the error variables analytically and asymptotically under the condition of normality, because of the block diagonal form of the Fisher-information matrix.

D-optimality for the product type designs for SUR models without intercepts by the marginal components is restricted by non block diagonal form for the information matrix and for homogeneous correlation structure, so the product type designs are D-optimal for some intervals for the correlation terms as in the examples of *(Changa et al. (2001))*, which include zero and their lengths are less than one in the absolute value, where the lengths for these intervals will be closer for growing number of components m. A theoretical background and justification for these restrictions on the D-optimality for some designs in the multivariate case, with respect to the correlation term for homogeneous correlation structure are generally enriched, with many simulations for different models and this result is valid for all multivariate problems, by which just the correlation term is included in the sensitivity functions for the D-optimality, i.e. that is valid just by homogeneous correlation structure and may be locally for heterogeneous structure.

IMSE-optimality criteria is determined in the multivariate case and the result of *(Kurotschka and Schwabe (1996))* due to the MANOVA-model is extended for it. The counterpart of D-optimality in the multivariate equivalence theorem is a weighted G-optimality due to the trace, so the D- and G-optimal design are not identical in the multivariate case and there is a covariance matrix for the prediction and no longer a variance function, therewith some G-optimal design criteria are suggested as the trace, maximal eigenvalue and the determinant, so the first step in this direction was the determining the upper bounds for these functions from the multivariate equivalence theorem for the D-optimality based on some inequality from the matrix theory for the product of the positive definite matrices, then these upper bounds are calculated for the MANOVA-models by the evaluation of the MANOVA-design and for SUR models by the evaluation of the product type designs due to the Gauß and OLS estimators. The OLS estimator is a limited-information estimator in many multivariate cases, see for example *(Amemiya (1985))*, but its advantage is, that it can be used, when the variance covariance matrix of the error variables is unknown, so the efficiencies for it vs the BLUE Gauß estimator with respect to the product type designs and D- and linear criteria in 4.1 are calculated, as well as the efficiencies for MANOVA-design vs product type designs due to the D- and linear criteria in 4.1 for the OLS and Gauß estimators. The reduction of the multivariate design problem for MANOVA-model with correlated components and individuals resp. observations on its corresponding problem for the corresponding univariate model with correlated observation is possible.

A short introduction for optimal design theory in the univariate case based on the one-factor linear models is introduced in the second chapter as well as some results for the optimal designs for univariate multi-factor models, which have been discussed by *(Schwabe (1996))*. The SUR model as general multivariate linear models and some related models as the homogeneous multivariate linear models (MANOVA), and heterogeneous multivariate linear models are presented in chapter three and in addition to that the correlation matrix and some of its prosperities and some estimators and their asymptotic properties are illustrated. Then some fundamentals of the optimal design theory in the multivariate case as the derivation of the IMSE-criterion are interpreted in chapter four and the optimal designs for SUR models with intercepts by the marginal components for general known and unknown variance covariance matrix of the error variables are explored. G-optimal designs and the efficiency of the OLS estimator are the main topics of the fifth chapter. Optimal designs for multivariate multi-factor models with and without nesting structure or for SUR models with multi-factor marginal components are discussed in chapter six. Optimal designs for SUR models without intercepts by the marginals as the multivariate version of the spring and chemical balance models are explored extensively in chapter seven. The work is then concluded with a Discussion, some extensions and possible future research.

2 Optimal Designs for Univariate Linear Regression Models

The necessary fundamentals and tools of the optimal design theory for this work are presented in this chapter based on an univariate linear, one-factor regression model. The most published books on the optimal design theory were introduced based on the univariate linear case and in the simplest case see for example *(Fedorov (1972))*, *(Pázman (1986))*, *(Bandemer and Bellmann (1994))* or *(Silvey (1980))*. This theory was introduced in some books for different estimators and with more details and examples in the approximate and exact cases, see for example *(Bandemer (1977)* and *(Bandemer and Näther(1980).* There are many theoretical results with respect to the approximate design in the works *(Kiefer (1959), (1961), (1974) and (1975))* and of (*Kiefer and Wolfowitz (1959) and (1960)*). The work (*(Schwabe (2008)*) can be also a very good short introduction for the optimal design theory. This chapter is ordered as follows, the univariate linear, one-factor regression models have been introduced in the first section , some fundamentals for the optimal design theory are defined based on the univariate case and in addition some useful examples in the second section. The univariate linear, multi-factor regression models are introduced and the most important results for optimal designs for these kinds of models from the work of *(Schwabe (1996)*) summarized and supported with two clear examples in the last section.

2.1 The Univariate Linear One-Factor Regression Model

The univariate linear regression model can be derived from an experiment based on the possible relationship between two variables, the first one is influenced by the second one linearly, thus the first one is called the response, dependent, outcome, or goal variable and the second one is the control, independent, income, explanatory, setting, or predictor variable. This relationship is weighted with parameters and bound with an error variable, which is uncorrelated for multiple observations. Thus, the model for one observation can look for the response function $\eta\ (x_i)$ as follows

$$Y_i = \eta\ (x_i) + \varepsilon_i,,\ i = 1, ..., n$$

for linear factorization of the response function $\eta\ (x_i) = \mathbf{f}(x_i)^\top \boldsymbol{\beta}$, where $\mathbf{f}(x_i)^\top = (f_1(x_1), ..., f_p((x_n))$ is the known regression function and $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)^\top$ is the unknown parameter vector for the model, the regression model is given as follows

$$Y_i = \mathbf{f}(x_i)^\top \boldsymbol{\beta} + \varepsilon_i,\ i = 1, ..., n$$

With the specifications $\mathrm{E}(\varepsilon_i) = 0$, $\mathrm{Var}(\varepsilon_i) = \sigma^2$, where $\sigma > 0$ and known, and $\mathrm{Cov}\,(\varepsilon_i, \varepsilon_k) = 0; i \neq k$. We can introduce the compact model in a vector form

$$\mathbf{Y} = \mathbf{F}\boldsymbol{\beta} + \boldsymbol{\varepsilon},\ \mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0},\ \mathrm{Cov}\,(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}_{n\times n}$$

Where $\mathbf{x} = (x_1, ..., x_n)$ is the experimental setting, $\mathbf{F} = (\mathbf{f}(x_1), ..., \mathbf{f}(x_n))^\top$ is the design matrix, which is assumed to be of the rank p, and the experimental setting x_i may be chosen from an experimental region $\mathcal{X}$. Denote by $\mathbf{Y} = (Y_1, ..., Y_n)^\top$, $\boldsymbol{\varepsilon} = (\varepsilon_1, ..., \varepsilon_n)^\top$ the vectors of observations and error terms, and $\mathbf{I}_{n\times n}$ is the identity matrix.
We can use the Gauß Markov theorem, see *(Rao et al. (2007))*, to estimate the parameter of the univariate regression model by the second order specifications and thus we get the Gauß Markov estimator which has the same form of the ordinary least square estimator (OLS) in the univariate case. When the error variable are normally distributed, then we can use the Maximum Likelihood estimator (ML), see *(Casella and Lehmann (1998))*. For the univariate case the three estimators are BLUE. lso, the three estimators (Gauß), (OLS) and (ML) are identical and have the next form

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^\top \mathbf{F})^{-1} \mathbf{F}^\top \mathbf{Y}$$

The variance covariance matrix of those estimators are given in the next form

$$\text{Cov}\,(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{F}^\top \mathbf{F})^{-1} \tag{2.1}$$

We call the inverse of the variance covariance matrix for BLUE estimators the information matrix because of the BLUE property for the considered estimators, which has the next form

$$\mathbf{M}(\mathbf{x}) = \frac{1}{\sigma^2} \mathbf{F}^\top(\mathbf{x}) \mathbf{F}(\mathbf{x}) = \frac{1}{\sigma^2} \sum_{i=1}^{n} \mathbf{f}(x_i) \mathbf{f}(x_i)^\top \tag{2.2}$$

The minimal loss of the information is given by the variance covariance matrix (2.1) resp. the maximum profit of the informations is given by the information matrix of (2.2) by the theory of the mathematical statistics. That is enough for some areas of science as the econometrics, because the analysis of the collected data is in focus. But it is an advantage to minimize the loss of the information resp. to maximize the profit of information with respect to the estimator, or in other words to determine the optimal control variables and observe the response variable by it for some technical areas of science, such as the pharmaceutical, resp. chemical industry or neuroscience, and that is the goal of the optimal design theory.

2.2 Optimal Design Theory's Fundamentals in the Univariate Case

The information matrix resp. the variance covariance matrix for an estimator can not be directly maximized resp. minimized generally, thus we have to apply some functions to it and then optimize it with respect to the control variables. also, a tuple of control variables $(x_1, ..., x_n)$ may be as equal form an exact design, other interpretations of the exact design can be formed based on q. Different design points and their corresponding frequencies or the weights of the design points as follows

$$d_n = \left\{ \begin{matrix} x_1, & \dots & , x_q \\ w_1, & \dots & , w_q \end{matrix} \right\} = [x_i, w_i]_{i=1}^{q}; w_i = \frac{n_i}{n}, n_i \text{ integer and } \sum_{i=1}^{q} n_i = n$$

In dealing with the exact or discrete designs, are in many cases difficult because of combinatorial problems, which are the problems by the discrete optimization,also, by generalization of the exact design by accepting real values of w_i we get the approximate resp. normalized design, which can be defined as follows

$$\xi = \left\{ \begin{array}{ccc} x_1, & \dots & ,x_q \\ w_1, & \dots & ,w_q \end{array} \right\} = [x_i, w_i]_{i=1}^q;\ w_i > 0,\ \sum_{i=1}^q w_i = 1$$

When the w_i are real positive values and the design points x_i are called the support points of the design, they belong to the experimental region $\mathcal{X}$. The approximate design ξ can be interpreted too as a probability measure and therefor is sometimes called in literature the continuous design. The information matrix with respect to the design ξ is called the standardized information matrix, which can be defined as follows

$$\mathbf{M}(\xi) = \sum_{i=1}^q w_i \mathbf{f}(x_i)\mathbf{f}(x_i)^\top, \quad \text{resp. } \mathbf{M}(\xi) = \int_{\mathcal{X}} \mathbf{f}(x)\mathbf{f}(x)^\top \xi(\mathrm{d}x)$$

The set of all possible approximate designs ξ in an experimental region $\mathcal{X}$ is Ξ and the corresponding set of all information matrices for the ξ are formed by the set $\mathcal{M}$, i.e. mathematically

$$\mathcal{M} = \mathbf{M}(\xi),\ \xi \in \Xi$$

The set of information matrices $\mathcal{M}$ corresponding to all possible approximate designs is convex, where the information matrix is positive semi definite see *(Fedorov (1972))*.

Remark 2.1. *Only positive definite and regular information matrices are considered in this work.*

2.2.1 Some Optimality criteria

To make the optimization on the convex set $\mathcal{M}$ convex, we have to apply to it convex functions as the $-\log\det$ or trace of the inverse of the information matrix , such functions Φ are called the optimality criteria and are used to minimize the variance covariance matrix of the estimator for the parameter Cov $(\hat{\boldsymbol{\beta}})$ resp. $\mathbf{M}^{-1}(\xi)$. For a good description of the geometrical meaning of the optimality criteria see *(Fedorov (1972))*, *(Pukelsheim (1993))* or *(Myers et al. (2010))*.

Def 2.1 (D-optimal criterion)**.** *An approximate design ξ_D^* is called D-optimal design, if it minimizes the determinant of the variance covariance matrix resp. maximizes the determinant of the information matrix or*

$$-\log(\det \mathbf{M}(\xi_D^*)) = \min_{\xi\in\Xi}[-\log(\det \mathbf{M}(\xi))]$$

Some criteria can be defined as linear criteria

Def 2.2. *A linear Criterion $\Phi : \mathcal{M}(\Xi) \longrightarrow \mathcal{R}$ is defined as follows*

$$\Phi(\mathbf{M}(\xi)) = \mathit{trace}\ [\mathbf{L}\mathbf{M}^{-1}(\xi)]$$

Where $\mathbf{L}$ is a positive semi definite Matrix and multipliable with $\mathbf{M}(\xi)$.

for example the A- and IMSE-criteria, which will be defined soon

Def 2.3 (A-optimal criterion)**.** *An approximate design ξ_A^* is called A-optimal design, if it minimizes the trace of the variance covariance matrix, i.e.*

$$\mathit{trace}\ (\mathbf{M}^{-1}(\xi_A^*)) = \min_{\xi\in\Xi}[\mathit{trace}\ (\mathbf{M}^{-1}(\xi))]$$

Def 2.4 (E-optimal criterion). *An approximate design ξ_E^* is called E-optimal design, if it minimizes the largest eigenvalue of the variance covariance matrix, i.e.*

$$\lambda_{\max}\mathbf{M}^{-1}(\xi_E^*) = \min_{\xi\in\Xi}[\lambda_{\max}\mathbf{M}^{-1}(\xi)]$$

Another interpretation with some geometrical meaning for the D-, A- and E-optimality criteria can be found in *(Kiefer (1975))*.
Not all optimal design criteria are implemented to improve the quality by estimating the parameters, rather, some optimality criteria are applied to improve the quality of the predictions or to minimize the variance of it, as the following two criteria show

Def 2.5 (G-optimal criterion in the univariate case). *An approximate design ξ_G^* is called G-optimal design, if it minimizes the maximum of the variance of the prediction in the design region $\mathcal{X}$, i.e.*

$$\max_{x\in\mathcal{X}}\left(\mathbf{f}(x)^\top\mathbf{M}^{-1}(\xi_G^*)\mathbf{f}(x)\right) = \min_{\xi\in\Xi}[\max_{x\in\mathcal{X}}(\mathbf{f}(x)^\top\mathbf{M}^{-1}(\xi)\mathbf{f}(x))]$$

Def 2.6 (IMSE in the Univariate Case). *The integrated mean square error is the integrated predictive variance with respect to the uniform measure $\mu(\mathrm{d}x)$ and is defined as follows*

$$\mathrm{IMSE} = \int_{\mathcal{X}} \mathrm{Var}\left(\mathbf{f}(x)^\top\hat{\boldsymbol{\beta}}\right)\mu(\mathrm{d}x) = \int \mathrm{E} \parallel \left(\mathbf{f}(x)^\top(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\right)\parallel^2\mu(\mathrm{d}x)$$

Where $\|.\|$ denotes the Euclidean norm.

Def 2.7 (IMSE-optimal criterion). *An approximate design ξ_{IMSE}^* is called IMSE-optimal design, if it minimizes the integrated predictive variance with respect to the uniform measure $\mu(\mathrm{d}x)$ in the design region $\mathcal{X}$, i.e.*

$$\int_{\mathcal{X}} \mathbf{f}(x)^\top\mathbf{M}^{-1}(\xi_{IMSE}^*)\mathbf{f}(x)\ \mu(\mathrm{d}x) = \min_{\xi\in\Xi}[\int_{\mathcal{X}} \mathbf{f}(x)^\top\mathbf{M}^{-1}(\xi)\mathbf{f}(x)\ \mu(\mathrm{d}x)]$$

Remark 2.2. *The IMSE-optimal criterion can be interpreted due to the weight matrix for the linear criteria* $\mathbf{L}$ *as follows*

$$\mathit{trace}\ (\mathbf{L}\mathbf{M}^{-1}(\xi_{IMSE}^*)) = \min_{\xi\in\Xi}[\mathit{trace}\ (\mathbf{L}\mathbf{M}^{-1}(\xi))],\ \mathbf{L} = \int_{\mathcal{X}} \mathbf{f}(x)\mathbf{f}(x)^\top\ \mu(\mathrm{d}x)$$

2.2.2 Some Tools by the Approximate Design Theory

Dealing with the optimality criteria is difficult in many situations, so some instruments can be developed based on the direction resp. Fréchet derivative as the equivalence theorem and
coefficients to check the optimality or the efficiency of a design with respect to a criterion Φ.

Def 2.8 (The directional Derivative). *Let $\mathbf{M}(\xi)$ be the information matrix for a design ξ and ϕ a criterion, then the directional derivative of it by $\mathbf{M}(\xi)$ in the direction of $\mathbf{M}(\acute{\xi})$ is*

$$F_\Phi(\mathbf{M}(\xi);\mathbf{M}(\acute{\xi})) = \lim_{\alpha\downarrow 0}\frac{1}{\alpha}(\Phi[(1-\alpha)\mathbf{M}(\xi)+\alpha\mathbf{M}(\acute{\xi})]-\Phi[\mathbf{M}(\xi)])$$

Thus $F_\Phi(\mathbf{M}(\xi);\mathbf{M}(\xi)) = 0$.
The negative signed direction derivative in the direction of the one-point design with weight equal to one, resp. $\xi = \{ \begin{matrix} x \\ 1 \end{matrix} \}$ and the corresponding information matrix $\mathbf{M}(x) = \mathbf{f}(x)\mathbf{f}(x)^\top$ is

a useful term for the equivalence theorem, so the sensitivity function can be defined as follows based on it

Def 2.9 (The Sensitivity Function). *The sensitivity function for differentiable criteria Φ is given with respect to one-point design as follows*

$$\varphi_{\ \Phi}(x;\xi) = -F_\Phi(\mathbf{M}(\xi);\mathbf{M}(x))$$

which shows which experimental settings are "most informative".

The next form of the equivalence theorem is given generally in the concave case by *(Silvey (1980))*, thus because of the duality between the concavity and convexity it can be reformulated as follows

Theorem 2.1 (The General Equivalence Theorem). *Let ξ_Φ^* be Φ-optimal design and Φ convex and differentiable on $\mathcal{M}$, then ξ_Φ^* minimizes the function $\Phi(\mathbf{M}(\xi))$ if and only if*

$$F_\Phi(\mathbf{M}(\xi_\Phi^*);\mathbf{M}(x)) \geq 0, \ \textit{resp.} \ \ \varphi_{\ \Phi}(x;\xi_\Phi^*) \leq 0, \quad \forall x \in \mathcal{X}$$

Graphically, it means that, the design ξ is Φ-optimal, when we can not improve it, wherever the direction for the derivative is. In figure 2.1, for the the direction derivative of the convex function $x^2 + y^2$ on $[-1,1]^2$ is clear, that the direction derivative by the minimum $(0,0)$ is still up.
In the work of *(Kiefer and Wolfowitz (1959))* is a common version of the equivalence theorem for the D- and G-optimality development , which have the following context

Theorem 2.2 (The Equivalence Theorem for the G- and D-Optimality). *The approximate design ξ^* is D-optimal or G-optimal in the univariate linear model if and only if*

$$\varphi_D(x;\xi_D^*) = \mathbf{f}(x)^\top\mathbf{M}^{-1}(\xi_D^*)\mathbf{f}(x) \leq p$$

for all $x \in \mathcal{X}$, where p is the number of parameters in the model, and the maximum for the sensitivity function for the D-optimality by ξ_D^* is given as follows

$$\max_{x\in\mathcal{X}} \mathbf{f}(x)^\top\mathbf{M}^{-1}(\xi_D^*)\mathbf{f}(x) = p$$

and that occurs by the support points of the optimal design.

A linear function $\Phi : \mathcal{M}(\Xi) \longrightarrow \mathcal{R}$ has the following property

$$\Phi[\gamma_1\mathbf{M}(\xi_1) + \gamma_2\mathbf{M}(\xi_2)] = \gamma_1\Phi[\mathbf{M}(\xi_1)] + \gamma_2\Phi[\mathbf{M}(\xi_2)]$$

This is the equivalence theorem for the linear criteria developed in the work of *(Fedorov (1972))*

Theorem 2.3. *The approximate design ξ_L^* is linear optimal in the univariate linear model if and only if*

$$\varphi_L(x;\xi_L^*) = \mathbf{f}(x)^\top\mathbf{M}(\xi^*)^{-1}\mathbf{L}\ \mathbf{M}(\xi_L^*)^{-1}\mathbf{f}(x) \leq \text{trace}\left(\mathbf{L}\ \mathbf{M}(\xi_L^*)^{-1}\right)$$

for all $x \in \mathcal{X}$.

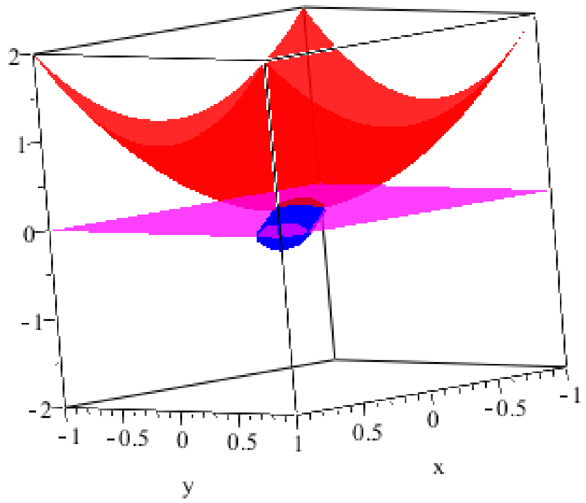

Figure 2.1: The directional Derivative for $x^2 + y^2$ by min $= (0,0)$ in the direction of $(1,1)$ is up

1. *For the A-optimality* $\mathbf{L} = \mathbf{I}_p$*, where* $\mathbf{I}_p$ *is the* $p \times p$ *identity matrix.*
2. *For the IMSE-optimality* $\mathbf{L} = \int_{\mathcal{X}} \mathbf{f}(x)\mathbf{f}(x)^\top \mu(\mathrm{d}x)$.

The quality of a competing approximate design ξ with respect to an Φ-optimality criteria can be measured in terms of its efficiency compared to the Φ- optimal design ξ^*_Φ.
So the D-efficiency for a design ξ is given as follows

$$\mathrm{eff}_D(\xi) = \left(\frac{\det \mathbf{M}(\xi)}{\det \mathbf{M}(\xi^*_D)}\right)^{1/p} = \left(\frac{\det \mathbf{M}^{-1}(\xi^*_D)}{\det \mathbf{M}^{-1}(\xi)}\right)^{1/p}$$

The G-efficiency is given as follows

$$\mathrm{eff}_G(\xi) = \frac{p}{\max_{x\in\mathcal{X}} \left(\mathbf{f}(x)^\top \mathbf{M}^{-1}(\xi^*_G)\mathbf{f}(x)\right)}$$

And the efficiency for the linear criteria is given as follows

$$\mathrm{eff}_L(\xi) = \frac{\mathrm{trace}\left(\mathbf{L}\ \mathbf{M}^{-1}(\xi^*_L)\right)}{\mathrm{trace}\left(\mathbf{L}\ \mathbf{M}^{-1}(\xi)\right)}$$

For the A-optimality $\mathbf{L} = \mathbf{I}_p$, where $\mathbf{I}_p$ is the $p \times p$ identity matrix and for the IMSE-optimality $\mathbf{L} = \int_{\mathcal{X}} \mathbf{f}(x)\mathbf{f}(x)^\top \mu(\mathrm{d}x)$. The efficiency states, how much less observations are required, when the Φ-optimal design ξ^*_Φ is used instead of ξ.
It is valid for the following examples, that $E(\varepsilon_i) = 0, var(\varepsilon_i) = \sigma^2$, $\sigma^2 = constant$, $\mathrm{Cov}\,(\varepsilon_i, \varepsilon_k) = 0; i \neq k$.

Example 2.1. *It is useful to learn, how the weights for the Φ-optimal designs can be determined, which is the topic for this example. Here is the simple regression model*

$$Y_i(x_i) = \beta_1 + \beta_2 x_i + \varepsilon_i,\ x_i \in \mathcal{X} = [0,1]$$

let the design be

$$\xi = \begin{pmatrix} 0 & 1 \\ 1-w & w \end{pmatrix},\quad \sum_{i=1}^{2} w_i = 1$$

The regression function is

$$\mathbf{f}(x)^\top = \begin{pmatrix} 1 & x \end{pmatrix}$$

and including the information matrix for the one-point design is given as follows

$$\mathbf{M}(x) = \mathbf{f}(x)\mathbf{f}(x)^\top = \begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix}$$

The standardized information matrix for ξ is

$$\mathbf{M}(\xi) = \sum_{i=1}^{2} w_i \mathbf{f}(x_i)\mathbf{f}(x_i)^\top = \begin{pmatrix} 1-w & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} w & w \\ w & w \end{pmatrix} = \begin{pmatrix} 1 & 1-w \\ 1-w & 1-w \end{pmatrix}$$

Then

$$\mathbf{M}^{-1}(\xi^*) = \frac{1}{w}\begin{pmatrix} 1 & -1 \\ -1 & \frac{1}{1-w} \end{pmatrix}$$

So the weights for the D-optimal design can be determined by minimizing the convex function $-\log\det\mathbf{M}(\xi)$, *i.e.*

$$-\log\det\mathbf{M}(\xi) = -\log(w - w^2) \longmapsto \min$$

Thus the resulted minimum is $w_D^* = \frac{1}{2}$, so the D- resp. G-optimal design is given as follows

$$\xi_{D;G}^* = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

For the A-optimal design the weights can be determined by minimizing the convex function $\operatorname{trace}\left(\mathbf{M}^{-1}(\xi)\right)$, *i.e.*

$$\operatorname{trace}\left(\mathbf{M}^{-1}(\xi)\right) = \frac{2-w}{w(1-w)} \longmapsto \min$$

Thus the resulting minimum is $w_A^* = 2-\sqrt{2}$, so the A-optimal design is given as follows

$$\xi_A^* = \begin{pmatrix} 0 & 1 \\ 2-\sqrt{2} & \sqrt{2}-1 \end{pmatrix}$$

(check the A-optimality equivalence theorem 2.3 as in example 2.3).
For the IMSE-optimal design the weights can be determined by minimizing the convex function $\operatorname{trace}\left(\mathbf{L}\ \mathbf{M}^{-1}(\xi^*)\right)$, *where*

$$\mathbf{L} = \int_0^1 \mathbf{f}(x)\mathbf{f}(x)^\top \mathrm{d}x = \int_0^1 \begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix} \mathrm{d}x = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}$$

So

$$\text{trace}\left(\mathbf{L}\ \mathbf{M}^{-1}(\xi^*)\right) = \frac{1}{3w(1-w)} \longmapsto \min$$

Thus the resulting minimum is $w^*_{IMSE} = \frac{1}{2}$*, so the IMSE-optimal design is given as follows*

$$\xi^*_{IMSE} = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

By a similar process we can illustrate, that the D- resp. G-optimal, A-optimal and IMSE-optimal designs for the same model in the design region $\mathcal{X} = [-1, 1]$ *are given as follows*

$$\xi^*_{D;G;A;IMSE} = \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Example 2.2. *It is also useful to learn, how the design points can be determined, so the boundary points should be considered as well as a kind of symmetry, may be just for univariate linear models with intercepts. For the cubic regression model*

$$Y_i(x_i) = \beta_1 + \beta_2 x_i + \beta_3 x_i^2 + \beta_4 x_i^3 + \varepsilon_i,\ x_i \in \mathcal{X} = [-1, 1]$$

a candidate design with respect to the D-optimality could be as follows

$$\xi = \begin{pmatrix} -1 & -x & x & 1 \\ \frac{w}{2} & \frac{(1-w)}{2} & \frac{(1-w)}{2} & \frac{w}{2} \end{pmatrix}$$

Thus the information matrix for one-point design has the following form

$$\mathbf{M}(x) = \mathbf{f}(x)\mathbf{f}(x)^\top = \begin{pmatrix} 1 & x & x^2 & x^3 \\ x & x^2 & x^3 & x^4 \\ x^2 & x^3 & x^4 & x^5 \\ x^3 & x^4 & x^5 & x^6 \end{pmatrix}$$

Then the standardized information matrix with respect to the four point design ξ *is*

$$\mathbf{M}(\xi^*) = \sum_{i=1}^{4} w_i \mathbf{f}(x_i)\mathbf{f}(x_i)^\top = \frac{w}{2}\begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{pmatrix} + \frac{(1-w)}{2}\begin{pmatrix} 1 & -x & x^2 & -x^3 \\ -x & x^2 & -x^3 & x^4 \\ x^2 & -x^3 & x^4 & -x^5 \\ -x^3 & x^4 & -x^5 & x^6 \end{pmatrix}$$
$$+\frac{(1-w)}{2}\begin{pmatrix} 1 & x & x^2 & x^3 \\ x & x^2 & x^3 & x^4 \\ x^2 & x^3 & x^4 & x^5 \\ x^3 & x^4 & x^5 & x^6 \end{pmatrix} + \frac{w}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & w+(1-w)x^2 & 0 \\ 0 & w+(1-w)x^2 & 0 & w+(1-w)x^4 \\ w+(1-w)x^2 & 0 & w+(1-w)x^4 & 0 \\ 0 & w+(1-w)x^4 & 0 & w+(1-w)x^6 \end{pmatrix}$$

So the points and the weights for the D-optimal design can be determined by minimizing the convex function
$-\log\det \mathbf{M}(\xi)$ *Thus the resulting minimum is* $x = \frac{1}{\sqrt{5}} \simeq 0.45, w = \frac{1}{4}$, *so the resulting design is*

$$\xi_D^* = \begin{pmatrix} -1 & -\frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

which is D- resp. G-optimal design.
To determine the support points of the A- and IMSE-optimal designs, we have to minimize resp. $\operatorname{trace}\left(\mathbf{M}(\xi)^{-1}\right)$ *and* $\operatorname{trace}\left(\mathbf{L}\,\mathbf{M}(\xi)^{-1}\right)$ *with respect ot w and x, so we get different design points and weights for the both designs, so*

$$\xi_A^* = \begin{pmatrix} -1 & -x & x & 1 \\ \frac{w}{2} & \frac{1-w}{2} & \frac{1-w}{2} & \frac{w}{2} \end{pmatrix}, \; x \simeq 0.46, \;\; w \simeq 0.3$$

$$\xi_{IMSE}^* = \begin{pmatrix} -1 & -x & x & 1 \\ \frac{w}{2} & \frac{1-w}{2} & \frac{1-w}{2} & \frac{w}{2} \end{pmatrix}, \; x \simeq 0.44, \;\; w \simeq 0.31$$

where the weight matrix has the following form for this model

$$\mathbf{L} = \begin{pmatrix} 2 & 0 & 2/3 & 0 \\ 0 & 2/3 & 0 & 2/5 \\ 2/3 & 0 & 2/5 & 0 \\ 0 & 2/5 & 0 & 2/7 \end{pmatrix}$$

Example 2.3. *For the quadratic regression model*

$$Y_i(x_i) = \beta_1 + \beta_2 x_i + \beta_3 x_i^2 + \varepsilon_i, \;\; x_i \in \mathcal{X} = [-1,1]$$

the design

$$\xi_{D;G}^* = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

satisfies the conditions of the corresponding equivalence theorem 2.2 and therewith it is D- resp. G-optimal design for this model, and that is illustrated by the figure 2.2 for the related following sensitivity function resp. the variance function

$$\varphi_D(x;\xi_D^*) = \mathbf{f}(x)^\top \mathbf{M}^{-1}(\xi_{D;G}^*)\mathbf{f}(x) = \frac{3}{2}(2 - 3x^2 + 3x^4)$$

The next design is A- and IMSE-optimal

$$\xi_{A;IMSE}^* = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

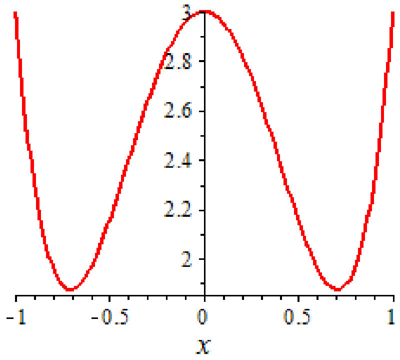

Figure 2.2: The D-sensitivity: $\frac{3}{2}(2 - 3x^2 + 3x^4)$

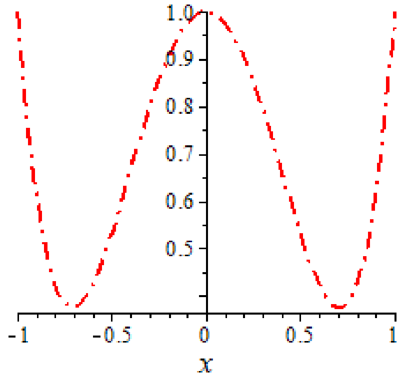

Figure 2.3: A-optimality: $1 - \frac{5}{2}x^2(1 - x^2)$

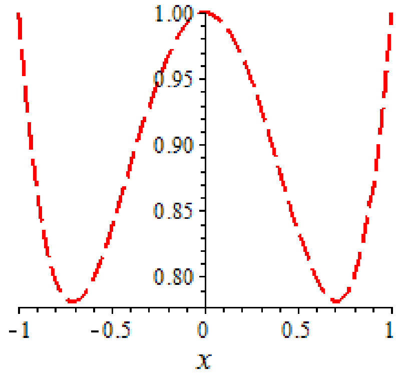

Figure 2.4: IMSE-optimality: $1 - \frac{7}{8}x^2(1 - x^2)$

satisfies the conditions of the corresponding equivalence theorem 2.3 and therewith it is A-optimal and that is illustrated by figure 2.3 for the function

$$\frac{\varphi_A(x;\xi_A^*)}{\operatorname{trace}\left(\ \mathbf{M}(\xi_A^*)^{-1}\right)} = \frac{\mathbf{f}(x)^\top \mathbf{M}^{-2}(\xi_A^*)\mathbf{f}(x)}{\operatorname{trace}\left(\ \mathbf{M}(\xi_A^*)^{-1}\right)} = 1 - \frac{5}{2}x^2(1 - x^2)$$

The weight matrix for this model due to the IMSE-optimality has the following form

$$\mathbf{L} = \begin{pmatrix} 2 & 0 & 2/3 \\ 0 & 2/3 & 0 \\ 2/3 & 0 & 2/5 \end{pmatrix}$$

figure 2.4 for the next function illustrate, that the candidate design satisfies the conditions of the equivalence theorem for the IMSE-optimality

$$\frac{\varphi_{IMSE}(x;\xi_{IMSE}^*)}{\operatorname{trace}\left(\mathbf{L}\ \mathbf{M}(\xi_{IMSE}^*)^{-1}\right)} = \frac{\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi_{IMSE}^*)^{-1}\mathbf{L}\ \mathbf{M}(\xi_{IMSE}^*)^{-1}\mathbf{f}(\mathbf{x})}{\operatorname{trace}\left(\mathbf{L}\ \mathbf{M}(\xi_{IMSE}^*)^{-1}\right)} = 1 - \frac{7}{8}x^2(1 - x^2)$$

By similar process we can illustrate, that the D- resp. G-optimal, A-optimal and IMSE-optimal designs for the same model in the design region $\mathcal{X} = [0, 1]$ are given as follows

$$\xi^* = \begin{pmatrix} 0 & \frac{1}{2} & 1 \\ w_0 & w_{\frac{1}{2}} & w_1 \end{pmatrix}$$

Where the weights $w_0 = w_{\frac{1}{2}} = w_1 = \frac{1}{3}$ for the D- resp. G-optimality, $w_0 = .322, w_{\frac{1}{2}} = 0.486, w_1 = 0.192$ for the A-optimality and $w_0 = w_1 = \frac{1}{4}, w_{\frac{1}{2}} = \frac{1}{2}$ for the IMSE-optimality.

2.3 The Univariate Linear Multi-Factor Regression Models

The most famous types of these models are the multiplicative resp. product type and additive models, where interactions between the different factors of the model exist through the multiplicative models, which do not exist for the additive models. May be under the multi-factor regression models can be understood, that many factors influence the response variable, which can be directly illustrated by the variance analysis models but within regression models meaning, that there is more than one input resp. control variable, which can belong to different experimental or design regions, effecting the response variable may be, with or without interactions. For example modeling the influence of two fever reducers with or without interactions between each other by the temperature of a patient. We will define the multiplicative models as well as the additive models as in *(Schwabe (1996))*.

Remark 2.3. *It is assumed in this section, that $E(\varepsilon_i) = 0$, $Var(\varepsilon_i) = \sigma^2$, where $\sigma > 0$ and is known, and $Cov(\varepsilon_i, \varepsilon_k) = 0; i \neq k$, $x_j \in \mathcal{X}_j$, $j = 1, ..., m$.*

Def 2.10 (The Multiplicative Models)**.** (Schwabe (1996)*) The multiplicative regression model can be defined as a regression model with more than one control variable, when its regression is the Kronecker product of two or more marginal regression functions. The m marginal single factor models are described by the corresponding marginal response functions as follows*

$$\eta_{\ j}(x_j) = \mathbf{f}^{(j)}(x_j)^\top \boldsymbol{\beta}^{(j)}$$

where $\mathbf{f}^{(j)}(x_j) : \mathcal{X}_j \longrightarrow \Re^{p_j}$, $\boldsymbol{\beta}^{(j)} \in \Re^{p_j}$, $j = 1, ..., m$. The response function for the resulting multiplicative model with j-factor is given as follows

$$\eta\ (x_1, ..., x_m) = \sum_{k_1=1}^{p_1} ... \sum_{k_m=1}^{p_m} \mathbf{f}_{k_1}^{(1)}(x_1)...\mathbf{f}_{k_m}^{(m)}(x_m)\boldsymbol{\beta}_{k_1,...,k_m} \tag{2.3}$$

The form 2.3 may be rewritten in a more intelligible way by use of the notation of Kronecker products, for the more informations about the Kronecker product see *(Zhang (1999)*)

$$\eta\ (x_1, ..., x_m) = (\mathbf{f}^{(1)}(x_1)^\top \otimes ... \otimes \mathbf{f}^{(m)}(x_m)^\top)\boldsymbol{\beta} = (\otimes_{j=1}^{m} \mathbf{f}^{(j)}(x_j)^\top)\boldsymbol{\beta} \tag{2.4}$$

where $\boldsymbol{\beta} \in \Re^p$ with $p = \prod_{j=1}^{m} p_j$. Hence $\boldsymbol{\beta}$ collects the unknown parameters $\boldsymbol{\beta}_{k_1,...,k_m}$, $k_j = 1, ..., p_j$, $j = 1, ..., m$, in lexicographic order $\boldsymbol{\beta} = (\boldsymbol{\beta}_{1,...,1,1}, \boldsymbol{\beta}_{1,...,1,2}, ..., \boldsymbol{\beta}_{1,...,1,p_j}, \boldsymbol{\beta}_{1,...,2,1}, ..., \boldsymbol{\beta}_{p_1,...,p_{j-1},p_j})$ and the response function (2.4) is parametrized by the regression function $\mathbf{f} : \mathcal{X}_1 \times \mathcal{X}_2 ... \times \mathcal{X}_m \longrightarrow \Re^p$ with $\mathbf{f}(x_1, ..., x_m) = \otimes_{j=1}^{m} \mathbf{f}^{(j)}(x_j)$. So the multiplicative regression model for one observation is specified as follows

$$Y(x_1, ..., x_m) = (\otimes_{j=1}^{m} \mathbf{f}^{(j)}(x_j)^\top)\boldsymbol{\beta} + \varepsilon \tag{2.5}$$

Def 2.11 (The Additive Models)**.** (Schwabe (1996)*) The additive regression model can be defined as a regression model with an intercept, which its regression is the intercept plus the union of two or more regression functions. With this restriction the response function of the additive model with m-factor is given by*

$$\eta\ (x_1, ..., x_m) = \mathbf{f}(x_1, ..., x_m)^\top \boldsymbol{\beta} = \beta_0 + \sum_{j=1}^{m} \mathbf{g}_j(x_j)^\top \boldsymbol{\beta}_j$$

$\mathbf{x} = (x_1, ..., x_m) \in \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 ... \times \mathcal{X}_m$, such that $\mathbf{f}(x_1, ..., x_m)^\top = (1, \mathbf{g}_1(x_1)^\top, ..., \mathbf{g}_m(x_m)^\top)^\top$, $\boldsymbol{\beta}_j \in \Re^{p_j - 1}$, and $\boldsymbol{\beta} = (\beta_0, \boldsymbol{\beta}_1^\top, ..., \boldsymbol{\beta}_m^\top) \in \Re^p$ with $p = \sum_{j=1}^m p_j - m + 1$. The corresponding regression functions for the marginal models have the forms $\mathbf{f}^{(j)}(x_j) = \begin{pmatrix} 1 \\ \mathbf{g}_j(x_j) \end{pmatrix}$, and response functions are given as follows

$$\eta_{\ j}(x_j) = \mathbf{f}^{(j)}(x_j)^\top \boldsymbol{\beta}^{(j)} = \beta_0 + \mathbf{g}_j(x_j)^\top \boldsymbol{\beta}_j,\ \boldsymbol{\beta}^{(j)} = (\beta_0, \boldsymbol{\beta}_j^\top)^\top \in \Re^{p_j}$$

So the additive regression model for one observation is specified as follows

$$Y(x_1, ..., x_m) = \beta_0 + \sum_{j=1}^{m} \mathbf{g}_j(x_j)^\top \boldsymbol{\beta}_j + \varepsilon_i \tag{2.6}$$

2.3.1 Optimal Designs for Multi-Factor Models

Optimal designs for multi-factor models were extensively explored and determined in the work of *(Schwabe (1996)*) and it has been proven, that the product type designs are D-optimal for the multiplicative and additive models, and linear optimal for multiplicative models without conditions and for additive models by block-diagonal information matrices. That means, we can reduce the problem of finding optimal designs for a multi-factor model on the problem of finding optimal designs for the corresponding one-factor models, which reformulate the multi-factor models.

Def 2.12 (The Product Type Design). *The support points for the product type design is the Cartesian product of the support points of approximate designs and its weights are the product of their weights, so it means that, for $\xi_1, \xi_2, \ldots, \xi_m$, which are defined as follows*

$$\xi_1 = \begin{pmatrix} x_{11} & \ldots & x_{1N_1} \\ w_{11} & \ldots & w_{1N_1} \end{pmatrix},\ x_{1i} \in \mathcal{X}_1,\ i_1 = 1, \ldots, N_1$$

$$\xi_2 = \begin{pmatrix} x_{21} & \ldots & x_{2N_2} \\ w_{21} & \ldots & w_{2N_2} \end{pmatrix},\ x_{2i} \in \mathcal{X}_2,\ i_2 = 1, \ldots, N_2$$

$$\vdots$$

$$\xi_m = \begin{pmatrix} x_{m1} & \ldots & x_{mN_m} \\ w_{m1} & \ldots & w_{mN_m} \end{pmatrix},\ x_{mi} \in \mathcal{X}_m,\ i_m = 1, \ldots, N_m$$

The product type design has the form

$$\xi = \xi_1 \otimes \xi_2 \otimes \ldots \otimes \xi_m = \begin{pmatrix} (x_{1i_1}, x_{2i_2}, \ldots, x_{mi_m}) \\ w_{1i_1} \cdot w_{2i_2} \cdot \ldots \cdot w_{mi_m} \end{pmatrix}_{i_1=1,\ldots,N_1, i_2=1,\ldots,N_2, i_m=1,\ldots,N_m}$$

Where $\otimes$ denotes the product measure operator and $(x_{1i_1}, x_{2i_2}, \ldots, x_{mi_m}) \in \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_m$, i.e. the product type design as measure ξ splits as product of the measures $\xi_1, \ldots, \xi_m$.

Remark 2.4. *The most important property of the product type design resp. the product measure is the topic of Fubini's theorem, which mainly emphasizes, that the computing of a double integral for an integrable function can be evaluated by iterated integrals and the order of integration can be changed, as follows*

$$\int_{\mathcal{X}_1 \times \mathcal{X}_2} f \mathrm{d}(\xi_1 \otimes \xi_2) = \int_{\mathcal{X}_2} \left(\int_{\mathcal{X}_1} f(x_1, x_2) \xi_1(\mathrm{d}x_1) \right) \xi_2(\mathrm{d}x_2) = \int_{\mathcal{X}_1} \left(\int_{\mathcal{X}_2} f(x_1, x_2) \xi_2(\mathrm{d}x_2) \right) \xi_1(\mathrm{d}x_1) \tag{2.7}$$

If $f(x_1, x_2) = f(x_1)f(x_2)$, then the form (2.7) can be reduced as follows

$$\int_{\mathcal{X}_1\times\mathcal{X}_2} f\mathrm{d}(\xi_1 \otimes \xi_2) = \int_{\mathcal{X}_1} f(x_1)\xi_1(\mathrm{d}x_1) \int_{\mathcal{X}_2} f(x_2)\xi_2(\mathrm{d}x_2) \tag{2.8}$$

For more information see for example *(Durrett (2010))* or *(Gaffke (2009/ 2010).*
Now some important results for optimal designs for multiplicative and additive models can be summarized by the following theorems. The considered linear criteria are the A- or the IMSE-criterion by the following theorems.

Theorem 2.4. (Schwabe (1996)*) Let ξ_j^* D- resp. G-optimal designs for the marginal model with the response function $\eta_{\ j}(x_j) = \mathbf{f}^{(j)}(x_j)^\top \boldsymbol{\beta}^{(j)}$ in the design region , then $\xi^* = \otimes_{j=1}^m \xi_j^*$ is D- resp. G-optimal design for the product-type model given in (2.5) in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$.*

Theorem 2.5. (Schwabe (1996)*) Let ξ_j^* linear optimal designs for the marginal model with the response function $\eta_{\ j}(x_j) = \mathbf{f}^{(j)}(x_j)^\top \boldsymbol{\beta}^{(j)}$ in the design region $\mathcal{X}_j$, with weight matrix $\mathbf{L}_j$, then $\xi^* = \otimes_{j=1}^m \xi_j^*$ is linear optimal design for the product-type model given in (2.5) in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$, with weight matrix $\mathbf{L} = \otimes_{j=1}^m \mathbf{L}_j$.*

Theorem 2.6. (Schwabe (1996)*) Let ξ_j^* D- resp. G-optimal designs designs for the marginal model with the response function $\eta_{\ j}(x_j) = \beta_0 + \mathbf{g}_j(x_j)^\top \boldsymbol{\beta}_j$ in the design region $\mathcal{X}_j$, then $\xi^* = \otimes_{j=1}^m \xi_j^*$ is D-resp. G-optimal design for the additive model given in (2.6) in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$.*

Theorem 2.7. (Schwabe (1996)*) Let ξ_j^* linear optimal designs for the marginal model with the response function $\eta_{\ j}(x_j) = \beta_0 + \mathbf{g}_j(x_j)^\top \boldsymbol{\beta}_j$ in the design region $\mathcal{X}_j$, with weight matrix $\mathbf{L}_j = \int_{\mathcal{X}} \mathbf{f}^{(j)\ \top} \mathbf{f}^{(j)} \mathrm{d}(\mu)$, and if $\int_{\mathcal{X}_j} \mathbf{g}_j(x_j)\ \xi_j^*(\mathrm{d}\mathbf{x}_j) = \mathbf{0},\ \ j = 2, \ldots, m$, then $\xi^* = \otimes_{j=1}^m \xi_j^*$ is linear optimal design for the additive model given in (2.6) in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$, with weight matrix*
$\mathbf{L} = \int (1, \mathbf{g}_1^\top, ..., \mathbf{g}_m^\top)^\top\ (1, \mathbf{g}_1^\top, ..., \mathbf{g}_m^\top)\mathrm{d}(\mu_m)$, where μ_m is the m-dimensional uniform measure.

Example 2.5. *We have two univariate linear one-factor quadratic regression models of the following form*

$$Y_{ij} = \beta_0 + \beta_{j1}x_{ij} + \beta_{j2}x_{ij}^2 + \varepsilon_i,\ j = 1, 2$$

Where the design regions are $\mathcal{X}_j = [-1, +1]$, thus from example 2.3 the D- resp. G-, A- and IMSE-optimal designs for those models are given as follows
$\xi_{1;D}^* = \xi_{2;D}^* = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$, $\xi_{1;A}^* = \xi_{2;A}^* = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$
$\xi_{1;IMSE}^* = \xi_{2;IMSE}^* = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$ *then for the multiplicative model*

$$Y_i = \beta_0 + \beta_1 x_{i2} + \beta_2 x_{i2}^2 + \beta_3 x_{i1} + \beta_4 x_{i1}x_{i2} + \beta_5 x_{i1}x_{i2}^2 + \beta_6 x_{i1}^2 + \beta_7 x_{i1}^2 x_{i2} + \beta_8 x_{i1}^2 x_{i2}^2 + \varepsilon_i,$$

the following product type designs are resp. D- resp. G-optimal and A- or IMSE-optimal designs because of theorems 2.4 and 2.5 respectively

$$\xi_D^* = \xi_{1;D}^* \otimes \xi_{2;D}^* = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} & \cdots & \begin{pmatrix} 1 \\ -1 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \begin{pmatrix} -1 \\ 0 \end{pmatrix} & \cdots & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \frac{1}{9} & \cdots & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \cdots & \frac{1}{9} \end{pmatrix}$$

$\xi^*_{A,IMSE} = \xi^*_{1;A,IMSE} \otimes \xi^*_{2;A,IMSE} =$

$$\left(\begin{pmatrix} 1 \\ 1 \\ \frac{1}{16} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ \frac{1}{16} \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ \frac{1}{16} \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ \frac{1}{16} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \frac{1}{4} \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ \frac{1}{8} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \frac{1}{8} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ \frac{1}{8} \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ \frac{1}{8} \end{pmatrix} \right)$$

where the weight matrix due to IMSE-optimality has the following form for this model

$$\mathbf{L} = \begin{pmatrix} 4 & 0 & \frac{4}{3} & 0 & 0 & 0 & \frac{4}{3} & 0 & \frac{4}{9} \\ 0 & \frac{4}{3} & 0 & 0 & 0 & 0 & 0 & \frac{4}{9} & 0 \\ \frac{4}{3} & 0 & \frac{4}{5} & 0 & 0 & 0 & \frac{4}{9} & 0 & \frac{4}{15} \\ 0 & 0 & 0 & \frac{4}{3} & 0 & \frac{4}{9} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{4}{9} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{4}{9} & 0 & \frac{4}{15} & 0 & 0 & 0 \\ \frac{4}{3} & 0 & \frac{4}{9} & 0 & 0 & 0 & \frac{4}{5} & 0 & \frac{4}{15} \\ 0 & \frac{4}{9} & 0 & 0 & 0 & 0 & 0 & \frac{4}{15} & 0 \\ \frac{4}{9} & 0 & \frac{4}{15} & 0 & 0 & 0 & \frac{4}{15} & 0 & \frac{4}{25} \end{pmatrix}$$

on $\mathcal{X} = \times_{j=1}^{2} \mathcal{X}_j = \times_{j=1}^{2}[-1, +1]$.
So, the conditions of the equivalence theorem for the D- resp. G-optimality are satisfied by evaluating $\xi^*_D = \xi^*_{1;D} \otimes \xi^*_{2;D}$ *as we can see by the figure 2.5 for the sensitivity function*

$$\varphi_D = \mathbf{f}(x_1, x_2)^\top \mathbf{M}^{-1}(\xi^*_D)\mathbf{f}(x_1, x_2) = \frac{81}{4}(\frac{2}{3} - x_1^2 + x_1^4)(\frac{2}{3} - x_2^2 + x_2^4) \leq 9$$

The conditions of the equivalence theorem for A-optimality are satisfied by evaluating $\xi^*_A = \xi^*_{1;A} \otimes \xi^*_{2;A}$*, as it is illustrated by figure 2.6 for the following function*

$$\frac{\varphi_A}{\text{trace}\left(\mathbf{M}(\xi^*_A)^{-1}\right)} = \frac{\mathbf{f}(x_1, x_2)^\top \mathbf{M}^{-2}(\xi^*_A)\mathbf{f}(x_1, x_2)}{\text{trace}\left(\mathbf{M}(\xi^*_A)^{-1}\right)} = \frac{25}{4}(\frac{2}{5} - x_1^2 + x_1^4)(\frac{2}{5} - x_2^2 + x_2^4) \leq 1$$

The conditions of the equivalence theorem for IMSE-optimality are satisfied by evaluating $\xi^*_{IMSE} = \xi^*_{1;IMSE} \otimes \xi^*_{2;IMSE}$*, as it is illustrated by figure 2.7 for the following function*

$$\frac{\varphi_{IMSE}}{\text{trace}\left(\mathbf{L}\ \mathbf{M}(\xi^*_{IMSE})^{-1}\right)} = \frac{\mathbf{f}(x_1, x_2)^\top \mathbf{M}^{-1}(\xi^*_{IMSE})\ \mathbf{L}\ \mathbf{M}^{-1}(\xi^*_{IMSE})\mathbf{f}(x_1, x_2)}{\text{trace}\left(\mathbf{L}\ \mathbf{M}^{-1}(\xi^*_{IMSE})\right)}$$
$$= \frac{49}{64}(\frac{8}{7} - x_1^2 + x_1^4)(\frac{8}{7} - x_2^2 + x_2^4) \leq 1$$

Example 2.6. *We have two univariate linear one-factor quadratic regression models of the following form*

$$Y_{ij} = \beta_0 + \beta_{j1}x_{ij} + \beta_{j2}x_{ij}^2 + \varepsilon_i,\ j = 1, 2$$

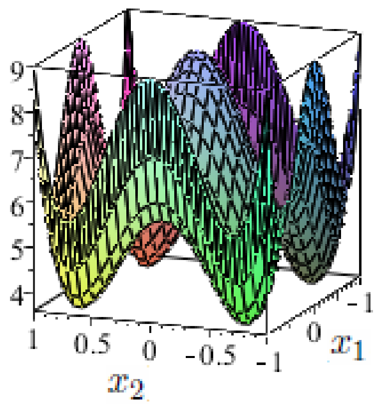

Figure 2.5: D-optimality: Multiplicative

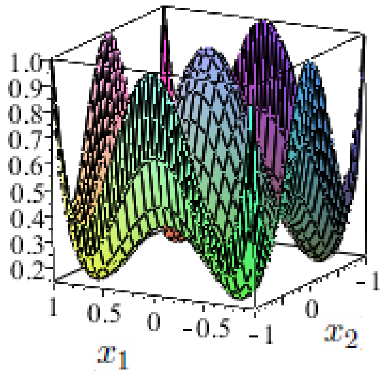

Figure 2.6: A-optimality : Multiplicative

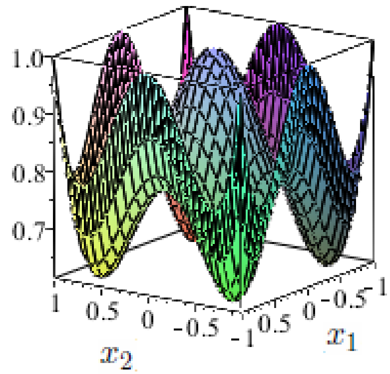

Figure 2.7: IMSE-optimality : Multiplicative

where the design regions are $\mathcal{X}_j = [-1, +1]$*, thus from example 2.3 the D- resp. G-, A- and IMSE-optimal designs for those models are given as follows*

$\xi^*_{1;D} = \xi^*_{2;D} = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$, $\xi^*_{1;A} = \xi^*_{2;A} = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$

$\xi^*_{1;IMSE} = \xi^*_{2;IMSE} = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$ *then for the additive model*

$$Y_i = \beta_0 + x_{i1}\beta_{11} + x_{i1}^2\beta_{12} + x_{i2}\beta_{21} + x_{i2}^2\beta_{22} + \varepsilon_i,$$

the following product type designs are D- resp. G-optimal designs but not A- or IMSE-optimal designs because of theorems 2.6 and 2.7

$$\xi^*_D = \xi^*_{1;D} \otimes \xi^*_{2;D} = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} & \cdots & \begin{pmatrix} 1 \\ -1 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \begin{pmatrix} -1 \\ 0 \end{pmatrix} & \cdots & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \frac{1}{9} & \cdots & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \cdots & \frac{1}{9} \end{pmatrix}$$

$$\xi_{A,IMSE} = \xi^*_{1;A,IMSE} \otimes \xi^*_{2;A,IMSE} =$$
$$\begin{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ -1 \end{pmatrix} & \begin{pmatrix} -1 \\ 1 \end{pmatrix} & \begin{pmatrix} -1 \\ -1 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \begin{pmatrix} -1 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 0 \\ -1 \end{pmatrix} \\ \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}$$

Where the weight matrix for this model due to IMSE-optimality has the form

$$\mathbf{L} = \begin{pmatrix} 4 & 0 & \frac{4}{3} & 0 & \frac{4}{3} \\ 0 & \frac{4}{3} & 0 & 0 & 0 \\ \frac{4}{3} & 0 & \frac{4}{5} & 0 & \frac{4}{9} \\ 0 & 0 & 0 & \frac{4}{3} & 0 \\ \frac{4}{3} & 0 & \frac{4}{9} & 0 & \frac{4}{5} \end{pmatrix}$$

So, the conditions of the equivalence theorem for D- resp. G-optimality are satisfied by evaluating the product type design $\xi_D^* = \xi_{1;D}^* \otimes \xi_{2;D}^*$, *as it is illustrated by figure 2.8 for the sensitivity function*

$$\varphi_D = \mathbf{f}(x_1, x_2)^\top \mathbf{M}^{-1}(\xi_D^*)\mathbf{f}(x_1, x_2) = 5 - \frac{9}{2}(x_1^2(1 - x_1^2) - x_2^2(1 - x_2^2)) \le 5$$

But the conditions of the conditions of the equivalence theorem for A-optimality are not satisfied by evaluating the product type design $\xi_{IMSE} = \xi_{1;IMSE}^* \otimes \xi_{2;IMSE}^*$, *as it is illustrated by figure 2.9 for the following function*

$$\frac{\varphi_A}{\text{trace}\left(\mathbf{M}(\xi_A^*)^{-1}\right)} = \frac{\mathbf{f}(x_1, x_2)^\top \mathbf{M}^{-2}(\xi_A)\mathbf{f}(x_1, x_2)}{\text{trace}\left(\mathbf{M}(\xi_A)^{-1}\right)} =$$
$$\frac{17}{15} + \frac{4}{3}(\cdot x_1^4 + x_2^4) - \frac{8}{5}(x_1^2 + x_2^2) + \frac{8}{15}x_1^2 x_2^2 > 1 \textit{ for } (x_1, x_2) = (0, 0)$$

and the conditions of the equivalence theorem for IMSE-optimality are not satisfied by evaluating the product type design $\xi_A = \xi_{1;A}^* \otimes \xi_{2;A}^*$, *as it is illustrated by the figure 2.10 for the following function*

$$\frac{\varphi_{IMSE}}{\text{trace}\left(\mathbf{LM}(\xi_{IMSE}^*)^{-1}\right)} = \frac{\mathbf{f}(x_1, x_2)^\top \mathbf{M}^{-1}(\xi_{IMSE})\ \mathbf{L}\ \mathbf{M}^{-1}(\xi_{IMSE})\mathbf{f}(x_1, x_2)}{\text{trace}\left(\mathbf{L}\ \mathbf{M}^{-1}(\xi_{IMSE})\right)}$$
$$= \frac{157}{147} + \frac{4}{7}(x_1^4 + x_2^4) - \frac{157}{147}(x_1^2 + x_2^2) + \frac{40}{147} \cdot x_1^2 x_2^2 > 1 \textit{ for } (x_1, x_2) = (0, 0)$$

It may be useful to mention, that the product type designs are linear optimal for additive models, which are reformulated from the simple linear regression models in the experimental region $\mathcal{X} = [-1, +1]$.

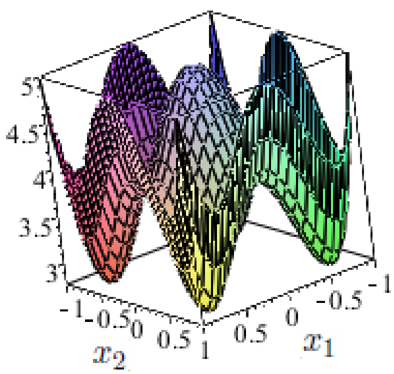

Figure 2.8: D-optimality : Additive

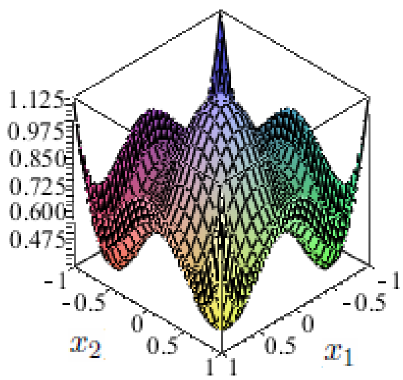

Figure 2.9: A-optimality : Additive

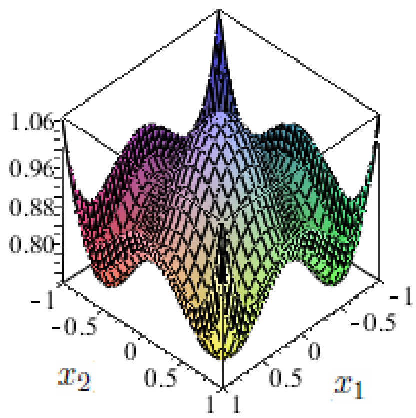

Figure 2.10: IMSE-optimality : Additive

3 The Seemingly Unrelated Regression (SUR) Models

The model of seemingly unrelated regression or the SUR model was introduced by *(Zellner (1962))* for modeling and analysis of a general multivariate linear regression model, and since then various types of it have played an important role in many areas of science. For more information and to have a deeper look into these kinds of models see *(Amemiya (1985))*, *(Hackl (2008))* and *(Srivastava and Giles (1987))*. These works can be also be useful, if one is interested in many resp. multi processes, which verify by the same subject and can model by many resp. multi equation models, where the time points resp. the control variables need not be identical for the measurements of the multi quantities within one subject. As the observations will be correlated within one unit, the data may be described by a multivariate linear model, which has the structure of a seemingly unrelated regression.
This chapter is organized as follows, the SUR model is interpreted in the first section, different estimators for the regression coefficients by known and unknown variance covariance matrix of the error variables, and the Fisher-information matrix are interpreted in the second section. The correlation matrix with some of its properties is interpreted in the third section. Other related multivariate linear models to the SUR models, such as the heterogeneous or homogeneous (MANOVA) multivariate linear models are illustrated in the fourth section . The Pharmacokinetic and Pharmacodynamic processes are modeled as a bivariate SUR model in the last section.

3.1 Model specification

This model is based on m-dimensional multivariate observations for n individuals. The components of the multivariate observations can be heterogeneous, which means that the response can be described by different regression functions and different experimental settings, which may be chosen from different experimental regions. Then the observation of the j-th component of individual i can be described by

$$Y_{ij} = \mathbf{f}_j(x_{ij})^\top \boldsymbol{\beta}_j + \varepsilon_{ij} = \sum_{l=1}^{p_j} f_{jl}(x_{ij})\beta_{jl} + \varepsilon_{ij},\ j = 1, ..., m$$

where $\mathbf{f}_j = (f_{j1}, ..., f_{jp_j})^\top$ are the known regression functions and $\boldsymbol{\beta}_j = (\beta_{j1}, ..., \beta_{jp_j})^\top$ the unknown parameter vectors for the j-th component and the experimental setting x_{ij} may be chosen from an experimental region $\mathcal{X}_j$. Denoted by $\mathbf{Y}_i = (Y_{i1}, ..., Y_{im})^\top$ and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, ..., \varepsilon_{im})^\top$ the multivariate vectors of observations and error terms, respectively, for individual i and correspondingly the block diagonal multivariate regression function

$$\mathbf{f}(\mathbf{x}) = \mathrm{diag}\left(\mathbf{f}_j(x_j)\right)_{j=1,...,m} = \begin{pmatrix} \mathbf{f}_1(x_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x_m) \end{pmatrix} \tag{3.1}$$

for the multivariate experimental setting $\mathbf{x} = (x_1, ..., x_m)^\top \in \mathcal{X} = \times_{j=1}^m \mathcal{X}_j$, which is different from the setting in the last chapter. Then the individual observation vector can be written as

$$\mathbf{Y}_i = \mathbf{f}(\mathbf{x}_i)^\top \boldsymbol{\beta} + \boldsymbol{\varepsilon}_i\,, \tag{3.2}$$

where $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, ..., \boldsymbol{\beta}_m^\top)^\top$ is the complete stacked parameter vector for all components. For the error vectors $\boldsymbol{\varepsilon}_i$ it is assumed that they have zero mean, are uncorrelated across the individuals and that they have a common positive definite variance covariance matrix of the error variables $\mathrm{Cov}\,(\boldsymbol{\varepsilon}_i) = \boldsymbol{\Sigma}$, and therefor $\mathrm{Cov}\,(\mathbf{Y}_i) = \boldsymbol{\Sigma}$, within the individuals, and their correlation components are heterogeneous.
Finally, denoted by $\mathbf{Y} = (\mathbf{Y}_1^\top, ..., \mathbf{Y}_n^\top)^\top$ and $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1^\top, ..., \boldsymbol{\varepsilon}_n^\top)^\top$ the stacked vectors of all observations and all error terms, respectively. Then we can write the complete observation vector as

$$\mathbf{Y} = \mathbf{F}\boldsymbol{\beta} + \varepsilon\,, \tag{3.3}$$

where $\mathbf{F} = (\mathbf{f}(\mathbf{x}_1), ..., \mathbf{f}(\mathbf{x}_n))^\top$ is the complete experiment design matrix. The complete observational error $\boldsymbol{\varepsilon}$, then its covariance matrix is $\mathbf{V} = \mathrm{Cov}(\boldsymbol{\varepsilon}) = \mathbf{I}_n \otimes \boldsymbol{\Sigma}$, where $\mathbf{I}_n$ is the $n \times n$ identity matrix.

Remark 3.1. *The univariate marginal models of the components have the following form*

$$\mathbf{Y}^{(j)} = \mathbf{F}^{(j)}\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}^{(j)}\,,\ j = 1, ..., m \tag{3.4}$$

where $\mathbf{Y}^{(j)} = (Y_{1j}, ..., Y_{nj})^\top$ and $\boldsymbol{\varepsilon}^{(j)} = (\varepsilon_{1j}, ..., \varepsilon_{nj})^\top$ are vectors of observations and errors for the j-th component, respectively, and $\mathbf{F}^{(j)} = (\mathbf{f}_j(x_{1j}), ..., \mathbf{f}_j(x_{nj}))^\top$ is the design matrix for the j-th marginal model. The corresponding error terms are uncorrelated and homoscedastic, $Cov\,(\boldsymbol{\varepsilon}^{(j)}) = \sigma_j^2\mathbf{I}_n$, where $\sigma_j^2 = \sigma_{jj}$ is the j-th diagonal entry of $\boldsymbol{\Sigma}$.

For more information about writing multivariate models in the individual or component approaches see *(Muller and Stewart (2006))*.

3.2 Estimating the parameter $\boldsymbol{\beta}$ with known and unknown variance covariance matrix

If we assume the variance covariance matrix of the error $\boldsymbol{\Sigma}$ and, hence $\mathbf{V}$ are known, we can estimate the parameter $\boldsymbol{\beta}$ efficiently by the Gauß-Markov estimator

$$\hat{\boldsymbol{\beta}}_{\mathrm{GM}} = (\mathbf{F}^\top\mathbf{V}^{-1}\mathbf{F})^{-1}\mathbf{F}^\top\mathbf{V}^{-1}\mathbf{Y}$$

where its variance covariance matrix is equal to the inverse of the corresponding following information matrix

$$\mathbf{M} = \mathbf{F}^\top\mathbf{V}^{-1}\mathbf{F} = \sum_{i=1}^{n} \mathbf{f}(\mathbf{x}_i)\boldsymbol{\Sigma}^{-1}\mathbf{f}(\mathbf{x}_i)^\top \tag{3.5}$$

which is the sum of the individual informations.
Alternatively, we can estimate the parameter $\boldsymbol{\beta}$ with less efficiency by the ordinary least squares estimator even in the case of unknown variance covariance matrix of the error variables

$$\hat{\boldsymbol{\beta}}_{\mathrm{OLS}} = \left(\mathbf{F}^\top\mathbf{F}\right)^{-1}\mathbf{F}^\top\mathbf{Y}$$

Which is minimizes the least squares

$$(\mathbf{Y} - \mathbf{F}\boldsymbol{\beta})^\top\,(\mathbf{Y} - \mathbf{F}\boldsymbol{\beta})$$

The corresponding variance covariance matrix for this estimator has the following form

$$\mathbf{cov}(\hat{\boldsymbol{\beta}}_{\mathrm{OLS}}) = \left(\mathbf{F}^\top\mathbf{F}\right)^{-1}\left(\mathbf{F}^\top\mathbf{V}\mathbf{F}\right)\left(\mathbf{F}^\top\mathbf{F}\right)^{-1}$$

Remark 3.2. *If* $\mathbf{V} \cdot \mathbf{F} = \mathbf{F} \cdot \mathbf{U}$*, for some* $\mathbf{U}$*, then the estimators* $\hat{\boldsymbol{\beta}}_{\mathrm{GM}}$ *and* $\hat{\boldsymbol{\beta}}_{\mathrm{OLS}}$ *are identical. This is valid at least for uncorrelated components, i.e. when the variance covariance matrix of the error* $\boldsymbol{\Sigma}$ *is diagonal, and for identical* $(\mathbf{f}_j(x_j))_{j=1,\ldots,m}$*, i.e. when we have multivariate regression model (MANOVA). For more informations see for example* (Baltagi (2011).

The general least squares estimator is given as follows

$$\hat{\boldsymbol{\beta}}_{\mathrm{GLS}} = (\mathbf{F}^\top \mathbf{V}^{-1} \mathbf{F})^{-1} \mathbf{F}^\top \mathbf{V}^{-1} \mathbf{Y}$$

Which minimizes the weighted least squares, resp. the following term

$$(\mathbf{Y} - \mathbf{F}\boldsymbol{\beta})^\top \; \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{F}\boldsymbol{\beta})$$

Remark 3.3. *The goal of the following discussion is to show, that the asymptotic information matrix is block diagonal with unknown variance covariance matrix, which is useful for a result in the next chapter.*

We can estimate the parameter $\boldsymbol{\beta}$ asymptotically efficient for unknown variance covariance matrix by the feasible general least squares estimator (FGLS) through replacing the variance covariance matrix of the error variables with the following estimator

$$\hat{V} = \frac{1}{n} (\mathbf{Y} - \mathbf{F}\hat{\boldsymbol{\beta}}_{OLS})^\top \; (\mathbf{Y} - \mathbf{F}\hat{\boldsymbol{\beta}}_{OLS}) \tag{3.6}$$

and estimating the parameter$\boldsymbol{\beta}$ by

$$\hat{\boldsymbol{\beta}}_{\mathrm{FGLS}} = (\mathbf{F}^\top \hat{\mathbf{V}}^{-1} \mathbf{F})^{-1} \mathbf{F}^\top \hat{\mathbf{V}}^{-1} \mathbf{Y} \tag{3.7}$$

see *(Magnus (1978))*, *(Amemiya (1985))*, or *(Srivastava and Giles (1987))*. The parameter $\boldsymbol{\beta}$ can be estimated too by the maximum likelihood method, when the error components underlie the normal distribution through replacing the variance covariance matrix of the error variables in the log-likelihood function with the following term

$$\mathbf{V} = \frac{1}{n} (\mathbf{Y} - \mathbf{F}\boldsymbol{\beta})^\top \; (\mathbf{Y} - \mathbf{F}\boldsymbol{\beta})$$

and maximizing it for the parameter $\boldsymbol{\beta}$ and then replacing it with $\hat{\boldsymbol{\beta}}_{ML}$ in (3.8) and therewith estimating $\mathbf{V}$ renewed. Where the likelihood function has the following form

$$\ell \; = (2\pi)^{-\frac{nm}{2}} \; (\det \mathbf{V}^{-\frac{1}{2}}) \; \exp\left(-\frac{1}{2} (\mathbf{Y} - \mathbf{F}\boldsymbol{\beta})^\top \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{F}\boldsymbol{\beta})\right)$$

And respectively the log-likelihood function as follows

$$\log \ell = -\frac{nm}{2} \log(2\pi) \; - \frac{1}{2} \log\det \mathbf{V} \; - \frac{1}{2} ((\mathbf{Y} - \mathbf{F}\boldsymbol{\beta})^\top \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{F}\boldsymbol{\beta})) \tag{3.8}$$

When the variance covariance matrix of the error variables is known, then $\hat{\boldsymbol{\beta}}_{ML}$, $\hat{\boldsymbol{\beta}}_{GLS}$ and $\boldsymbol{\beta}_{GM}$ coincide with each other. The Use of $\hat{\boldsymbol{\beta}}_{ML}$ and $\hat{\boldsymbol{\beta}}_{FGLS}$ have the advantage of asymptomatic efficiency for big n

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \sim N(\mathbf{0}, \mathbf{M}_{\boldsymbol{\theta}}^{-1}), \; \boldsymbol{\theta} = (\boldsymbol{\beta}; \mathbf{V})$$

Where $\mathbf{M}_{\boldsymbol{\theta}}$ is the Fisher information matrix, which exists, is finite, regulare and can be calculated by the expression

$$\mathrm{E}\left(\frac{\partial \log \ell}{\partial \boldsymbol{\theta}}.\frac{\partial \log \ell}{\partial \boldsymbol{\theta}}^{\top}\right), \text{ and its corresponding elements by } \mathrm{E}\left(\frac{\partial \log \ell}{\partial \theta_{\jmath}}.\frac{\partial \log \ell}{\partial \theta_q}^{\top}\right)$$

Where $\jmath, q = 1, ..., r$ and $r = p + m(m+1)/2$ is the dimension of the Fisher information matrix, respectively the dimension of the parameter $\boldsymbol{\theta}$. The Fisher information matrix has the following form for the SUR model for $\boldsymbol{\nu} = \mathrm{vec}(\boldsymbol{\Sigma}) = \mathrm{vec}((u_{j\acute{j}})_{j,\acute{j}=1,...,m})$, where the vec$(\boldsymbol{\Sigma})$ is a column vector of the columns of $\boldsymbol{\Sigma}$.

$$\mathbf{M}_{\boldsymbol{\beta};\mathbf{V}} = \begin{pmatrix} \mathbf{M}_{\boldsymbol{\beta}} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\mathbf{V}} \end{pmatrix};\ \mathbf{M}_{\boldsymbol{\beta}} = \mathbf{F}^{\top}\mathbf{V}^{-1}\mathbf{F};\ \mathbf{M}_{\mathbf{V}} = \frac{1}{2}\frac{\partial \mathrm{vec}\boldsymbol{\Sigma}}{\partial \boldsymbol{\nu}}^{\top}\left(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\right)\frac{\partial \mathrm{vec}\boldsymbol{\Sigma}}{\partial \boldsymbol{\nu}} \tag{3.9}$$

Where $\mathbf{V} = \mathbf{I}_{n\times n}\otimes\boldsymbol{\Sigma}$.
The $\jmath q th$ elements of the matrix $\mathbf{M}_{\mathbf{V}}$ has the following forms because of symmetry

$$\left(\mathrm{vec}\frac{\partial \boldsymbol{\Sigma}}{\partial u_{\jmath}}\right)^{\top}\left(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\right)\left(\mathrm{vec}\frac{\partial \boldsymbol{\Sigma}}{\partial u_q}\right) = \mathrm{trace}\left(\frac{\partial \boldsymbol{\Sigma}}{\partial u_{\jmath}}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\delta u_q}\boldsymbol{\Sigma}^{-1}\right)$$

Where u interprets the different terms of the variances and correlations of the variance covariance matrix of the error variables.
We can calculate the elements of the Fisher-information matrix for $\boldsymbol{\eta} = \mathbf{F}\boldsymbol{\beta}$ by the next simple form

$$(\mathbf{M}_{\boldsymbol{\theta}})_{\jmath q} = \mathrm{E}\left(\frac{\partial \boldsymbol{\eta}_{\boldsymbol{\theta}}}{\partial \theta_{\jmath}}^{\top}\mathbf{V}^{-1}.\frac{\partial \boldsymbol{\eta}_{\boldsymbol{\theta}}}{\partial \theta_q}\right) + \frac{1}{2}\mathrm{trace}\left(\frac{\partial \boldsymbol{\Sigma}}{\partial u_{\jmath}}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\delta u_q}\boldsymbol{\Sigma}^{-1}\right)$$

Similarly, the Fisher information matrix for one individual model (3.2) looks as follows

$$\mathbf{M}_{\boldsymbol{\theta}} = \begin{pmatrix} \mathbf{M}_{\boldsymbol{\beta}} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\mathbf{V}} \end{pmatrix};\ \mathbf{M}_{\boldsymbol{\beta}} = \sum_{i=1}^{n}\mathbf{f}(\mathbf{x}_i)\boldsymbol{\Sigma}^{-1}\mathbf{f}(\mathbf{x}_i)^{\top};\ \mathbf{M}_{\mathbf{V}} = \frac{1}{2}\frac{\partial \mathrm{vec}\boldsymbol{\Sigma}}{\partial \boldsymbol{\nu}}^{\top}\left(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\right)\frac{\partial \mathrm{vec}\boldsymbol{\Sigma}}{\partial \boldsymbol{\nu}} \tag{3.10}$$

For more information see *(Turkington (2002))*, *(Rao (1973))*, *(Magnus (1978))*, *(Amemiya (1985))*, or *(Srivastava and Giles (1987))*.

3.3 The correlation matrix and its properties

Remark 3.4. *When the variance covariance matrix of the error variables* $\boldsymbol{\Sigma}$ *is given as follows*

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \cdots & \sigma_1\sigma_m\rho_{1m} \\ \vdots & \ddots & \vdots \\ \sigma_1\sigma_m\rho_{1m} & \cdots & \sigma_m^2 \end{pmatrix} \tag{3.11}$$

then we can deduce the corresponding correlation matrix from the variance covariance matrix of the error variables $\boldsymbol{\Sigma}$ *as follows*

$$\mathbf{cor} = diag\,(\frac{1}{\sigma_j})_{j=1,...,m}\ \boldsymbol{\Sigma}\ diag\,(\frac{1}{\sigma_j})_{j=1,...,m} = \begin{pmatrix} 1 & \cdots & \rho_{1m} \\ \vdots & \ddots & \vdots \\ \rho_{1m} & \cdots & 1 \end{pmatrix} \tag{3.12}$$

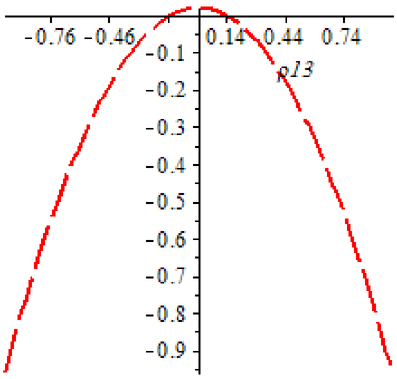

Figure 3.1: The determinant of the matrix $\mathbf{R}$ for $\rho_{12} = 0.99$ and $\rho_{23} = 0$.

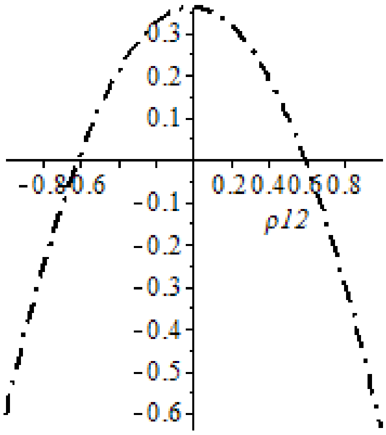

Figure 3.2: The determinant of the matrix $\mathbf{R}$ for$\rho_{13} = 0.01$ and $\rho_{23} = -0.8$.

To estimate the parameter by the Gauß-Markov estimator, the variance covariance matrix of the error variables should be regular, i.e. that all $\sigma_j^2 > 0$ and the correlation matrix positive definite

Theorem 3.1 (*(Zhang (1999))*, (6.2)). *Let* $\mathbf{H}$ *be an* $m-$*square symmetrical matrix. Then*

1. $\mathbf{H}$ *is positive definite if and only if the determinant of every leading principal submatrix (i.e., Minor) of* $\mathbf{H}$ *is positive.*

2. $\mathbf{H}$ *is positive definite if and only if the determinant of every (not only leading) principal submatrix of* $\mathbf{H}$ *is nonnegative.*

Theorem 3.2 (*(Peterson and Pederson (2008))*). *Let* $\mathbf{H}$ *be an* $m-$*square complex matrix. Then* $\mathbf{A}$ *is positive definite if and only if all eigenvalues of it,* λ *are positive, i.e.* $\lambda_i > 0$*, where* $i = 1 \ldots m$*.*

Example 3.1. *Let* $\mathbf{R} = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}_{3\times 3}$

then the determinant leading Minor are positive, if $\det(1) = 1 > 0$, $\det \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix} = 1 - \rho_{12}^2 > 0 \Rightarrow \rho_{12}^2 < 1$, $\det \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix} = 1 - \rho_{12}^2 - \rho_{13}^2 - \rho_{23}^2 + 2 \cdot \rho_{12} \cdot \rho_{13} \cdot \rho_{23} > 0$
Thus for $\rho_{13} = 0.01$*,* $\rho_{23} = -0.8$ *must be* $-0.61 < \rho_{12} < 0.59$ *and therewith the correlation matrix is positive definite. For* $\rho_{12} = 0.99$*,* $\rho_{23} = 0$*, if* $-0.14 \leq \rho_{13} \leq 0.14$ *then the matrix* $\mathbf{R}$ *is a correlation matrix. The next graphics illustrate the curves for the determinant of the correlation matrix*

Remark 3.5. *We can remark by the example of (3.1) that the determinant of the correlation matrix is less than or equal to the determinants for the corresponding submatrices with diagonal elements equal* *to* 1*, also, for* $c = \det(\mathbf{R})$*,* $c_{12} = \det\begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix} = 1 - \rho_{12}^2$*,* $c_{13} = \det\begin{pmatrix} 1 & \rho_{13} \\ \rho_{13} & 1 \end{pmatrix} = 1 - \rho_{13}^2$ *and* $c_{23} = \det\begin{pmatrix} 1 & \rho_{23} \\ \rho_{23} & 1 \end{pmatrix} = 1 - \rho_{23}^2$*, we have*

$$c = 1 - \rho_{12}^2 - \rho_{13}^2 - \rho_{23}^2 + 2 \cdot \rho_{12} \cdot \rho_{13} \cdot \rho_{23} = c_{12} + c_{13} + c_{23} - 2(1 - \rho_{12}\rho_{13}\rho_{23}) > 0 \text{ then}$$
$$0 < \frac{c}{c_{12}} = 1 + \frac{c_{13}}{c_{12}} + \frac{c_{23}}{c_{12}} - 2\frac{(1 - \rho_{12}\rho_{13}\rho_{23})}{c_{12}} \text{ then } \frac{c}{c_{12}} \leq \max(1 + \frac{c_{13}}{c_{12}} + \frac{c_{23}}{c_{12}} - 2\frac{(1 - \rho_{12}\rho_{13}\rho_{23})}{c_{12}}) \text{ then}$$
$$0 < \frac{c}{c_{12}} \leq 1 \text{ thus by the same method } 0 < \frac{c}{c_{13}} \leq 1, 0 < \frac{c}{c_{23}} \leq 1 \text{ where } \max \text{ is the maximum.} \tag{3.13}$$

That is valid in general.

3.4 The Derivation of Some Multivariate Linear Models from The SUR Models

The SUR models are general multivariate linear models with different regression functions and control variables for the marginal components see *(Reinsel and Velu (1998))*, so for different regression functions with the same control variables for the marginal components we get the so called heterogeneous multivariate linear models see *(Kurotschka and Schwabe (1996))* or *(Krafft and Schaefer (1992))*, thus the block diagonal multivariate regression function for such models has the special form of 6.1 in the following form

$$\mathbf{f}(x) = \text{diag}\left(\mathbf{f}_j(x)\right)_{j=1,\dots,m} = \begin{pmatrix} \mathbf{f}_1(x) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x) \end{pmatrix} \tag{3.14}$$

The so called homogeneous multivariate linear models or (MANOVA) can be held from the SUR model for same regression functions and control variables for the marginal components see *(Kurotschka and Schwabe (1996))*, *(Chang (1994))* or *(Christensen (2001))*, thus the block diagonal multivariate regression function for such models has the special form of 6.1 in the following form

$$\mathbf{f}(x) = \text{diag}\left(\mathbf{f}_0(x)\right)_{j=1,\dots,m} = \begin{pmatrix} \mathbf{f}_0(x) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_0(x) \end{pmatrix} = \mathbf{I}_m \otimes \mathbf{f}_0(x) \tag{3.15}$$

And it is known, that the OLS and Gauß estimator are identical for this multivariate linear model and the variance covariance matrices for both estimators are identical and have the following form

$$\mathbf{Cov} = \left(\mathbf{F}^\top \mathbf{V}^{-1} \mathbf{F}\right)^{-1} \text{ where } \left(\mathbf{F}^\top \mathbf{V}^{-1} \mathbf{F}\right)^{-1} =$$
$$\left((\mathbf{I}_m \otimes \mathbf{f}_0(x_1), ..., \mathbf{I}_m \otimes \mathbf{f}_0(x_n))(\mathbf{I}_n \otimes \Sigma^{-1})(\mathbf{I}_m \otimes \mathbf{f}_0(x_1), ..., \mathbf{I}_m \otimes \mathbf{f}_0(x_n))^\top\right)^{-1}$$
$$= \mathbf{\Sigma} \otimes \left(\sum_{i=1}^{n} \mathbf{f}_0(x_i)\mathbf{f}_0(x_i)^\top\right)^{-1} \quad \text{then}$$

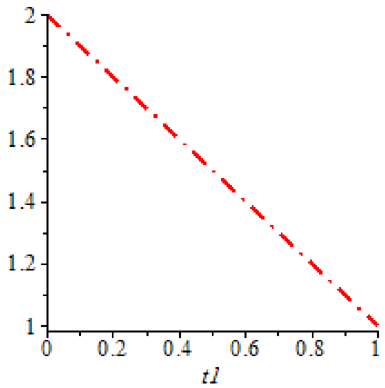

Figure 3.3: The logarithmic concentration

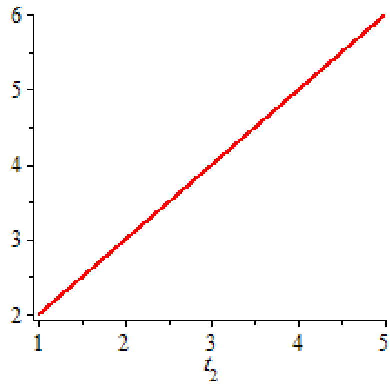

Figure 3.4: The logarithmic efficiency

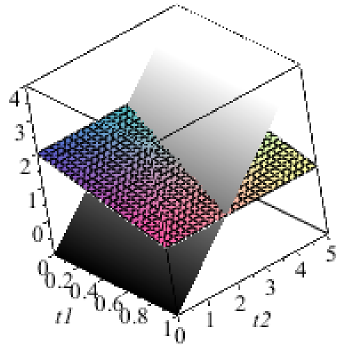

Figure 3.5: The both processes

$$\mathbf{Cov} = \mathbf{\Sigma} \otimes \mathbf{M}_0^{-1}(\mathbf{x}),\ \mathbf{M}_0(\mathbf{x}) = \sum_{i=1}^{n} \mathbf{f}_0(x_i)\mathbf{f}_0(x_i)^\top \tag{3.16}$$

and the information matrix has the next form

$$\mathbf{M}(\mathbf{x}) = \mathbf{\Sigma}^{-1} \otimes \mathbf{M}_0(\mathbf{x}) \tag{3.17}$$

3.5 The Modeling of Pharmacokinetic and Pharmacodynamic Processes as a SUR Model

If one is interested in both pharmacokinetics and pharmacodynamics, where the pharmacokinetics (PK) is the study of the influence of body on the drug resp. the study of the concentration and the pharmacodynamics (PD) is the study of the influence of the drug on the body resp. the efficiency of it, the time points need not be identical for the measurements of the two quantities within one subject. As the observations will be correlated within one unit, the data may be described by a bivariate model, which has the structure of a bivariate seemingly unrelated regression. A simple example for such modeling by one dose can be found in the work of *(Bertrand and Mentré (2008))* in the exponential form, thus after linearizing both equations by applying the logarithmic functions to them a bivariate SUR model or a bivariate straight line regression models with different control variables can be obtained, i.e. the model has the following form for n observations

$$Y_{ij} = \beta_{j0} + \beta_{j1}t_{ij} + \varepsilon_{ij}\ ,\ \operatorname{Cov}(\varepsilon_{ij}, \varepsilon_{ik}) = \rho_{jk}\sigma_j\sigma_k, j \neq k, j, k = 1, 2,\ j = 1 \to PK, j = 2 \to PD.$$

So if we assume that the estimated logarithmic predictions for PK is $\ln(\mathbf{Y}_1) = 2 - \mathbf{t}_1$ and for PD the form $\ln(\mathbf{Y}_2) = -1 + \mathbf{t}_2$, the both processes and the paradox of both marginal models are respectively illustrated by figures (3.3), (3.4) and (3.5).

4 Optimal Designs for SUR Models wih One-Factor and Intercepts in their Components

The optimal design theory in the multivariate case has been well developed, as *(Fedorov (1972))* established the equivalence theorem in the multivariate case for the D- and linear optimality. There were different works in this research area, so the D-optimality for a kind of heterogeneous multivariate linear model with grown hierarchical nesting components was explored in the work of *(Krafft and Schaefer (1992))*. And it is shown , that a jointly D-optimal design for all marginal components models is D-optimal design for the reformulated multivariate model. Where these D-optimal designs can be determined by a developed method in the work of *(Dette (1990))*. It has been found in the work of *(Chang (1994))*, that the D-optimal design's problem for the homogeneous multivariate linear models (MANOVA) can be reduced to their corresponding univariate marginal component models. I.e. the D-optimal design for one of the marginal components is the D-optimal design for the reformulated MANOVA-model by those components. Both these results have been generalized by the work of *(Kurotschka and Schwabe (1996))*, so the result of *(Chang (1994))* were extended for A-, C- and E-optimality. And another proof for the result of *(Krafft and Schaefer (1992))* were illustrated. The same result due to D-optimality was enhanced for heterogeneous multivariate linear models with different marginal components under strict conditions, which made the corresponding information matrix block-diagonal. There are different control variables for the considered SUR models and therewith the property (2.8) due to Fubini's theorem resp. the corresponding remark 2.4 are valid, and the strict conditions for the heterogeneous multivariate linear models are disappeared for SUR models with intercepts by the corresponding marginal models, and made the optimal design problem for SUR models much easier.
The optimal design theory in the multivariate case, as the IMSE-criterion are introduced in the first section. The D- and some linear optimal as the A- and IMSE-optimal designs for the considered SUR models are established in the second section. And it has been shown, that the product type designs are D-, A-, IMSE-optimal designs as well as linear optimal if the weight matrix is block diagonal(4.1) , for the SUR models with marginal components, which have one-factor resp. one control variable and an intercept by each component. These results are valid for a known variance covariance matrix of the error variables. The optimality of the product type designs is to be held but asymptoticly for an unknown variance covariance matrix of the error variables. That is the mean topic of this chapter. The results are illustrated by a clear example at the end of this chapter.

4.1 Optimal Design Theory in the Multivariate Case

We can define an experimental design in the multivariate case with support points as vectors as follows $\xi = \begin{pmatrix} \mathbf{x}_1 & \dots & \mathbf{x}_k \\ w_1 & \dots & w_k \end{pmatrix}$ by the set of all different experimental settings $\mathbf{x}_i = (x_{i1}, \dots, x_{im})$, $i = 1, ..., k$, which belong to the design region $\mathcal{X} = \times_{j=1}^{m} \mathcal{X}_j$, with the corresponding relative frequencies $w_i = \frac{k_i}{k}$, where k_i is the number of replications at $\mathbf{x}_i$. When $w_i, i = 1, ...k$ don't depend on the sample size k and are just real numbers between zero and one, then the exact and approximative design are not identical. The corresponding standardized information matrix resp. the variance covariance matrix for the Gauß estimator for the SUR models can be obtained as

$$\mathbf{M}(\xi) = \sum_{i=1}^{k} w_i \mathbf{f}(\mathbf{x}_i)\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x}_i)^\top, \ \mathbf{Cov}(\hat{\boldsymbol{\beta}}) = \mathbf{M}^{-1}(\xi).$$

and respectively the limited information matrix for the OLS estimator has the following form

$$\mathbf{M}_{OLS}(\xi) = \left(\sum_{i=1}^{k} w_i \mathbf{f}(\mathbf{x}_i)\mathbf{f}(\mathbf{x}_i)^\top\right)\left(\sum_{i=1}^{k} w_i \mathbf{f}(\mathbf{x}_i)\mathbf{\Sigma}\mathbf{f}(\mathbf{x}_i)^\top\right)^{-1}\left(\sum_{i=1}^{k} w_i \mathbf{f}(\mathbf{x}_i)\mathbf{f}(\mathbf{x}_i)^\top\right).$$

We consider approximate designs for analytical purposes , see for example *(Kiefer (1974)*, for which the weights $w_i \geq 0$ need not be multiples of $\frac{1}{n}$, but only have to satisfy $\sum_{i=1}^{k} w_i = 1$.

4.1.1 A-, D- and IMSE Criteria in the Multivariate Case

The A- and D-optimality criteria is defined similarly to the univariate case, as well as the IMSE-criterion, which can be illustrated by the following lemma

Def 4.1 (IMSE). *The integrated mean square error is the integrated predictive covariance with respect to the uniform measure $\mu(\mathrm{d}\mathbf{x})$ and defined as follows*

$$\mathbf{IMSE} = \int_{\mathcal{X}} trace\ (\mathbf{Cov}\left(\mathbf{f}(\mathbf{x})^\top\hat{\boldsymbol{\beta}}\right))\mu(\mathrm{d}\mathbf{x}) = \int_{\mathcal{X}} \mathbf{E}(\|\mathbf{f}(\mathbf{x})^\top(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\|^2)\mu(\mathrm{d}\mathbf{x})$$

Where $\|.\|$ denotes the euclidean norm.

Lemma 4.1. *An approximate design ξ^*_{IMSE} is called IMSE-optimal design in the multivariate case, if it minimizes the integrated predictive variance with respect to the uniform measure $\mu(\mathrm{d}x)$ and is equivalent to*

$$trace\ (\mathbf{L}\mathbf{M}^{-1}(\xi^*_{IMSE})) = \min_{\xi\in\Xi}[trace\ (\mathbf{L}\mathbf{M}^{-1}(\xi))], \ \mathbf{L} = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top\ \mu(\mathrm{d}\mathbf{x}) \tag{4.1}$$

Proof:

$$\begin{aligned}
\mathbf{IMSE} &= \int_{\mathcal{X}} \mathbf{E}(\|\mathbf{f}(\mathbf{x})^\top(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\|^2)\mu(\mathrm{d}\mathbf{x}) \\
&= \int_{\mathcal{X}} \mathbf{E}(\mathrm{trace}\left((\mathbf{f}(\mathbf{x})^\top(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}))^\top\mathbf{f}(\mathbf{x})^\top(\hat{\beta} - \beta)\right)\mu(\mathrm{d}\mathbf{x}) \\
&= \int_{\mathcal{X}} \mathbf{E}(\mathrm{trace}\left(\mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^\top\right))\mu(\mathrm{d}\mathbf{x}) \\
&= \int_{\mathcal{X}} \mathrm{trace}\left(\mathbf{E}[\mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top]\mathbf{Cov}(\hat{\boldsymbol{\beta}})\right)\mu(\mathrm{d}\mathbf{x}) \\
&= \int_{\mathcal{X}} \mathrm{trace}\left(\mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top\mathbf{M}(\xi)^{-1}\right)\mu(\mathrm{d}\mathbf{x}) \\
&= \mathrm{trace}\left(\mathbf{L}\ \mathbf{M}(\xi)^{-1}\right), \ \mathbf{L} = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top\ \mu(\mathrm{d}\mathbf{x})
\end{aligned}$$

Thus the IMSE-criterion in the multivariate case is given in (4.1).

4.1.2 The Equivalence Theorems for D- and Linear-Optimality in the Multivariate Case

Useful tools for checking the performance of a given candidate design are the multivariate equivalence theorems for the D- and linear criteria see *(Fedorov (1972))*), theorems 5.2.1, 5.3.1):

Theorem 4.1. *The approximate design ξ^* is D-optimal in the multivariate linear model if and only if*

$$\text{trace}\left(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top\mathbf{M}(\xi^*)^{-1}\mathbf{f}(\mathbf{x})\right) \leq p \tag{4.2}$$

for all $x \in \mathcal{X}$, where $p = \sum_{j=1}^{m} p_j$ is the number of parameters in the model.

Theorem 4.2. *The approximate design ξ^* is linear optimal in the multivariate linear model if and only if*

$$\varphi_L(x;\xi_L^*) = \text{trace}\left(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top\mathbf{M}(\xi^*)^{-1}\mathbf{L}\ \mathbf{M}(\xi^*)^{-1}\mathbf{f}(\mathbf{x})\right) \leq \text{trace}\left(\mathbf{L}\ \mathbf{M}(\xi^*)^{-1}\right) \tag{4.3}$$

for all $x \in \mathcal{X}$.

1. *For A-optimality $\mathbf{L} = \mathbf{I}_p$, where $\mathbf{I}_p$ is the $p \times p$ identity matrix.*
2. *For IMSE-optimality $\mathbf{L} = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})^\top\mathbf{f}(\mathbf{x})\mu(\mathrm{d}\mathbf{x})$.*

4.2 D- and Linear-Optimal Designs for SUR Models

Remark 4.1. *The considered linear optimal designs in this work are A-, IMSE- and special linear optimal designs by block diagonal form of the weight matrix $\mathbf{L}$, where its diagonal blocks are the weight matrices for the marginal components $\mathbf{L}_j, j = 1, ..., m$, i.e. $\mathbf{L} = \text{block}-\text{diag}\,(\mathbf{L}_j)_{j=1,\dots,m}$ in general. For A-optimality the weight matrix $\mathbf{L}$ is equal to the identity matrix with the dimension p $\mathbf{L} = \mathbf{I}_p = \text{block}-\text{diag}\,(\mathbf{I}_{p_j})_{j=1,\dots,m}$, $\mathbf{I}_{p_j}$ is the identity matrix with the dimension p_j, and for IMSE-optimality $\mathbf{L} = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})^\top\mathbf{f}(\mathbf{x})\mu(\mathrm{d}\mathbf{x}) = \text{block}-\text{diag}\,(\int_{\mathcal{X}_j} \mathbf{f}(x_j)^\top\mathbf{f}(x_j)\mu(\mathrm{d}x_j))_{j=1,\dots,m}$. The same meaning can be extended for the univariate case.*

To obtain a complete characterization of the D-optimal and linear optimal designs in 4.1 we have to require that all marginal models related to the components contain an intercept, $f_{j1}(x_j) \equiv 1$. And some auxiliary lemmas are to be introduced before the main results

Lemma 4.2. *The information matrix for one-point designs $\xi = \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}$ for the SUR model (3.3) has the following form*

$$\mathbf{M}(\mathbf{x}) = \left[\sigma^{(kj)}\mathbf{f}_k(x_k)\mathbf{f}_j(x_j)^\top\right]_{(k,j=1,\dots,m)}, \quad \textit{where} \quad \left(\sigma^{(kj)}_{k,j=1,\dots,m} = \mathbf{\Sigma}^{-1}\right) \tag{4.4}$$

Proof: Because of the form of the regression function in (3.14) for the considered SUR model (3.3) the information matrix for a one-point design has the following form

$$\begin{aligned} \mathbf{M}(\mathbf{x}) &= \mathbf{f}(\mathbf{x})\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top = \text{block}-\text{diag}\,(\mathbf{f}_k(x_k))\ \mathbf{\Sigma}^{-1}\ \text{block}-\text{diag}\,(\mathbf{f}_j(x_j)^\top) \\ &= \left[\sigma^{(kj)}\mathbf{f}_k(x_k)\mathbf{f}_j(x_j)^\top\right]_{(k,j=1,\dots,m)} \end{aligned}$$

and therewith the lemma has been proven.

Lemma 4.3. *The information matrix for the SUR model (3.3) by the product design $\xi = \otimes_{j=1}^{m}\xi_j$ in the design region $\mathcal{X} = \times_{j=1}^{m}\mathcal{X}_j$ has the following form*

$$\mathbf{M}(\xi) = \text{block} - \text{diag}\,(\sigma^{(jj)}\ \breve{\mathbf{M}}_j(\xi_j)) + \mathbf{m}(\xi)\ \mathbf{\Sigma}^{-1}\ \mathbf{m}(\xi)^\top \tag{4.5}$$

$$\textit{where } \breve{\mathbf{M}}_j(\xi_j) = \mathbf{M}_j(\xi_j) - \mathbf{m}_j(\xi_j)\mathbf{m}_j(\xi_j)^\top,\ \mathbf{M}_j(\xi_j) = \int_{\mathcal{X}_j}\mathbf{f}_j(x_j)\mathbf{f}_j(x_j)^\top\xi_j(\mathrm{d}x_j) \tag{4.6}$$

$$\mathbf{m}_j(\xi_j) = \int_{\mathcal{X}_j}\mathbf{f}_j(x_j)\xi_j(\mathrm{d}x_j),\ \mathbf{m}(\xi) = \text{block} - \text{diag}\,(\mathbf{m}_j(\xi_j)) \tag{4.7}$$

$$\textit{and } j = 1, ..., m,\quad \left(\sigma^{(kj)}_{k,j=1,\dots,m} = \mathbf{\Sigma}^{-1}\right)$$

Proof: the proof can be implemented by the integral of the information matrix for the one-point design due to the product type design $\xi = \otimes_{j=1}^{m}\xi_j$ in the design region $\mathcal{X} = \times_{j=1}^{m}\mathcal{X}_j$ so

$$\begin{aligned}
\mathbf{M}(\xi) &= \int_{\mathcal{X}}\mathbf{M}(\mathbf{x})\xi(\mathrm{d}\mathbf{x}) = \int_{\mathcal{X}}\mathbf{f}(\mathbf{x})\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top\xi(\mathrm{d}\mathbf{x}) \\
&= \left[\sigma^{(kj)}\int_{\mathcal{X}_k\times\mathcal{X}_j}\mathbf{f}_k(x_k)\mathbf{f}_j(x_j)^\top(\xi_k(\mathrm{d}x_k)\otimes\xi_j(\mathrm{d}x_j))\right]_{(k,j=1,\dots,m)} \\
&= \text{block} - \text{diag}\,(\sigma^{(jj)}(\mathbf{M}_j(\xi_j) - \mathbf{m}_j(\xi_j)\mathbf{m}_j(\xi_j)^\top)_{j=1,\dots,m} + \left[\sigma^{(kj)}\mathbf{m}_k(\xi_k)\mathbf{m}_j(\xi_j)^\top\right]_{(k,j=1,\dots,m)} \\
&= \text{block} - \text{diag}\,(\sigma^{(jj)}(\mathbf{M}_j(\xi_j) - \mathbf{m}_j(\xi_j)\mathbf{m}_j(\xi_j)^\top)) \\
&\quad + \text{block} - \text{diag}\,(\mathbf{m}_j(\xi_j))\ \mathbf{\Sigma}^{-1}\ \text{block} - \text{diag}\,(\mathbf{m}_j(\xi_j)^\top) \\
&= \text{block} - \text{diag}\,(\sigma^{(jj)}\ \breve{\mathbf{M}}_j(\xi_j)) + \mathbf{m}(\xi)\ \mathbf{\Sigma}^{-1}\ \mathbf{m}(\xi)^\top
\end{aligned}$$

and therewith the lemma has been proven. □

Remark 4.2. *Just regular information and positive definite variance covariance matrices of the error variables are considered by this work, i.e.* $\mathbf{M}_j(\xi), j = 1, ...,$*, resp.* $\mathbf{M}^{-1}(\xi)$ *are regular and* $\Sigma > \mathbf{0}$.

The next results are to hold, if all marginal models related to the components contain an intercept, i.e $\mathbf{f}_{j1} \equiv 1$, or

$$\mathbf{f}_j(x_j) = \begin{pmatrix} 1 \\ \mathbf{g}_j(x_j) \end{pmatrix} \tag{4.8}$$

Lemma 4.4. *The inverse of the information matrix for the SUR model (3.3), which marginal regression functions contain constant terms, i.e.* $\mathbf{f}_{j1} \equiv 1$ *for* $j = 1, ..., m$*, by the product design* $\xi = \otimes_{j=1}^{m}\xi_j$ *in the design region* $\mathcal{X} = \times_{j=1}^{m}\mathcal{X}_j$ *has the following form*

$$\mathbf{M}^{-1}(\xi) = \text{block} - \text{diag}\,(\frac{1}{\sigma^{(jj)}}(\mathbf{M}_j^{-1}(\xi_j) - \mathbf{e}_j\mathbf{e}_j^\top)) + \mathbf{e}\ \mathbf{\Sigma}\ \mathbf{e}^\top \tag{4.9}$$

$$\mathbf{e} = \text{block} - \text{diag}\,(\mathbf{e}_j),\ \mathbf{e}_j \textit{ is the } (p_j \times 1) \textit{ first identity vector},\ \ j = 1, ..., m$$

Proof: the proof can be implemented by the multiplication of the information matrix in (4.5) and its inverse in (4.9), so the identity matrix is obtained

$$\mathbf{M}(\xi)\ \mathbf{M}^{-1}(\xi)$$
$$= \Big(\text{block}-\text{diag}\,(\sigma^{(jj)}\check{\mathbf{M}}_j(\xi_j)) + \mathbf{m}(\xi)\ \boldsymbol{\Sigma}^{-1}\ \mathbf{m}(\xi)^\top\Big)\ \Big(\text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}}(\mathbf{M}_j^{-1}(\xi_j) - \mathbf{e}_j\mathbf{e}_j^\top)) + \mathbf{e}\ \boldsymbol{\Sigma}\ \mathbf{e}^\top\Big)$$
$$= \mathbf{A}_1 + \mathbf{A}_2 + \mathbf{A}_3 + \mathbf{A}_4$$

where

$$\mathbf{A}_1 = \Big(\text{block}-\text{diag}\,(\sigma^{(jj)}(\mathbf{M}_j(\xi_j) - \mathbf{m}_j(\xi_j)\mathbf{m}_j(\xi_j)^\top))\ \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}}(\mathbf{M}_j^{-1}(\xi_j) - \mathbf{e}_j\mathbf{e}_j^\top))\Big)$$
$$\mathbf{A}_2 = \Big(\text{block}-\text{diag}\,(\sigma^{(jj)}(\mathbf{M}_j(\xi_j) - \mathbf{m}_j(\xi_j)\mathbf{m}_j(\xi_j)^\top))\ \text{block}-\text{diag}\,(\mathbf{e}_j)\ \boldsymbol{\Sigma}\ \text{block}-\text{diag}\,(\mathbf{e}_j^\top)\Big)$$
$$\mathbf{A}_3 = \Big(\text{block}-\text{diag}\,(\mathbf{m}_j(\xi_j))\ \boldsymbol{\Sigma}^{-1}\ \text{block}-\text{diag}\,(\mathbf{m}_j(\xi_j)^\top)\ \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}}(\mathbf{M}_j^{-1}(\xi_j) - \mathbf{e}_j\mathbf{e}_j^\top))\Big)$$
$$\mathbf{A}_4 = \Big(\mathbf{m}(\xi)\ \boldsymbol{\Sigma}^{-1}\ \mathbf{m}(\xi)^\top\ \mathbf{e}\ \boldsymbol{\Sigma}\ \mathbf{e}^\top\Big)$$

then because of $\mathbf{M}_j(\xi_j)\mathbf{e}_j = \mathbf{m}_j(\xi_j),\ \mathbf{m}_j(\xi_j)^\top\mathbf{e}_j = \mathbf{e}_j^\top\mathbf{e}_j = 1$
$\mathbf{M}_j^{-1}(\xi_j)\mathbf{m}_j(\xi_j) = \mathbf{e}_j,$ because $\mathbf{m}_j$ is the first column of $\mathbf{M}_j,\ j = 1, ..., m$
the four terms by the last sum have ordered the four next forms

$$\begin{aligned}
\mathbf{A}_1 &= \text{block}-\text{diag}\,(\mathbf{I}_{p_j} - \mathbf{m}_j(\xi_j)\mathbf{e}_j^\top - \mathbf{m}_j(\xi_j)\mathbf{m}_j(\xi_j)^\top\mathbf{M}_j^{-1}(\xi_j) + \mathbf{m}_j(\xi_j)\mathbf{e}_j^\top) \\
&= \text{block}-\text{diag}\,(\mathbf{I}_{p_j} - \mathbf{m}_j(\xi_j)\mathbf{m}_j(\xi_j)^\top\mathbf{M}_j^{-1}(\xi_j)) = \text{block}-\text{diag}\,(\mathbf{I}_{p_j} - \mathbf{m}_j(\xi_j)\mathbf{e}_j^\top) \\
\mathbf{A}_2 &= \mathbf{0},\ \text{because of } (\mathbf{M}_j(\xi_j) - \mathbf{m}_j(\xi_j)\mathbf{m}_j(\xi_j)^\top)\ \mathbf{e}_j = \mathbf{m}_j(\xi_j) - \mathbf{m}_j(\xi_j) = \mathbf{0} \\
\mathbf{A}_3 &= \mathbf{0},\ \text{because of } \mathbf{m}_j(\xi_j)^\top\ (\mathbf{M}_j^{-1}(\xi_j) - \mathbf{e}_j\mathbf{e}_j^\top) = \mathbf{e}_j^\top - \mathbf{e}_j^\top = \mathbf{0} \\
\mathbf{A}_4 &= \mathbf{m}(\xi)\ \mathbf{e}^\top
\end{aligned} \tag{4.10}$$

therewith $\mathbf{M}(\xi)\ \mathbf{M}^{-1}(\xi) = \mathbf{I}_p$ and the lemma has been proven. □

Now, the D- and linear -optimality in 4.1 of the product type designs can be proven based on the calculation of the inverse of the information matrix for the SUR model (3.3) for the product type design □.

Theorem 4.3. *Let ξ_j^* be D-optimal for the j-th marginal component (3.4) in the marginal design region $\mathcal{X}_j$ with an intercept included, $j = 1, ..., m$, then the product type design*

$$\xi^* = \otimes_{j=1}^m \xi_j^*$$

is D-optimal for the SUR model (3.3) in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$.
The sensitivity function φ_D does not depend on $\boldsymbol{\Sigma}$.

Proof: the proof can be implemented by the equivalence theorem (4.1), where the terms of the equivalence theorem are calculated by the help lemmas 4.2 and 4.4, so because of (4.4) and (4.9) has the sensitivity function for the D-optimality in (4.2) the following form

$$\text{trace}\Big(\boldsymbol{\Sigma}^{-1}\ \mathbf{f}(\mathbf{x})^\top\ \mathbf{M}^{-1}(\xi^*)\ \mathbf{f}(\mathbf{x})\Big) = \text{trace}\Big(\boldsymbol{\Sigma}^{-1}\ \mathbf{f}(\mathbf{x})^\top\ \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}}(\mathbf{M}_j^{-1}(\xi_j^*)-\mathbf{e}_j\mathbf{e}_j^\top))\ \mathbf{f}(\mathbf{x})\Big) + \text{trace}\Big(\boldsymbol{\Sigma}^{-1}\ \mathbf{f}(\mathbf{x})^\top\ \mathbf{e}\ \boldsymbol{\Sigma}\ \mathbf{e}^\top\ \mathbf{f}(\mathbf{x})\Big) \quad (4.11)$$

so because of

$$\mathbf{f}_j(x_j)^\top\mathbf{e}_j = \mathbf{e}_j^\top\mathbf{f}_j(x_j) \equiv 1 \text{ then } \mathbf{f}(\mathbf{x})^\top\ \mathbf{e} = \mathbf{e}^\top\ \mathbf{f}(\mathbf{x}) = \mathbf{I}_m \quad (4.12)$$

we obtain

$$\begin{aligned}\text{trace}\Big(\boldsymbol{\Sigma}^{-1}\ \mathbf{f}(\mathbf{x})^\top\ \mathbf{M}^{-1}(\xi^*)\ \mathbf{f}(\mathbf{x})\Big) &= \text{trace}\Big(\boldsymbol{\Sigma}^{-1}\ \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}}\mathbf{f}_j(x_j)^\top\ \mathbf{M}_j^{-1}(\xi_j^*)\mathbf{f}_j(x_j))\Big)\\ &\quad -\text{trace}\Big(\boldsymbol{\Sigma}^{-1}\ \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}})\Big) + \text{trace}(\mathbf{I}_m)\\ &= \sum_{j=1}^{m}\mathbf{f}_j(x_j)^\top\ \mathbf{M}_j^{-1}(\xi_j^*)\mathbf{f}_j(x_j) \le \sum_{j=1}^{m} p_j = p \quad (4.13)\end{aligned}$$

and therewith the theorem has been proven. □

Theorem 4.4. *Let ξ_j^* be linear optimal designs with the weight matrix $\mathbf{L}_j$ for the j-th marginal component (3.4) in the marginal design region $\mathcal{X}_j$ with an intercept included, $j = 1, ..., m$, then the product type design*

$$\xi^* = \otimes_{j=1}^m \xi_j^*$$

is linear optimal design for the SUR model (3.3) in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$, if the weight matrix $\mathbf{L} = \text{block}-\text{diag}\,(\mathbf{L}_j)$ (4.1).

Proof: The weight matrix for the considered linear criteria is block diagonal and have the next form

$$\mathbf{L} = \text{block}-\text{diag}\,(\mathbf{L}_j),\ \text{where because of (4.8)}\ \mathbf{L}_j = \begin{pmatrix} L_{j11} & \mathbf{L}_{j12}^\top \\ \mathbf{L}_{j12} & \mathbf{L}_{j22}\end{pmatrix}, j = 1, ..., m \quad (4.14)$$

So because of (4.9) and (4.12)

$$\begin{aligned}\mathbf{M}(\xi^*)^{-1}\mathbf{f}(\mathbf{x}) &= \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}}(\mathbf{M}_j^{-1}(\xi_j^*)\mathbf{f}(x_j)-\mathbf{e}_j)) + \mathbf{e}\boldsymbol{\Sigma}\\ &= \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}}\mathbf{M}_j^{-1}(\xi_j^*)\mathbf{f}(x_j)) + \mathbf{e}\Big(\boldsymbol{\Sigma}-\text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}})\Big) \quad (4.15)\end{aligned}$$

Then

$$\mathbf{f}(\mathbf{x})^\top\mathbf{M}(\xi^*)^{-1} = \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}}\mathbf{f}(x_j)^\top\mathbf{M}_j^{-1}(\xi_j^*)) + \Big(\boldsymbol{\Sigma}-\text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}})\Big)\,\mathbf{e}^\top$$

So the the left side of the equivalence theorem for the linear criteria has the following form

$$\text{trace}\left(\mathbf{\Sigma}^{-1}\ \mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi^*)^{-1}\mathbf{L}\ \mathbf{M}(\xi^*)^{-1}\mathbf{f}(\mathbf{x})\right) =$$

$$\text{trace}\left(\mathbf{\Sigma}^{-1}\ \text{block}-\text{diag}\,(\frac{1}{(\sigma^{(jj)})^2}\mathbf{f}_j(x_j)^\top\ \mathbf{M}_j^{-1}(\xi_j^*)\ \mathbf{L}_j\ \mathbf{M}_j^{-1}(\xi_j^*))\ \mathbf{f}_j(x_j)\right) + a_1 + a_2 \text{ where}$$

$$a_1 = \text{trace}\left(\mathbf{\Sigma}^{-1}\ \left(\mathbf{\Sigma} - \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}})\right)\mathbf{e}^\top\ \mathbf{L}\ \mathbf{e}\ \left(\mathbf{\Sigma} - \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}})\right)\right)$$

$$a_2 = 2\text{trace}\left(\mathbf{\Sigma}^{-1}\ \left(\mathbf{\Sigma} - \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}})\right)\mathbf{e}^\top\ \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}}\mathbf{L}_j\mathbf{M}_j^{-1}(\xi_j^*)\mathbf{f}_j(x_j))\right)$$

So because of the diagonal entries of the matrix $\mathbf{\Sigma}^{-1}\ \left(\mathbf{\Sigma} - \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}})\right)$ are equal to zero

$$\begin{aligned}
a_1 &= \text{trace}\left(\left(\mathbf{\Sigma} - \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}})\right)\left(\mathbf{I}_m - \mathbf{\Sigma}^{-1}\ \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}})\right)\right)\text{block}-\text{diag}\,(L_{j11}) \\
&= \text{trace}\left(\mathbf{\Sigma}\ .\text{block}-\text{diag}\,(L_{j11}) - \text{block}-\text{diag}\,(\frac{2}{\sigma^{(jj)}})\ .\text{block}-\text{diag}\,(L_{j11})\right) \\
&\quad+\text{trace}\left(\text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}})\ .\mathbf{\Sigma}^{-1}\ .\ \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}})\ .\text{block}-\text{diag}\,(L_{j11})\right) \\
&= \sum_{j=1}^m L_{j11}(\sigma_j^2 - \frac{1}{\sigma^{(jj)}}) \qquad (4.16)
\end{aligned}$$

and the term $a_2 = 0$ and therewith

$$\begin{aligned}
&\text{trace}\left(\mathbf{\Sigma}^{-1}\ \mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi^*)^{-1}\mathbf{L}\ \mathbf{M}(\xi^*)^{-1}\mathbf{f}(\mathbf{x})\right) \\
&= \sum_{j=1}^m \frac{1}{\sigma^{(jj)}}\mathbf{f}_j(x_j)^\top \mathbf{M}_j^{-1}(\xi_j^*)\mathbf{L}_j\mathbf{M}_j^{-1}(\xi_j^*)\mathbf{f}_j(x_j) + \sum_{j=1}^m L_{j11}(\sigma_j^2 - \frac{1}{\sigma^{(jj)}}) \qquad (4.17)
\end{aligned}$$

The right side of the equivalence theorem for the Linear criteria can be calculated by integral the right side due to the optimal design ξ^*, so

$$\int_{\mathcal{X}} \mathbf{M}(\mathbf{x})\mathbf{M}(\xi^*)^{-1}\mathbf{L}\ \mathbf{M}(\xi^*)^{-1}\xi^*(\mathrm{d}\mathbf{x}) = \mathbf{M}(\xi^*)\mathbf{M}(\xi^*)^{-1}\mathbf{L}\ \mathbf{M}(\xi^*)^{-1} = \mathbf{L}\ \mathbf{M}(\xi^*)^{-1}$$

Also,

$$\text{trace}\left(\mathbf{L}\ \mathbf{M}(\xi^*)^{-1}\right) = \sum_{j=1}^m \frac{1}{\sigma^{(jj)}}\mathbf{L}_j\mathbf{M}_j^{-1}(\xi_j^*) + \sum_{j=1}^m L_{j11}(\sigma_j^2 - \frac{1}{\sigma^{(jj)}}) \qquad (4.18)$$

It is shown by comparing (4.17) and (4.18) that

$$\text{trace}\left(\mathbf{\Sigma}^{-1}\ \mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi^*)^{-1}\mathbf{L}\ \mathbf{M}(\xi^*)^{-1}\mathbf{f}(\mathbf{x})\right) \leq \text{trace}\left(\mathbf{L}\ \mathbf{M}(\xi^*)^{-1}\right)$$

And therewith the theorem has been proven. □

The next result is about the asymptotic (i.e. for $n \to \infty$ and normality) D- and linear optimality by block diagonal weight matrix $\mathbf{L}$ of the product designs for the SUR model (3.3) with respect to estimators by unknown variance covariance matrix of the error variables as the Maximum Likelihood and the feasible general least squares estimators by the normal distribution of the error variables.

Lemma 4.5. *Theorems 4.3 and 4.4 are asymptotically valid, under the asymptotic normality and regularity conditions as the convergence of the information matrix to an existend, finite and regulare information matrx, by unknown variance covariance matrix of the error variables with respect to propitious estimators as the Maximum Likelihood and feasible general least squares estimators.*

Proof: The validity of theorem 4.3 follows from the independence between the parameter $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ and subsequently because of the block-diagonal form of the Fisher-information matrix $\mathbf{M}(\boldsymbol{\beta}, \mathbf{V})$ in (3.9) and the independence of $\mathbf{M_V}$ of the control variables ($\mathbf{x}$). The validity of theorem 4.4 follows from the independence between the parameter $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ and subsequently because of the block diagonal form of the inverse Fisher-information matrix res. the variance covariance matrix of the considered estimators .

This lemma is valid for SUR models with multi-factorial marginal regression functions as in chapter six or marginal regression functions without intercepts as in chapter seven. □

4.3 Example: SUR model with three components

To illustrate the results we consider the SUR model with simple straight line, quadratic, and cubic regression models, for the components,

$$
\begin{aligned}
Y_{i1} &= \beta_{10} + \beta_{11}x_{i1} + \varepsilon_{i1} \\
Y_{i2} &= \beta_{20} + \beta_{21}x_{i2} + \beta_{22}x_{i2}^2 + \varepsilon_{i2} \\
Y_{i3} &= \beta_{30} + \beta_{31}x_{i3} + \beta_{32}x_{i3}^2 + \beta_{33}x_{i3}^3 + \varepsilon_{i3}
\end{aligned} \tag{4.19}
$$

in the unit intervals $\mathcal{X}_1 = [0, 1]$, $\mathcal{X}_2 = \mathcal{X}_3 = [-1, 1]$ as experimental regions. Then it is well-known or see the examples in the second chapter, that the D-and IMSE-optimal designs for the first marginal model are $\xi^*_{D;IMSE,1} = \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}$, which assign equal weights to each of the endpoint of the interval. The A-optimal design for the first marginal model is $\xi^*_{A,1} = \begin{pmatrix} 0 & 1 \\ 2-\sqrt{2} & \sqrt{2}-1 \end{pmatrix}$

The D-optimal designs for the second marginal model is $\xi^*_{D,2} = \begin{pmatrix} -1 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$, which assign equal weights to each of the design points. The A-and IMSE-optimal design for the second marginal model are $\xi^*_{A;IMSE,2} = \begin{pmatrix} -1 & 0 & 1 \\ 1/4 & 1/2 & 1/4 \end{pmatrix}$

The D-optimal designs for the third marginal model is $\xi^*_{D,3} = \begin{pmatrix} -1 & -1/\sqrt{5} & 1/\sqrt{5} & 1 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$ assign equal weights to each of the design points.

The IMSE-optimal design for the third marginal model is
$\xi^*_{IMSE,3} = \begin{pmatrix} -1 & -x & x & 1 \\ w_I/2 & (1-w_I)/2 & (1-w_I)/2 & w_I/2 \end{pmatrix}$ where $x \simeq 0.44$, $w_I \simeq 0.31$

The A-optimal design for the third marginal model is
$\xi^*_{A,3} = \begin{pmatrix} -1 & -x & x & 1 \\ w_A/2 & (1-w_A)/2 & (1-w_A)/2 & w_A/2 \end{pmatrix}$ where $x \simeq 0.46,\ w_A = 0.3$.
Then By Theorem 4.3 the product type design

$$\xi^*_D = \xi^*_{D,1} \otimes \xi^*_{D,2} \otimes \xi^*_{D,3}$$

is D-optimal for the SUR model (4.19) in $\mathcal{X} = [0,1] \times [-1,1]^2$.
By the Theorem 4.4 the product type design

$$\xi^*_{IMSE} = \xi^*_{IMSE,1} \otimes \xi^*_{IMSE,2} \otimes \xi^*_{IMSE,3}$$

is IMSE-optimal for the SUR model (4.19) in $\mathcal{X} = [0,1] \times [-1,1]^2$.
And by the Theorem 4.4 the product type design

$$\xi^*_A = \xi^*_{A,1} \otimes \xi^*_{A,2} \otimes \xi^*_{A,3}$$

is A-optimal for the SUR model (4.19) in $\mathcal{X} = [0,1] \times [-1,1]^2$.
The corresponding sensitivity function for the D-optimality by $x_1 = 1$

$$\begin{aligned} \varphi_D(\mathbf{x}; \xi^*_D) &= \text{trace}(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi^*_D)^{-1}\mathbf{f}(\mathbf{x})) \\ &= \frac{3}{2}(\frac{11}{2} - 3\,x_2{}^2 + 3\,x_2{}^4 + \frac{11}{2}\,x_3{}^2 - \frac{35}{2}\,x_3{}^4 + \frac{25}{2}\,x_3{}^6) \end{aligned} \tag{4.20}$$

is plotted in figure 4.1.
It can be easily seen that the sensitivity function is independent on $\mathbf{\Sigma}$ and satisfies the condition $\varphi_D(\mathbf{x}; \xi^*_D) \le p = 4$ for all $\mathbf{x} \in \mathcal{X}$.
The corresponding sensitivity function for IMSE-optimality is equal or smaller than
$\text{trace}(\mathbf{LM}(\xi^*_{IMSE})^{-1})$ and the function

$$\frac{\varphi_{IMSE}(\mathbf{x}; \xi^*_{IMSE})}{\text{trace}\left(\mathbf{L}\ \mathbf{M}(\xi^*_{IMSE})^{-1}\right)} = \frac{\text{trace}(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi^*_{IMSE})^{-1}\mathbf{LM}(\xi^*_{IMSE})^{-1}\mathbf{f}(\mathbf{x}))}{\text{trace}(\mathbf{LM}(\xi^*_{IMSE})^{-1})} \tag{4.21}$$

is plotted in figure 4.2, for $\sigma_1 = \sigma_2 = \sigma_3 = 1, \rho_{12} = 0.99, \rho_{13} = 0.14, \rho_{23} = 0$.
It can be easily seen that the sensitivity function for the IMSE-optimality satisfies the condition $\varphi_{IMSE}(\mathbf{x}; \xi^*_{IMSE}) \le \text{trace}(\mathbf{LM}(\xi^*_{IMSE})^{-1})$ for all $\mathbf{x} \in \mathcal{X}$. The corresponding sensitivity function for A-optimality is equal to or less than
$\text{trace}(\mathbf{M}(\xi^*_A)^{-1})$ and the function

$$\frac{\varphi_A(\mathbf{x}; \xi^*_A)}{\text{trace}\left(\mathbf{M}(\xi^*_A)^{-1}\right)} = \frac{\text{trace}(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi^*_A)^{-1}\mathbf{M}(\xi^*_A)^{-1}\mathbf{f}(\mathbf{x}))}{\text{trace}(\mathbf{M}(\xi^*_A)^{-1})} \tag{4.22}$$

is plotted in figure 4.3, for $\sigma_1 = \sigma_2 = \sigma_3 = 1, \rho_{12} = 0.58, \rho_{13} = 0.01$ and $\rho_{23} = -0.8$.
It can be easily seen that the sensitivity function for the A-optimality satisfies the condition $\varphi_A(\mathbf{x}; \xi^*) \le \text{trace}(\mathbf{M}(\xi^*_A)^{-1})$ for all $\mathbf{x} \in \mathcal{X}$. The product designs ξ^*_D, ξ^*_{IMSE} and ξ^*_A are asymptotic optimal for the SUR model (6.34) because of lemma (4.5) and the asymptotic Fisher Information matrix has the following form for the given inverse of the variance covariance matrix of the error variables in (5.9)

$$\mathbf{M}(\boldsymbol{\beta}, \mathbf{\Sigma}) = \begin{pmatrix} \mathbf{M}_{\boldsymbol{\beta}} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\mathbf{\Sigma}} \end{pmatrix}$$

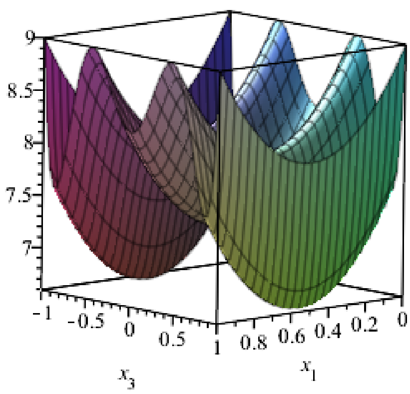

Figure 4.1: Function 4.20 for ξ_D^* (SUR with one-factor marginals)

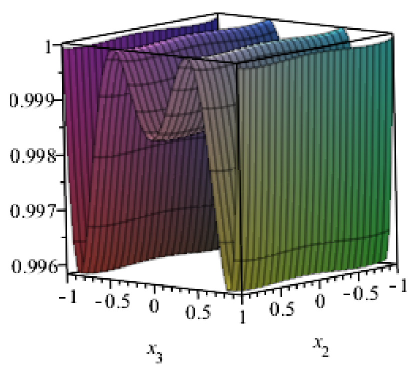

Figure 4.2: Function 4.21 for ξ_{IMSE}^* (SUR with one-factor marginals)

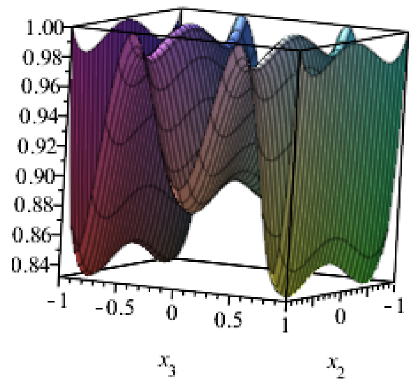

Figure 4.3: Function 4.22 for ξ_A^* (SUR with one-factor marginals)

5 G-Optimal Designs and Some Design Efficiencies for Multivariate Linear Models

The G-optimal design criterion in the univariate case was the entrance for the optimal design theory by the work of *(Smith (1918))*. The equivalence between the approximate optimality for D- and G-designs under some conditions have been proven in the work of *(Kiefer and Wolfowitz (1960))*. The equivalence between G- and D-optimality have been discussed in the works of *(Wong (1993), (1994) and (1995))*, *(King, Wong (1998))*, *(Brown and Wong (2000))* and *(Chen et al. (2008))* in more generalized cases . The multivariate equivalence theorem for D-optimality by *(Fedorov (1972))* supplies a standardized or weighted G-optimal design due to the trace of the inverse of the variance covariance matrix for the error variables multiplied with the covariance matrix for the prediction. That was positively assessed and may be logic in the multivariate sense, however the research of the non-weighted G-optimal design problem in the multivariate case can not be loss. The G-optimal designs problem in the multivariate case are brought to light in this work, so different function as the trace, maximum eigenvalue and the determinant are applied to the covariance matrix for the prediction and the upper bounds for the applied functions on the covariance matrix of the prediction for multivariate linear models in general are determined with respect to the D- resp. weighted G-optimal design and due to some inequalities for the multiplication of the positive definite matrices. Upper bounds for the applied functions on the covariance matrix of the prediction for the SUR models with respect to the product type designs are determined too due to the Gauß and OLS estimator as well as the upper bounds for MANOVA-models by the MANOVA-design.
The convex optimal design theory for limited information resp. non efficient estimators such as the OLS estimator has not been established, but the OLS estimator can be used by an unknown variance covariance matrix of the error variables, so it is useful, to calculate the efficiency of the OLS estimator for product type design and MANOVA-design in comparison to the BLUE Gauß Markov estimator for different optimality criteria, which are the topics of the second part of this chapter.
It has been shown, that the the reduction of the IMSE-optimal design problem for MANOVA-model is possible as in the work of *(Kurotschka and Schwabe (1996))* for other criteria.
Three different G-optimal design's criteria are introduced and their upper boundaries for D- resp. weighted G-optimal determined for the multivariate models in general, for SUR models by Gauß and OLS estimator and for MANOVA-models by MANOVA-design in the first section. The efficiency for the OLS estimator against the Gauß estimator by the product type design and the efficiency of the MANOVA-design against the product type design by the Gauß and OLS estimator for D- and linear criteria in 4.1 are determined in the second section. There is a clear example in the bivariate case for illustrating the theoretical results in the last section.

5.1 G-Optimal Designs for Multivariate Linear Models

The approximate D-optimal design is G-optimal in the univariate case because of the equivalence theorem (2.2), but the equivalence theorem for D-optimality in the multivariate case (4.1) supplies a weighted G-optimal design with the inverse of the variance covariance matrix of the error variables, so there is no longer elementary variance function but a covariance matrix for the prediction estimate, thus some convex function are to apply on this matrix to define the min max criterion, so we get different G-optimal criteria according to the implemented functions, for example the trace, det or $\lambda_{\max}$ can be applied to the covariance matrix $\mathbf{Cov}\left(\mathbf{f}(\mathbf{x})^\top\hat{\boldsymbol{\beta}}\right)$, so the G-optimal design criteria can be defined as follows

Def 5.1 (G-optimal Criteria in the Multivariate Case)**.** *For the following defined covariance matrix for the prediction estimate*

$$\mathbf{COV}(\mathbf{x};\xi) \;= \mathbf{Cov}\left(\mathbf{f}(\mathbf{x})^\top\hat{\boldsymbol{\beta}}\right) \;= \mathbf{f}(\mathbf{x})^\top\ \mathbf{M}^{-1}\ (\xi)\ \mathbf{f}(\mathbf{x}) \tag{5.1}$$

the min max *criterion can have the following three different types*

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}}\operatorname{trace}\left(\mathbf{COV}(\mathbf{x};\xi)\right)\right]$$

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}}\boldsymbol{\lambda}_{\ \max}\left(\mathbf{COV}(\mathbf{x};\xi)\right)\right]$$

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}}\det\left(\mathbf{COV}(\mathbf{x};\xi)\right)\right]$$

Some additional reforms to the equivalence theorem 4.1 are done as a first step to apply the suggested G-optimal criteria based on some inequalities for the product of positive definite matrices, which are applied to the trace, determinant, and the maximum eigenvalue for the product of the inverse of the variance covariance matrix of the error variables multiplied with the covariance of the prediction, also, at first we want to present these inequalities and their proofs, which are located in the solution manual of *(Zhang (1999)*) or in *(Yang (2000)*) and *(YANG and FENG (2002)*), and then the equivalence theorem with the bounds for the suggested G-optimal design criterion in the multivariate case due to the D-optimal design.

Lemma 5.1. *When the matrices $\mathbf{H}_{q\times q}$ and $\mathbf{B}_{q\times q}$ are respectively positive semi definite and positive definite, then*

$$\operatorname{trace}\left(\mathbf{B}^{-1}\mathbf{H}\right) \geq \frac{\operatorname{trace}\left(\mathbf{H}\right)}{\operatorname{trace}\left(\mathbf{B}\right)} \tag{5.2}$$

$$\left[\det\left(\mathbf{H}\right)\right]^{\frac{1}{q}} \leq \frac{\operatorname{trace}\left(\mathbf{H}\right)}{q} \tag{5.3}$$

Proof: The proof of (5.2) is as follows: Suppose $\mathbf{B} = \mathbf{U}diag(\lambda_1...\lambda_q)\mathbf{U}^\top$ for some unitary matrix $\mathbf{U}$, where λ are the eigenvalues of $\mathbf{B}$. Denote the diagonal entries of $\mathbf{U}^\top\mathbf{H}\mathbf{U}$ by $h_1...h_q$, each $h_i \geq 0$. Then

$$\operatorname{trace}\left(\mathbf{B}^{-1}\mathbf{H}\right) = \operatorname{trace}\left(\mathbf{U}diag(\lambda_1^{-1}...\lambda_q^{-1})\mathbf{U}^\top\mathbf{H}\right) = \operatorname{trace}\left(diag(\lambda_1^{-1}...\lambda_q^{-1})\mathbf{U}^\top\mathbf{H}\mathbf{U}\right)$$
$$= \lambda_1^{-1}h_1 + ... + \lambda_q^{-1}h_q \geq (\lambda_{\max}(\mathbf{B}))^{-1}(h_1 + ... + h_q) = \frac{\operatorname{trace}\left(\mathbf{U}^\top\mathbf{H}\mathbf{U}\right)}{\lambda_{\max}(\mathbf{B})} = \frac{\operatorname{trace}\left(\mathbf{H}\right)}{\operatorname{trace}\left(\mathbf{B}\right)}$$

The proof of (5.3) is as follows: Let the eigenvalues of $\mathbf{H}$ be $\lambda_1, \dots, \lambda_q$. Then all λ are nonnegative. By the arithmetic mean-geometric inequality, we have

$$[\det(\mathbf{H})]^{\frac{1}{q}} = (\lambda_1 ... \lambda_q)^{\frac{1}{q}} \leq \frac{\lambda_1 + \ldots + \lambda_q}{q} = \frac{\text{trace}(\mathbf{H})}{q} \quad \square$$

The next theorem gives upper bounds for the trace, the maximum eigenvalue and the determinant for the covariance matrix of the prediction due to the D-optimal resp. the weighted G-optimal design in the general multivariate case

Theorem 5.1. *The following three statements are equivalent :*

1. ξ_D^* *minimizes* $\det\left(\mathbf{M}^{-1}(\xi)\right)$
2. ξ_D^* *minimizes* $\max_{\mathbf{x}\in\mathcal{X}} \text{trace}\left(\mathbf{\Sigma}^{-1}\ \mathbf{f}(\mathbf{x})^\top\ \mathbf{M}^{-1}\ (\xi)\ \mathbf{f}(\mathbf{x})\right)$
3. $\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \text{trace}\left(\mathbf{\Sigma}^{-1}\ \mathbf{COV}(\mathbf{x};\xi)\right)\right] = p$

And because of lemma (5.1) [3] $\Longrightarrow$

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \text{trace}\left(\mathbf{COV}(\mathbf{x};\xi)\right)\right] \leq\ p\ \text{trace}\left(\mathbf{\Sigma}\right) \tag{5.4}$$

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \boldsymbol{\lambda}_{\max}\left(\mathbf{COV}(\mathbf{x};\xi)\right)\right] <\ p\ \text{trace}\left(\mathbf{\Sigma}\right) \tag{5.5}$$

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \det\left(\mathbf{COV}(\mathbf{x};\xi)\right)\right] \leq\ \left[\frac{p}{m}\text{trace}\left(\mathbf{\Sigma}\right)\right]^m \tag{5.6}$$

where the covariance matrix for the prediction $\mathbf{COV}(\mathbf{x};\xi)$ *is given in (5.1)*

Proof: (5.4) follows from [3] due to (5.2) by replacing $\mathbf{B} = \mathbf{\Sigma}$ and $\mathbf{H} = \mathbf{COV}(\mathbf{x};\xi)$, (5.5) follows from (5.4), because the maximum eigenvalue of a positive definite matrix less than its trace, and (5.6) follows from (5.4) due to (5.3). $\square$

Lemma 5.2. *The design* $\xi = \xi_{0;D}^*$ *gives the next upper bounds respectively according to the trace, the maximum eigenvalue and the determinant for the covariance matrix of the prediction for the MANOVA-model (introduced in 3.4) with the regression function (3.15) on the design region* $\mathcal{X}_0$

1. $\min_{\xi\in\Xi}\left[\max_{x\in\mathcal{X}_0} \text{trace}\left(\mathbf{COV}(x;\xi)\right)\right] =\ p_0\ \text{trace}\left(\mathbf{\Sigma}\right) \leq\ p\ \text{trace}\left(\mathbf{\Sigma}\right)$
2. $\min_{\xi\in\Xi}\left[\max_{x\in\mathcal{X}_0} \boldsymbol{\lambda}_{\max}\left(\mathbf{COV}(x;\xi)\right)\right] =\ p_0\ \boldsymbol{\lambda}_{\max}\left(\mathbf{\Sigma}\right) < p\ \text{trace}\left(\mathbf{\Sigma}\right)$
3. $\min_{\xi\in\Xi}\left[\max_{x\in\mathcal{X}_0} \det\left(\mathbf{COV}(x;\xi)\right)\right] = \left[\ (p_0\ )^m \det\mathbf{\Sigma}\right] \leq \left[p_0\ \text{trace}\left(\mathbf{\Sigma}\right)\right]^m$

Where $p_0 = \frac{p}{m}$ *is the number of the parameter by one univariate component model.*

Proof: the information matrix has for an approximate design ξ_0, because of (3.16) the form
$\mathbf{M}(\xi_0) = \mathbf{\Sigma}^{-1} \otimes \left(\int \mathbf{f}_0(x_0)\mathbf{f}_0(x_0)^\top d\xi_0\right) = \mathbf{\Sigma}^{-1} \otimes \mathbf{M}_0(\xi_0)$
i.e the inverse of it is as follows

$$\mathbf{M}^{-1}(\xi_0) = \mathbf{\Sigma} \otimes \mathbf{M}_0^{-1}(\xi_0)$$

Then the covariance matrix for the prediction with respect to ξ_0 has for this model the following form

$$\mathbf{COV}(x;\xi_0) = \left(\mathbf{I}_{m\times m} \otimes \mathbf{f}_0(x)^\top\right) \left[\mathbf{\Sigma} \otimes \mathbf{M}_0^{-1}(\xi_0)\right] \left(\mathbf{I}_{m\times m} \otimes \mathbf{f}_0(x)\right)$$

$$= \mathbf{\Sigma} \otimes \left[\mathbf{f}_0(x)^\top \mathbf{M}_0^{-1}(\xi_0)\mathbf{f}_0(x)\right] = \mathbf{\Sigma} \otimes \mathrm{Var}(x;\xi_0) \tag{5.7}$$

where $\mathrm{Var}(x;\xi) = \mathbf{f}(x)^\top \mathbf{M}^{-1}(\xi)\mathbf{f}(x)$ (5.8)

is the variance function for the prediction due to the design ξ in the univariate case.

The upper bound for the trace as function has the following form for ξ_0^{D*} and because of (5.7) and some properties of the Kronecker product

$$\begin{aligned}
&\min_{\xi\in\Xi}\left[\max_{x\in\mathcal{X}_0} \mathrm{trace}\left(\mathbf{COV}(x;\xi)\right)\right] = \max_{x\in\mathcal{X}_0} \mathrm{trace}\left(\mathbf{\Sigma} \otimes \mathrm{Var}(x;\xi_0^{D*})\right)\\
&\mathrm{trace}\left(\mathbf{\Sigma}\right). \max_{x\in\mathcal{X}_0} \mathrm{trace}\left(\mathrm{Var}(x;\xi_0^{D*})\right) \ = \mathrm{trace}\mathbf{\Sigma}. \max_{x\in\mathcal{X}_0} \mathrm{Var}(x;\xi_0^{D*})\\
&= \ p_0 \ \mathrm{trace}\left(\mathbf{\Sigma}\right) \leq \ p \ \mathrm{trace}\mathbf{\Sigma}
\end{aligned}$$

Where the equality occurs by $\mathbf{\Sigma} = \mathbf{I}_m$, because of $\mathrm{trace}\mathbf{\Sigma} = \mathrm{trace}(\mathbf{I}_m) = m$ and $p = mp_0$.
for the maximum eigenvalue

$$\begin{aligned}
&\min_{\xi\in\Xi}\left[\max_{x\in\mathcal{X}_0} \boldsymbol{\lambda}_{\max}\left(\mathbf{COV}(x;\xi)\right)\right] = \max_{x\in\mathcal{X}_0} \boldsymbol{\lambda}_{\max}\left(\mathbf{\Sigma} \otimes \mathrm{Var}(x;\xi_0^{D*})\right)\\
&\boldsymbol{\lambda}_{\max}\left(\mathbf{\Sigma}\right). \max_{x\in\mathcal{X}_0} \boldsymbol{\lambda}_{\max}\left(\mathrm{Var}(x;\xi_0^{D*})\right) \ = \boldsymbol{\lambda}_{\max}\left(\mathbf{\Sigma}\right). \max_{x\in\mathcal{X}_0} \mathrm{Var}(x;\xi_0^{D*})\\
&= \ p_0 \ \boldsymbol{\lambda}_{\max}\left(\mathbf{\Sigma}\right) < \ p \ \boldsymbol{\lambda}_{\max}\left(\mathbf{\Sigma}\right)
\end{aligned}$$

and in the same way for the determinant and due to the inequality (5.3), we have

$$\begin{aligned}
&\min_{\xi\in\Xi}\left[\max_{x\in\mathcal{X}_0} \det\left(\mathbf{COV}(x;\xi)\right)\right] = \max_{x\in\mathcal{X}_0} \det\left(\mathbf{\Sigma} \otimes \mathrm{Var}(x;\xi_0^{D*})\right)\\
&\det \mathbf{\Sigma}. \max_{x\in\mathcal{X}_0} \left(\det \mathrm{Var}(x;\xi_0^{D*})\right)^{p_0} \ = \det \mathbf{\Sigma}. \max_{x\in\mathcal{X}_0} \left(\mathrm{Var}(x;\xi_0^{D*})\right)^{m}\\
&= \ (p_0)^m \ \det \mathbf{\Sigma} \leq \ (p_0)^m \ (\mathrm{trace}\mathbf{\Sigma})^m < \ (p_o \ \mathrm{trace}\mathbf{\Sigma})^m \quad \square
\end{aligned}$$

Lemma 5.3. *The linear optimal design $\xi = \xi^*_{0;L}$ for one of the corresponding marginal models with the regression function $\mathbf{f}_0(x_0)$ is linear optimal design by block diagonal weight matrix $\mathbf{L} = \mathrm{block-diag}\,(\mathbf{L}_j)$ (4.1), for the MANOVA-model with the regression function (3.15) on the design region $\mathcal{X}_0$.*

Proof: The proof is included in the proof of theorem 5.5 .
It have been proven in chapter three, that the product designs of the D- resp. weighted G- and linear optimal designs for the marginal components are the linear optimal in 4.1 designs for the SUR models, which their marginal components related to the components contain an intercept, $f_{j1}(\mathbf{x}) \equiv 1$, also, for the same SUR models and under the class of product designs we want present the next theorems, which illustrate that the bounds due to the product type designs less than the resulted bounds by the theorem 5.1 for general multivariate linear models. $\square$

Remark 5.1. *For example the inverse of the variance covariance matrix of the error variables* $\mathbf{\Sigma}_{3\times 3}$ *is given because of remark 3.5 as follows*

$$\mathbf{\Sigma}^{-1} = \begin{pmatrix} \sigma^{(11)} & \sigma^{(12)} & \sigma^{(13)} \\ \sigma^{(12)} & \sigma^{(22)} & \sigma^{(23)} \\ \sigma^{(13)} & \sigma^{(23)} & \sigma^{(33)} \end{pmatrix} = \begin{pmatrix} \frac{c^1}{\sigma_1^2 c} & -\frac{\rho_{12}-\rho_{13}\rho_{23}}{\sigma_1\sigma_2 c} & \frac{\rho_{12}\rho_{23}-\rho_{13}}{\sigma_1\sigma_3 c} \\ -\frac{\rho_{12}-\rho_{13}\rho_{23}}{\sigma_1\sigma_2 c} & \frac{c^2}{\sigma_2^2 c} & \frac{\rho_{12}\rho_{13}-\rho_{23}}{\sigma_2\sigma_3 c} \\ \frac{\rho_{12}\rho_{23}-\rho_{13}}{\sigma_1\sigma_3 c} & \frac{\rho_{12}\rho_{13}-\rho_{23}}{\sigma_2\sigma_3 c} & \frac{c^3}{\sigma_3^2 c} \end{pmatrix} \tag{5.9}$$

where

$$c_j = c_{\ell q},\ \ \ell, q \neq j,\ \ \ell, q = 1, 2, 3 \tag{5.10}$$

is the determinant for the minor of the correlation matrix, which do not contains the column and line j.
So similarly and without loss of generality the diagonal entries of the inverse of the variance covariance matrix of the error variables $\mathbf{\Sigma}$ *have the forms*

$$\sigma^{(jj)} = \frac{c_j}{\sigma_j^2\ c},\ \ j = 1, ..., m \tag{5.11}$$

Lemma 5.4. *By considering SUR model (3.3). The covariance matrix for the prediction with respect to the Gauß estimator and the product type design* $\xi = \otimes_{j=1}^{m} \xi_j$ *has the following form*

$$\mathbf{COV}(x;\xi) = \ \text{block}-\text{diag}\,(\frac{\sigma_j^2\ c}{c_j}(\text{Var}(x_j;\xi_j)-1)) + \ \mathbf{\Sigma} \tag{5.12}$$

where c_j *is the determinant for the minor of the correlation matrix, which do not contains the column and line* j.

Proof: because of form of the inverse of the information matrix 4.9 due to the Gauß estimator with respect to the product type design by the lemma 4.4 has the covariance matrix for the prediction the following form

$$\mathbf{COV}(x;\xi) = \ \text{block}-\text{diag}\,(\mathbf{f}_j(x_j)^\top)$$
$$\left[\text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}}(\mathbf{M}_j^{-1}(\xi_j) - \mathbf{e}_j\mathbf{e}_j^\top) + \text{block}-\text{diag}\,(\mathbf{e}_j)\ \mathbf{\Sigma}\ \text{block}-\text{diag}\,(\mathbf{e}_j^\top)\right]\ \text{block}-\text{diag}\,(\mathbf{f}_j(x_j))$$
$$= \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}}(\mathbf{f}_j(x_j)^\top\ \mathbf{M}_j^{-1}(\xi_j)\ \mathbf{f}_j(x_j) - \mathbf{f}_j(x_j)^\top\ \mathbf{e}_j\ \mathbf{e}_j^\top\ \mathbf{f}_j(x_j)))$$
$$+ \text{block}-\text{diag}\,(\mathbf{f}_j(x_j)^\top\ \mathbf{e}_j)\ \mathbf{\Sigma}\ \text{block}-\text{diag}\,(\mathbf{e}_j^\top\ \mathbf{f}_j(x_j))$$

Then because of $\mathbf{f}_j(\xi_j)^\top\ \mathbf{e}_j = \mathbf{e}_j^\top\ \mathbf{f}_j(\xi_j) = 1,\ j = 1, ..., m,$ then $\text{block}-\text{diag}\,(\mathbf{f}_j(x_j)^\top\ \mathbf{e}_j) = \mathbf{I}_m,$ thus

$$\mathbf{COV}(x;\xi)\ = \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}}(\mathbf{f}_j(x_j)^\top\ \mathbf{M}_j^{-1}(\xi_j)\ \mathbf{f}_j(x_j) -\ 1)) + \ \mathbf{\Sigma}$$

And because of (5.11) $\dfrac{1}{\sigma^{(jj)}} = \dfrac{\sigma_j^2\ c}{c_j}$ and (5.8)

$$\mathbf{COV}(x;\xi)\ = \text{block}-\text{diag}\,(\frac{\sigma_j^2\ c}{c_j}(\text{Var}(x_j;\xi_j) -\ 1)) + \ \mathbf{\Sigma}$$

where c_j is the determinant for the minor of the correlation matrix, which do not contains the column and line j. □

Lemma 5.5. *By considering SUR model (3.3). The min max covariance matrix for the prediction with respect to the Gauß estimator and the product type design of the D- resp. G- optimal designs for the corresponding marginal components models , i.e.* $\xi_D^* = \otimes_{j=1}^m \xi_{D,j}^*$ *has the following form*

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \ \mathbf{COV}(x;\xi)\right] = \text{block}-\text{diag}\,(\frac{\sigma_j^2\ c}{c_j}(p_j-1)) + \ \mathbf{\Sigma} \tag{5.13}$$

Proof: because of (5.12)

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \ \mathbf{COV}(x;\xi)\right] = \min_{\xi_j\in\Xi_j}\left[\max_{x_j\in\mathcal{X}_j} \ \text{block}-\text{diag}\,(\frac{\sigma_j^2\ c}{c_j}(\text{Var}(x_j;\xi_j)-1)) + \ \mathbf{\Sigma}\right]$$
$$= \max_{x_j\in\mathcal{X}_j}\left[\text{block}-\text{diag}\,(\frac{\sigma_j^2\ c}{c_j}(\text{Var}(x_j;\xi_{D,j}^*)-1)) + \ \mathbf{\Sigma}\right] = \text{block}-\text{diag}\,(\frac{\sigma_j^2\ c}{c_j}(p_j-1)) + \ \mathbf{\Sigma}$$

The matrix in (5.13) has the form block − diag (p_j) for $\mathbf{\Sigma} = \mathbf{I}_m$. □

Remark 5.2. *For SUR model (3.3). The min max for the trace, maximum eigenvalue, determinant of the covariance matrix for the prediction with respect to the Gauß estimator and the product type design of the D- resp. G- optimal designs for the corresponding marginal components models , i.e.* $\xi_D^* = \otimes_{j=1}^m \xi_{D,j}^*$ *have because of the form (5.13) by the lemma 5.5 the following forms*

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \text{trace}\,(\mathbf{COV}(\mathbf{x};\xi))\right] = \text{trace}\left(\text{block}-\text{diag}\,(\frac{\sigma_j^2\ c}{c_j}(p_j-1)) + \ \mathbf{\Sigma}\right) \tag{5.14}$$

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \boldsymbol{\lambda}_{\max}\,(\mathbf{COV}(\mathbf{x};\xi))\right] = \boldsymbol{\lambda}_{\max}\left(\text{block}-\text{diag}\,(\frac{\sigma_j^2\ c}{c_j}(p_j-1)) + \ \mathbf{\Sigma}\right) \tag{5.15}$$

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \det\,(\mathbf{COV}(\mathbf{x};\xi))\right] = \det\left(\text{block}-\text{diag}\,(\frac{\sigma_j^2\ c}{c_j}(p_j-1)) + \ \mathbf{\Sigma}\right) \tag{5.16}$$

For $\mathbf{\Sigma} = \mathbf{I}_m$, $\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \text{trace}\,(\mathbf{COV}(\mathbf{x};\xi))\right] = \sum_{j=1}^m p_j = p$.

Theorem 5.2. *Let* $\xi_{j;D}^*$ *be D- resp. G-optimal for the j-th marginal component (3.4) on the marginal design region* $\mathcal{X}_j$ *with an intercept included,* $j = 1, ..., m$, *then the product type design*

$$\xi^* = \otimes_{j=1}^m \xi_{j,D}^*$$

gives according to the trace, the maximum eigenvalue and the determinant for the covariance matrix of the prediction for the SUR model (3.3) on the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$ *with respect to the Gauß estimator, respectively the next upper boundaries*

1. $\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \text{trace}\,(\mathbf{COV}(\mathbf{x};\xi))\right] = \sum_{j=1}^m \sigma_j^2\ TC_j \leq\ p\ \text{trace}\,(\mathbf{\Sigma})$
2. $\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \boldsymbol{\lambda}_{\max}\,(\mathbf{COV}(\mathbf{x};\xi))\right] < \ \sum_{j=1}^m \sigma_j^2\ TC_j\ < p\ \text{trace}\,(\mathbf{\Sigma})$
3. $\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \det\,(\mathbf{COV}(\mathbf{x};\xi))\right] \leq \left[\frac{\sum_{j=1}^m \sigma_j^2\ TC_j}{m}\right]^m \ < \left[\frac{p}{m}\text{trace}\,(\mathbf{\Sigma})\right]^m$

Where

$$TC_j = 1 + (p_j - 1) \cdot \frac{c}{c^j}$$

where c_j is the determinant for the minor of the correlation matrix, which do not contains the column and line j.

Proof: The trace of the covariance matrix for the prediction due to $\otimes_{j=1}^{m} \xi_{j,D}^{*}$ and the Gauß estimator because of the form (5.14) by the remark 5.2 the following form

$$\min_{\xi \in \Xi} \left[\max_{\mathbf{x} \in \mathcal{X}} \text{trace} \left(\mathbf{f}(\mathbf{x})^{\top} \mathbf{M}(\xi)^{-1} \mathbf{f}(\mathbf{x}) \right) \right] = \text{trace} \left(\text{block} - \text{diag} \left(\frac{\sigma_j^2 \; c}{c_j} (p_j - 1) \right) \right) + \text{trace} \left(\; \mathbf{\Sigma} \right)$$
$$\sum_{j=1}^{m} \frac{\sigma_j^2 \; c}{c_j} (p_j - 1) + \sum_{j=1}^{m} \sigma_j^2 \; = \sum_{j=1}^{m} \sigma_j^2 \; TC_j, \; TC_j = 1 + (p_j - 1) \frac{c}{c^j}$$

$$\text{Thus} \quad \min_{\xi \in \Xi} \left[\max_{\mathbf{x} \in \mathcal{X}} \text{trace} \left(\mathbf{f}(\mathbf{x})^{\top} \mathbf{M}(\xi)^{-1} \mathbf{f}(\mathbf{x}) \right) \right] = \sum_{j=1}^{3} \sigma_j^2 \; TC_j \leq \sum_{j=1}^{m} \sigma_j^2 \; p_j \leq p \; \text{trace} \mathbf{\Sigma} \tag{5.17}$$

Where the equality occurs by $\mathbf{\Sigma} = \mathbf{I}_m$, because of trace$(\mathbf{I}_m) = m$ and $p = mp_0$. and therewith we have proven part (1) of the theorem.
The proof the part (2) of the theorem follows from applying the inequality with respect to the positive definite matrices, that the maximum eigenvalue of a positive definite matrix is less than its trace, on (5.17) , i.e.

$$\min_{\xi \in \Xi} \left[\max_{\mathbf{x} \in \mathcal{X}} \boldsymbol{\lambda}_{\max} \left(\mathbf{f}(\mathbf{x})^{\top} \mathbf{M}(\xi)^{-1} \mathbf{f}(\mathbf{x}) \right) \right] < \sum_{j=1}^{m} \sigma_j^2 \; TC_j \leq \sum_{j=1}^{m} \sigma_j^2 \; p_j < p \; \text{trace} \mathbf{\Sigma} \tag{5.18}$$

Due to applying the inequality (5.3) on (5.17) follows the proof of the part (3) of the theorem, i.e.

$$\min_{\xi \in \Xi} \left[\max_{\mathbf{x} \in \mathcal{X}} \det \left(\mathbf{f}(\mathbf{x})^{\top} \mathbf{M}(\xi)^{-1} \mathbf{f}(\mathbf{x}) \right) \right] \leq \left[\frac{\sum_{j=1}^{m} \sigma_j^2 TC_j}{m} \right]^m < \left[\frac{p}{m} \; \text{trace} \left(\mathbf{\Sigma} \right) \right]^m \quad \square \tag{5.19}$$

Before we determine the upper bounds for the applied functions to the covariance matrix for the prediction due to the $\otimes_{j=1}^{m} \xi_{j,D}^{*}$ and the limited information OLS estimator, we need th form of the covariance matrix of it due to the product type design, so that is the topics of the next lemmas, which can be proved similarly their contraries for the Gauß estimator.

Lemma 5.6. *The covariance matrix for one-point designs* $\mathbf{x}$ *for the SUR model (3.3) due to the OLS estimator has the following form*

$$\mathbf{Cov}_{OLS}(\mathbf{x}) = \left[\sigma_{jk} \; \mathbf{M}_k^{-1}(x_k) (\mathbf{f}_k(x_k) \mathbf{f}_j(x_j)^{\top}) \; \mathbf{M}_j^{-1}(x_j) \right]_{(j,k=1,\ldots,m)}, \; ((\sigma_{jk})_{j,k=1,\ldots,m} = \mathbf{\Sigma}) \tag{5.20}$$

and the block diagonal of it have the following form

$$[\mathbf{Cov}_{OLS}(\mathbf{x})]_{jj} = \left[\sigma_{jj} \; \mathbf{M}_j^{-1}(x_j) \right]_{(j=1,\ldots,m)} = \left[\sigma_j^2 \; \mathbf{M}_j^{-1}(x_j) \right]_{(j=1,\ldots,m)} \tag{5.21}$$

Proof: Because of the form of the regression function in (3.14) for the considered SUR model (3.3) has the information matrix for one-point design the following form

$$\mathbf{Cov}_{OLS}(\mathbf{x}) = \left(\mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top\right)^{-1} \left(\mathbf{f}(\mathbf{x})\boldsymbol{\Sigma}\mathbf{f}(\mathbf{x})^\top\right) \left(\mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top\right)^{-1}$$
$$= \left(\text{block} - \text{diag}\,(\mathbf{f}_j(x_j))\text{block} - \text{diag}\,(\mathbf{f}_j(x_j)^\top)\right)^{-1} \left(\text{block} - \text{diag}\,(\mathbf{f}_j(x_j))\boldsymbol{\Sigma}\text{block} - \text{diag}\,(\mathbf{f}_j(x_j)^\top\right)$$
$$\left(\text{block} - \text{diag}\,(\mathbf{f}_j(x_j))\text{block} - \text{diag}\,(\mathbf{f}_j(x_j)^\top)\right)^{-1}$$
$$= \text{block} - \text{diag}\,(\mathbf{M}_j^{-1}(x_j)) \left[\sigma_{jk}\ \mathbf{f}_k(x_k)\mathbf{f}_j(x_j)^\top\right] \text{block} - \text{diag}\,(\mathbf{M}_j^{-1}(x_j))$$
$$= \left[\sigma_{jk}\ \mathbf{M}_k^{-1}(x_k)(\mathbf{f}_k(x_k)\mathbf{f}_j(x_j)^\top)\ \mathbf{M}_j^{-1}(x_j)\right]_{(j,k=1,\ldots,m)}$$

So for $j = k$ or the the diagonal blocks have the following form

$$[\mathbf{Cov}_{OLS}(\mathbf{x})]_{jj} = \left[\sigma_{jj}\ \mathbf{M}_j^{-1}(x_j)(\mathbf{f}_j(x_j)\mathbf{f}_j(x_j)^\top)\ \mathbf{M}_j^{-1}(x_j)\right]_{(j=1,\ldots,m)} = \left[\sigma_j^2\ \mathbf{M}_j^{-1}(x_j)\right]_{(j=1,\ldots,m)}$$

and therewith the lemma has been proven. □

Lemma 5.7. *For SUR model (3.3). The covariance matrix for the product type design* $\xi = \otimes_{j=1}^m \xi_j$ *due to the OLS estimator has the following form*

$$\mathbf{Cov}_{OLS}(\xi) = \left[\sigma_{jk}\ \mathbf{M}_k^{-1}(\xi_k)(\mathbf{m}_k(\xi_k)\mathbf{m}_j(\xi_j))^\top\ \mathbf{M}_j^{-1}(\xi_j)\right]_{(j,k=1,\ldots,m)},\ \mathbf{m}_j(\xi_j) = \int_{\mathcal{X}_j} \mathbf{f}_j(x_j)\xi_j(\mathrm{d}x_j) \tag{5.22}$$

and the block diagonal of it have the following form

$$[\mathbf{Cov}_{OLS}(\xi)]_{jj} = \left[\sigma_{jj}\ \mathbf{M}_j^{-1}(\xi_j)\right]_{(j=1,\ldots,m)} = \left[\sigma_j^2\ \mathbf{M}_j^{-1}(\xi_j)\right]_{(j=1,\ldots,m)} \tag{5.23}$$

Proof: By integral of the variance covariance matrix for one-point design with respect to the product type design on the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$ in the form (5.20) by the lemma 5.6

$$\mathbf{Cov}_{OLS}(\xi) = \int_{\mathcal{X}} \mathbf{Cov}_{OLS}(\mathbf{x})\xi(\mathrm{d}\mathbf{x})$$
$$= \left[\sigma_{jk} \int_{\mathcal{X}_k\times\mathcal{X}_j} \mathbf{M}_k^{-1}(x_k)(\mathbf{f}_k(x_k)\mathbf{f}_j(x_j)^\top)\ \mathbf{M}_j^{-1}(x_j)d(\xi_k\otimes\xi_j)\right]_{(j,k=1,\ldots,m)}$$

then because of Fubini's theorem resp. (2.8)

$$\mathbf{Cov}_{OLS}(\xi) = \left[\sigma_{jk} \int_{\mathcal{X}_k} \mathbf{M}_k^{-1}(x_k)\ d\xi_k(\int_{\mathcal{X}_k} \mathbf{f}_k(x_k)\ d\xi_k \int_{\mathcal{X}_j} \mathbf{f}_j(x_j)^\top\ d\xi_j) \int_{\mathcal{X}_j} \mathbf{M}_j^{-1}(x_j)\ d\xi_j\right]_{(j,k=1,\ldots,m)} \tag{5.24}$$

$$= \left[\sigma_{jk}\ \mathbf{M}_k^{-1}(\xi_k)(\mathbf{m}_k(\xi_k)\mathbf{m}_j(\xi_j)^\top)\ \mathbf{M}_j^{-1}(\xi_j)\right]_{(j,k=1,\ldots,m)}$$

So the diagonal blocks have the following form

$$
\begin{aligned}
&[\mathbf{Cov}_{OLS}(\xi)]_{jj} = \\
&= \left[\sigma_{jj} \int_{\mathcal{X}_j} \mathbf{M}_j^{-1}(x_j)(\mathbf{f}_j(x_j)\mathbf{f}_j(x_j)^\top)\ \mathbf{M}_j^{-1}(x_j)d(\xi_j)\right]_{(j=1,...,m)} \\
&= \left[\sigma_{jj}\ \mathbf{M}_j^{-1}(\xi_j)\mathbf{M}_j(\xi_j)\ \mathbf{M}_j^{-1}(\xi_j)\right]_{(j=1,...,m)} \\
&= \left[\sigma_{jj}\ \mathbf{M}_j^{-1}(\xi_j)\right]_{(j=1,...,m)} = \left[\sigma_j^2\ \mathbf{M}_j^{-1}(\xi_j)\right]_{(j=1,...,m)}
\end{aligned}
\tag{5.25}
$$

and therewith the lemma has been proven. □

Theorem 5.3. *Let $\xi^*_{j;D}$ be D- resp. G-optimal for the j-th marginal component (3.4) on the marginal design region $\mathcal{X}_j$ with an intercept included, $j = 1, ..., m$, then the product type design*

$$\xi^*_D = \otimes_{j=1}^m \xi^*_{j,D}$$

gives according to the trace, the maximum eigenvalue and the determinant for the covariance matrix of the prediction for the SUR model (3.3) on the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$ with respect to the OLS-estimator, respectively the next upper boundaries

1. $\min_{\xi\in\Xi}[\max_{\mathbf{x}\in\mathcal{X}} \text{trace}\,(\mathbf{COV}(\mathbf{x};\xi))] = \ \sum_{j=1}^m \sigma_j^2\ p_j \leq \ p \ \text{trace}\,(\mathbf{\Sigma})$
2. $\min_{\xi\in\Xi}[\max_{\mathbf{x}\in\mathcal{X}} \boldsymbol{\lambda}_{\max}(\mathbf{COV}(\mathbf{x};\xi))] < \ \sum_{j=1}^m \sigma_j^2\ p_j \ < p\ \text{trace}\,(\mathbf{\Sigma})$
3. $\min_{\xi\in\Xi}[\max_{\mathbf{x}\in\mathcal{X}} \det(\mathbf{COV}(\mathbf{x};\xi))] \leq \ \left[\frac{\sum_{j=1}^m \sigma_j^2\ p_j}{m}\right]^m \ < \left[\frac{p}{m}\ \text{trace}\,(\mathbf{\Sigma})\right]^m$

Proof: the covariance matrix for the prediction has due to the $\xi^*_D = \otimes_{j=1}^m \xi^*_{j,D}$ and to the OLS-estimator because of (5.22) and (5.23) by the lemma 5.6

$$
\begin{aligned}
&\mathbf{f}(\mathbf{x})^\top \mathbf{Cov}_{OLS}(\xi)\mathbf{f}(\mathbf{x}) = \text{block} - \text{diag}\,(\mathbf{f}_j(x_j)^\top)_{(j=1,...,m)} \\
&\left[\sigma_{jk}\ \mathbf{M}_k^{-1}(\xi_k)(\mathbf{m}_k(\xi_k)\mathbf{m}_j(\xi_j)^\top)\ \mathbf{M}_j^{-1}(\xi_j)\right]_{(j,k=1,...,m)} \\
&\text{block} - \text{diag}\,(\mathbf{f}_j(x_j))_{(j=1,...,m)} \\
&\ \text{Then because of (5.23)}\ \min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}} \text{trace}\left(\mathbf{f}(\mathbf{x})^\top \mathbf{Cov}_{OLS}(\xi)\mathbf{f}(\mathbf{x})\right)\right] \\
&= \sum_{j=1}^m \sigma_j^2 \min_{\xi_j\in\Xi}\left[\max_{x_j\in\mathcal{X}j}\left(\mathbf{f}_j(x_j)^\top\ \mathbf{M}_j^{-1}(\xi_j)\ \mathbf{f}_j(x_j)^\top\right)\right] \\
&= \sum_{j=1}^m \sigma_j^2 \max_{x_j\in\mathcal{X}j}\left[\mathbf{f}_j(x_j)^\top\ \mathbf{M}_j^{-1}(\xi^*_{j,D})\ \mathbf{f}_j(x_j)^\top\right] = \ \sum_{j=1}^m \sigma_j^2 p_j
\end{aligned}
\tag{5.26}
$$

and therewith we have proven part (1) of the theorem.
The proof for part (2) of the theorem follows from applying the inequality with respect to the positive

definite matrices, that the maximum eigenvalue of a positive definite matrix is less than its trace, on (5.17) , i.e.

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}}\boldsymbol{\lambda}_{\max}\left(\mathbf{f}(\mathbf{x})^\top\mathbf{Cov}_{OLS}(\xi^*)\mathbf{f}(\mathbf{x})\right)\right]<\sum_{j=1}^{3}\sigma_j^2\;p_j<p\;\mathrm{trace}\boldsymbol{\Sigma} \tag{5.27}$$

Due to applying of the inequality (5.3) on (5.17) follows the proof for part (3) of the theorem, i.e.

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}}\det\left(\mathbf{f}(\mathbf{x})^\top\mathbf{Cov}_{OLS}(\xi^*)\mathbf{f}(\mathbf{x})\right)\right]\leq\left[\frac{\sum_{j=1}^{3}\sigma_j^2p_j}{m}\right]^m<\left[\frac{p}{m}\;\mathrm{trace}\,(\boldsymbol{\Sigma})\right]^m \tag{5.28}$$

(5.26), (5.27) and (5.28) can be generalized similarly for arbitrary m and therewith the theorem has been proven. □

Remark 5.3. *For* $\boldsymbol{\Sigma}=\mathbf{I}_m$*,* $\sum_{j=1}^{m}\sigma_j^2\;p_j=p$*, so in this case* $\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}}\mathrm{trace}\,(\mathbf{COV}(\mathbf{x};\xi))\right]=p$*.*

Remark 5.4. *We can remark by lemmas 5.1, 5.2, 5.3 and 5.2, that the upper bounds of the trace for the covariance matrix of the prediction are equal to the number of parameters p, if the variance covariance matrix is equal to the identity matrix, but that is not the situation for the determinant, thus may be the G-criterion due to the determinant should be standardized as follows*

$$\min_{\xi\in\Xi}\left[\max_{\mathbf{x}\in\mathcal{X}}\det\,(m\mathbf{COV}(\mathbf{x};\xi))^{\frac{1}{m}}\right] \tag{5.29}$$

So this standardized criterion have the same upper bound by applying the trace on the covariance matrix of the prediction, because of the inequality of (5.3).

5.2 Efficiency of the OLS Estimator and the MANOVA-Design

The efficiencies for the OLS estimator against the Gauß estimator due to the product type designs for the SUR models with respect to D- and linear optimal in 4.1 have been measured by theorem 5.4. As well as the efficiencies for the MANOVA-design against the product type design for the SUR models with respect to D- and linear optimal in 4.1 due to the Gauß estimator and OLS estimator respectively by theorems 5.5 and 5.6 in this section.
Another interpretations of variance covariance matrices for the Gauß and OLS estimators are helpful by calculating resp. determination the efficiencies with respect to the D-optimality, and that is the topics of the next lemmas. So the next lemma supplies the form of the transformed variance covariance matrix for the Gauß estimator with respect to the product design, and illustrates the invariance of this linear transformation with respect to the Determinant resp. the D-optimality.

Lemma 5.8. *The D-optimal criterion is invariant against the linear transformation for the SUR model (3.3), which regression functions for the corresponding marginal components models have the forms (4.8), with respect to the product type design and the Gauß estimator on the design region* $\mathcal{X}=\times_{j=1}^{m}\mathcal{X}_j$ *, if the transformations matrix have the form*

$$\mathbf{A}(\xi)=\mathrm{block}-\mathrm{diag}\,(\mathbf{A}_j(\xi_j))_{j=1,\ldots,m},\;\mathbf{A}_j(\xi_j)=\begin{pmatrix}1&\mathbf{0}\\-\mathbf{a}_j&\mathbf{I}_{p_j-1}\end{pmatrix},\;\mathbf{a}_j=\int\mathbf{g}_j(x_j)\xi_j(\mathrm{d}x_j) \tag{5.30}$$

And the inverse for the transformed information matrix has the form

$$\tilde{\mathbf{M}}_{GM}^{-1}(\xi) = \begin{pmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \text{block} - \text{diag}\,(\frac{1}{\sigma^{(jj)}}\mathbf{Q}_j^{-1}(\xi_j))_{j=1,...,m} \end{pmatrix}, \; \mathbf{Q}_j = \int \mathbf{g}_j(x_j)\mathbf{g}_j(x_j)^\top \xi_j(\mathrm{d}x_j) - \mathbf{a}_j\mathbf{a}_j^\top \tag{5.31}$$

Proof: The information matrix in (4.4) by lemma 4.2 with respect to the product design $\xi_k \otimes \xi_j$ on the design region $\mathcal{X}_k \times \mathcal{X}_j$, $k, j = 1, ..., m$ can be seen as follows

$$\mathbf{M}(\xi) = \left[\sigma^{(kj)} \int_{\mathcal{X}_k \times \mathcal{X}_j} \mathbf{f}_k(x_k)\mathbf{f}_j(x_j)^\top d(\xi_k \otimes \xi_j)\right]_{(k,j=1,...,m)}$$

So for the transformed regression function by the multiplication with the following transformation matrix of the form

$$\mathbf{A}_j(\xi_j) = \begin{pmatrix} 1 & \mathbf{0} \\ -\mathbf{a}_j & \mathbf{I}_{p_j-1} \end{pmatrix}, \; \mathbf{a}_j = \int \mathbf{g}_j(x_j)\xi_j(\mathrm{d}x_j)$$

can we orthogonalized the regression functions $\mathbf{g}_j$ for a given design ξ_j with respect to $\mathbf{f}_{j1} = 1$, so that $\tilde{\mathbf{g}}_j(x_j) = \mathbf{g}_j(x_j) - \int_{\mathcal{X}_j} \mathbf{g}_j \xi_j(\mathrm{d}x_j)$, Also, for one component look the orthogonalized regression function for $\mathbf{a}_j(\xi_j) = \int_{\mathcal{X}_j} \mathbf{g}_j \mathrm{d}xi_j$ as follows

$$\tilde{\mathbf{f}}_j(x_j) = \mathbf{A}_j(\xi_j)\mathbf{f}_j(x_j) = \begin{pmatrix} 1 & \mathbf{0} \\ -\mathbf{a}_j(\xi_j) & \mathbf{I}_{p_j-1} \end{pmatrix} \cdot \mathbf{f}_j(x_j) = \begin{pmatrix} 1 \\ \tilde{\mathbf{g}}_j(x_j) \end{pmatrix}$$

Then the diagonal blocks for the transformed information matrix for the corresponding SUR model with respect to the product type design have the following form because of $\int_{\mathcal{X}_j} \tilde{\mathbf{g}}_j(x_j) = \mathbf{0}$, $j = 1, ..., m$

$$\begin{aligned}
&\text{block} - \text{diag}\,(\sigma^{(jj)} \int_{\mathcal{X}_j} \tilde{\mathbf{f}}_j(x_j)\tilde{\mathbf{f}}_j(x_j)^\top \xi_j(\mathrm{d}x_j))_{(j=1,...,m)} \\
&= \text{block} - \text{diag}\left(\sigma^{(jj)} \int_{\mathcal{X}_j} \begin{pmatrix} 1 \\ \tilde{\mathbf{g}}_j(x_j) \end{pmatrix} \left(1 \; (\tilde{\mathbf{g}}_j(x_j))^\top\right) d(\xi_j)\right)_{(j=1,...,m)} \\
&= \text{block} - \text{diag}\left(\sigma^{(jj)} \int_{\mathcal{X}_j} \begin{pmatrix} 1 & \tilde{\mathbf{g}}_j(x_j)^\top \\ \tilde{\mathbf{g}}_j(x_j) & \tilde{\mathbf{g}}_j(x_j)\,\tilde{\mathbf{g}}_j(x_j)^\top \end{pmatrix} d(\xi_j)\right)_{(j=1,...,m)} \\
&= \text{block} - \text{diag}\left(\sigma^{(jj)} \begin{pmatrix} 1 & \int_{\mathcal{X}_j} \tilde{\mathbf{g}}_j(x_j)^\top d(\xi_j) \\ \int_{\mathcal{X}_j} \tilde{\mathbf{g}}_j(x_j) d(\xi_j) & \int_{\mathcal{X}_j} \tilde{\mathbf{g}}_j(x_j)\,\tilde{\mathbf{g}}_j(x_j)^\top d(\xi_j) \end{pmatrix}\right)_{(j=1,...,m)} \\
&= \text{block} - \text{diag}\left(\sigma^{(jj)} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \int_{\mathcal{X}_j} \mathbf{g}_j(x_j)\mathbf{g}_j(x_j)^\top d\xi_j - \mathbf{a}_j(\xi_j)\mathbf{a}_j(\xi_j)^\top \end{pmatrix}\right)_{(j=1,...,m)} \\
&= \text{block} - \text{diag}\left(\sigma^{(jj)} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_j(\xi_j) \end{pmatrix}\right)_{(j=1,...,m)} = \text{block} - \text{diag}\left(\sigma^{(jj)}\tilde{\mathbf{M}}_j(\xi_j)\right)_{(j=1,...,m)}
\end{aligned} \tag{5.32}$$

And the non-diagonal blocks, i.e. $k \neq j, k, j = 1, ..., m$ have the following form

$$\text{block}-\text{diag}\,(\sigma^{(kj)}\int_{\mathcal{X}_k\times\mathcal{X}_j}\tilde{\mathbf{f}}_k(x_k)\tilde{\mathbf{f}}_j(x_j)^\top d(\xi_k\otimes\xi_j))_{(k,j=1,\ldots,m)}$$

$$=\text{block}-\text{diag}\left(\sigma^{(kj)}\int_{\mathcal{X}_k\times\mathcal{X}_j}\begin{pmatrix}1\\ \tilde{\mathbf{g}}_k(x_k)\end{pmatrix}\left(1\ (\tilde{\mathbf{g}}_j(x_j))^\top\right)d(\xi_k\otimes\xi_j)\right)_{(k,j=1,\ldots,m)}$$

$$=\text{block}-\text{diag}\left(\sigma^{(kj)}\int_{\mathcal{X}_k\times\mathcal{X}_j}\begin{pmatrix}1 & \tilde{\mathbf{g}}_j(x_j)^\top\\ \tilde{\mathbf{g}}_k(x_k) & \tilde{\mathbf{g}}_k(x_k)\,\tilde{\mathbf{g}}_j(x_j)^\top\end{pmatrix}d(\xi_k\otimes\xi_j)\right)_{(k,j=1,\ldots,m)}$$

$$=\text{block}-\text{diag}\left(\sigma^{(kj)}\begin{pmatrix}1 & \int_{\mathcal{X}_j}\tilde{\mathbf{g}}_j(x_j)^\top d(\xi_j)\\ \int_{\mathcal{X}_k}\tilde{\mathbf{g}}_k(x_k)d(\xi_k) & \int_{\mathcal{X}_k}\tilde{\mathbf{g}}_k(x_k)\ \int_{\mathcal{X}_j}\tilde{\mathbf{g}}_j(x_j)^\top\end{pmatrix}\right)_{(k,j=1,\ldots,m)}$$

$$=\text{block}-\text{diag}\left(\begin{pmatrix}\sigma^{(kj)} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\right)_{(k,j=1,\ldots,m)} \tag{5.33}$$

So the transformed information matrix by applying some rows and columns change can be seen because of (5.32) and (5.33) as follows

$$\tilde{\mathbf{M}}(\xi)=\begin{pmatrix}\Sigma^{-1} & \mathbf{0}\\ \mathbf{0} & \text{block}-\text{diag}\,(\sigma^{(jj)}\mathbf{Q}_j(\xi_j))_{j=1,\ldots,m}\end{pmatrix},\ \mathbf{Q}_j=\int\mathbf{g}_j(x_j)\mathbf{g}_j(x_j)^\top\xi_j(\mathrm{d}x_j)-\mathbf{a}_j\mathbf{a}_j^\top$$

and therewith

$$\tilde{\mathbf{M}}_{GM}^{-1}(\xi)=\begin{pmatrix}\Sigma & \mathbf{0}\\ \mathbf{0} & \text{block}-\text{diag}\,(\frac{1}{\sigma^{(jj)}}\mathbf{Q}_j^{-1}(\xi_j))_{j=1,\ldots,m}\end{pmatrix}$$

And because of $\tilde{\mathbf{M}}(\xi)=\mathbf{A}(\xi)\ \mathbf{M}(\xi)\ \mathbf{A}^\top(\xi)$ then $\det\tilde{\mathbf{M}}(\xi)=\det\mathbf{A}(\xi)\ \det\mathbf{M}(\xi)\ \det\mathbf{A}(\xi)^\top$ so because of $\det\mathbf{A}(\xi)=1$ then $\det\tilde{\mathbf{M}}(\xi)=\ \det\mathbf{M}(\xi)$ and therewith the D-optimal design is invariant against the applied transformation and the lemma has been proven. □

The next lemma supplies the form of the transformed variance covariance matrix for the OLS estimator with respect to the product design, and illustrates the invariance of this linear transformation with respect to the Determinant resp. the D-optimality.

Lemma 5.9. *The D-optimal criterion is invariant against the linear transformation for the SUR model (3.3), which regression functions for the corresponding marginal components models have the forms (4.8), with respect to the product type design and the OLS estimator on the design region* $\mathcal{X}=\times_{j=1}^m\mathcal{X}_j$ *, if the transformations matrix have the form*

$$\mathbf{A}(\xi)=\text{block}-\text{diag}\,(\mathbf{A}_j(\xi_j))_{j=1,\ldots,m},\ \mathbf{A}_j(\xi_j)=\begin{pmatrix}1 & \mathbf{0}\\ -\mathbf{a}_j & \mathbf{I}_{p_j-1}\end{pmatrix},\ \mathbf{a}_j=\int\mathbf{g}_j(x_j)\xi_j(\mathrm{d}x_j) \tag{5.34}$$

And the transformed variance covariance matrix due to the OLS estimator have the form

$$\tilde{\mathbf{Cov}}_{OLS}(\xi)=\begin{pmatrix}\Sigma & \mathbf{0}\\ \mathbf{0} & \text{block}-\text{diag}\,(\sigma_j^2\mathbf{Q}_j^{-1}(\xi_j))_{j=1,\ldots,m}\end{pmatrix},\ \mathbf{Q}_j=\int\mathbf{g}_j(x_j)\mathbf{g}_j(x_j)^\top d\xi_j-\mathbf{a}_j\mathbf{a}_j^\top \tag{5.35}$$

Proof: The diagonal blocks for the variance covariance matrix with respect to the product type design $\xi=\otimes_{j=1}^m\xi_j$ for the SUR model (3.3) due to the OLS estimator are given in the form (5.23) so the diagonal blocks for the transformed covariance matrix have because of (5.32) the following forms

$$\begin{aligned}\left[\tilde{\mathbf{Cov}}_{OLS}(\xi)\right]_{jj} &= \text{block} - \text{diag}\,(\sigma_j^2\ \tilde{\mathbf{M}}_j^{-1}(\xi_j))_{(j=1,\ldots,m)} \\ &= \left[\begin{pmatrix} \sigma_j^2 & \mathbf{0} \\ \mathbf{0} & \text{block} - \text{diag}\,(\sigma_j^2 \mathbf{Q}_j^{-1}(\xi_j))_{j=1,\ldots,m} \end{pmatrix}\right]_{(j=1,\ldots,m)}\end{aligned} \tag{5.36}$$

The non-diagonal blocks for the transformed variance covariance matrix can be seen with respect to the OLS estimator in (5.24) by lemma 5.7 with respect to the product design $\xi_k \otimes \xi_j$ on the design region $\mathcal{X}_k \times \mathcal{X}_j$, $k \neq j, k, j = 1, \ldots, m$, as follows

$$\left[\tilde{\mathbf{Cov}}_{OLS}(\xi)\right]_{k,j,k\neq j} =$$

$$\left[\sigma_{jk} \int_{\mathcal{X}_k \times \mathcal{X}_j} \tilde{\mathbf{M}}_k^{-1}(x_k)(\tilde{\mathbf{f}}_k(x_k)\tilde{\mathbf{f}}_j(x_j))^\top\ \tilde{\mathbf{M}}_j^{-1}(x_j) d(\xi_k \otimes \xi_j)\right]_{(j,k=1,\ldots,m)} =$$

$$\left[\sigma_{jk} \int_{\mathcal{X}_k} \tilde{\mathbf{M}}_k^{-1}(x_k) d(\xi_k) \left(\int_{\mathcal{X}_k} \tilde{\mathbf{f}}_k(x_k) d(\xi_k) \int_{\mathcal{X}_j} \tilde{\mathbf{f}}_j(x_j)^\top d(\xi_j)\right) \int_{\mathcal{X}_j} \tilde{\mathbf{M}}_j^{-1}(x_j) d(\xi_j) d(\xi_j)\right]$$

Then because of (5.32) and (5.33)

$$\begin{aligned}&\left[\tilde{\mathbf{Cov}}_{OLS}(\xi)\right]_{k,j,\ k\neq j} \\ &= \left[\sigma_{jk} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_j^{-1}(\xi_j) \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_k^{-1}(\xi_k) \end{pmatrix}\right] \\ &= \left[\sigma_{jk} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\right]_{(j,k=1,\ldots,m)}\end{aligned} \tag{5.37}$$

So the transformed variance covariance matrix for the OLS estimator can be seen, by applying some rows and columns change, because of (5.36) and (5.37) as follows

$$\tilde{\mathbf{Cov}}_{OLS}(\xi) = \begin{pmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \text{block} - \text{diag}\,(\sigma_j^2 \mathbf{Q}_j^{-1}(\xi_j))_{j=1,\ldots,m} \end{pmatrix}$$

and therewith the limited information matrix for the OLS estimator has the following form

$$\tilde{\mathbf{M}}_{OLS}(\xi) = \begin{pmatrix} \Sigma^{-1} & \mathbf{0} \\ \mathbf{0} & \text{block} - \text{diag}\,(\frac{1}{\sigma_j^2} \mathbf{Q}_j(\xi_j))_{j=1,\ldots,m} \end{pmatrix}$$

And because of $\tilde{\mathbf{M}}_{OLS}(\xi) = \mathbf{A}(\xi)\,\mathbf{M}_{OLS}(\xi)\,\mathbf{A}(\xi)^\top$ then $\det \tilde{\mathbf{M}}_{OLS}(\xi) = \det \mathbf{A}(\xi)\ \det \mathbf{M}_{OLS}(\xi)\ \det \mathbf{A}(\xi)^\top$ so because of $\det \mathbf{A}(\xi) = 1$ then $\det \tilde{\mathbf{M}}_{OLS}(\xi) = \det \mathbf{M}_{OLS}(\xi)$ and therewith the D-optimal design is invariant against the applied transformation with respect to the OLS estimator and the lemma has been proven. □

Theorem 5.4. *The efficiencies of the OLS estimator versus the Gauß estimator for the SUR model (3.3) in the design region* $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$ *for* $\xi = \otimes_{j=1}^m \xi_j$ *are according to D-, IMSE- and A-optimal criteria respectively as follows*

- *eff*$_D(\hat{\boldsymbol{\beta}}_{OLS}(\xi)) = \left(\frac{c^m}{\prod_{j=1}^m c^j}\right)^{1/p}$
- *eff*$_{IMSE}(\hat{\boldsymbol{\beta}}_{OLS}(\xi)) = \frac{\sum_{j=1}^m \sigma_j^2(1-\frac{c}{c_j})L_{j11}+\sigma_j^2\frac{c}{c_j}\text{trace}\left(\mathbf{L}_j\mathbf{M}_j^{-1}(\xi_j^*)\right)}{\sum_{j=1}^m \sigma_j^2\text{trace}(\mathbf{L}_j\mathbf{M}^{-1}(\xi_j))}$
- *eff*$_A(\hat{\boldsymbol{\beta}}_{OLS}(\xi)) = \frac{\sum_{j=1}^m \sigma_j^2(1-\frac{c}{c_j})+\sigma_j^2\frac{c}{c_j}\text{trace}\left(\mathbf{M}_j^{-1}(\xi_j^*)\right)}{\sum_{j=1}^m \sigma_j^2\text{trace}\mathbf{M}(\xi_j)^{-1}}$

Where $\xi = \otimes_{j=1}^m \xi_j$ *and* c_j *is the determinant for the minor of the correlation matrix, which do not contains the column and line* j.

Proof: the D-efficiency for the OLS estimator versus the Gauß estimator has the following form for $\xi = \otimes_{j=1}^m \xi_j$, because of the forms of the transformed variance covariance matrices for the Gauß estimator in (5.31) and the OLS estimator in (5.35)

$$\begin{aligned}\text{eff}_D(\hat{\boldsymbol{\beta}}_{OLS}(\xi)) &= \left(\frac{\det \mathbf{M}_{GM}^{-1}(\xi)}{\det \mathbf{M}_{OLS}^{-1}(\xi)}\right)^{1/p} = \left(\frac{\det \boldsymbol{\Sigma}\prod_{j=1}^m \frac{1}{\sigma^{(jj)}}\det \mathbf{Q}_j^{-1}(\xi_j)}{\det \boldsymbol{\Sigma}\prod_{j=1}^m \sigma_j^2 \det \mathbf{Q}_j^{-1}(\xi_j)}\right)^{1/p} \\ &= \left(\frac{1}{\prod_{j=1}^m \sigma_j^2\sigma^{(jj)}}\right)^{1/p} \quad \text{so because of (5.11)}\end{aligned}$$

Then

$$\text{eff}_D(\hat{\boldsymbol{\beta}}_{OLS}(\xi)) = \left(\frac{1}{\frac{(\sigma_1^2...\sigma_m^2 c^1...c^m)}{(\sigma_1^2...\sigma_m^2(c)^m)}}\right)^{1/p} = \left(\frac{c^m}{\prod_{j=1}^m c^j}\right)^{1/p}$$

The efficiencies according to linear criteria with respect to the Gauß estimator because of (4.18) have the form trace $\left(\mathbf{L}\ \mathbf{M}(\xi^*)_{GM}^{-1}\right) = \sum_{j=1}^m \frac{1}{\sigma^{(jj)}}\mathbf{L}_j\mathbf{M}_j^{-1}(\xi_j) + \sum_{j=1}^m L_{j11}(\sigma_j^2 - \frac{1}{\sigma^{(jj)}})$ We have the following term, because of the block diagonal form of the weight matrix $\mathbf{L}$ given in (4.14) and the form of the diagonal blocks of the variance covariance matrix for the OLS estimator given in (5.21)

$$\text{trace}\left(\mathbf{L}\ \mathbf{M}_{OLS}^{-1}(\xi)\right) = \text{trace}\left(\text{block}-\text{diag}\left(\mathbf{L}_j\right)_{(j=1,...,m)}\ \mathbf{M}_{OLS}^{-1}(\xi)\right) = \sum_{j=1}^m \sigma_j\ \mathbf{L}_j\ \mathbf{M}_j^{-1}(\xi_j) \qquad (5.38)$$

Know, we can calculate the IMSE-optimality efficiency for the OLS estimator versus the Gauß estimator with respect to the product type designs, due to (4.18) and (5.38), as follows

$$\text{eff}_{IMSE}(\hat{\boldsymbol{\beta}}_{OLS}(\xi)) = \frac{\text{trace}\left(\mathbf{L}\ \mathbf{M}_{GM}^{-1}(\xi)\right)}{\text{trace}\left(\mathbf{L}\ \mathbf{M}_{OLS}^{-1}(\xi)\right)} = \frac{\sum_{j=1}^m \sigma_j^2(1-\frac{c}{c_j})L_{j11} + \sigma_j^2\frac{c}{c_j}\text{trace}\left(\mathbf{L}_j\mathbf{M}_j^{-1}(\xi_j^*)\right)}{\sum_{j=1}^m \sigma_j^2\text{trace}\left(\mathbf{L}_j\mathbf{M}^{-1}(\xi_j)\right)} \qquad (5.39)$$

we can obtain the A-optimality efficiency for the OLS estimator versus the Gauß estimator, due to replacing $\mathbf{L}_j = \mathbf{I}_{p_j\times p_j}$ and $L_{j11} = 1$ by (5.39), as follows

$$\text{eff}_A(\hat{\boldsymbol{\beta}}_{OLS}(\xi)) = \frac{\text{trace}\mathbf{M}_{GM}^{-1}(\xi)}{\text{trace}\mathbf{M}_{OLS}^{-1}(\xi)} = \frac{\sum_{j=1}^m \sigma_j^2(1-\frac{c}{c_j}) + \sigma_j^2\frac{c}{c_j}\text{trace}\left(\mathbf{M}_j^{-1}(\xi_j^*)\right)}{\sum_{j=1}^m \sigma_j^2\text{trace}\mathbf{M}(\xi_j)^{-1}}$$

The product type designs for SUR models with intercepts by the marginal models are D- and linear optimal in 4.1, as it has been proven in chapter four. So the next result illustrates the efficiencies for the MANOVA-design vs the product type design due to the Gauß estimator and with respect to the D- and linear criteria in 4.1

Def 5.2. $\xi_{MANOVA}(\mathbf{x}) = \xi_0(x)$ *for* $\mathbf{x} = (x, ..., x)$ *[all components are equal] and* $\xi_{MANOVA}(\mathbf{x}) = 0$ *else.*

Theorem 5.5. *Let* $\mathbf{f}_{j,j=1,...,m} = \mathbf{f}_0$ *be the marginal regression functions for the components for the SUR model (3.3) , which coincide with the regression function for one component of the MANOVA model see (3.4), in the design region* $\mathcal{X} = \times_{j=1}^{m}\mathcal{X}_0$*, then the efficiencies of the MANOVA-design* ξ_0 *in the design region* $\mathcal{X}$ *versus the Product type design* $\xi = \otimes_{j=1}^{m}\xi_j = \otimes_{j=1}^{m}\xi_0$*, according to D-, IMSE- and A-optimal criteria and due to the Gauß estimator are respectively as follows*

- *eff*$_D(\xi_0) = \frac{(c)^{\frac{m(m+1)-p}{pm}} \left(\prod_{j=1}^{m}\sigma_j^2\right)^{\frac{2m-p}{pm}}}{\left(\prod_{j=1}^{m} c_j\right)^{1/p}}$
- *eff*$_{IMSE}(\xi_0) = \frac{\sum_{j=1}^{m}\sigma_j^2(1-\frac{c}{c_j})L_{011}+\sigma_j^2\frac{c}{c_j}\mathrm{trace}(\mathbf{L}_0\mathbf{M}_0^{-1}(\xi_0^*))}{\mathrm{trace}\boldsymbol{\Sigma}\ \mathrm{trace}(\mathbf{L}_0\ \mathbf{M}_0^{-1}(\xi_0))}$
- *eff*$_A(\xi_0) = \frac{\sum_{j=1}^{m}\sigma_j^2(1-\frac{c}{c_j})+\sigma_j^2\frac{c}{c_j}\mathrm{trace}(\mathbf{M}_0^{-1}(\xi_0^*))}{\mathrm{trace}\boldsymbol{\Sigma}\ \mathrm{trace}\ \mathbf{M}_0^{-1}(\xi_0)}$

Where $\mathbf{M}_0 = \int \mathbf{f}_0\mathbf{f}_0^\top \mathrm{d}\xi_0$*,* $\mathbf{L}_0 = \int_{\mathcal{X}_0} \mathbf{f}_0(x)\mathbf{f}_0(x)^\top \mu(\mathrm{d}x)$ *and* L_{011} *is the first diagonal element of the matrix* $\mathbf{L}_0$*. And* c_j *is the determinant for the minor of the correlation matrix, which do not contains the column and line* j*.*

Proof: The variance covariance matrix for the MANOVA-model resp. for MANOVA-design ξ_0 for regression function with intercept, i.e.

$$\mathbf{f}_0(x_0) = \begin{pmatrix} 1 \\ \mathbf{g}_0(x_0) \end{pmatrix}$$

can be calculated by the integral of the information matrix for the one-point design in (3.17) with respect to design ξ_0 so

$$\mathbf{M}(\xi_0) = \boldsymbol{\Sigma}^{-1} \otimes \mathbf{M}_0(\xi_0) \tag{5.40}$$

Then

$$\mathbf{M}^{-1}(\xi_0) = \boldsymbol{\Sigma} \otimes \mathbf{M}_0^{-1}(\xi_0) \tag{5.41}$$

then because of (5.31)

$$\tilde{\mathbf{M}}^{-1}(\xi_0) = \boldsymbol{\Sigma} \otimes \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_0^{-1}(\xi_0) \end{pmatrix}, \text{ Where } \mathbf{Q}_0(\xi_0) = \int \mathbf{g}_0(x_0)\mathbf{g}_0(x_0)^\top \xi_0(\mathrm{d}x_0) - \mathbf{a}_0\mathbf{a}_0^\top \tag{5.42}$$

Also, for SUR model with the control variables $x_{j,j=1,...,m} = x_0$ and the marginal regression functions for the components, which are $\mathbf{f}_{j,j=1,...,m} = \mathbf{f}_0$ has the transformed covariance matrix for the Gauß estimator with respect to the product design $\xi = \otimes_{j=1}^m \xi_j = \otimes_{j=1}^m \xi_0$ because of (5.31) the following form

$$\tilde{\mathbf{M}}_{GM}^{-1}(\xi) = \begin{pmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \text{block} - \text{diag}\,(\frac{1}{\sigma^{(jj)}}\mathbf{Q}_0^{-1}(\xi_0))_{j=1,\ldots,m} \end{pmatrix}, \quad \mathbf{Q}_0 = \int \mathbf{g}_0(x_0)\mathbf{g}_0(x_0)^\top \xi_0(\mathrm{d}x_0) - \mathbf{a}_0\mathbf{a}_0^\top \tag{5.43}$$

and the variance covariance matrix for it because of the form (4.9) has the following form

$$\mathbf{M}_{GM}^{-1}(\xi) = \text{block} - \text{diag}\,(\frac{1}{\sigma^{(jj)}}(\mathbf{M}_j^{-1}(\xi_0) - \mathbf{e}_j\mathbf{e}_j^\top) + \mathbf{e}\,\boldsymbol{\Sigma}\,\mathbf{e}^\top \tag{5.44}$$

So because of (5.9) and (5.11) $\frac{1}{\sigma^{(jj)}} = \frac{\sigma_j^2 \det \mathbf{C}}{c_j}$, and because of $\det \boldsymbol{\Sigma} = \prod_{j=1}^m \sigma_j^2\ c$, and because of (5.42) and (5.43) have the D-efficiency for the MANVOVA-design against the product type design with respect to the Gauß estimator the following form

$$\begin{aligned}
\text{eff}_D(\xi_0) &= \left(\frac{\det \mathbf{M}_{GM}^{-1}(\xi)}{\det \mathbf{M}^{-1}(\xi_0)}\right)^{1/p} \\
&= \left(\frac{\det \tilde{\mathbf{M}}_{GM}^{-1}(\xi)}{\det \tilde{\mathbf{M}}^{-1}(\xi_0)}\right)^{1/p} \\
&= \left(\frac{\det \boldsymbol{\Sigma} \prod_{j=1}^m \frac{1}{\sigma^{(jj)}} \det \mathbf{Q}_0^{-1}(\xi_0)}{(\det \boldsymbol{\Sigma})^{\frac{p}{m}} \left(\det \mathbf{Q}_0^{-1}(\xi_0)\right)^m}\right)^{1/p} \\
&= \left(\frac{c\left(\prod_{j=1}^m \sigma_j^2\right) \prod_{j=1}^m \frac{\sigma_j^2 c}{c_j} \det \mathbf{Q}_0^{-1}(\xi_0)}{\left(c\prod_{j=1}^m \sigma_j^2\right)^{\frac{p}{m}} \left(\det \mathbf{Q}_0^{-1}(\xi_0)\right)^m}\right)^{1/p} \\
&= \left(\frac{(c)^{m+1}\left(\det \mathbf{Q}_0^{-1}(\xi_0)\right)^m \left(\prod_{j=1}^m \sigma_j^2\right)^2 \prod_{j=1}^m \frac{1}{c_j}}{(c)^{\frac{p}{m}} \left(\prod_{j=1}^m \sigma_j^2\right)^{\frac{p}{m}} \left(\det \mathbf{Q}_0^{-1}(\xi_0)\right)^m}\right)^{1/p} \\
&= \left(\frac{(c)^{m+1}\left(\prod_{j=1}^m \sigma_j^2\right)^2}{\left(\prod_{j=1}^m c_j\right)(c)^{\frac{p}{m}} \left(\prod_{j=1}^m \sigma_j^2\right)^{\frac{p}{m}}}\right)^{1/p} \\
&= \left(\frac{(c)^{\frac{m(m+1)-p}{m}} \left(\prod_{j=1}^m \sigma_j^2\right)^{\frac{2m-p}{m}}}{\prod_{j=1}^m c_j}\right)^{1/p} \\
&= \frac{(c)^{\frac{m(m+1)-p}{pm}} \left(\prod_{j=1}^m \sigma_j^2\right)^{\frac{2m-p}{pm}}}{\left(\prod_{j=1}^m c_j\right)^{1/p}}
\end{aligned}$$

Also, for $m = 2, c^1 = c^2 = 1$, $\text{eff}_D(\xi_0) = (c)^{\frac{6-p}{2p}} \left(\prod_{j=1}^2 \sigma_j^2\right)^{\frac{4-p}{2p}}$
To calculate the IMSE-efficiency for the MANOVA-design ξ_0, we should calculate the matrix $\mathbf{L}$ for the

MANOVA-model, also

$$\mathbf{L} = \mathbf{I}_{m\times m} \otimes \int_{\mathcal{X}_0} \mathbf{f}_0(x_0)\mathbf{f}_0(x_0)^\top \mu(\mathrm{d}x_0) = \mathbf{I}_{m\times m} \otimes \mathbf{L}_0 \ , \ \mathbf{L}_0 = \begin{pmatrix} L_{011} & \mathbf{L}_{012}^\top \\ \mathbf{L}_{012} & \mathbf{L}_{022} \end{pmatrix} \tag{5.45}$$

We get the next term because of (5.41) and by replacing $\mathbf{L}_j = \mathbf{L}_0$ by (4.14)

$$\begin{aligned} &\text{trace}\left(\mathbf{L}\mathbf{M}^{-1}(\xi_0)\right) = \text{trace}\left((\mathbf{I}_{m\times m} \otimes \mathbf{L}_0)\left(\mathbf{\Sigma} \otimes \mathbf{M}_0^{-1}(\xi_0)\right)\right) \\ &= \text{trace}\left(\mathbf{\Sigma} \otimes \mathbf{L}_0\mathbf{M}_0^{-1}(\xi_0)\right) \\ &= \text{trace}\mathbf{\Sigma}\ \text{trace}\left(\mathbf{L}_0\mathbf{M}_0^{-1}(\xi_0)\right) \end{aligned} \tag{5.46}$$

For SUR model with respect to Gauß estimator has the term trace $\left(\mathbf{L}\mathbf{M}_{GM}^{-1}(\otimes_{j=1}^m \xi_0)\right)$, where $\mathbf{L} = \text{diag}\,(\mathbf{L}_0)_{j=1,\dots,m}$ and $\mathbf{L}_0$ is defined in (5.45) and because of (4.18)

$$\text{trace}\left(\mathbf{L}\ \mathbf{M}(\xi)^{-1}\right) = \sum_{j=1}^m \frac{1}{\sigma^{(jj)}}\mathbf{L}_0\mathbf{M}_0^{-1}(\xi_0) + \sum_{j=1}^m L_{011}(\sigma_j^2 - \frac{1}{\sigma^{(jj)}}) \tag{5.47}$$

then because of (5.46) and (5.47)

$$\text{eff}_{IMSE}(\xi_0) = \frac{\sum_{j=1}^m \sigma_j^2(1-\frac{c}{c_j})L_{011} + \sigma_j^2\frac{c}{c_j}\text{trace}\left(\mathbf{L}_0\mathbf{M}_0^{-1}(\xi_0^*)\right)}{\text{trace}\mathbf{\Sigma}\ \text{trace}\left(\mathbf{L}_0\ \mathbf{M}_0^{-1}(\xi_0)\right)} \tag{5.48}$$

Due to replacing $\mathbf{L}_0 = \mathbf{I}_{\frac{p}{m}\times\frac{p}{m}}$ and $L_{011} = 1$ by (5.48) we can obtain the A-optimality efficiency for ξ_0 versus the product design $\otimes_{j=1}^m \xi_0$ with respect to the Gauß estimator as follows

$$\text{eff}_A(\xi_0) = \frac{\sum_{j=1}^m \sigma_j^2(1-\frac{c}{c_j}) + \sigma_j^2\frac{c}{c_j}\text{trace}\left(\mathbf{M}_0^{-1}(\xi_0^*)\right)}{\text{trace}\mathbf{\Sigma}\ \text{trace}\ \mathbf{M}_0^{-1}(\xi_0)}$$

and therewith we have ended the proof. □

By the next theorem can be shown, that the MANOVA-design is more efficient than the product type design due to the OLS estimator with respect to the D-optimality, and the both designs have the same efficiency for the linear optimality in 4.1 due to the OLS-estimator.

Theorem 5.6. *Let $\mathbf{f}_{j,j=1,\dots,m} = \mathbf{f}_0$ be the marginal regression functions for the components for the SUR model (3.3) , which coincide with the regression function for one component of the MANOVA model, in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_0$, then the efficiencies of the MANOVA-design ξ_0 in the design region $\mathcal{X}_0$ versus the Product type design $\xi = \otimes_{j=1}^m \xi_j = \otimes_{j=1}^m \xi_0$, according to D-, IMSE- and A-optimal criteria and due to the OLS estimator are respectively as follows*

- $\mathit{eff}_D(\xi_0) = \left(\prod_{j=1}^m \sigma_j^2\right)^{\frac{2m-p}{pm}} (c)^{\frac{m-p}{pm}}$
- $\mathit{eff}_{IMSE}(\xi_0) = 1$
- $\mathit{eff}_A(\xi_0) = 1$

Proof: The proof of the theorem 5.6 is analogous to the proof of theorem 5.5 with the distinction by the terms of the variance covariance matrices for OLS estimator and the Gauß estimator, also, for the product design $\xi = \otimes_{j=1}^{m}\xi_j = \otimes_{j=1}^{m}\xi_0$ the transformed variance covariance matrix for the OLS estimator has the following form, because of (5.35)

$$\tilde{\mathbf{Cov}}_{OLS}(\xi) = \begin{pmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \text{block} - \text{diag}\,(\sigma_j^2 \mathbf{Q}_0^{-1}(\xi_0))_{j=1,\dots,m} \end{pmatrix}$$

then

$$\det \tilde{\mathbf{Cov}}_{OLS}^{-1}(\xi) = \det \mathbf{\Sigma} \left(\det \mathbf{Q}_0^{-1}(\xi_0)\right)^m \prod_{j=1}^{m} \sigma_j^2 \tag{5.49}$$

Then because of (5.43), (5.49) and (5.9) has the efficiency of ξ_0 versus the product design $\xi = \otimes_{j=1}^{m}\xi_j = \otimes_{j=1}^{m}\xi_0$ with respect to the OLS-estimator the next form

$$\begin{aligned}
\text{eff}_D(\xi_0) &= \left(\frac{\det \mathbf{Cov}_{OLS}^{-1}(\xi)}{\det \mathbf{M}^{-1}(\xi_0)}\right)^{1/p} = \left(\frac{\det \tilde{\mathbf{Cov}}_{OLS}^{-1}(\xi)}{\det \tilde{\mathbf{M}}^{-1}(\xi_0)}\right)^{1/p} \\
&= \left(\frac{\det \mathbf{\Sigma}\ \left(\det \mathbf{Q}_0^{-1}(\xi_0)\right)^m \prod_{j=1}^{m} \sigma_j^2}{(\det \mathbf{\Sigma})^{\frac{p}{m}} \left(\det \mathbf{Q}_0^{-1}(\xi_0)\right)^m}\right)^{1/p} \\
&= \left(\frac{\prod_{j=1}^{m} \sigma_j^2}{(\det \mathbf{\Sigma})^{\frac{p-m}{m}}}\right)^{1/p} = \left(\frac{\prod_{j=1}^{m} \sigma_j^2}{\left(c \prod_{j=1}^{m} \sigma_j^2\right)^{\frac{p-m}{m}}}\right)^{1/p} \\
&= \left(\frac{\left(\prod_{j=1}^{m} \sigma_j^2\right)^{\frac{2m-p}{m}}}{(c)^{\frac{p-m}{m}}}\right)^{1/p} = \left(\prod_{j=1}^{m} \sigma_j^2\right)^{\frac{2m-p}{pm}} (c)^{\frac{m-p}{pm}}
\end{aligned}$$

Also, for $m = 2, \text{eff}_D(\xi_0) = \left(\prod_{j=1}^{2} \sigma_j^2\right)^{\frac{4-p}{2p}} (c)^{\frac{2-p}{2p}}$

To calculate the IMSE-efficiency for the MANOVA-design ξ_0 versus the product design $\otimes_{j=1}^{m}\xi_0$, we have for the SUR model with respect to the OLS estimator the term trace $\left(\mathbf{L}\mathbf{M}_{OLS}^{-1}(\otimes_{j=1}^{m}\xi_0)\right)$, where $\mathbf{L} = \text{diag}\,(\mathbf{L}_0)_{j=1,\dots,m}$ and $\mathbf{L}_0$ is defined in (5.45) because of (5.38) and (5.46) the following form

$$\text{trace}\left(\mathbf{L}\ \mathbf{Cov}_{OLS}^{-1}(\xi)\right) = \sum_{j=1}^{m} \sigma_j^2 \text{trace}\left(\mathbf{L}_0\mathbf{M}_0^{-1}(\xi_0)\right) = \text{trace}\mathbf{\Sigma}\ \text{trace}\left(\mathbf{L}_0\mathbf{M}_0^{-1}(\xi_0)\right) \tag{5.50}$$

then because of (5.46) and (5.50)

$$\text{eff}_{IMSE}(\xi_0) = \frac{\text{trace}\mathbf{\Sigma}\ \text{trace}\left(\mathbf{L}_0\mathbf{M}_0(\xi_0)^{-1}\right)}{\text{trace}\mathbf{\Sigma}\ \text{trace}\left(\mathbf{L}_0\mathbf{M}_0(\xi_0)^{-1}\right)} = 1 \tag{5.51}$$

By replacing $\mathbf{L}_0 = \mathbf{I}_{\frac{p}{m}\times\frac{p}{m}}$ by (5.51), we obtain the A-optimality efficiency for ξ_0 versus the product design $\otimes_{j=1}^{m}\xi_0$ with respect to the OLS estimator as follows

$$\text{eff}_A(\xi_0) = \frac{\text{trace } \mathbf{M}_0^{-1}(\xi_0)}{\text{trace } \mathbf{M}_0^{-1}(\xi_0)} = 1$$

and therewith we have ended the proof. □

Remark 5.5. *The results for the IMSE-optimality can be extended for linear optimality in 4.1.*

Remark 5.6. *The MANOVA-design ξ_0 is more efficient than product type design $\otimes_{j=1}^{m}\xi_0$ with respect to the OLS estimator according to the D-criterion, for example, by $m = p = 2$ and $\sigma_j^2 = 4$ and $\rho = 0$ then $c = 1$ has the efficiency the next term $\text{eff}_D(\xi_0) = (4)^{\frac{1}{2}} = 2$. And that is logic, because the OLS estimator is limited information or limited efficient estimator. But the both estimators are equal efficient with respect to the linear criteria in 4.1.*

5.3 Example: Bivariate straight line regression

To illustrate the results, we can consider the SUR model with simple straight line regression models for the components,

$$Y_{ij} = \beta_{j0} + \beta_{j1}x_{ij} + \varepsilon_{ij}\,. \tag{5.52}$$

in the unit interval $\mathcal{X}_1 = \mathcal{X}_2 = [0,1]$ as experimental regions. Then it is well-known that the D-, IMSE- and A-optimal designs for the marginal models $\xi_1^* = \xi_2^* = \begin{pmatrix} 0 & 1 \\ w_1 & w_2 \end{pmatrix}$ where the weights $w_1 = w_2 = 1/2$ for the D- and IMSE-optimal designs and $w_1 = 2-\sqrt{2}, w_2 = 1/2$ for the A-optimal design. Because of the theorems 4.3 and 4.4 in the previous chapter, the product type designs

$$\xi^* = \xi_1^* \otimes \xi_2^* = \begin{pmatrix} (1,1) & (0,0) & (1,0) & (0,1) \\ w_1^2 & w_1.w_2 & w_1.w_2 & w_2^2 \end{pmatrix}$$

are D-, IMSE-, and A-optimal designs for the SUR model (6.34) on $\mathcal{X} = [0,1]^2$.
An obvious alternative for the product type design would be a multivariate linear regression design

$$\xi_0 = \begin{pmatrix} (1,1) & (0,0) \\ w_1 & w_2 \end{pmatrix},$$

where the weights $w_1 = w_2 = 1/2$ for the D- and IMSE-optimal designs and $w_1 = 2-\sqrt{2}, w_2 = 1/2$ for the A-optimal design and $x_1 = x_2$ are required and the corresponding marginals of ξ_0 are optimal in the marginal models. The statistical analysis is simplified for such a design, because the Gauß-Markov estimator reduces to ordinary least squares for any $\mathbf{\Sigma}$. At first we want to illustrate the results for the upper bounds of the variance covariance matrix for the prediction by Lemma (5.2) and the theorems (5.2) and (5.3), where in the bivariate case $c^1 = c^2 = 1$ and also for Lemma (5.2) the graphics for the next functions, which are on the left side

1. $\frac{\max_{x\in\mathcal{X}} \text{trace}(\mathbf{COV}(x;\xi_{0;D}^*))}{\frac{p}{m}\,\text{trace}(\mathbf{\Sigma})} \leq 1$ for $\sigma_1 = \sigma_2 = 5$

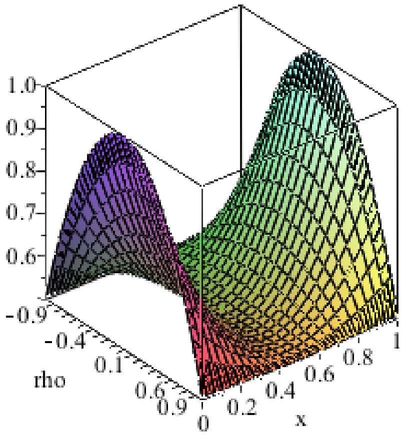

Figure 5.1: Max.= 1 with respect to the trace (MANOVA)

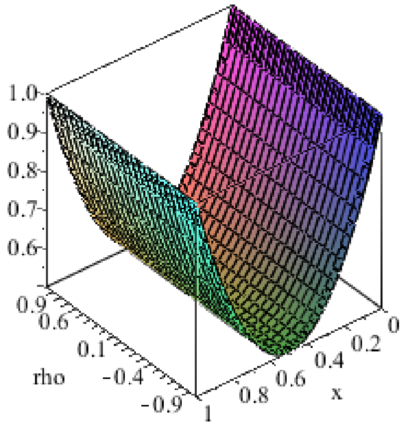

Figure 5.2: Max.= 1 with respect to the maximum eigenvalue (MANOVA)

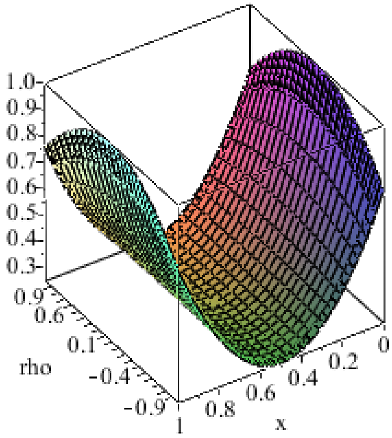

Figure 5.3: Max.= 1 with respect to the determinant (MANOVA)

2. $\dfrac{\max_{x\in\mathcal{X}_0} \boldsymbol{\lambda}_{\max}\big(\mathbf{COV}(x;\xi^*_{0;D})\big)}{\frac{p}{m}\boldsymbol{\lambda}_{\max}(\boldsymbol{\Sigma})} \leq 1$ for $\sigma_1 = 0.1, \sigma_2 = 10$

3. $\dfrac{\max_{x\in\mathcal{X}_0} \det\big(\mathbf{COV}(x;\xi^*_{0;D})\big)}{\left(\frac{p}{m}\right)^m \det \boldsymbol{\Sigma}} =$
$\left(-4\,\rho^2+4\right)x^4+\left(8\,\rho^2-8\right)x^3+\left(-6\,\rho^2+8\right)x^2+\left(2\,\rho^2-4\right)x-1/4\,\rho^2+1 \leq 1$

have respectively a maximum of 1, by which can be illustrated by figures 5.1, 5.2 and 5.3. respectively the graphics for the next functions with respect to the theorem (5.2) , which are on the left side

1. $\dfrac{\max_{\mathbf{x}\in\mathcal{X}} \operatorname{trace}\big(\mathbf{COV}(\mathbf{x};\xi^*_D)\big)}{\sum_{j=1}^m \sigma_j^2 TC_j} \leq 1$ for $x_2 = 1, \sigma_1 = 0.1, \sigma_2 = 10$

2. $\dfrac{\max_{\mathbf{x}\in\mathcal{X}} \boldsymbol{\lambda}_{\max}\big(\mathbf{COV}(\mathbf{x};\xi^*_D)\big)}{\sum_{j=1}^m \sigma_j^2 TC_j} < 1$ for $\rho = 0.99, \sigma_1 = 0.1, \sigma_2 = 10$

3. $\dfrac{\max_{\mathbf{x}\in\mathcal{X}} \det\big(\mathbf{COV}(\mathbf{x};\xi^*_D)\big)}{\left[\frac{\sum_{j=1}^m \sigma_j^2 TC_j}{m}\right]^m} \leq 1$ for $x_2 = 1, \sigma_1 = 5, \sigma_2 = 5$

have respectively a maximum of 1, by which can be illustrated by the figures 5.4, 5.5 and 5.6. respectively the graphics for the next functions with respect to the theorem (5.3), which are on the left side

1. $\dfrac{\max_{\mathbf{x}\in\mathcal{X}} \operatorname{trace}\big(\mathbf{COV}(\mathbf{x};\xi^*_D)\big)}{\sum_{j=1}^m \sigma_j^2 p_j} \leq 1$ for $\rho = 0, \sigma_1 = \sigma_2 = 5$

2. $\dfrac{\max_{\mathbf{x}\in\mathcal{X}} \boldsymbol{\lambda}_{\max}\big(\mathbf{COV}(\mathbf{x};\xi^*_D)\big)}{\sum_{j=1}^m \sigma_j^2 p_j} < 1$ for $\rho = 0, \sigma_1 = 10, \sigma_2 = 0.1$

3. $\dfrac{\max_{\mathbf{x}\in\mathcal{X}} \det\big(\mathbf{COV}(\mathbf{x};\xi^*_D)\big)}{\left[\sum_{j=1}^m \sigma_j^2 p_j\right]^m} \leq 1$ for $\rho = 0, \sigma_1 = \sigma_2 = 5$

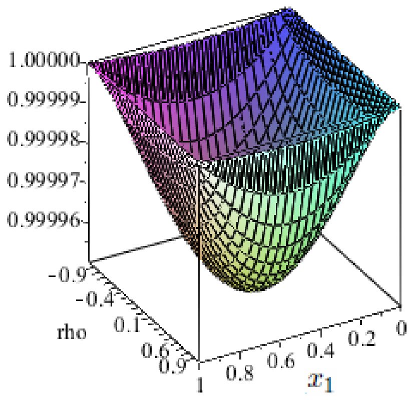

Figure 5.4: Max.= 1 with respect to the trace (Product type design by Gauß Markov)

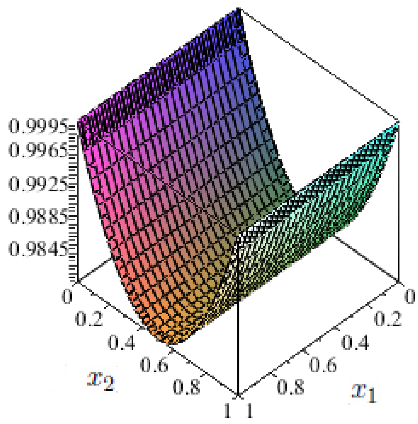

Figure 5.5: Max.< 1 with respect to the maximum eigenvalue (Product type design by Gauß Markov)

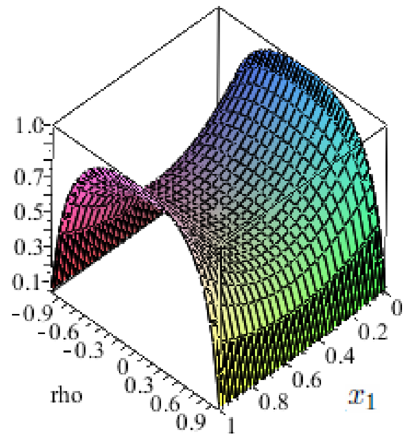

Figure 5.6: Max.= 1 with respect to the determinant (Product type design by Gauß Markov)

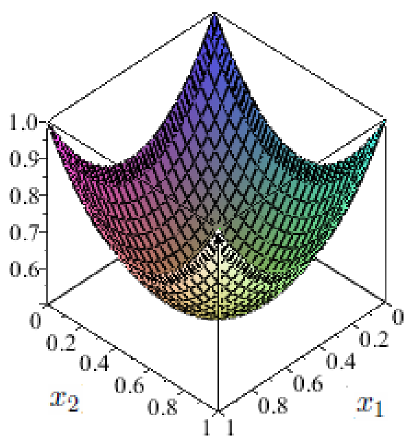

Figure 5.7: Max.= 1 with respect to the trace (Product type design by OLS)

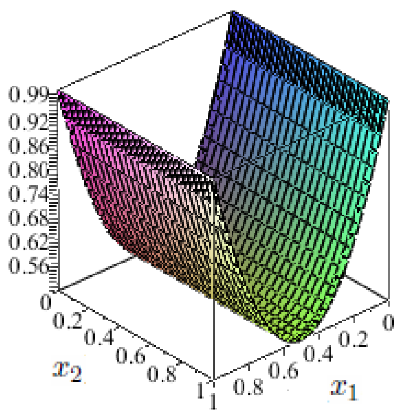

Figure 5.8: Max.< 1 with respect to the maximum eigenvalue (Product type design by OLS)

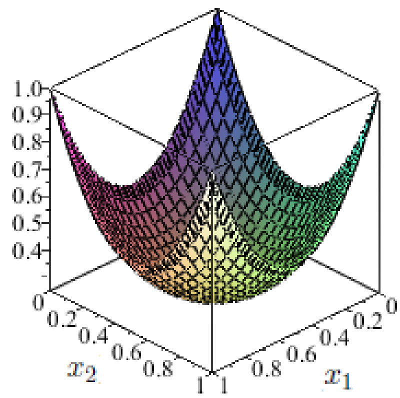

Figure 5.9: Max.= 1 with respect to the determinant (Product type design by OLS)

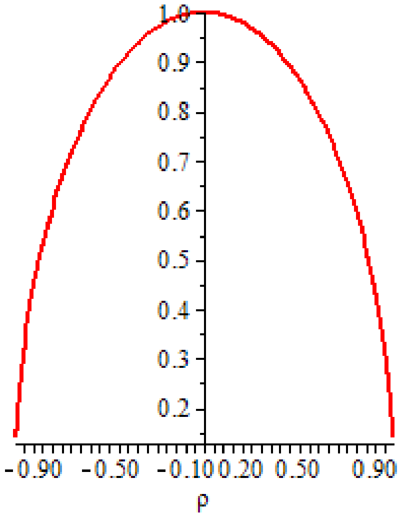

Figure 5.10: The efficiency for OLS vs Gauß Markov with respect to the D-optimality by product type design

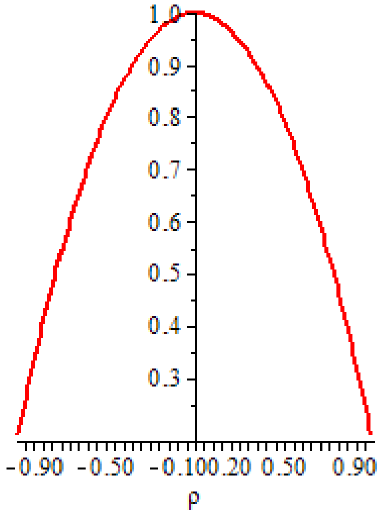

Figure 5.11: The efficiency for OLS vs Gauß Markov with respect to the A-optimality by product type design

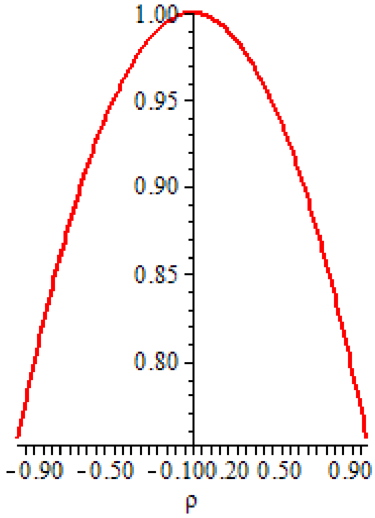

Figure 5.12: The efficiency for OLS vs Gauß Markov with respect to the IMSE-optimality by product type design

have respectively a maximum of 1, by which can be illustrated the figures 5.7, 5.8 and 5.9. respectively the terms of the efficiency for the OLS versus the Gauß Markov by the theorem 5.4 have under the class of product type designs the following forms $\text{eff}_D(\hat{\boldsymbol{\beta}}_{OLS}(\xi)) = (1-\rho^2)^{1/2}$, $\text{eff}_A(\hat{\boldsymbol{\beta}}_{OLS}(\xi)) = 1.000019570 - 0.8284433380\,\rho^2$ and $\text{eff}_{IMSE}(\hat{\boldsymbol{\beta}}_{OLS}(\xi)) = 1 - 1/4\,\rho^2$ where the corresponding behaviors are respectively depicted in the figures 5.10, 5.11 and 5.12. the terms of the efficiency for MANOVA-desgin versus the product type design with respect to the Gauß Markov estimator by the theorem (5.5) have the following forms $\text{eff}_D(\xi_0) = (1-\rho^2)^{1/4}$, $\text{eff}_A(\xi_0) = 1 + 2(1-\sqrt{2})\,\rho^2$ and $\text{eff}_{IMSE}(\xi_0) = 1 - 1/4\,\rho^2$ where the corresponding behaviors are respectively depicted in the figures 5.13, 5.14 and 5.15. respectively the terms of the efficiency for MANOVA-desgin versus the product type design with respect to the OLS estimator by the theorem (5.6) have the following forms $\text{eff}_D(\xi_0) = (1-\rho^2)^{-1/4}$, $\text{eff}_A(\xi_0) = 1$ and $\text{eff}_{IMSE}(\xi_0) = 1$ where the corresponding behavior for the D-optimality is depicted in the figure 5.16 and illustrates the topics of the remark 5.6.

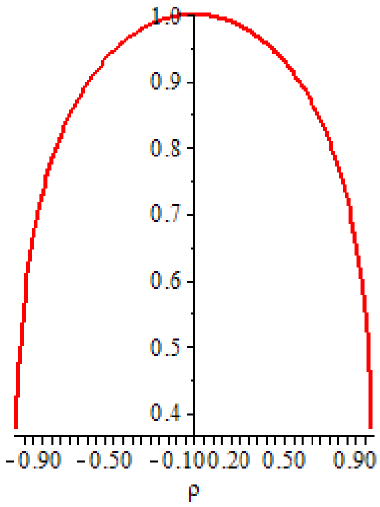

Figure 5.13: The efficiency for MANOVA vs product type design due to D-optimality by Gauß Markov

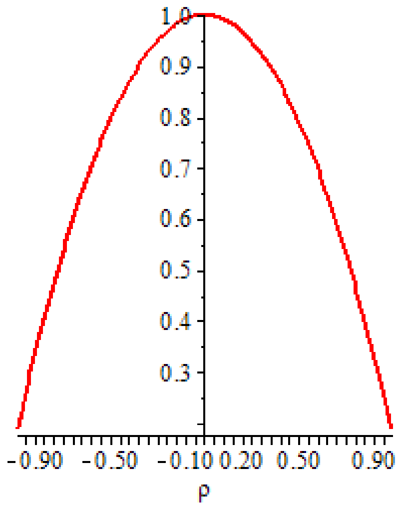

Figure 5.14: The efficiency for MANOVA vs product type design due to A-optimality by Gauß Markov

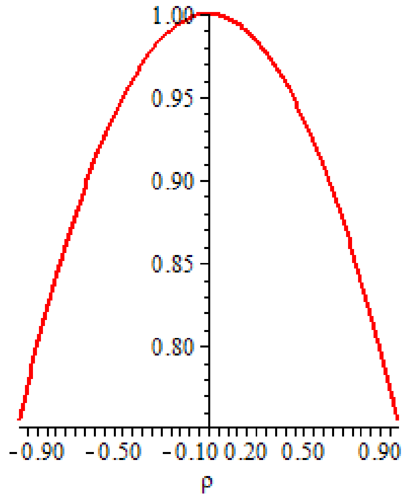

Figure 5.15: The efficiency for MANOVA vs product type design due to IMSE-optimality by Gauß Markov

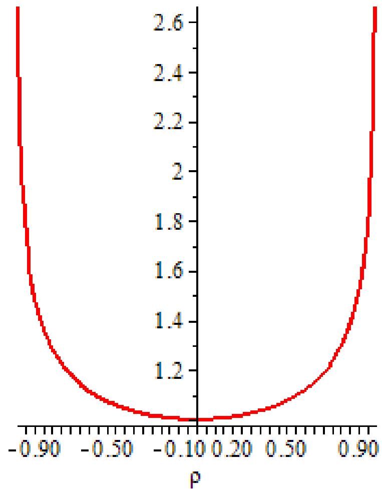

Figure 5.16: The efficiency for MANOVA vs. product type design due to D-optimality by OLS

6 Optimal Designs for SUR Models with Multi-Factor and Nested Components

The multi-factor experiments similar to additive models and product-type models have many applications as well as their multivariate version, which can be presented as general multivariate linear models resp. as SUR models with multi-factor models nested or not as marginals. The first one who spoke about SUR models was *(Zellner (1962))*, where he used a bivariate additive model, their marginals had the same additive regression functions and the same control variables, which belonged to the same design region that analyzed annual investment data, between 1935-1954, for two firms. Optimal designs for multivariate multi-factor models by a Kronecker product form for the variance covariance matrix of the error were explored in the work of *(Schwabe (1996))*.
Optimal designs for different multivariate linear multi-factor models by different nesting forms are explored in this work. So, D- and some linear optimal designs for multivariate multi-factor models with a nesting structure are explored in the first section, where one marginal component is nested multiplicatively or additively in the other marginal components, and shows, that the product type designs are D- and linear optimal in 4.1, for the SUR model with the multiplicative nesting case, D-optimal for the SUR model with the the additively nesting case without conditions and linear optimal in 4.1 by block diagonal information matrices. For growing complexity or hierarchically nesting multi-factor models, where the first component is nested in the second multiplicatively or additively, and the second in the third, and so on are the results in the first section valid. When a new component is nested multiplicatively or additively in all components of the SUR-model with one-factor marginals or a new different component is nested in each component of the SUR model with one-factor marginals stay the results in the first section valid, under similar conditions, that is shown in the third section. There are two clear examples at the end of this chapter, which illustrate the theoretical results.

6.1 Optimal Designs for Multivariate Multi-Factor Models for the Simplest Nesting form

The multivariate multi-factor models can be presented as SUR models by the same assumption as the presented SUR model (3.3) in chapter two, but with a different forms for the block diagonal regression function, which can be seen in one of the nesting form, where the first component is nested in the other components, generally as follows

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \mathbf{f}_1(x_1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2(x_1, x_2) & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_m(x_1, x_m) \end{pmatrix} \tag{6.1}$$

for the multivariate experimental setting $\mathbf{x} = (x_1, ..., x_m)^\top \in \mathcal{X} = \times_{j=1}^{n} \mathcal{X}_j$.
To obtain a complete characterization of D- and linear optimal designs 4.1, it is required that all

marginal models related to the components contain an intercept, $f_{j1}(\mathbf{x}) \equiv 1$, with respect to SUR models with marginals as additive models as well as product-type models. We will sort the results according to the models, and to the structures of nesting.

6.1.1 Optimal Designs for Multiplicative Nesting

The topics of the next two results are the D- and linear optimality in 4.1 of the product type designs for the SUR model with multiplicative marginal components models by regression functions for the components have the next forms

$$\mathbf{f}_j(x_1, x_j)_{j=2,\dots,m} = \mathbf{f}_1(x_1) \otimes \mathbf{f}_j(x_j) \tag{6.2}$$

i.e. the regression function for the considered SUR model has the following form

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \mathbf{f}_1(x_1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_1(x_1) \otimes \mathbf{f}_2(x_2) & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_1(x_1) \otimes \mathbf{f}_m(x_m) \end{pmatrix} \tag{6.3}$$

Theorem 6.1. *Let ξ_j^* be D-optimal for the j-th marginal component with the regression function given in (6.2) in the marginal design region $\mathcal{X}_j$, $j = 1, ..., m$, then the product type design*

$$\xi^* = \otimes_{j=1}^m \xi_j^*$$

is D-optimal for the SUR model (3.3) in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$, where the block diagonal multivariate regression function for the considered SUR model is given in (6.3).
The sensitivity function φ_D does not depend on $\mathbf{\Sigma}$.

Theorem 6.2. *Let ξ_j^* be linear optimal with weight matrix $\mathbf{L}_j$, for the j-th marginal component with the regression function given in (6.2) in the marginal design region $\mathcal{X}_j$, $j = 1, ..., m$, then the product type design*

$$\xi^* = \otimes_{j=1}^m \xi_j^*$$

is linear optimal for the SUR model (3.3) in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$, if the weight matrix $\mathbf{L} = \text{block} - \text{diag}\,(\mathbf{L}_j)$ (4.1), where the block diagonal multivariate regression function for the considered SUR model is given in (6.3).

Proof: With respect to the nesting structure for the regression function of the SUR model given in (6.3) has the block regression function for the SUR model the following form

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}_1(x_1) \otimes \begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2(x_2) & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_m(x_m) \end{pmatrix}, \text{ and suppose, that } \Sigma^{-1} = \begin{pmatrix} \sigma^{(11)} & \cdots & \sigma^{(1m)} \\ \vdots & \ddots & \vdots \\ \sigma^{(1m)} & \cdots & \sigma^{(mm)} \end{pmatrix}$$

Then the information matrix for one-point design has the following form for the considered SUR model

$$\mathbf{M}(\mathbf{x}) = \mathbf{f}(\mathbf{x})\Sigma^{-1}\mathbf{f}(\mathbf{x})^\top = \ \mathbf{f}_1(x_1) \otimes \begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2(x_2) & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_m(x_m) \end{pmatrix}$$

$$\begin{pmatrix} \sigma^{(11)} & \cdots & \sigma^{(1m)} \\ \vdots & \ddots & \vdots \\ \sigma^{(1m)} & \cdots & \sigma^{(mm)} \end{pmatrix} \ \mathbf{f}_1^\top(x_1) \otimes \begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2(x_2)^\top & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_m(x_m)^\top \end{pmatrix}$$

Then

$$\mathbf{M}(\mathbf{x}) = \mathbf{f}_1(x_1)\mathbf{f}_1(x_1)^\top \otimes \begin{pmatrix} \sigma^{(11)} & \sigma^{(12)}\mathbf{f}_2(x_2) & \cdots & \sigma^{(1m)}\mathbf{f}_1(x_1) \\ \sigma^{(12)}\mathbf{f}_2(x_2) & \sigma^{(22)}\mathbf{f}_2(x_2) & \cdots & \sigma^{(2m)}\mathbf{f}_2(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^{(1m)}\mathbf{f}_m(x_m) & \sigma^{(2m)}\mathbf{f}_m(x_m) & \cdots & \sigma^{(mm)}\mathbf{f}_m(x_m) \end{pmatrix}$$

$$\begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2(x_2)^\top & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_m(x_m)^\top \end{pmatrix}$$

Also, the information matrix for the considered SUR model has the following form for $j = 1, ..., m$, $\mathbf{M}_j(x_j) = \mathbf{f}_j(x_j)\mathbf{f}_j(x_j)^\top$ and

$$\mathbf{M}(\mathbf{x}) = \mathbf{M}_1(x_1) \otimes \begin{pmatrix} \sigma^{(11)} & \sigma^{(12)}\mathbf{f}_2(x_2)^\top & \cdots & \sigma^{(1m)}\mathbf{f}_m(x_m)^\top \\ \sigma^{(12)}\mathbf{f}_2(x_2) & \sigma^{(22)}\mathbf{M}_2(x_2) & \cdots & \sigma^{(2m)}\mathbf{f}_2(x_2)^\top\mathbf{f}_m(x_m)^\top \\ \vdots & \cdots & \ddots & \vdots \\ \sigma^{(1m)}\mathbf{f}_m(x_m) & \cdots & \cdots & \sigma^{(mm)}\mathbf{M}_m(x_m) \end{pmatrix} \tag{6.4}$$

Then the information matrix for $\xi = \otimes_{j=1}^m \xi_j$, is $\mathbf{M}_j(x_j) = \int \mathbf{f}_j(x_j)\mathbf{f}_j(x_j)^\top \xi_j(\mathrm{d}x_j)$

$$\mathbf{m}_j(\xi_j) = \int \mathbf{f}_j(x_j)\xi_j(\mathrm{d}x_j)$$

$$\mathbf{M}(\xi) = \mathbf{M}_1(\xi_1) \otimes \begin{pmatrix} \sigma^{(11)} & \sigma^{(12)}\mathbf{m}_2(\xi_2)^\top & \cdots & \sigma^{(1m)}\mathbf{m}_m(\xi_m)^\top \\ \sigma^{(12)}\mathbf{m}_2(\xi_2) & \sigma^{(22)}\mathbf{M}_2(\xi_2) & \cdots & \sigma^{(2m)}\mathbf{m}_2(\xi_2)^\top\mathbf{m}_m(\xi_m)^\top \\ \vdots & \cdots & \ddots & \vdots \\ \sigma^{(1m)}\mathbf{m}_m(\xi_m) & \cdots & \cdots & \sigma^{(mm)}\mathbf{M}_m(\xi_m) \end{pmatrix}$$

$$= \mathbf{M}_1(\xi_1) \otimes \left(\mathbf{M}_{jk}(\xi_j \otimes \xi_k)\right)_{j,k=1,...,m} \tag{6.5}$$

The second part of the information matrix in (6.5) is an information matrix for a SUR model with m

components and the block regression function of it is given as follows

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} 1 & 0 & 0 \\ \mathbf{0} & \mathbf{f}_2(x_2) & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x_m) \end{pmatrix}$$

where their D- and linear optimal in 4.1 designs are discussed explicitly in chapter three of this work, and it is shown, that the product type designs are D- and linear optimal in 4.1 for these kinds of SUR models, when its marginal regression function contains an intercept, So the equivalence theorem for the D-optimal design $\xi^* = \otimes_{j=1}^m \xi_j^*$ because of theorem (4.3) has the following form

$$\begin{aligned} &\text{trace}\left(\mathbf{M}(\mathbf{x})\ \mathbf{M}^{-1}(\xi_D^*)\right) = \text{trace}\left(\left(\mathbf{M}_1(x_1) \otimes \mathbf{M}_{jk}(x_j, x_k)\ \mathbf{M}_1^{-1}(\xi_1^*) \otimes \mathbf{M}_{jk}^{-1}(\xi_j^* \otimes \xi_k^*)\right)_{j,k=1,\ldots,m}\right) \\ &= \text{trace}\left(\mathbf{M}_1(x_1)\ \mathbf{M}_1^{-1}(\xi_1^*)\ \otimes\ \left(\mathbf{M}_{jk}(x_j, x_k)\ \mathbf{M}_{jk}^{-1}(\xi_j^* \otimes \xi_k^*)\right)_{j,k=1,\ldots,m}\right) \\ &= \text{trace}\left(\mathbf{M}_1(x_1)\ \mathbf{M}_1^{-1}(\xi_1^*)\right)\ \text{trace}\left(\left(\mathbf{M}_{jk}(x_j \otimes x_k)\ \mathbf{M}_{jk}^{-1}(\xi_j^* \otimes \xi_k^*)\right)_{j,k=1,\ldots,m}\right) \le p_1\ (1 + \sum_{j=2}^{m} p_j) = p \end{aligned}$$

And for positive definite block diagonal matrix $\mathbf{L}$, where for $j, k = 1, ..., m$, $\mathbf{L} = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})\ \mathbf{f}(\mathbf{x})^\top \mu(\mathrm{d}\mathbf{x}) = \mathbf{L}_1 \otimes \mathbf{L}_{jk}$

$$\mathbf{L}_j = \begin{pmatrix} L_{111} & \mathbf{L}_{j12}^\top \\ \mathbf{L}_{j12} & \mathbf{L}_{j22} \end{pmatrix},\ \mathbf{L}_{jk} = \begin{pmatrix} L_{111} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_2 & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{L}_m \end{pmatrix}$$

also, the left side of the equivalence theorem for linear criteria has the following form, because of the form of the matrix $\mathbf{L}$ and the information matrices in (6.4) and (6.5)

$$\begin{aligned} &\text{trace}\left(\mathbf{M}(\mathbf{x})\ \mathbf{M}^{-1}(\xi_L^*)\ \mathbf{L}\ \mathbf{M}^{-1}(\xi_L^*)\right) \\ &= \text{trace}\left(\mathbf{M}_1(x_1) \otimes \mathbf{M}_{jk}(x_j \otimes x_k)\ \mathbf{M}_1^{-1}(\xi_1^*) \otimes \mathbf{M}_{jk}^{-1}(\xi_j^* \otimes \xi_k^*)\mathbf{L}_1 \otimes \mathbf{L}_{jk}\mathbf{M}_1^{-1}(\xi_1^*) \otimes \mathbf{M}_{jk}^{-1}(\xi_j^* \otimes \xi_k^*)\right) \\ &= \text{trace}\left(\left(\mathbf{M}_1(x_1)\ \mathbf{M}_1^{-1}(\xi_1^*)\mathbf{L}_1\ \mathbf{M}_1^{-1}(\xi_1^*)\right) \otimes \left(\mathbf{M}_{jk}(x_j \otimes x_k)\ \mathbf{M}_{jk}^{-1}(\xi_j^* \otimes \xi_k^*)\mathbf{L}_{jk}\mathbf{M}_{jk}^{-1}(\xi_j^* \otimes \xi_k^*)\right)\right) \\ &= \text{trace}\left(\mathbf{M}_1(x_1)\ \mathbf{M}_1^{-1}(\xi_1^*)\mathbf{L}_1\ \mathbf{M}_1^{-1}(\xi_1^*)\right) \text{trace}\left(\mathbf{M}_{jk}(x_j \otimes x_k)\ \mathbf{M}_{jk}^{-1}(\xi_j^* \otimes \xi_k^*)\mathbf{L}_{jk}\mathbf{M}_{jk}^{-1}(\xi_j^* \otimes \xi_k^*)\right) \end{aligned}$$

So because of theorem (4.4)

$$\begin{aligned} &\text{trace}\left(\mathbf{M}_1(x_1)\ \mathbf{M}_1^{-1}(\xi_1^*)\mathbf{L}_1\ \mathbf{M}_1^{-1}(\xi_1^*)\right) \text{trace}\left(\mathbf{M}_{jk}(x_j \otimes x_k)\ \mathbf{M}_{jk}^{-1}(\xi_j^* \otimes \xi_k^*)\mathbf{L}_{jk}\mathbf{M}_{jk}^{-1}(\xi_j^* \otimes \xi_k^*)\right) \\ &\le \text{trace}\left(\mathbf{L}_1\mathbf{M}_1^{-1}(\xi_1^*)\right)\ \text{trace}\left(\mathbf{L}_{jk}\ \mathbf{M}_{jk}^{-1}(\xi_j^* \otimes \xi_k^*)\right) \end{aligned}$$

And therewith the product type designs are linear optimal in 4.1 and the theorem has been proven.

6.1.2 Optimal Designs for Additive Nesting

The product type designs are D-optimal without conditions and linear optimal in 4.1, by block diagonal information matrices, or when the product type designs optimal for the univariate additive models. These results are valid for the multivariate additive linear models with the next forms of the regression functions for the components $\mathbf{f}_1(x_1) = \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) \end{pmatrix}$, $\mathbf{f}_j(x_1, x_j)_{j=2,\dots,m} = \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) \\ \mathbf{g}_j(x_j) \end{pmatrix}$ i.e. the regression function for the considered SUR model has the following form

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) \end{pmatrix} & \cdots & \mathbf{0} \\ \mathbf{0} & \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) \\ \mathbf{g}_2(x_2) \end{pmatrix} & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) \\ \mathbf{g}_m(x_m) \end{pmatrix} \end{pmatrix} \tag{6.6}$$

Theorem 6.3. *Let ξ_j^* be D-optimal for the j-th marginal component with the regression function*

$$\mathbf{f}_1(x_1) = \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) \end{pmatrix}, \ \mathbf{f}_j(x_1, x_j)_{j=2,\dots,m} = \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) \\ \mathbf{g}_j(x_j) \end{pmatrix} \tag{6.7}$$

in the marginal design region $\mathcal{X}_j$, $j = 1, ..., m$, then the product type design

$$\xi^* = \otimes_{j=1}^m \xi_j^*$$

is D-optimal for the SUR model (3.3) in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$, where the block diagonal multivariate regression function for the considered SUR model is given in (6.6).
The sensitivity function φ_D does not depend on $\mathbf{\Sigma}$.

Proof: the proof can be implemented step by step by some auxiliary lemmas and theorems as follows

Lemma 6.1. *The D-criterion is invariant by application a linear transformation.*

Proof: the transformed regression function has the form $\tilde{\mathbf{f}} = \mathbf{Af}$ for a constant and regular matrix $\mathbf{A}_{p\times p}$, and resp. the transformed information matrix has the form

$$\tilde{\mathbf{M}}(\xi^*) = \int \tilde{\mathbf{f}}(\mathbf{x})\mathbf{\Sigma}^{-1}\tilde{\mathbf{f}}(\mathbf{x})^\top \xi^*(d\mathbf{x}) = \mathbf{A} \int \mathbf{f}(\mathbf{x})\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \xi^*(d\mathbf{x})\mathbf{A}^\top = \mathbf{AM}(\xi^*)\mathbf{A}^\top \tag{6.8}$$

So the left side of the multivariate equivalence theorem for the D-optimality 4.1 by the transformed regression function and resp. information matrix has the following form

$$\begin{aligned} &\text{trace}\left(\mathbf{\Sigma}^{-1}\tilde{\mathbf{f}}(\mathbf{x})^\top \tilde{\mathbf{M}}(\xi^*)^{-1}\tilde{\mathbf{f}}(\mathbf{x})\right) \\ &= \text{trace}\left(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{A}^\top \mathbf{A}^{-\top}\mathbf{M}(\xi^*)^{-1}\mathbf{A}^{-1}\mathbf{Af}(\mathbf{x})\right) \\ &= \text{trace}\left(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi^*)^{-1}\mathbf{f}(\mathbf{x})\right) \end{aligned}$$

i.e. the D-criterion due to its equivalence theorem is invariant to the applied linear transformation. The next theorem is a help tool and it have the same topic of the goal result, but it is proven for a special case, which is the block diagonal form of the information matrix for the corresponding SUR model. □

Theorem 6.4. *Let ξ_j^* be D-optimal for the j-th marginal component (6.7) in the marginal design region $\mathcal{X}_j$ with an intercept included, $j = 1, ..., m$, if*

$$\int_{\mathcal{X}_j} \mathbf{g}_j(x_j)\xi_j^*(\mathrm{d}x_j) = \mathbf{0},\ \ j = 1, ..., m \tag{6.9}$$

then the product type design

$$\xi^* = \otimes_{j=1}^m \xi_j^* \tag{6.10}$$

is D-optimal for the SUR model (3.3) in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$, where the block diagonal multivariate regression function for the considered SUR model is given in (6.6).
The sensitivity function φ_D does not depend on $\mathbf{\Sigma}$.

Proof: The diagonal blocks of the information matrix for one-point design have the following form because of the form of the block diagonal form for the multivariate regression function, given in the form (6.6)

$$\mathbf{M}(\mathbf{x})_{jj} = (\mathbf{f}(\mathbf{x})\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top)_{jj} = \mathrm{block-diag}\,(\sigma^{(11)}\mathbf{M}_1(x_1), \sigma^{(jj)}\mathbf{M}_{1j}(x_1, x_j))_{j=1,...,m} \tag{6.11}$$
$$\text{where}\ \ \mathbf{M}_{1j}(x_j) = \mathbf{f}_j(x_1, x_j)\mathbf{f}_j(x_1, x_j)^\top$$

Then the information matrix with respect to the product type design $\xi_D^* = \otimes_{j=1}^m \xi_{j,D}^*$ in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$ has the next block diagonal form because of 6.9

$$\mathbf{M}(\xi_D^*) = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \xi_D^*(\mathrm{d}\mathbf{x}) = \begin{pmatrix} \sigma^{(11)}\mathbf{M}_1(\xi_1^*) & \cdots & \mathbf{0} \\ \mathbf{0} & \sigma^{(22)}\mathbf{M}_{12}(\xi_1^* \otimes \xi_2^*) & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \sigma^{(mm)}\mathbf{M}_{1m}(\xi_1^* \otimes \xi_m^*) \end{pmatrix}$$

$$\mathbf{M}_{1j}(\xi_1^* \otimes \xi_m^*) = \int_{\mathcal{X}_1 \times \mathcal{X}_j} \mathbf{f}_j(x_1, x_j)\mathbf{f}_j(x_1, x_j)^\top d(\xi_1^* \otimes \xi_j^*) = \begin{pmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_1(\xi_1^*) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_j(\xi_j^*) \end{pmatrix}$$

$$\mathbf{G}_j(\xi_j^*) = \int_{\mathcal{X}_j} \mathbf{g}_j(x_j)\mathbf{g}_j(x_j)^\top d(\xi_j^*)\, j = 1, ..., m$$

So the inverse of the information matrix has the next form

$$\mathbf{M}^{-1}(\xi_D^*) = \begin{pmatrix} \frac{1}{\sigma^{(11)}}\mathbf{M}_1^{-1}(\xi_1^*) & \cdots & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sigma^{(22)}}\mathbf{M}_{12}^{-1}(\xi_1^* \otimes \xi_2^*) & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{\sigma^{(mm)}}\mathbf{M}_{1m}^{-1}(\xi_1^* \otimes \xi_m^*) \end{pmatrix} \tag{6.12}$$

Thus the next terms are to be hold by replacing (6.11) and (6.12) int the multivariate equivalence theorem for D-optimality 4.1, so

$$\mathbf{M}(\mathbf{x})\mathbf{M}(\xi^*)^{-1} = \begin{pmatrix} \sigma^{(11)}\mathbf{M}_1(x_1) & \cdots & . \\ . & \sigma^{(22)}\mathbf{M}_{12}(x_1,x_2) & \cdots \\ \vdots & \ddots & \vdots \\ . & \cdots & \sigma^{(mm)}\mathbf{M}_{1m}(x_1,x_m) \end{pmatrix}$$

$$\begin{pmatrix} \frac{1}{\sigma^{(11)}}\mathbf{M}_1^{-1}(\xi_1^*) & \cdots & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sigma^{(22)}}\mathbf{M}_{12}^{-1}(\xi_1^*\otimes\xi_2^*) & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{\sigma^{(mm)}}\mathbf{M}_{1m}^{-1}(\xi_1^*\otimes\xi_m^*) \end{pmatrix}$$

Thus trace $(\mathbf{M}(\mathbf{x})\mathbf{M}^{-1}(\xi_D^*)) =$

trace $(\mathbf{M}_1(x_1)\mathbf{M}_1^{-1}(\xi_1^*)) +$ trace $(\mathbf{M}_{12}(x_1,x_2)\mathbf{M}_{12}^{-1}(\xi_1^*\otimes\xi_2^*))$

$+\ ...+$ trace $(\mathbf{M}_{1m}(x_1,x_m)\mathbf{M}_{1m}^{-1}(\xi_1^*\otimes\xi_m^*))$

$\leq p_1 + (p_1+p_2-1) + ... + (p_1+p_m-1) = mp_1 + p_2 + ... + p_m - (m-1) = p$

So the conditions of the equivalence theorem are valid and the product type design $\xi_D^* = \otimes_{j=1}^m \xi_{j,D}^*$ is D-optimal by block diagonal information matrix. And the sensitivity function for D-optimality is independent on the variance covariance terms. □

By the next help lemma can be illustrated, that the D-optimal design for a SUR model is the same for the transformed SUR model.

Lemma 6.2. *The D-optimal design for the SUR model stays D-optimal for the linearly transformed SUR model.*

Proof: The transformation matrix and its diagonal block matrices for $j = 1,...,m$ have the next forms

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{12} & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{A}_{1m} \end{pmatrix} \tag{6.13}$$

$$\mathbf{A}_1 = \begin{pmatrix} 1 & \mathbf{0} \\ -\mathbf{b}_1 & I_{(p_1-1)\times(p_1-1)} \end{pmatrix},\ \mathbf{A}_{1j} = \begin{pmatrix} 1 & \mathbf{0} & \mathbf{0} \\ -\mathbf{b}_1 & \mathbf{I}_{p_1-1} & \mathbf{0} \\ -\mathbf{b}_j & \mathbf{0} & \mathbf{I}_{p_j-1} \end{pmatrix},\ \mathbf{b}_j = \int \mathbf{g}_j(x_j)\xi_j(\mathrm{d}x_j) \tag{6.14}$$

i.e., the transformed regression function seems for the first component as follows

$$\tilde{\mathbf{f}}_1(x_1) = \mathbf{A}_1(\xi_1)\mathbf{f}_1(x_1) = \begin{pmatrix} 1 & \mathbf{0} \\ -\mathbf{b}_1(\xi_1) & \mathbf{I}_{p_1-1} \end{pmatrix}\mathbf{f}_1(x_1) = \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) - \mathbf{b}_1 \end{pmatrix}$$

and respectively for the $m-1$ other components

$$\tilde{\mathbf{f}}_j(x_1,x_j) = \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) - \mathbf{b}_1 \\ \mathbf{g}_j(x_j) - \mathbf{b}_j \end{pmatrix}$$

So for $\tilde{\mathbf{g}}_j = \mathbf{g}_j - \int \mathbf{g}_j \xi_j(\mathrm{d}x_j)$ then $\int \tilde{\mathbf{g}}_j d\xi_j = \int \mathbf{g}_j \xi_j(\mathrm{d}x_j) - \int \mathbf{g}_j \xi_j(\mathrm{d}x_j) = \mathbf{0}$, and therewith

$$\int \tilde{\mathbf{f}}_j(x_1) d\xi_1 = \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}, \ \int \tilde{\mathbf{f}}_j(x_1, x_j) d(\xi_1 \otimes \xi_j) = \begin{pmatrix} 1 \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}$$

And the information matrix for the transformed model has the following block diagonal form for the product type design $\xi^* = \otimes_{j=1}^m \xi_j^*$

$$\tilde{\mathbf{M}}(\xi^*) = \begin{pmatrix} \frac{1}{\sigma^{(11)}} \tilde{\mathbf{M}}_1^{-1}(\xi_1^*) & \cdots & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sigma^{(22)}} \tilde{\mathbf{M}}_{12}^{-1}(\xi_1^* \otimes \xi_2^*) & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{\sigma^{(mm)}} \tilde{\mathbf{M}}_{1m}^{-1}(\xi_1^* \otimes \xi_m^*) \end{pmatrix}$$

Where the transformed regression function has the following form

$$\tilde{\mathbf{f}}(\mathbf{x}) = \begin{pmatrix} \mathbf{A}_1 & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{12} & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{A}_{1m} \end{pmatrix} \begin{pmatrix} \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) \end{pmatrix} & \cdots & \mathbf{0} \\ \mathbf{0} & \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) \\ \mathbf{g}_2(x_2) \end{pmatrix} & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) \\ \mathbf{g}_m(x_m) \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} \begin{pmatrix} 1 \\ \tilde{\mathbf{g}}_1(x_1) \end{pmatrix} & \cdots & \mathbf{0} \\ \mathbf{0} & \begin{pmatrix} 1 \\ \tilde{\mathbf{g}}_1(x_1) \\ \tilde{\mathbf{g}}_2(x_2) \end{pmatrix} & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \begin{pmatrix} 1 \\ \tilde{\mathbf{g}}_1(x_1) \\ \tilde{\mathbf{g}}_m(x_m) \end{pmatrix} \end{pmatrix}$$

So that is the same special case of theorem 6.4, by it the product type design is D-optimal for the considered SUR model for block diagonal information matrix. Then the product type design $\xi_D^* = \otimes_{j=1}^m \xi_{j,D}^*$ is D-optimal for the SUR model by the following non transformed regression function because of the

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \mathbf{A}_1^{-1} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{12}^{-1} & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{A}_{1m}^{-1} \end{pmatrix} \begin{pmatrix} \begin{pmatrix} 1 \\ \tilde{\mathbf{g}}_1(x_1) \end{pmatrix} & \cdots & \mathbf{0} \\ \mathbf{0} & \begin{pmatrix} 1 \\ \tilde{\mathbf{g}}_1(x_1) \\ \tilde{\mathbf{g}}_2(x_2) \end{pmatrix} & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \begin{pmatrix} 1 \\ \tilde{\mathbf{g}}_1(x_1) \\ \tilde{\mathbf{g}}_m(x_m) \end{pmatrix} \end{pmatrix} \quad \square \tag{6.15}$$

because of lemma 6.1 and theorem 6.4, the D-optimality of the product type design as well as the independence of the corresponding sensitivity function on the variance covariance matrix of the error variables and the theorem 4.3 has been proven. □

Theorem 6.5. *Let ξ_j^* be linear optimal, by block diagonal weight matrix $\mathbf{L}_j$, for the j-th marginal component (3.4) in the marginal design region $\mathcal{X}_j$, $j = 1, ..., m$. If*

$$\int_{\mathcal{X}_j} \mathbf{g}_j(x_j)\xi_j^*(\mathrm{d}x_j) = \mathbf{0}, \; j = 2, ..., m \tag{6.16}$$

then the product type design

$$\xi^* = \otimes_{j=1}^m \xi_j^*$$

is linear optimal, if the weight matrix $\mathbf{L} = \text{block} - \text{diag}\,(\mathbf{L}_j)$ (4.1), for the SUR model (3.3) in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$, where the block diagonal multivariate regression function for the considered SUR model is given in (6.6).

Proof: the weight matrix for the considered SUR model has the following matrix

$$\mathbf{L} = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top \mu(\mathrm{d}\mathbf{x}) = \begin{pmatrix} \mathbf{L}_1 & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_{12} & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{L}_{1m} \end{pmatrix} \tag{6.17}$$

$$\mathbf{L}_1 = \begin{pmatrix} L_{111} & \mathbf{L}_{112}^\top \\ \mathbf{L}_{112} & \mathbf{L}_{122} \end{pmatrix}, \; \mathbf{L}_{1j} = \begin{pmatrix} L_{111} & \mathbf{L}_{112}^\top & \mathbf{L}_{j13}^\top \\ \mathbf{L}_{112} & \mathbf{L}_{122} & \mathbf{L}_{j23}^\top \\ \mathbf{L}_{j13} & \mathbf{L}_{j23} & \mathbf{L}_{j33} \end{pmatrix}$$

The inverse of the information matrix has the following block diagonal form for the product type design

$\xi_L^* = \otimes_{j=1}^m \xi_{j;L}^*$, under the conditions (6.28)

$$\mathbf{M}^{-1}(\xi_L^*) = \begin{pmatrix} \frac{1}{\sigma^{(11)}}\mathbf{M}_1^{-1}(\xi_1^*) & \cdots & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sigma^{(22)}}\mathbf{M}_{12}^{-1}(\xi_1^* \otimes \xi_2^*) & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{\sigma^{(mm)}}\mathbf{M}_{1m}^{-1}(\xi_1^* \otimes \xi_m^*) \end{pmatrix}, \quad \text{Where} \tag{6.18}$$

$$\mathbf{M}_{1j}^{-1}(\xi_1^* \otimes \xi_j^*) = \begin{pmatrix} \mathbf{M}_1^{-1}(\xi_1^*) & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_j^{-1}(\xi_j^*) \end{pmatrix},$$

$$\mathbf{M}_1^{-1}(\xi_1^*) = \int \mathbf{f}_1(x_1)\mathbf{f}_1(x_1)^\top \xi_1^*(\mathrm{d}x_1),\ \mathbf{G}_j^{-1}(\xi_j^*) = \int \mathbf{g}_1(x_j)\mathbf{g}_j(x_j)^\top \xi_j^*(\mathrm{d}x_j) \tag{6.19}$$

Then because of (6.17) and (6.18)

$$\mathbf{M}^{-1}(\xi_L^*)\ \mathbf{L}\ \mathbf{M}^{-1}(\xi_L^*) =$$

$$\begin{pmatrix} \frac{1}{(\sigma^{(11)})^2}\mathbf{M}_1^{-1}\mathbf{L}_1\mathbf{M}_1^{-1} & \cdots & \mathbf{0} \\ \mathbf{0} & \frac{1}{(\sigma^{(22)})^2}\mathbf{M}_{12}^{-1}(\xi_1^* \otimes \xi_2^*)\mathbf{L}_{12}\mathbf{M}_{12}^{-1}(\xi_1^* \otimes \xi_2^*) & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{(\sigma^{(mm)})^2}\mathbf{M}_{1m}^{-1}\mathbf{L}_{1m}\mathbf{M}_{1m}^{-1} \end{pmatrix} \tag{6.20}$$

So the sensitivity function for linear optimality has the following form because of (6.11) and because of theorem 2.7

$$\begin{aligned} &\operatorname{trace}\left(\mathbf{M}(\mathbf{x})\ \mathbf{M}^{-1}(\xi_L^*)\ \mathbf{L}\ \mathbf{M}^{-1}(\xi_L^*)\right) \\ &= \frac{1}{\sigma^{(11)}}\operatorname{trace}\left(\mathbf{M}_1(x_1)\mathbf{M}_1^{-1}(\xi_1^*)\mathbf{L}_1\mathbf{M}_1^{-1}(\xi_1^*)\right) \\ &+\frac{1}{\sigma^{(22)}}\operatorname{trace}\left(\mathbf{M}_{12}(x_1,x_2)\mathbf{M}_{12}^{-1}(\xi_1^* \otimes \xi_2^*)\mathbf{L}_{12}\mathbf{M}_{12}^{-1}(\xi_1^* \otimes \xi_2^*)\right) + \\ &\vdots \\ &+\frac{1}{\sigma^{(mm)}}\operatorname{trace}\left(\mathbf{M}_{1m}(x_1,x_m)\mathbf{M}_{1m}^{-1}(\xi_1^* \otimes \xi_m^*)\mathbf{L}_{1m}\mathbf{M}_{1m}^{-1}(\xi_1^* \otimes \xi_m^*)\right) \\ &\le \frac{1}{\sigma^{(11)}}\operatorname{trace}\left(\mathbf{L}_1\mathbf{M}_1^{-1}(\xi_1^*)\right) + \frac{1}{\sigma^{(22)}}\operatorname{trace}\left(\mathbf{L}_{12}\mathbf{M}_{12}^{-1}(\xi_1^* \otimes \xi_2^*)\right) \\ &+... + \frac{1}{\sigma^{(mm)}}\operatorname{trace}\left(\mathbf{L}_{1m}\mathbf{M}_{1m}^{-1}(\xi_1^* \otimes \xi_m^*)\right) \\ &= \operatorname{trace}\left(\mathbf{L}\ \mathbf{M}^{-1}(\xi_L^*)\right) \end{aligned}$$

And therewith, the product type design $\xi_L^* = \otimes_{j=1}^m \xi_{j;L}^*$, is linear optimal in 4.1 for the considered SUR model and the theorem has been proven. □

6.2 Optimal Designs for Multivariate Multi-Factor Models with Growing Hierarchically Nesting

theorems 6.1, 6.2, 6.3) and 6.5 can be generalized, when we consider the multivariate linear model with growing nesting of the components additively and multiplicatively, i.e the first component is nested through the second, and the second through the third, etc., for different control variables for the different components. The D-optimal design was founded for one control variables for the different components, i.e. for a special kind of heterogeneous multivariate model by *(Krafft and Schaefer (1992))* , and later by *(Kurotschka and Schwabe (1996))*, only in the additive case.

Theorem 6.6. *Theorems 6.1 and 6.2 are valid for the SUR models, which their components have the regression functions of the form*

$$\mathbf{f}_1(x_1),\ \mathbf{f}_2(x_1,x_2)=\mathbf{f}_1(x_1)\otimes\mathbf{f}_2(x_2),\ \ldots,\ \mathbf{f}_m(x_1,\ldots,x_m)=\mathbf{f}_1(x_1)\otimes\ \ldots\ \otimes\mathbf{f}_m(x_m)$$

Proof: the proof can be implemented totally analogous to the proofs of theorems 6.1 and 6.2 by replacing the following calculated quantities with their counterparts there.
The multivariate regression function has in this case the following form

$$\begin{aligned}\mathbf{f}(\mathbf{x}) &= \begin{pmatrix} \mathbf{f}_1(x_1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_1(x_1)\otimes\mathbf{f}_2(x_2) & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_1(x_1)\otimes\ \ldots\ \otimes\mathbf{f}_m(x_m) \end{pmatrix} \\ &= \mathbf{f}_1(x_1)\otimes \begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2(x_2) & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_2(x_2) \end{pmatrix} \otimes \ldots \otimes \begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & 1 & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_m(x_m) \end{pmatrix}\end{aligned}$$

then the information matrix has the following form for the one-point design

$$\begin{aligned}\mathbf{M}(\mathbf{x}) &= \mathbf{f}(\mathbf{x})\Sigma^{-1}\mathbf{f}(\mathbf{x})^\top \\ &= \mathbf{f}_1(x_1)\otimes \begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2(x_2) & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_2(x_2) \end{pmatrix} \otimes \ldots \otimes \begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & 1 & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_m(x_m) \end{pmatrix} \\ &\quad \begin{pmatrix} \sigma^{(11)} & \cdots & \sigma^{(1m)} \\ \vdots & \ddots & \vdots \\ \sigma^{(1m)} & \cdots & \sigma^{(1m)} \end{pmatrix} \\ &\quad \mathbf{f}_1(x_1)^\top \otimes \begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2^\top(x_2) & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_2^\top(x_2) \end{pmatrix} \otimes \ldots \otimes \begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & 1 & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_m^\top(x_m) \end{pmatrix} \\ &= \mathbf{M}_1(x_1)\otimes\mathbf{M}_{12}(x_1,x_2)\otimes \ldots \otimes\mathbf{M}_{1m}(x_1,\ldots,x_m)\end{aligned}$$

Thus the information matrix has the following form for the product type design $\xi=\otimes_{j=1}^m\xi_j$

$$\begin{aligned}\mathbf{M}(\xi) &= \int \mathbf{M}_1(x_1)\otimes\mathbf{M}_{12}(x_1,x_2)\otimes \ldots \otimes\mathbf{M}_{1m}(x_1,\ldots,x_m)\xi(\mathrm{d}x_1,\ldots,\mathrm{d}x_m) \\ &= \mathbf{M}_1(\xi_1)\otimes\mathbf{M}_{12}(\xi_2)\otimes \ldots \otimes\mathbf{M}_{1m}(\otimes\xi_m)\end{aligned}$$

where

$$\begin{aligned}
\mathbf{M}_1(\xi_1) &= \int \mathbf{f}_1\mathbf{f}_1^\top \mathrm{d}\xi_1 \\
\mathbf{M}_{12}(\xi_2) &= \int \begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2 & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_2 \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2^\top & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_2^\top \end{pmatrix} \mathrm{d}(\xi_1 \otimes \xi_2) \\
&\vdots \\
\mathbf{M}_{1m}(\xi_m) &= \int \begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & 1 & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_m \end{pmatrix} \begin{pmatrix} \sigma^{(11)} & \cdots & \sigma^{(1m)} \\ \vdots & \ddots & \vdots \\ \sigma^{(1m)} & \cdots & \sigma^{(mm)} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & 1 & \cdots & \mathbf{0} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_m^\top \end{pmatrix} \mathrm{d}\xi
\end{aligned}$$

The weight matrix can be calculated similarly, so it is the Kronecker product of the m block diagonal weight matrices for the marginals and has the following form

$$\mathbf{L} = \int \mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top \mu(\mathrm{d}\mathbf{x}) = \mathbf{L}_1 \otimes \mathbf{L}_{12} \otimes ... \otimes \mathbf{L}_{1m}$$

Where

$$\mathbf{L}_1 = \begin{pmatrix} L_{111} & \mathbf{L}_{112}^\top \\ \mathbf{L}_{112} & \mathbf{L}_{122} \end{pmatrix}, \ \mathbf{L}_{12} = \begin{pmatrix} L_{111} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_2 & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{L}_2 \end{pmatrix}, \ ..., \ \mathbf{L}_{1m} = \begin{pmatrix} L_{111} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_m & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{L}_m \end{pmatrix} \quad \square$$

Theorem 6.7. *Theorem 6.3 is valid for the SUR models, which their components have the regression functions of the form*

$$\mathbf{f}_j(\mathbf{x}_j) = \mathbf{C}^j \cdot \mathbf{f}_{j+1}(\mathbf{x}_j, \mathbf{x}_{j+1}), \ \mathbf{C}_j \in \mathrm{R}^{p_j \times p_{j+1}}, \ \mathbf{C}_j = \begin{pmatrix} \mathbf{I}_{p_j \times p_j} & \mathbf{0} \end{pmatrix}, \ j = 1, \ldots, m-1 \tag{6.21}$$

Or more explicitly

$$\mathbf{f}_1(x_1) = \begin{pmatrix} 1 \\ \mathbf{g}_1(\mathbf{x}_1) \end{pmatrix}, \mathbf{f}_2(x_1, x_2) = \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) \\ \mathbf{g}_2(x_2) \end{pmatrix}, \ldots, \mathbf{f}_m(x_1, \ldots, x_m) = \begin{pmatrix} 1 \\ \mathbf{g}_1(x_1) \\ \vdots \\ \mathbf{g}_m(x_m) \end{pmatrix} \tag{6.22}$$

Theorem 6.8. *If*

$$\int_{\mathcal{X}_j} \mathbf{g}_j(x_j)\xi_j^*(\mathrm{d}x_j) = \mathbf{0}, \ j = 2, ..., m \tag{6.23}$$

then theorem 6.5 is valid for the SUR model, its components have the regression functions given in (6.21) or (6.22).

Proof: The proofs of theorems 6.7 and 6.8 can be implemented similarly to the proofs of theorems 6.3 and 6.5. The difference by the proof is caused by the different nesting form and resp. marginal regression functions, thus we have to use other transformation matrices for the marginal models resp. another transformation matrix for the corresponding SUR model, other information matrix and weight matrix , which are given respectively as follows

$$\mathbf{f}(\mathbf{x}) = \mathrm{diag}\left(\mathbf{f}_1(x_1), \mathbf{f}_2(x_1, x_2), ..., \mathbf{f}_m(x_1, \ldots, x_m)\right)$$

And therewith the corresponding transformation matrix has the following form for $j = 2, ..., m$

$$\begin{aligned}
&\mathbf{A} = \mathrm{block-diag}\left(\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_m\right), \text{ where for } j = 2, ..., m \\
&\mathbf{b}_j = \int \mathbf{g}_j(x_j)\xi_j(\mathrm{d}x_j), \ \mathbf{G}_j(\xi_j) = \int \mathbf{g}_j(x_j)\mathbf{g}_j(x_j)^\top \xi_j(\mathrm{d}x_j) \\
&\mathbf{A}_1 = \mathbf{I}_{p_1}, \ \mathbf{A}_j = \begin{pmatrix} 1 & \mathbf{0} \\ -\mathbf{b}_j\mathbf{b}_{j-1}\ \mathbf{G}_{j-1}^{-1}(\xi_{j-1}) & \mathbf{I}_{p_j - p_{j-1} - 1} \end{pmatrix}
\end{aligned}$$

The diagonal blocks for the information matrix for the one-point design have the following forms

$$\begin{aligned}
&\mathbf{M}(\mathbf{x})_{jj} = \\
&\mathrm{block-diag}\left(\sigma^{(11)}\mathbf{M}_1(x_1), \sigma^{(22)}\mathbf{M}_1(x_1), \sigma^{(22)}\mathbf{G}_2(x_2), ..., \sigma^{(mm)}\mathbf{M}_1(x_1), \sigma^{(mm)}\mathbf{G}_2(x_2), ..., \sigma^{(mm)}\mathbf{G}_m(x_m)\right) \\
&\mathbf{M}_1(x_1) = \mathbf{f}_1(x_1)\mathbf{f}_1(x_1)^\top, \ \mathbf{G}_j(x_j) = \mathbf{g}_j(x_j)\mathbf{g}_j(x_j)^\top, \ j = 1, ..., m
\end{aligned}$$

The information matrix for the product type design has the following block diagonal form under conditions (6.23)

$$\begin{aligned}
&\mathbf{M}(\xi) = \\
&\mathrm{block-diag}\left(\sigma^{(11)}\mathbf{M}_1(\xi_1), \sigma^{(22)}\mathbf{M}_1(\xi_1), \sigma^{(22)}\mathbf{G}_2(\xi_2), ..., \sigma^{(mm)}\mathbf{M}_1(\xi_1), \sigma^{(mm)}\mathbf{G}_2(\xi_2), ..., \sigma^{(mm)}\mathbf{G}_m(\xi_m)\right) \\
&\mathbf{M}_1(\xi_1) = \int \mathbf{f}_1(x_1)\mathbf{f}_1(x_1)^\top \xi_1(\mathrm{d}x_1)
\end{aligned}$$

And the weight matrix has the following block diagonal form

$$\mathbf{L} = \mathrm{block-diag}\left(\mathbf{L}_1,\ \mathbf{L}_1,\ \mathbf{L}_2,\ ...,\ \mathbf{L}_1,\ \mathbf{L}_2,\ ...,\ \mathbf{L}_m\right)$$

Thus, theorem (6.8) can be proven analogous to theorems 6.3 and 6.5, by replacing these quantities with their counterparts there. □

6.3 Optimal Designs for More General Nesting Structures

6.3.1 Nesting of a New Component in all Other Components Simultaneously

When we have a nested SUR model, where a new different component is nested through all m components of the SUR model then the regression function of the SUR model given in (6.1) has the following form

$$\mathbf{f}(\mathbf{x}) = \text{block}-\text{diag}\,(\mathbf{f}_j(x_j, x_{m+1}))_{j=1,\ldots,m} = \begin{pmatrix} \mathbf{f}_1(x_1, x_{m+1}) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x_m, x_{m+1}) \end{pmatrix} \tag{6.24}$$

Where the regression functions for the marginal nested components have in general the next form for product-type models

$$\mathbf{f}_j(x_j, x_{m+1})_{j=1,\ldots,m} = \mathbf{f}_{m+1}(x_{m+1}) \otimes \mathbf{f}_j(x_j) \tag{6.25}$$

And respectively for additive models, where the regression functions must have intercepts

$$\mathbf{f}_j(x_j, x_{m+1})_{j=1,\ldots,m} = \begin{pmatrix} 1 \\ \mathbf{g}_j(x_j) \\ \mathbf{g}_{m+1}(x_{m+1}) \end{pmatrix} \tag{6.26}$$

Theorem 6.9. *Let ξ_j^* be D- or linear optimal by block diagonal weight matrix $\mathbf{L}_j$, for the j-th marginal component with the regression function given in (6.25) in the marginal design region $\mathcal{X}_j$, $j = 1, ..., m+1$, then the product type design*

$$\xi^* = \otimes_{j=1}^{m+1} \xi_j^*$$

is D- or linear optimal, if the weight matrix $\mathbf{L} = \text{block}-\text{diag}\,(\mathbf{L}_j)$ (4.1), for SUR model (3.3) with the following regression function

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}_{m+1}(x_{m+1}) \;\otimes\; \text{block}-\text{diag}\;\left(\mathbf{f}_j(x_j)\right)_{j=1,\ldots,m}$$

in the design region $\mathcal{X} = \times_{j=1}^{m+1} \mathcal{X}_j$.
The sensitivity function φ_D does not depend on $\mathbf{\Sigma}$.

Proof: D- or linear optimality in 4.1 of the product type design $\otimes_{j=1}^{m+1} \xi_j$ can be proven similarly to the proof of theorems 6.2 and 6.1, because of theorems 2.4, 2.5, by replacing the following different quantities with their counterparts.
The information matrix has the following form for one-point design and SUR model with the multiplicative marginals given in the form (6.25)

$$\mathbf{M}_1(\mathbf{x}) = \;\mathbf{M}_{m+1}(x_{m+1}) \;\otimes\; \left(\mathbf{M}_{jk}(x_j \otimes x_k)\right)_{j,k=1,\ldots,m}$$

The information matrix has with respect to the product type design $\otimes_{j=1}^{m+1} \xi_j$ the following form

$$\mathbf{M}_1(\xi) = \;\mathbf{M}_{m+1}(\xi_{m+1}) \;\otimes\; \left(\mathbf{M}_{jk}(\xi_j \otimes \xi_k)\right)_{j,k=1,\ldots,m}$$

The weight matrix has the following form

$$\mathbf{L} = \mathbf{L}_{m+1} \otimes \text{block}-\text{diag}\;(\mathbf{L}_j)_{j=1,\ldots,m}$$

Theorem 6.10. *Let ξ_j^* be D-optimal for the j-th marginal component with the regression function given in (6.26) in the marginal design region $\mathcal{X}_j$, $j = 1, ..., m+1$, then the product type design*

$$\xi^* = \otimes_{j=1}^{m+1} \xi_j^*$$

is D-optimal for the SUR model (3.3) with the following regression function given

$$\mathbf{f}(\mathbf{x}) = \text{ block} - \text{diag } \left(\begin{pmatrix} 1 \\ \mathbf{g}_j(x_j) \\ \mathbf{g}_{m+1}(x_{m+1}) \end{pmatrix} \right)_{j=1,...,m} \tag{6.27}$$

in the design region $\mathcal{X} = \times_{j=1}^{m+1} \mathcal{X}_j$.
The sensitivity function φ_D *does not depend on* $\mathbf{\Sigma}$.

Theorem 6.11. *Let* ξ_j^* *be linear optimal by block diagonal weight matrix* $\mathbf{L}_j$ *for the* j*-th marginal component with the regression function given in (6.26) in the marginal design region* $\mathcal{X}_j$, $j = 1, ..., m+1$. *If*

$$\int_{\mathcal{X}_j} \mathbf{g}_j(x_j)\xi_j^*(\mathrm{d}x_j) = \mathbf{0}, \ j = 1, 2, ..., m \tag{6.28}$$

the product type design

$$\xi^* = \otimes_{j=1}^{m+1} \xi_j^*$$

is linear optimal, if the weight matrix $\mathbf{L} = \text{block} - \text{diag}\,(\mathbf{L}_j)$ *(4.1), for the SUR model (3.3) with the regression function given in (6.27) in the design region* $\mathcal{X} = \times_{j=1}^{m+1} \mathcal{X}_j$.

Proof: D- and linear optimality in 4.1 of the product type design $\otimes_{j=1}^{m+1} \xi_j$ can be proven similarly to the proof of theorems 6.2 and 6.1, by replacing the the following different quantities with their counterparts there. □
The corresponding transformation matrix has the following form for $j = 2, ..., m$

$$\mathbf{A} = \text{block} - \text{diag}\,(\mathbf{A}_{j(m+1)})_{j=1,\ ...,\ m}, \text{ where}$$
$$\mathbf{A}_{j(m+1)} = \begin{pmatrix} 1 & \mathbf{0} & \mathbf{0} \\ -\mathbf{b}_j & \mathbf{I}_{p_j-1} & \mathbf{0} \\ -\mathbf{b}_{m+1} & \mathbf{0} & \mathbf{I}_{p_{m+1}-1} \end{pmatrix}, \ \mathbf{b}_j = \int \mathbf{g}_j(x_j)\xi_j(\mathrm{d}x_j), \ j = 1, \ ..., \ m+1$$

The diagonal blocks for the information matrix for the one-point design have the following forms

$$\mathbf{M}(\mathbf{x})_{jj} = \text{block} - \text{diag }(\sigma^{(11)}\mathbf{M}_1(x_1, x_{m+1}), \ \sigma^{(22)}\mathbf{M}_{12}(x_2, x_{m+1}), \ ..., \ \sigma^{(mm)}\mathbf{M}_{1m}(x_m, x_{m+1}))$$
$$\mathbf{M}_{jm}(x_j, x_{m+1}) = \mathbf{f}_j(x_j, x_{m+1})\mathbf{f}_j(x_j, x_{m+1})^\top, \ j = 1, ..., m$$

The information matrix for the product type design has the following block diagonal form under the conditions 6.28

$$\mathbf{M}(\xi) = \text{block} - \text{diag}\,(\sigma^{(11)}\mathbf{M}_1(\xi_1, \xi_{m+1}), \ \sigma^{(22)}\mathbf{M}_{12}(\xi_2, \xi_{m+1}), \ ..., \ \sigma^{(mm)}\mathbf{M}_{1m}(\xi_m, \xi_{m+1}))$$
$$\mathbf{M}_{jm}(\xi_j, \xi_{m+1}) = \int \mathbf{f}_1(x_j, x_{m+1})\mathbf{f}_1(x_j, x_{m+1})^\top(\xi_j(\mathrm{d}x_j) \otimes \xi_{m+1}(\mathrm{d}x_{m+1})), \ j = 1, ..., m$$

And the weight matrix has the following block diagonal form

$$\mathbf{L} = \text{block} - \text{diag}\,(\mathbf{L}_{1(m+1)}, \ \mathbf{L}_{2(m+1)}, ..., \ \mathbf{L}_{m(m+1)}), \ \mathbf{L}_{j(m+1)} = \begin{pmatrix} L_{j11} & \mathbf{L}_{j12}^\top & \mathbf{L}_{(m+1)13}^\top \\ \mathbf{L}_{j12} & \mathbf{L}_{j22} & \mathbf{L}_{(m+1)23}^\top \\ \mathbf{L}_{(m+1)13} & \mathbf{L}_{(m+1)23} & \mathbf{L}_{(m+1)33} \end{pmatrix} \quad \square$$

6.3.2 Nesting of a New Different Component in each Component

When we have a nested SUR model, where a new different component is nested through each component of the SUR model then the regression function of the SUR model given in (6.1) has the following form

$$\mathbf{f}(\mathbf{x}) = \text{block} - \text{diag}\,(\mathbf{f}_j(x_j, x_{k_j}))_{j=1,...,m} = \begin{pmatrix} \mathbf{f}_1(x_1, x_{k_1}) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x_m, x_{k_m}) \end{pmatrix} \tag{6.29}$$

Where the regression functions for the marginal nested components have in general the next form for product-type models

$$\mathbf{f}_j(x_j, x_{k_j})_{j=1,...,m} = \mathbf{f}_j(x_j) \otimes \mathbf{f}_{k_j}(x_{k_j}) \tag{6.30}$$

And respectively for additive models, where the regression functions must have intercepts

$$\mathbf{f}_j(x_j, x_{k_j})_{j=1,...,m} = \begin{pmatrix} 1 \\ \mathbf{g}_j(x_j) \\ \mathbf{g}_{k_j}(x_{k_j}) \end{pmatrix} \tag{6.31}$$

Corollary 6.1. *Let $\xi_j^* \otimes \xi_{k_j}^*$ be D- or linear optimal, by block diagonal weight matrix $\mathbf{L}_j$, for the j th marginal component (6.30) in the marginal design region $\mathcal{X}_j \times \mathcal{X}_{k_j}$, $j = 1, ..., m$, then the product type design*

$$\xi^* = \otimes_{j=1}^m \xi_j^* \otimes \xi_{k_j}^*$$

is D- or linear optimal , if the weight matrix $\mathbf{L} = \text{block} - \text{diag}\,(\mathbf{L}_j)$ (4.1), for SUR model (3.3) with the following regression function

$$\mathbf{f}(\mathbf{x}) = \text{block} - \text{diag}\ \left(\mathbf{f}_j(x_j) \otimes \mathbf{f}_{k_j}(x_{k_j})\right)_{j=1,...,m} \tag{6.32}$$

in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j \times \mathcal{X}_{k_j}$.
The sensitivity function φ_D does not depend on $\mathbf{\Sigma}$.

Corollary 6.2. *Let $\xi_j^* \otimes \xi_{k_j}^*$ be D- or linear optimal, by block diagonal weight matrix $\mathbf{L}_j$ for the j-th marginal component with the regression function given in (6.31) in the marginal design region $\mathcal{X}_j \times \mathcal{X}_{k_j}$, $j = 1, ..., m$, then the product type design*

$$\xi^* = \otimes_{j=1}^m \xi_j^* \otimes \xi_{k_j}^*$$

is D- or linear optimal , if the weight matrix $\mathbf{L} = \text{block} - \text{diag}\,(\mathbf{L}_j)$ (4.1), for SUR model (3.3) with the following regression function

$$\mathbf{f}(\mathbf{x}) = \text{block} - \text{diag}\ \left(\begin{pmatrix} 1 \\ \mathbf{g}_j(x_j) \\ \mathbf{g}_{k_j}(x_{k_j}) \end{pmatrix}\right)_{j=1,...,m} \tag{6.33}$$

in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j \times \mathcal{X}_{k_j}$.
The sensitivity function φ_D does not depend on $\mathbf{\Sigma}$.

Proof: The considered marginals are multi-factor but different for each components, so that can be considered as a special case of the topics of theorems 4.3 and 4.4 but with different multi-factor marginals, as follows $\mathbf{x}_j = (x_j, x_{k_j})_{j=1,\dots,m} \in \mathcal{X}_j = \mathcal{X}_j \times \mathcal{X}_{k_j}$, so the regression function for the considered SUR model has the form block $-$ diag $(\mathbf{f}_j(\mathbf{x}_j))_{j=1,\dots,m}$, thus the proofs can be implemented analogous to the proofs of theorems 4.3 and 4.4 in the third chapter about the D- and linear optimality in 4.1 of the product type designs for the considered SUR model and because of theorems 2.4, 2.5, 2.6 and 2.7, the corollaries 6.1, and 6.2 are proven. □

The regression functions for the SUR model (6.1) with the marginal regression functions given in (6.2) and (6.7) and the corresponding results can be generalized complicatedly but analogous, so more than one component is nested through the other components and we can do that analogous for the regression functions for the SUR model (6.24) and resp. (6.25) and (6.26) and the regression functions for the SUR model (6.29) and resp. (6.30) and (6.31).

Remark 6.1. *Theorems 6.1, 6.2 , 6.3, 6.5, 6.6, 6.8, 6.9, and 6.10 , and corollaries 6.1, and 6.2 may fail to hold, if the regression functions of the marginal components do not contain an intercept.*

Remark 6.2. *When all regression functions and the experiments regions are equal for all components, then the designs problem for these kinds of multivariate linear models by known variance covariance matrices of the error variables are the same in the corresponding univariate multi-factor models. The proof is fully analogous to the proof of the reduction the problem with respect to homogeneous components (MANOVA) in the work of (* Kurotschka and Schwabe (1996)*).*

Example 6.3 (A Multiplicative SUR Model)**.** *To illustrate the results for SUR models, which are nested through product-type models we consider the SUR model with nested product-type models through the same factor for the components,*

$$Y_{ij} = \beta_{j0} + \beta_{j1}x_{ij} + \beta_{j2}x_{i3} + \beta_{j3}x_{ij}x_{i3} + \varepsilon_{ij}\,. \tag{6.34}$$

in the unit intervals $\mathcal{X}_1 = \mathcal{X}_2 = [0,1], \mathcal{X}_3 = [-1,1]$ as experimental regions. Then the IMSE- and D-optimal designs for the product-type marginal models $1, 2$ are the product designs of the IMSE- and D-optimal designs of the corresponding one-factor models because of theorems 2.4 and 2.5

$$\xi^*_{D;IMSE,1} \otimes \xi^*_{D;IMSE,3} = \xi^*_2 \otimes \xi^*_3 = \begin{pmatrix} (1,1) & (1,-1) & (0,1) & (0,-1) \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$$

And the A-optimal design

$$\xi^*_{A,1} \otimes \xi^*_{A,2} = \begin{pmatrix} (1,1) & (1,-1) & (0,1) & (0,-1) \\ \frac{\sqrt{2}-1}{2} & \frac{\sqrt{2}-1}{2} & \frac{2-\sqrt{2}}{2} & \frac{2-\sqrt{2}}{2} \end{pmatrix}$$

Also, the product designs of the following form are respectively D-, IMSE-, and A-optimal for the considered model, because of theorems 6.1 and 6.2, i.e. $\xi^*_{D;IMSE} = \xi^*_{D;IMSE,1} \otimes \xi^*_{D;IMSE,2} \otimes \xi^*_{D;IMSE,3}$, or

$$\xi^*_{D;IMSE} = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \end{pmatrix}$$

For A-optimality

$$\xi^{*A} = \left(\begin{array}{cccccccc} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ w_1^2/4 & w_1^2/4 & w_3/4 & w_3/4 & w_3/4 & w_3/4 & w_2^2/4 & w_2^2/4 \end{array} \right)$$

Where $w_1 = \sqrt{2} - 1$, $w_2 = 2 - \sqrt{2}$ *and* $w_3 = w_1.w_2$ *are respectively D- resp. IMSE- and A-optimal. The corresponding sensitivity function for D-optimality*

$$\varphi_D(\mathbf{x}; \xi_D^*) = \text{trace}(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi_D^*)^{-1}\mathbf{f}(\mathbf{x})) = 4(1 + x_3^2)(1 + x_1^2 - x_1 + x_2^2 - x_2) \tag{6.35}$$

is plotted for $x_2 = 1$ *in figure 6.1. It can be easily seen that the sensitivity function is independent on* $\mathbf{\Sigma}$ *and satisfies the condition* $\varphi_D(\mathbf{x}; \xi_D^*) \leq p = 8$ *for all* $\mathbf{x} \in \mathcal{X}$.
The corresponding sensitivity function for IMSE-optimality is equal to or less than $\text{trace}(\mathbf{LM}(\xi_{IMSE}^*)^{-1})$

$$\varphi_{IMSE}(\mathbf{x}; \xi_{IMSE}^*) = \text{trace}(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi_{IMSE}^*)^{-1}\mathbf{LM}(\xi_{IMSE}^*)^{-1}\mathbf{f}(\mathbf{x})) \leq \text{trace}(\mathbf{LM}(\xi_{IMSE}^*)^{-1})$$

and the function

$$\frac{\varphi_{IMSE}(\mathbf{x}; \xi_{IMSE}^*)}{\text{trace}\left(\mathbf{L}\ \mathbf{M}(\xi_{IMSE}^*)^{-1}\right)} = \frac{\text{trace}(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi_{IMSE}^*)^{-1}\mathbf{LM}(\xi_{IMSE}^*)^{-1}\mathbf{f}(\mathbf{x}))}{\text{trace}(\mathbf{LM}(\xi_{IMSE}^*)^{-1})} = \frac{1}{4}(x_3^2 + 3) \tag{6.36}$$

for $\sigma_1 = \sigma_2 = 1, \rho = 0.99, x_2 = 1$ *and it is plotted in figure 6.2. It can be easily seen that the sensitivity function for IMSE-optimality satisfies the condition* $\varphi_{IMSE}(\mathbf{x}; \xi_{IMSE}^*) \leq \text{trace}(\mathbf{LM}(\xi_{IMSE}^*)^{-1})$ *for all* $\mathbf{x} \in \mathcal{X}$. *The corresponding sensitivity function for A-optimality is equal to or less than* $\text{trace}(\mathbf{M}(\xi_A^*)^{-1})$

$$\varphi_A(\mathbf{x}; \xi_A^*) = \text{trace}(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi_A^*)^{-1}\mathbf{M}(\xi_A^*)^{-1}\mathbf{f}(\mathbf{x})) \leq \text{trace}(\mathbf{M}(\xi_A^*)^{-1}) \tag{6.37}$$

and the function

$$\frac{\varphi_A(\mathbf{x}; \xi_A^*)}{\text{trace}\left(\mathbf{M}(\xi_A^*)^{-1}\right)} = \frac{\text{trace}(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi_A^*)^{-1}\mathbf{M}(\xi_A^*)^{-1}\mathbf{f}(\mathbf{x}))}{\text{trace}(\mathbf{M}(\xi_A^*)^{-1})} \tag{6.38}$$

for $\sigma_1 = 0.1, \sigma_2 = 10, \rho = 0.2, x_2 = 1$ *is plotted in figure 6.3. It can be easily seen that the sensitivity function for A-optimality satisfies the condition* $\varphi_A(\mathbf{x}; \xi_A^*) \leq \text{trace}(\mathbf{M}(\xi_A^*)^{-1})$ *for all* $\mathbf{x} \in \mathcal{X}$.

Example 6.4 (An Additive SUR Model)**.** *To illustrate the results for SUR models, which are nested through additive models we consider the SUR model with two components, where one of the components is nested through the second components as an additive model*

$$\begin{aligned} Y_{i1} &= \beta_{10} + \beta_{11}x_{i1} + \beta_{22}x_{i2}^2 + \varepsilon_{i1} \\ Y_{i2} &= \beta_{20} + \beta_{21}x_{i2}^2 + \varepsilon_{i2} \end{aligned} \tag{6.39}$$

in the unit intervals $\mathcal{X}_1 = \mathcal{X}_2 = [0, 1]$ *as experimental regions. Then the D-optimal design for the additive model 1 is the product design of the D-optimal designs of the corresponding one-factor models because of theorem 2.6 , but that is not the case for A- and IMSE-Optimality because of theorem 2.7*

$$\xi^* = \xi_1^* \otimes \xi_2^* = \begin{pmatrix} \textit{Points} & (1,1) & (1,0) & (0,1) & (0,0) \\ D & 1/4 & 1/4 & 1/4 & 1/4 \\ \textit{IMSE} & 1/4 & 1/4 & 1/4 & 1/4 \\ A & (\sqrt{2}-1)^2 & (\sqrt{2}-1)(2-\sqrt{2}) & (\sqrt{2}-1)(2-\sqrt{2}) & (2-\sqrt{2})^2 \end{pmatrix}$$

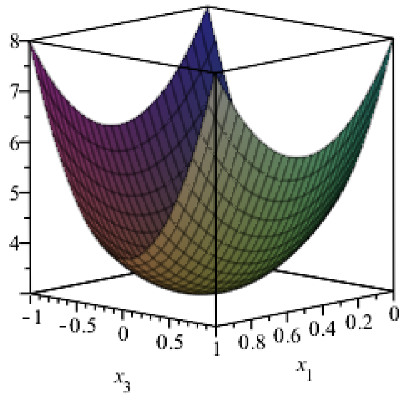

Figure 6.1: Function 6.35 for ξ_D^* (SUR with multiplicative marginals)

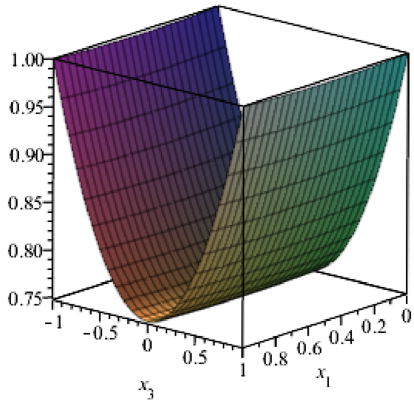

Figure 6.2: Function 6.36 for ξ_{IMSE}^* (SUR with multiplicative marginals)

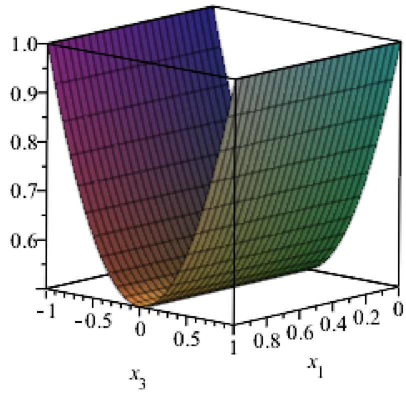

Figure 6.3: Function 6.38 for ξ_A^* (SUR with multiplicative marginals)

Also, the product designs ξ with the corresponding weights are respectively D- but not A- and IMSE-optimal for the considered SUR model because of theorems 6.3 and 6.5.
The corresponding sensitivity function for D-optimality

$$\varphi_D(\mathbf{x};\xi_D^*) = \mathrm{trace}(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi_D^*)^{-1}\mathbf{f}(\mathbf{x})) = 5 - 4\,x_1 - 8\,{x_2}^2 + 4\,{x_1}^2 + 8\,{x_2}^4 \tag{6.40}$$

is plotted in figure 6.4. It can be easily seen that the sensitivity function is independent on $\mathbf{\Sigma}$ and satisfies the condition $\varphi(\mathbf{x};\xi_D^*) \leq p = 5$ for all $\mathbf{x} \in \mathcal{X}$.
The corresponding sensitivity function for IMSE-optimality is in some cases bigger than $\mathrm{trace}(\mathbf{LM}(\xi_{IMSE}^*)^{-1})$

$$\varphi_{IMSE}(\mathbf{x};\xi_{IMSE}^*) = \mathrm{trace}(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi_{IMSE}^*)^{-1}\mathbf{LM}(\xi_{IMSE}^*)^{-1}\mathbf{f}(\mathbf{x})) \nleq \mathrm{trace}(\mathbf{LM}(\xi_{IMSE}^*)^{-1})$$

and the function

$$\frac{\mathrm{trace}(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi_{IMSE}^*)^{-1}\mathbf{LM}(\xi_{IMSE}^*)^{-1}\mathbf{f}(\mathbf{x}))}{\mathrm{trace}(\mathbf{LM}(\xi_{IMSE}^*)^{-1})} =$$
$$1.454545249 + 0.000001808909177\,{x_1}^2 - 0.000001808909177\,x_1 + 1.272726697\,{x_2}^4 - 2.181817195\,{x_2}^2 \tag{6.41}$$

for $\sigma_1 = 0.1, \sigma_2 = 10, \rho = 0.99$ and it is plotted in figure 6.5. It can be easily seen that the sensitivity function for the product design of the marginal doesn,t satisfy the condition $\varphi_{IMSE}(\mathbf{x};\xi_{IMSE}^*) \leq$ $\mathrm{trace}(\mathbf{LM}(\xi^*)^{-1})$ *, also, $\xi_{1;IMSE} \otimes \xi_{2;IMSE}$ is not IMSE-optimal. The corresponding sensitivity function for A-optimality is equal to or less than* $\mathrm{trace}(\mathbf{M}(\xi_{IMSE}^*)^{-1})$

$$\varphi_A(\mathbf{x};\xi_A^*) = \mathrm{trace}(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi_A^*)^{-1}\mathbf{M}(\xi_A^*)^{-1}\mathbf{f}(\mathbf{x})) \nleq \mathrm{trace}(\mathbf{M}(\xi_A^*)^{-1}) \tag{6.42}$$

and the function

$$\frac{\mathrm{trace}(\mathbf{\Sigma}^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi_A^*)^{-1}\mathbf{M}(\xi_A^*)^{-1}\mathbf{f}(\mathbf{x}))}{\mathrm{trace}(\mathbf{M}(\xi)^{-1})} =$$
$$1.0610 + 0.35366\,x_1{x_2}^2 - 1.3537\,x_1 - 2.5606\,{x_2}^2 + 1.2073\,{x_1}^2 + 2.4146\,{x_2}^4 \tag{6.43}$$

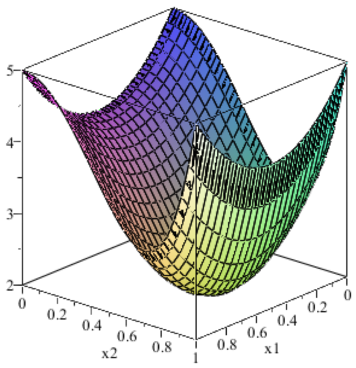

Figure 6.4: Function 6.40 for ξ_D^* (SUR with additive marginals)

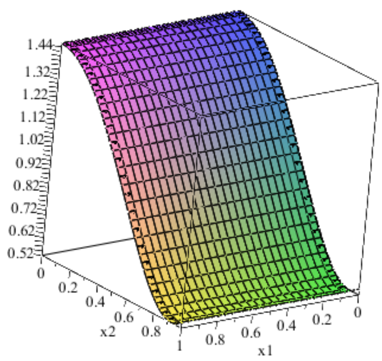

Figure 6.5: Function 6.41 for ξ_{IMSE} (SUR with additive marginals)

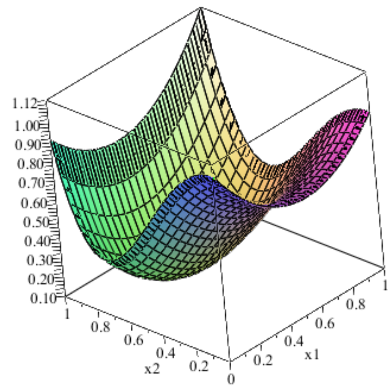

Figure 6.6: Function 6.43 for ξ_A (SUR with additive marginals)

for $\sigma_1 = \sigma_2 = 5, \rho = 0$ *is plotted in figure 6.6. It can be easily seen that the sensitivity function for A-optimality doesn,t satisfy the condition* $\varphi_A(\mathbf{x}; \xi_A^*) \not\leq \mathrm{trace}(\mathbf{M}(\xi_A^*)^{-1})$ *for all* $\mathbf{x} \in \mathcal{X}$.

7 Optimal Designs for Some SUR Models without Intercepts in their Components

Univariate linear regression models without intercepts are explored in many works see *(Eisenhauer (2003))* and *(Casella (1983))*, as well as their multivariate versions in the work *(Peng et al. (2010))*. Optimal designs for univariate linear models without intercepts are explored for example in *(Huang et al. (1995))* and *(Chang and Heiligers (1996))*. The practical model used in these works was a regression equation without an intercept for the speed of a car with respect to the distance needed to stop. Another physical or mechanical model is the equation for the speed of the car with respect to the time. Thus, if the both equations are to be observed for a car by the distance as a control variable for the first marginal model and the time as a control variable for the second marginal model , then they follow a bivariate SUR model without intercepts by the components. The optimal designs for such SUR models resp. multivariate linear models without intercepts in their components are not explored in the literature for correlated components, but this problem is explored for multivariate linear models by correlated observations for practical cases as the spring weighing resp. chemical balance regression models see for example *(Ceranka and Katulska (1987))*. Additional problems like the multi or multivariate spring resp. chemical balance weighing can be modeled as SUR models, where the marginal models are univariate spring resp. chemical balance weighing for different objects resp. different number of it. Optimal designs for such univariate models have been discussed, for example by *(Huda and Mukerjee (1988))* or *(Schwabe (1996))*.
This chapter is organized as follows, the D- and linear optimal designs in 4.1 for SUR models with different marginal models by non and block diagonal information matrices are explored in the first section. It is shown, that the product type designs are D- and linear optimal designs in 4.1 for block diagonal information matrices, and the multivariate chemical balance regression models are a practical case for such models. Where there is a relationship between the chemical balance designs and and the balance incomplete block designs see *(Ceranka and Katulska (1987))*. It is concluded, that the product type designs are D-optimal for SUR models without intercepts by the marginals for non block diagonal or arbitrary form of the information matrices, , when the correlation term belongs to an interval around zero and the interval length is monotonously falling due to the number of the components of the model m, i.e. the D-optimality for the product type designs in this situation are restricted. Such intervals for the correlation term are explored for optimal designs for a bivariate multivariate linear model by some examples in the work of *(Changa et al. (2001))*. A practical case for such SUR models are the multivariate spring weighing regression models, where there are local couplings between the D-optimal designs for the SUR models without intercepts by the marginals and the multi-factor (additively) or without interactions univariate model, which are formulated from the same marginal models of the SUR models, because of the relationship between the information matrices resp. the sensitivity functions due to D-optimality for the both models. The results of the first section are illustrated and supported by many simulations and examples in the second section . So the reduction of the optimal design problem from the multivariate case for SUR models without intercepts, to its corresponding univariate case by the optimality of the product type design is restricted for D-optimality by non block

diagonal information matrix, but by a simple example can be shown, that the product type designs are not linear optimal in 4.1 designs for SUR models without intercepts by non block diagonal information matrices.

7.1 Optimal Designs for Spring Weighing and Chemical Balance Regression Models

The multiple regression model is a special case of the additive models, and multiple regression without an intercept can be defined as the spring weighing regression model when the control variables take the values $0, 1$ and as chemical balance regression model when the control variables take the values $-1, 0, 1$. It is assumed for the next two definitions, that $\mathrm{E}(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2$, where $\sigma > 0$ and is known, and $\mathrm{Cov}\,(\varepsilon_i, \varepsilon_k) = 0; i \neq k$, $x_j \in \mathcal{X}_j$.

Def 7.1 (The Spring weighing regression Models). (Schwabe (1996)*) The spring weighing regression model describes the weighing m objects on an unbiased spring balance. If we denote the weight of the j-th object by β_j, $j = 1, ..., m$, the experiment can be modeled according to*

$$\eta\ (x_1, ..., x_m) = \sum_{j=1}^{m} x_j \beta_j$$

$x_j \in \{0, 1\}$, $j = 1, ..., m$, where x_j equals one or zero corresponding to whether the j-th object lies on the pan or not. As the balance is assumed to be unbiased no constant term occurs.

The marginal models are all identical

$$\eta\ _j(x_j) = x_j \beta_j \tag{7.1}$$

$x_j \in \{0, 1\}$, and represent the experimental situation for weighing one object on a spring balance. So the Spring weighing regression model for for $m-$objects is specified as follows

$$Y(x_1, ..., x_m) = \sum_{j=1}^{m} x_j \beta_j + \varepsilon \tag{7.2}$$

Def 7.2 (The Chemical Balance regression Models). (Schwabe (1996)*) The chemical balance regression models can be defined similarly to (7.1), but now on an unbiased chemical balance. If β_j denotes again the weight of the j-th object , $j = 1, ..., m$, the experiment can be modeled according to*

$$\eta\ (x_1, ..., x_m) = \sum_{j=1}^{m} x_j \beta_j$$

as in the definition (7.1) for the spring balance, but with a different design space $\mathcal{X}_j = \{-1, 0, 1\}$, $j = 1, ..., m$.

The design points have the following interpretation: x_j equals to 1 if the j-th object is in the left pan, to -1 if it is in the right pan, and to 0 if it is not present in the weighing arrangement. Again no constant term occurs since the balance is assumed to be unbiased. The marginal models are all identical

$$\eta_{\ j}(x_j) = x_j \beta_j, \ \ i = 1, ..., n$$

$x_j \in \{-1, 0, 1\}$, and represent the experimental situation for weighing one object on a chemical balance. So the Spring weighing regression model for $m-$objects is specified as follows

$$Y(x_1, ..., x_m) = \sum_{j=1}^{m} x_j \beta_j + \varepsilon \tag{7.3}$$

The spring weighing resp. the chemical balance regression models were considered foremost in the discrete resp. exact case, but dealing with the approximate case is easier, see *(Huda and Mukerjee (1988))* or *(Schwabe (1996))*. For the spring weighing regression models the product type designs are not D- resp. G- or linear optimal even for the simplest example for two objects because the information matrix is non-diagonal. But the product type designs are D- resp. G- and linear optimal because of the block diagonal form of the corresponding information matrix for the chemical balance regression models, see *(Schwabe (1996))*.

7.2 The Theoretical Results and Practical Cases

The proofs of the following results depend on the validity of the conditions of the corresponding equivalence theorem and some techniques from the matrix theory

Theorem 7.1. *Let ξ_j^* be Φ-optimal for the j-th marginal component with or without intercepts in the marginal design region $\mathcal{X}_j$, $j = 1, ..., m$. If the marginal components are independedet, ie. all correlation terms $\rho_{j\acute{j}} = 0, j, \acute{j} = 1, ..., m$, or the regression functions are orthogonal to a constant with respect to the Φ-optimal design, i.e.*

$$\int_{\mathcal{X}_j} \mathbf{f}_j(x_j) \xi_j^*(\mathrm{d}x_j) = \mathbf{0}, \ \ j = 1, ..., m \tag{7.4}$$

then the product type design

$$\xi^* = \otimes_{j=1}^{m} \xi_j^*$$

is Φ -optimal for SUR model (3.3) in the design region $\mathcal{X} = \times_{j=1}^{m} \mathcal{X}_j$. Where Φ-optimal can be D- or linear optimal criterion by block diagonal weight matrix $\mathbf{L} = \mathrm{block-diag}\,(\mathbf{L}_j)$ (4.1).
The sensitivity function φ_D does not depend on $\mathbf{\Sigma}$.

Proof: If $\rho_{j\acute{j}} = 0, j, \acute{j} = 1, ..., m$ or $\int_{\mathcal{X}_j} \mathbf{f}_j(x_j) \xi_j^*(\mathrm{d}x_j) = \mathbf{0}$, $j = 1, ..., m$, the information matrix has the following block diagonal form under the conditions (7.4)

$$\mathbf{M}(\xi^*) = \begin{pmatrix} \sigma^{(11)} \mathbf{M}_1(\xi_1^*) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \sigma^{(mm)} \mathbf{M}_m(\xi_m^*) \end{pmatrix}$$

Then the inverse of the information matrix has the following block-diagonal form, for $j = 1, ..., m$

$$\mathbf{M}^{-1}(\xi^*) = \begin{pmatrix} \frac{1}{\sigma^{(11)}}\mathbf{M}_1^{-1}(\xi_1^*) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{\sigma^{(mm)}}\mathbf{M}_m^{-1}(\xi_m^*) \end{pmatrix} \tag{7.5}$$

And resp. the equivalence theorem for D-optimality has the following form

$$\begin{aligned} &\mathrm{trace}\left(\Sigma^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi^*)^{-1}\mathbf{f}(\mathbf{x})\right) = \mathrm{trace} \begin{pmatrix} \sigma^{(11)} & \cdots & \sigma^{(1m)} \\ \vdots & \ddots & \vdots \\ \sigma^{(1m)} & \cdots & \sigma^{(mm)} \end{pmatrix} \\ &\begin{pmatrix} \mathbf{f}_1(x_1)^\top & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x_m)^\top \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma^{(11)}}\mathbf{M}_1^{-1}(\xi_1^*) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{\sigma^{(mm)}}\mathbf{M}_m^{-1}(\xi_m^*) \end{pmatrix} \begin{pmatrix} \mathbf{f}_1(x_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x_m) \end{pmatrix} \\ &= \mathrm{trace} \begin{pmatrix} \sigma^{(11)}\mathbf{f}_1(x_1)^\top & \cdots & \sigma^{(1m)}\mathbf{f}_m(x_m)^\top \\ \vdots & \ddots & \vdots \\ \sigma^{(1m)}\mathbf{f}_1(x_1)^\top & \cdots & \sigma^{(mm)}\mathbf{f}_m(x_m)^\top \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma^{(11)}}\mathbf{M}_1^{-1}(\xi_1^*)\mathbf{f}_1(x_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{\sigma^{(mm)}}\mathbf{M}_m^{-1}(\xi_m^*)\mathbf{f}_m(x_m) \end{pmatrix} \\ &= \mathrm{trace} \begin{pmatrix} \mathbf{f}_1(x_1)^\top \mathbf{M}_1^{-1}(\xi_1^*)\mathbf{f}_1(x_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_2(x_2)^\top \mathbf{M}_2^{-1}(\xi_2^*)\mathbf{f}_2(x_2) \end{pmatrix} \\ &= \mathbf{f}_1(x_1)^\top \mathbf{M}_1^{-1}(\xi_1^*)\mathbf{f}_1(x_1) + \cdots + \mathbf{f}_m(x_m)^\top \mathbf{M}_m^{-1}(\xi_m^*)\mathbf{f}_m(x_m) \le p_1 + \cdots + p_m = p \end{aligned} \tag{7.6}$$

And with there the product type design are D-optimal and the sensitivity function for the D-optimality does not contain any covariance terms.
The equivalence theorem for linear optimality in 4.1 has the following form for arbitrary variance covariance matrix of the error variables.

$$\begin{aligned} &\mathrm{trace}\left(\Sigma^{-1}\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi^*)^{-1}\mathbf{L}\mathbf{M}(\xi^*)^{-1}\mathbf{f}(\mathbf{x})\right) = \mathrm{trace} \begin{pmatrix} \sigma^{(11)} & \cdots & \sigma^{(1m)} \\ \vdots & \ddots & \vdots \\ \sigma^{(1m)} & \cdots & \sigma^{(mm)} \end{pmatrix} \begin{pmatrix} \mathbf{f}_1(x_1)^\top & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x_m)^\top \end{pmatrix} \\ &\begin{pmatrix} \frac{1}{\sigma^{(11)}}\mathbf{M}_1^{-1}(\xi_1^*) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{\sigma^{(mm)}}\mathbf{M}_m^{-1}(\xi_m^*) \end{pmatrix} \begin{pmatrix} \mathbf{L}_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{L}_m \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma^{(11)}}\mathbf{M}_1^{-1}(\xi_1^*) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{\sigma^{(mm)}}\mathbf{M}_m^{-1}(\xi_m^*) \end{pmatrix} \\ &\begin{pmatrix} \mathbf{f}_1(x_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x_m) \end{pmatrix} = \mathrm{trace} \begin{pmatrix} \sigma^{(11)}\mathbf{f}_1(x_1)^\top & \cdots & \sigma^{(1m)}\mathbf{f}_m(x_m)^\top \\ \vdots & \ddots & \vdots \\ \sigma^{(1m)}\mathbf{f}_1(x_1)^\top & \cdots & \sigma^{(mm)}\mathbf{f}_m(x_m)^\top \end{pmatrix}. \end{aligned} \tag{7.7}$$

$$\begin{pmatrix} \frac{1}{(\sigma^{(11)})^2}\mathbf{M}_1^{-1}(\xi_1^*)\mathbf{L}_1\mathbf{M}_1^{-1}(\xi_1^*)\mathbf{f}_1(x_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{(\sigma^{(mm)})^2}\mathbf{M}_m^{-1}(\xi_m^*)\mathbf{L}_m\mathbf{M}_m^{-1}(\xi_m^*)\mathbf{f}_m(x_m) \end{pmatrix}$$

$$= \frac{1}{\sigma^{(11)}}\mathbf{f}_1(x_1)^\top\mathbf{M}_1^{-1}(\xi_1^*)\mathbf{L}_1\mathbf{M}_1^{-1}(\xi_1^*)\mathbf{f}_1(x_1) + \cdots + \frac{1}{\sigma^{(mm)}}\mathbf{f}_m(x_m)^\top\mathbf{M}_m^{-1}(\xi_m^*)\mathbf{L}_m\mathbf{M}_m^{-1}(\xi_m^*)\mathbf{f}_m(x_m) \tag{7.8}$$

The left side of the equivalence theorem for the linear criteria in 4.1 has the following form

$$\begin{aligned} \text{trace}\left(\mathbf{L}\mathbf{M}(\xi^*)^{-1}\right) &= \text{trace}\left(\begin{pmatrix} \mathbf{L}_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{L}_m \end{pmatrix}\begin{pmatrix} \frac{1}{\sigma^{(11)}}\mathbf{M}_1^{-1}(\xi_1^*) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{\sigma^{(mm)}}\mathbf{M}_m^{-1}(\xi_m^*) \end{pmatrix}\right) \\ &= \frac{1}{\sigma^{(11)}}\text{trace}\left(\mathbf{L}_1\mathbf{M}_1^{-1}(\xi_1^*)\right) + \cdots + \frac{1}{\sigma^{(mm)}}\text{trace}\left(\mathbf{L}_m\mathbf{M}_m^{-1}(\xi_m^*)\right) \end{aligned} \tag{7.9}$$

the conditions of the equivalence theorem for the linear optimality in 4.1 are satisfied, by comparing of (7.8) with (7.9), and therewith the product type designs are linear optimal in 4.1. □
With respect to the theorem (7.1) , when the correlation is non Zero or the conditions (7.4) are not valid, then the product type designs may be D-optimal designs for SUR models without intercepts by the marginal components. The correlation term plays also a main role by the D-optimality of the product-type-designs. The product type designs are not D-optimal for correlation term in absolute terms close to one, but they are D-optimal for correlation term in an interval containing zero as $[\alpha_m, \gamma_m]$, where $-1 < \alpha_m < 0 < \gamma_m < 1$. This role of the correlation term derived from the dependency of the sensitivity function with respect to the equivalence theorem for D-optimality on the correlation term. We have considered homogeneous correlation structure by the next results to research the intervals of the correlation term, by which the product-type-designs are D-optimal.
We can derive the correlation matrix as shown in the following remark.

Remark 7.1. *The correlation matrix is given as follows*

$$\mathbf{C}_m = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \ddots & \ddots & \vdots \\ \rho & \cdots & \rho & 1 \end{pmatrix} \tag{7.10}$$

this matrix is positive definite, when $-\frac{1}{m-1} < \rho < 1$. *It is not difficult to show, that for the larger dimension m became the lesser determinant of the correlation matrix, also* $\det \mathbf{C}_{3\times3} \leq \det \mathbf{C}_{2\times2}$ *and so on.*

The next lemma supplies the inverse of the correlation matrix by the homogeneous structure, and we can remark, that the determinant of the correlation matrix is $\det \mathbf{C}_m = (1-\rho)^{m-1}((m-1)\rho+1)$.

Lemma 7.1. *The inverse of the correlation matrix 7.10 has the next form*

$$\mathbf{C}_m^{-1} = \frac{(1-\rho)^{m-2}}{(1-\rho)^{m-1}((m-1)\rho+1)}\begin{pmatrix} ((m-2)\rho+1) & -\rho & \cdots & -\rho \\ -\rho & (m-2)\rho+1 & \cdots & -\rho \\ \vdots & \vdots & \ddots & \vdots \\ -\rho & \cdots & -\rho & (m-2)\rho+1 \end{pmatrix}$$

or

$$invcor\mathbf{C}_m^{-1} = \frac{1}{(1-\rho)((m-1)\rho+1)}\begin{pmatrix} (m-2)\rho+1 & -\rho & \cdots & -\rho \\ -\rho & (m-2)\rho+1 & \cdots & -\rho \\ \vdots & \vdots & \ddots & \vdots \\ -\rho & \cdots & -\rho & (m-2)\rho+1 \end{pmatrix} \quad (7.11)$$

Proof: to prove the lemma, should the multiplication of the matices in 7.10 and 7.11 equal to the identity matrix, so

$$\begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \ddots & \ddots & \vdots \\ \rho & \cdots & \rho & 1 \end{pmatrix} \cdot \frac{1}{(1-\rho)((m-1)\rho+1)}\begin{pmatrix} (m-2)\rho+1 & -\rho & \cdots & -\rho \\ -\rho & (m-2)\rho+1 & \cdots & -\rho \\ \vdots & \vdots & \ddots & \vdots \\ -\rho & \cdots & -\rho & (m-2)\rho+1 \end{pmatrix}$$

the diagonal elements of the resulted matrix are equal to

$$\frac{(m-2)\rho+1-(m-1)\rho}{(1-\rho)((m-1)\rho+1)} = \frac{(1-\rho)((m-1)\rho+1)}{(1-\rho)((m-1)\rho+1)} = 1$$

the non-diagonal elements of the resulted matrix are equal to

$$\frac{-\rho+(m-2)\rho^2+\rho-(m-2)^2\rho^2}{(1-\rho)((m-1)\rho+1)} = 0$$

and therewith the lemma is proven. □
The following conjecture is discussed in the bivariate case to overcome the calculating the inverse of the information matrix in the multivariate case, so the determinant of the correlation matrix is just a scaler by the form of the information matrxi and therewith the sensitivit function for the D-optimaltiy do not content it. The illustrated examples for $m \geq 2$ in the next section do not contrast this Conjecture.

Conjecture 7.1. *Let $\xi_{j,D}^*$ be D-optimal for the j-th marginal component without an intercept included in the marginal design region $\mathcal{X}_j$, $j = 1, ..., m$.*
The product type design

$$\xi_D^* = \otimes_{j=1}^m \xi_{j,D}^* \quad (7.12)$$

is D-optimal for the SUR model (3.3) in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$.
If and only if $\alpha \leq \rho \leq \gamma$, where $-1 < \alpha_m < 0 < \gamma_m < 1$

Proof: to prove the optimality of the product-designs corresponding to theorems (7.1) and (7.1) for the SUR model, we need the form of the information matrices for marginal models, also, the information matrix for the univariate marginal models has the following form $\mathbf{M}_j(\xi_j) = \int \mathbf{f}_j\mathbf{f}_j^\top \xi_j^*(\mathrm{d}x_j)$.
The regression function of the SUR model in the individual approach has the following form for $m = 2$ because of 6.1

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \mathbf{f}_1(x_1) & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2(x_2) \end{pmatrix} \quad (7.13)$$

then the information matrix with respect to the product type design $\xi = \otimes_{j=1}^2 \xi_j$ has the following form

$$\mathbf{M}(\xi) = \int \mathbf{f}\Sigma^{-1}\mathbf{f}^\top d\xi = \begin{pmatrix} \sigma^{(11)} \int \mathbf{f}_1\ \mathbf{f}_1^\top d\xi_1 & \sigma^{(12)} \int \mathbf{f}_1 d\xi_1 \cdot \int \mathbf{f}_2^\top d\xi_2 \\ \sigma^{(12)} \int \mathbf{f}_2 d\xi_2 \cdot \int \mathbf{f}_1^\top d\xi_1 & \sigma^{(22)} \int \mathbf{f}_2\ \mathbf{f}_2^\top d\xi_2 \end{pmatrix} \quad (7.14)$$

where $\sigma^{(jk)}$ are the elements of $\mathbf{U} = \Sigma^{-1}$. And for $\mathbf{M}_j(\xi_j) = \int \mathbf{f}_j\, \mathbf{f}_j^\top d\xi_j$ and $\mathbf{m}_j(\xi_j) = \int \mathbf{f}_j d\xi_j$, $jk = 1, 2$, the information matrix has the summarized form

$$\mathbf{M}(\xi) = \int \mathbf{f}\Sigma^{-1}\mathbf{f}^\top d\xi = \begin{pmatrix} \sigma^{(11)}\mathbf{M}_1(\xi_1) & \sigma^{(12)}\mathbf{m}_1(\xi_1)\mathbf{m}_2(\xi_2)^\top \\ \sigma^{(12)}\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top & \sigma^{(22)}\mathbf{M}_2(\xi_2) \end{pmatrix} \tag{7.15}$$

the form of the inverse of the block matrices is to be used, to invert the information matrix, see for example *(Peterson and Pederson (2008)*, so the inverse of a block matrix is given as follows

$$\begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix} \tag{7.16}$$

$$\text{where } \mathbf{S}_{11} = \left(\mathbf{B}_{11} - \mathbf{B}_{12}\mathbf{B}_{22}^{-1}\mathbf{B}_{21}\right)^{-1}, \ \mathbf{S}_{22} = \left(\mathbf{B}_{22} - \mathbf{B}_{21}\mathbf{B}_{11}^{-1}\mathbf{B}_{12}\right)^{-1}$$

$$\mathbf{S}_{12} = -\mathbf{B}_{11}^{-1}\mathbf{B}_{12}\left(\mathbf{B}_{22} - \mathbf{B}_{21}\mathbf{B}_{11}^{-1}\mathbf{B}_{12}\right)^{-1} = -\mathbf{B}_{11}^{-1}\mathbf{B}_{12}\mathbf{S}_{22}$$

$$\mathbf{S}_{21} = -\mathbf{B}_{22}^{-1}\mathbf{B}_{21}\left(\mathbf{B}_{11} - \mathbf{B}_{12}\mathbf{B}_{22}^{-1}\mathbf{B}_{21}\right)^{-1} = -\mathbf{B}_{22}^{-1}\mathbf{B}_{21}\mathbf{S}_{11}$$

So the inverse of the information matrix has the following form because of (7.16)

$$\mathbf{M}^{-1}(\xi) = \begin{pmatrix} \sigma^{(11)}\mathbf{Z}_{11}(\xi_1, \xi_2) & -\sigma^{(12)}\mathbf{Z}_{12}(\xi_1, \xi_2) \\ -\sigma^{(12)}\mathbf{Z}_{21}(\xi_1, \xi_2) & \sigma^{(22)}\mathbf{Z}_2(\xi_1, \xi_2) \end{pmatrix} \tag{7.17}$$

$$\text{where } \mathbf{Z}_{11}(\xi_1, \xi_2) = \left(\sigma^{(11)}\sigma^{(22)}\mathbf{M}_1(\xi_1) - (\sigma^{(12)})^2\mathbf{m}_1(\xi_1)\mathbf{m}_2(\xi_2)^\top\mathbf{M}_2^{-1}(\xi_2)\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\right)^{-1} \tag{7.18}$$

$$\mathbf{Z}_{22}(\xi_1, \xi_2) = \left(\sigma^{(11)}\sigma^{(22)}\mathbf{M}_2(\xi_2) - (\sigma^{(12)})^2\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\mathbf{M}_1^{-1}(\xi_1)\mathbf{m}_1(\xi_1)\mathbf{m}_2(\xi_2)^\top\right)^{-1} \tag{7.19}$$

$$\mathbf{Z}_{12}(\xi_1, \xi_2) = \mathbf{M}_1^{-1}(\xi_1)\mathbf{m}_1(\xi_1)\mathbf{m}_2(\xi_2)^\top\mathbf{Z}_{22}(\xi_1, \xi_2) \tag{7.20}$$

$$\mathbf{Z}_{21}(\xi_1, \xi_2) = \mathbf{M}_2^{-1}(\xi_2)\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\mathbf{Z}_{11}(\xi_1, \xi_2) \tag{7.21}$$

To prove the theorem (7.1) we calculate the right side of the equivalence theorem for D-optimality, where the inverse of the information matrix is given in (7.17), also, when the product type design $\xi^* = \otimes_{j=1}^2 \xi^*$ is D-optimal, then

$$\begin{aligned}
& \text{trace}\left(\Sigma^{-1}\mathbf{f}(\mathbf{x})^\top\mathbf{M}(\xi^*)^{-1}\mathbf{f}(\mathbf{x})\right) \\
= \ & \text{trace}\begin{pmatrix} \sigma^{(11)} & \sigma^{(12)} \\ \sigma^{(12)} & \sigma^{(22)} \end{pmatrix}\begin{pmatrix} \mathbf{f}_1(x_1)^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2(x_2)^\top \end{pmatrix} \\
& \cdot\begin{pmatrix} \sigma^{(11)}\mathbf{Z}_{11}^{-1}(\xi_1^*, \xi_2^*) & -\sigma^{(12)}\mathbf{Z}_{12}^{-\top}(\xi_1^*, \xi_2^*) \\ -\sigma^{(12)}\mathbf{Z}_{21}^{-1}(\xi_1^*, \xi_2^*) & \sigma^{(22)}\mathbf{Z}_{22}^{-1}(\xi_1^*, \xi_2^*) \end{pmatrix}\begin{pmatrix} \mathbf{f}_1(x_1) & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2(x_2) \end{pmatrix} \\
= \ & \text{trace}\begin{pmatrix} \sigma^{(11)}\mathbf{f}_1(x_1)^\top & \sigma^{(12)}\mathbf{f}_2(x_2)^\top \\ \sigma^{(12)}\mathbf{f}_1(x_1)^\top & \sigma^{(22)}\mathbf{f}_2(x_2)^\top \end{pmatrix}\begin{pmatrix} \sigma^{(11)}\mathbf{Z}_{11}^{-1}(\xi_1^*, \xi_2^*)\mathbf{f}_1(x_1) & -\sigma^{(12)}\mathbf{Z}_{12}^{-\top}(\xi_1^*, \xi_2^*)\mathbf{f}_2(x_2) \\ -\sigma^{(12)}\mathbf{Z}_{12}^{-1}(\xi_1^*, \xi_2^*)\mathbf{f}_1(x_1) & \sigma^{(12)}\mathbf{Z}_{22}^{-1}(\xi_1^*, \xi_2^*)\mathbf{f}_2(x_2) \end{pmatrix} \\
= \ & \text{trace}\left((\sigma^{(11)})^2\mathbf{f}_1(x_1)^\top\mathbf{Z}_{11}^{-1}(\xi_1^*, \xi_2^*)\mathbf{f}_1(x_1) - (\sigma^{(12)})^2\mathbf{f}_2(x_2)^\top\mathbf{Z}_{12}^{-1}(\xi_1^*, \xi_2^*)^\top\mathbf{f}_1(x_1)\right) \\
& +\text{trace}\left(-(\sigma^{(12)})^2\mathbf{f}_1(x_1)^\top\mathbf{Z}_{21}^{-1}(\xi_1^*, \xi_2^*)\mathbf{f}_2(x_2) + \sigma^{(11)}\sigma^{(22)}\mathbf{f}_2(x_2)^\top\mathbf{Z}_{22}^{-1}(\xi_1^*, \xi_2^*)\mathbf{f}_2(x_2)\right)
\end{aligned} \tag{7.22}$$

Then because of $\sigma^{(11)} = \frac{\sigma_{22}}{\sigma_{11}\sigma_{22}-\sigma_{12}^2} = \frac{1}{\sigma_1^2(1-\rho^2)}, \sigma^{(22)} = \frac{\sigma_{11}}{\sigma_{11}\sigma_{22}-\sigma_{12}^2} = \frac{1}{\sigma_2^2(1-\rho^2)}$ $\sigma^{(12)} = \frac{-\sigma_{12}}{\sigma_{11}\sigma_{22}-\sigma_{12}^2} =$

$\frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)}$ the terms in (7.18), (7.19),(7.20) and (7.21) have the following forms

$$\mathbf{Z}_{11}(\xi_1,\xi_2) = \sigma_1^2\sigma_2^2(1-\rho^2)\left(\mathbf{M}_1(\xi_1) - \rho^2\mathbf{m}_1(\xi_1)\mathbf{m}_2(\xi_2)^\top\mathbf{M}_2^{-1}(\xi_2)\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\right)^{-1}$$
$$\mathbf{Z}_{22}(\xi_1,\xi_2) = \sigma_1^2\sigma_2^2(1-\rho^2)\left(\mathbf{M}_2(\xi_2) - \rho^2\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\mathbf{M}_1^{-1}(\xi_1)\mathbf{m}_1(\xi_1)\mathbf{m}_2(\xi_2)^\top\right)^{-1}$$
$$\mathbf{Z}_{12}(\xi_1,\xi_2) =$$
$$\sigma_1^2\sigma_2^2(1-\rho^2)\mathbf{M}_1^{-1}(\xi_1)\mathbf{m}_1(\xi_1)\mathbf{m}_2(\xi_2)^\top\left(\mathbf{M}_2(\xi_2) - \rho^2\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\mathbf{M}_1^{-1}(\xi_1)\mathbf{m}_1(\xi_1)\mathbf{m}_2(\xi_2)^\top\right)^{-1}$$
$$\mathbf{Z}_{21}(\xi_1,\xi_2) =$$
$$\sigma_1^2\sigma_2^2(1-\rho^2)\mathbf{M}_2^{-1}(\xi_2)\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\left(\mathbf{M}_1(\xi_1) - \rho^2\mathbf{m}_1(\xi_1)\mathbf{m}_2(\xi_2)^\top\mathbf{M}_2^{-1}(\xi_2)\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\right)^{-1}$$

Then

$$\begin{aligned}
&\text{trace}\left(\Sigma^{-1}\mathbf{f}(\mathbf{x})^\top\mathbf{M}(\xi^*)^{-1}\mathbf{f}(\mathbf{x})\right) = \\
&\mathbf{f}_1(x_1)^\top\left(\mathbf{M}_1(\xi_1) - \rho^2\mathbf{m}_1(\xi_1)\mathbf{m}_2(\xi_2)^\top\mathbf{M}_2^{-1}(\xi_2)\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\right)^{-1}\mathbf{f}_1(x_1) \\
&-\rho^2\mathbf{f}_1(x_1)^\top\mathbf{M}_2^{-1}(\xi_2)\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\left(\mathbf{M}_1(\xi_1) - \rho^2\mathbf{m}_1(\xi_1)\mathbf{m}_2(\xi_2)^\top\mathbf{M}_2^{-1}(\xi_2)\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\right)^{-1}\mathbf{f}_2(x_2) \\
&-\rho^2\mathbf{f}_2(x_2)^\top\mathbf{M}_2^{-1}(\xi_2)\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\left(\mathbf{M}_1(\xi_1) - \rho^2\mathbf{m}_1(\xi_1)\mathbf{m}_2(\xi_2)^\top\mathbf{M}_2^{-1}(\xi_2)\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\right)^{-1}\mathbf{f}_1(x_1) \\
&+\mathbf{f}_2(x_2)^\top\left(\mathbf{M}_2(\xi_2) - \rho^2\mathbf{m}_2(\xi_2)\mathbf{m}_1(\xi_1)^\top\mathbf{M}_1^{-1}(\xi_1)\mathbf{m}_1(\xi_1)\mathbf{m}_2(\xi_2)^\top\right)^{-1}\mathbf{f}_2(x_2) \qquad (7.23)
\end{aligned}$$

We can remark, that the conditions $\max\varphi_D(\mathbf{x},\xi^*) = p$ could be rewritten as an equality $h(\rho) = 0$, where h is a quadratic function in ρ, then there exist α_m and γ_m such that $\max\varphi_D(\mathbf{x},\xi^*) = p$ for all $\rho \in [\alpha_m,\gamma_m]$ and $\max\varphi_D(\mathbf{x},\xi^*) > p$ for $\rho \notin [\alpha_m,\gamma_m]$.the constants α_m and γ_m are dependent on the model, because we have to maximize $\max\varphi_D(\mathbf{x},\xi^*)$ with respect to control variables $\mathbf{x}$ and the form of the sensitivity function is for each model different with respect to $\mathbf{x}$, but the equation stays quadratic in ρ because of the form of the inverse of the correlation matrix 7.11 by the lemma 7.1. □
The next corollary could give the reader an simple possible explanation about the role of the number of components of the SUR model m resp. the dimension of the correlation matrix by the inversely proportionality property of m to the ρ -intervals length as well as the form of the marginal regression functions , which is to remark in (7.1).

Corollary 7.1. *For the SUR models without an intercept included by the marginals and for the product type design has the sensitivity function for the D-optimality the following form*

$$\varphi_D(\mathbf{x};\xi) = \textit{trace}\,(\mathbf{C}^{-1}\mathbf{f}(\mathbf{x})^\top(\mathbf{M}(\xi))_{\mathbf{C}}^{-1}\mathbf{f}(\mathbf{x}));\;\; (\mathbf{M}(\xi))_{\mathbf{C}} = \sum_{i=1}^{k} w_i\mathbf{f}(\mathbf{x}_i)\mathbf{C}^{-1}\mathbf{f}(\mathbf{x}_i)^\top$$

Thus

$$\textit{trace}\,(\mathbf{f}(\mathbf{x})^\top(\mathbf{M}(\xi^*))_{\mathbf{C}}^{-1}\mathbf{f}(\mathbf{x}))\;\; \le pm;\;\; \xi^* = \otimes_{j=1}^{m}\xi_j^*$$

Thus the determinant of the correlation matrix has the upper bounds

$$\det \mathbf{C} \le pm(1-\rho)^{m-2}((m-2)\rho+1) \min_{\mathbf{x}_j \in \mathcal{X}_j} \frac{\sum_{j=1}^{m} \mathit{trace}\, \mathbf{M}_j(\xi_j^*)}{\sum_{j=1}^{m} \mathit{trace}\, \mathbf{M}_j(x_j)} \tag{7.24}$$

$$= pm(1-\rho)^{m-2}((m-2)\rho+1) \frac{\sum_{j=1}^{m} \mathit{trace}\, \mathbf{M}_j(\xi_j^*)}{\max_{\mathbf{x}_j \in \mathcal{X}_j} \sum_{j=1}^{m} \mathit{trace}\, \mathbf{M}_j(x_j)} \tag{7.25}$$

Proof: To prove the first part of corollary (7.1), it is enough to look at the form of the sensitivity function for D-optimality with respect to the product type design in (7.23), i.e
$\varphi_D(\mathbf{x};\xi) = \text{trace}\,(\mathbf{C}^{-1}\mathbf{f}(\mathbf{x})^\top(\mathbf{M}(\xi))_{\mathbf{C}}^{-1}\mathbf{f}(\mathbf{x}))$ and because of the lemma 5.1 in chapter three

$$\text{trace}\left(\mathbf{B}^{-1}\mathbf{A}\right) \ge \frac{\text{trace}\,(\mathbf{A})}{\text{trace}\,(\mathbf{B})}$$

for the matrices $\mathbf{A}_{q\times q}$ and $\mathbf{B}_{q\times q}$ respectively positive semi definite and positive definite. Then, for $\mathbf{A} = \mathbf{f}(\mathbf{x})^\top(\mathbf{M}(\xi^*))_{\mathbf{C}}^{-1}\mathbf{f}(\mathbf{x})$ and $\mathbf{B} = \mathbf{C}$, thus

$$\text{trace}\,(\mathbf{f}(\mathbf{x})^\top(\mathbf{M}(\xi^*))_{\mathbf{C}}^{-1}\mathbf{f}(\mathbf{x})) \ \le pm;\ \xi^* = \otimes_{j=1}^{m}\xi_j^* \tag{7.26}$$

because of $\text{trace}\,(\mathbf{C}) = m$ and therewith we have proven the second part of the corrolary.
From (7.26) we have

$$\text{trace}\,(\mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top(\mathbf{M}(\xi^*))_{\mathbf{C}}^{-1}) \ \le pm$$

Also, for $\mathbf{A} = \mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top$ and $\mathbf{B} = (\mathbf{M}(\xi^*))_{\mathbf{C}}^{-1} \Longrightarrow$

$$\frac{\text{trace}\,(\mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top)}{\text{trace}\,(\mathbf{M}(\xi^*)_{\mathbf{C}})} \ \le \text{trace}\,(\mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top(\mathbf{M}(\xi^*))_{\mathbf{C}}^{-1}) \ \le pm$$

then

$$\text{trace}\,(\mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top) \ \le pm\ \text{trace}\,(\mathbf{M}(\xi^*)_{\mathbf{C}}) \tag{7.27}$$

so

$$\text{trace}\,(\mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top) = \begin{pmatrix} \mathbf{f}_1(x_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x_m) \end{pmatrix} \begin{pmatrix} \mathbf{f}_1(x_1)^\top & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x_m)^\top \end{pmatrix} = \sum_{j=1}^{m=2} \text{trace}\,\mathbf{M}_j(x_j) \tag{7.28}$$

And because of 7.11

$$\mathbf{M}(\xi^*)_{\mathbf{C}} = \frac{(1-\rho)^{m-2}}{\det \mathbf{C}} \int \begin{pmatrix} \mathbf{f}_1(x_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x_m) \end{pmatrix}$$

$$\cdot \begin{pmatrix} (m-2)\rho+1 & -\rho & \cdots & -\rho \\ -\rho & (m-2)\rho+1 & \cdots & -\rho \\ \vdots & \vdots & \ddots & \vdots \\ -\rho & \cdots & -\rho & (m-2)\rho+1 \end{pmatrix} \begin{pmatrix} \mathbf{f}_1(x_1)^\top & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x_m)^\top \end{pmatrix} d(\otimes_{j=1}^{m}\xi_m^*)$$

then

$$\begin{aligned}
\text{trace}\,(\mathbf{M}(\xi^*)_{\mathbf{C}}) &= \frac{(1-\rho)^{m-2}((m-2)\rho+1)}{\det \mathbf{C}} \sum_{j=1}^{m} \text{trace}\,\mathbf{M}_j(\xi_j^*) \\
&= \frac{(1-\rho)^{m-2}((m-2)\rho+1)}{\det \mathbf{C}} \sum_{j=1}^{m} \text{trace}\,\mathbf{M}_j(\xi_j^*) \qquad (7.29)
\end{aligned}$$

By replacing (7.28) and (7.29) in (7.27) we obtain

$$\begin{aligned}
\sum_{j=1}^{m} \text{trace}\,\mathbf{M}_j(x_j) &\leq \frac{pm(1-\rho)^{m-2}((m-2)\rho+1)}{\det \mathbf{C}} \sum_{j=1}^{m} \text{trace}\,\mathbf{M}_j(\xi_j^*) \\
\Longrightarrow \ \det \mathbf{C} \ &\leq pm(1-\rho)^{m-2}((m-2)\rho+1) \frac{\sum_{j=1}^{m} \text{trace}\,\mathbf{M}_j(\xi_j^*)}{\sum_{j=1}^{m} \text{trace}\,\mathbf{M}_j(x_j)}
\end{aligned}$$

Thus

$$\begin{aligned}
&\det \mathbf{C} \leq pm(1-\rho)^{m-2}((m-2)\rho+1) \min_{\mathbf{x}_j \in \mathcal{X}_j} \frac{\sum_{j=1}^{m} \text{trace}\,\mathbf{M}_j(\xi_j^*)}{\sum_{j=1}^{m} \text{trace}\,\mathbf{M}_j(x_j)} \\
&= pm(1-\rho)^{m-2}((m-2)\rho+1) \frac{\sum_{j=1}^{m} \text{trace}\,\mathbf{M}_j(\xi_j^*)}{\max_{\mathbf{x}_j \in \mathcal{X}_j} \sum_{j=1}^{m} \text{trace}\,\mathbf{M}_j(x_j)}
\end{aligned}$$

And therewith the corollary is proven. □.
The next theorem presents the optimality of the product type designs for different criteria for SUR models with hierarchical nested marginals, also, the regression function for $j-$th marginal component contains the regression function for the $j-1-$th marginal components, where this kind of multivariate linear models is in the works of *(Krafft and Schaefer (1992)*) and *(Kurotschka and Schwabe (1996)*) but for the same control variables for the different marginal regression functions of the components

Theorem 7.2. *Let ξ_j^* be Φ-optimal for the j-th marginal component without intercepts included, in the marginal design region $\mathcal{X}_j$, $j = 1, ..., m$ and the regression functions of the marginal components have the following form*

$$\mathbf{f}_1(x_1) = \mathbf{g}_1(x_1), \mathbf{f}_2(x_1, x_2) = \begin{pmatrix} \mathbf{g}_1(x_1) \\ \mathbf{g}_2(x_2) \end{pmatrix}, ..., \mathbf{f}_m(x_1, ..., x_m) = \begin{pmatrix} \mathbf{g}_1(x_1) \\ \vdots \\ \mathbf{g}_m(x_m) \end{pmatrix}, \ j = 1, \dots, m \qquad (7.30)$$

Or

$$\mathbf{f}_j(\mathbf{x}_j) = \mathbf{C}^j \cdot \mathbf{f}_{j+1}(\mathbf{x}_j, \mathbf{x}_{j+1}), \ \mathbf{C}^j \in \mathrm{R}^{r_j \times r_{j+1}}, \ j = 1, \dots, m-1 \qquad (7.31)$$

where

$$\mathbf{C}^j = \begin{pmatrix} \mathbf{I}_{r_j \times r_j} & \mathbf{0} \end{pmatrix}, \ j = 1, \dots, m-1 \qquad (7.32)$$

if

$$\int_{\mathcal{X}_j} \mathbf{g}_j(x_j) \xi_j^*(\mathrm{d}x_j) = \mathbf{0}, \ j = 2, \dots, m \qquad (7.33)$$

then the product type design

$$\xi^* = \otimes_{j=1}^m \xi_j^*$$

is Φ *-optimal for SUR model (3.3) in the design region* $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$*. Where* Φ*-optimal can be D- or linear optimal criterion by block diagonal weight matrix* $\mathbf{L} = \text{block} - \text{diag}\,(\mathbf{L}_j)$*.*
The sensitivity function φ_D *does not depend on* $\mathbf{\Sigma}$*.*

Proof: To prove the optimality of the product-designs corresponding the theorems (7.2) for the SUR model, we need the form of the information matrices for marginal models, also, the information matrix for the univariate marginal models has the following form $\mathbf{M}_j(\xi_j) = \int \mathbf{f}_j \mathbf{f}_j^\top d\xi_j$. For the SUR model, We implement the proofs for arbitrary m, where the regression function for the model in the individual approach has the following form

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \mathbf{f}_1(x_1) & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2(x_1, x_2) & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{f}_m(x_1, ..., x_m) \end{pmatrix} = \begin{pmatrix} \mathbf{g}_1(x_1) & \cdots & \mathbf{0} \\ \mathbf{0} & \begin{pmatrix} \mathbf{g}_1(x_1) \\ \mathbf{g}_2(x_2) \end{pmatrix} & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \begin{pmatrix} \mathbf{g}_1(x_1) \\ \vdots \\ \mathbf{g}_m(x_m) \end{pmatrix} \end{pmatrix}$$

then the diagonal blocks of the information matrix for one-point design and has the following form for $j = 1, ..., m$

$$(\mathbf{M}(\mathbf{x}))_{jj} = \begin{pmatrix} \sigma^{(11)}\mathbf{M}_1(x_1) & \cdots & . \\ . & \sigma^{(22)}\mathbf{M}_{12}(x_1, x_2) & \cdots \\ \vdots & \ddots & \vdots \\ . & \cdots & \sigma^{(mm)}\mathbf{M}_{1m}(x_1, ..., x_m) \end{pmatrix} \tag{7.34}$$

$$\mathbf{M}_{1j}(x_j) = \mathbf{f}_j(x_1, ..., x_j)\mathbf{f}_j(x_1, ..., x_j)^\top \tag{7.35}$$

So, the information matrix for the product type design $\xi = \otimes_{j=1}^m \xi_j$ has the following block diagonal form because of the condition (7.41) and by integral the information matrix for the one-point design in (7.34)

$$\mathbf{M}(\xi) = \int \mathbf{f}\ \Sigma^{-1}\ \mathbf{f}^\top d\xi =$$

$$\begin{pmatrix} \sigma^{(11)}\mathbf{M}_1(\xi_1) & \cdots & \mathbf{0} \\ \mathbf{0} & \sigma^{(22)}\mathbf{M}_{12}(\xi_1 \otimes \xi_2) & \cdots \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \sigma^{(mm)}\mathbf{M}_{1m}(\otimes_{j=1}^m \xi_j) \end{pmatrix} \tag{7.36}$$

where the block diagonal because of 7.41

$$\mathbf{M}_{1j}(\xi) = \begin{pmatrix} \mathbf{G}_1(\xi_1) & \cdots & \mathbf{0} \\ \mathbf{0} & \sigma^{(22)}\mathbf{G}_2(\xi_2) & \cdots \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{G}_j(\xi_j) \end{pmatrix}, \mathbf{G}_j(\xi_j) = \int \mathbf{g}_j\ \mathbf{g}_j^\top d\xi_j \tag{7.37}$$

The resulted sensitivity function has the following form, by replacing the both matrices in (7.34) and the inverse of the information matrix (7.36) in the equivalence theorem for D-optimality (4.1)

$$\begin{aligned}&\operatorname{trace}(\mathbf{M}(\mathbf{x})\mathbf{M}^{-1}(\xi_D^*))\\&= m \operatorname{trace}(\mathbf{G}_1(x_1)\mathbf{G}_1^{-1}(\xi_1^*)) + (m-1)\operatorname{trace}(\mathbf{G}_2(x_2)\mathbf{G}_2^{-1}(\xi_2^*)) + ... + \operatorname{trace}(\mathbf{G}_m(x_m)\mathbf{G}_m^{-1}(\xi_m^*))\\&\leq mp_1 + (m-1)p_2 + ... + p_m = p\end{aligned}$$

By similar way to the proof of the D-optimality in this case or to the proof of theorem 7.1 , the linear optimality in 4.1 of the product type designs can be proven because of theorem 2.7, where the weight matrix has the following form in this case

$$\mathbf{L} = \begin{pmatrix} \mathbf{L}_1 & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_{12} & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{L}_{1m} \end{pmatrix}, \ \mathbf{L}_{1m} = \int_{\mathcal{X}} \mathbf{f}_j(x_1,...,x_j)\mathbf{f}_j(x_1,...,x_j)^\top \mu(\mathrm{d}x_1,...,\mathrm{d}x_j) \quad \square$$

Remark 7.2. *The product type designs may be not D-optimal for SUR models with the nested form of the regression functions in (7.30), for the non-block diagonal form of the information matrix, if $\rho = 0$, because the diagonal block information matrices for the marginal components given in (7.37) are no longer block diagonal.*

A direct corollary of theorems 7.2 is the optimality of the product type designs for the multivariate hierarchical nesting chemical balance models resp. multiple or multivariate chemical balance models, when the functions $\mathbf{g}_{jh_j} \equiv 1$ are the identity functions, that means the control variables from the experimental region $\mathcal{X}_j = [-1, 1]$, $j = 1, ..., m$.

Corollary 7.2 (Optimal multivariate chemical balance designs). *Let ξ_j^* be Φ-optimal for the j-th marginal component without intercepts included, in the marginal design region $\mathcal{X}_{h_j}$, $j = 1, ..., m$ and the regression functions of the marginal components have the following form*

$$\mathbf{f}_1(x_1) = x_1, \mathbf{f}_2(x_1,x_2) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, ..., \mathbf{f}_m(x_1,...,x_m) = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, \ j = 1,\ldots,m \qquad (7.38)$$

if

$$\int_{\mathcal{X}_j} x_j \xi_j^*(\mathrm{d}x_j) = 0, \ j = 2,\ldots,m \qquad (7.39)$$

hen the product type design

$$\xi^* = \otimes_{j=1}^m \xi_j^*$$

is Φ -optimal for SUR model (3.3) in the design region $\mathcal{X} = \times_{j=1}^m \mathcal{X}_{h_j}$. Where Φ-optimal can be D- or linear optimal criterion by block diagonal weight matrix $\mathbf{L} = \mathrm{block} - \mathrm{diag}\,(\mathbf{L}_j)$.
The sensitivity function φ_D does not depend on $\mathbf{\Sigma}$.

Proof: this corollary is a special case of theorem 7.2, also, it can be proven analogous to the proofs of theorems 7.2. □

The topics of the next theorems are the optimal designs for a multivariate chemical balance regression model with h_j-objects by each marginal component, $j = 1, ..., m$

Corollary 7.3. *Let ξ_k^* be Φ-optimal for the marginal univariate one factor regression model with the response function given in (7.1) in the marginal design region $\mathcal{X}_k = [-1,1]$, $k = 1, ..., h_1, h_1 + 1, ..., h_2, ..., h_{m-1} + 1, ..., h_m$ and the regression functions of the $m-$ marginal components have the following form*

$$\mathbf{f}_1(x_{11}, ..., x_{1h_1}) = \begin{pmatrix} x_{11} \\ \vdots \\ x_{1h_1} \end{pmatrix}, \; ..., \mathbf{f}_m(x_{m1}, ..., x_{mh_m}) = \begin{pmatrix} x_{m1} \\ \vdots \\ x_{mh_m} \end{pmatrix} \tag{7.40}$$

and if

$$\int_{[-1,1]} x_{1k} d\xi_k = 0, \; k = 1, ..., h_1, h_1 + 1, ..., h_2, ..., h_{m-1} + 1, ..., h_m \tag{7.41}$$

the product type design

$$\xi^* = \otimes_{k=1}^{h_m} \xi_k^*$$

is Φ -optimal for the considered SUR model with the marginal regression functions is (7.40) in the design region $\mathcal{X} = \times_{k=1}^{h_m} \mathcal{X}_k$. Where Φ-optimal can be D- or linear optimal criterion by block diagonal weight matrix $\mathbf{L} = \text{block} - \text{diag}\,(\mathbf{L}_j)$.
The sensitivity function φ_D does not depend on $\mathbf{\Sigma}$.

The proofs of corollary 7.3 can be similarly implemented to the proofs of theorems 7.1 resp. 7.2, because of the diagonal form of the corresponding information matrix under the conditions (7.41). Where the weight matrix has the following block diagonal form for this model

$$\mathbf{L} = \text{block} - \text{diag}\,(\mathbf{L}_j)_{j=1,\ldots,m}, \; \mathbf{L}_j = \text{block} - \text{diag}\,(L_{k11})_{k=1,\ldots,h_j} \quad \square$$

Remark 7.3 (Optimal multivariate spring balance designs). *When the the conditions by theorem 7.3 and the corollary 7.2 are invalid, then the sensitivity function for the D-optimality depends on the correlation ρ, thus we can proof the D-optimality of the product type design for the multivariate spring balance models for some intervals or values of ρ's similarly to the proof of 7.1.*

Remark 7.4 (Optimal designs for SUR models without intercepts for more generalized nesting structures). *Some results due to optimal designs for the considered SUR models in chapter six can be generalized for SUR models without intercepts by the marginal regression functions, i.e. $f_{j1}(\mathbf{x}) \not\equiv 1$. So when the first component is nested multiplicatively through the other components as in (6.2), then if*

$$\int_{\mathcal{X}_j} \mathbf{f}_j(x_j)\xi_j^*(\mathrm{d}x_j) = \mathbf{0}, \; j = 2, \ldots, m$$

then theorems 6.1 and 6.2 are valid for the SUR model with the regression function given in (6.3), but $f_{j1}(\mathbf{x}) \not\equiv 1$.
When the first component is nested additive -wise through the other components as in (6.6), and if

$$\int_{\mathcal{X}_j} \mathbf{f}_j(x_j)\xi_j^*(\mathrm{d}x_j) = \mathbf{0}, \; j = 2, \ldots, m$$

then theorems 6.3 and 6.5 are valid for the SUR model with the regression function given in (6.6), but $f_{j1}(\mathbf{x}) \not\equiv 1$.
When a new component is nested multiplicatively through all other $m-$*components as in (6.25), and if*

$$\int_{\mathcal{X}_j} \mathbf{f}_j(x_j)\xi_j^*(\mathrm{d}x_j) = \mathbf{0},\ j = 1, \ldots, m$$

then theorem 6.9 is valid for the SUR model with the regression function given in (6.25), but $f_{j1}(\mathbf{x}) \not\equiv 1$.
When a new component is nested additive -wise through all other $m-$*components as in (6.26), and if*

$$\int_{\mathcal{X}_j} \mathbf{f}_j(x_j)\xi_j^*(\mathrm{d}x_j) = \mathbf{0},\ j = 1, \ldots, m$$

theorems 6.10 and 6.11 are valid for the SUR model with the regression function given in (6.27), but $f_{j1}(\mathbf{x}) \not\equiv 1$.
When a new different component is nested multiplicatively through each $j-$*component,* $j = 1, ..., m$ *as in (6.30), then corollary 6.1 is valid for the SUR model with the regression function given in (6.32), but* $f_{j1}(\mathbf{x}) \not\equiv 1$.
When a new different component is nested additive -wise through each $j-$*component,* $j = 1, ..., m$ *as in (6.31), then corollary 6.2 is valid for the SUR model with the regression function given in (6.33), but* $f_{j1}(\mathbf{x}) \not\equiv 1$.

Remark 7.5. *The inverse of the information matrix with respect to the product type design has under conditions 7.4 for SUR models without intercepts by the one-factor marginal regression functions the block diagonal form in 7.5, so the covariance matrix for prediction has the following block diagonal form due to the Gauß estimator*

$$\mathbf{f}(\mathbf{x})^\top \mathbf{M}_{GM}^{-1}(\xi)\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \frac{1}{\sigma^{(11)}}\mathbf{f}_1(x_1)^\top \mathbf{M}_1^{-1}(\xi_1)\mathbf{f}_1(x_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{\sigma^{(mm)}}\mathbf{f}_m(x_m)^\top \mathbf{M}_m^{-1}(\xi_m)\mathbf{f}_m(x_m) \end{pmatrix}$$

the variance covariance matrix for the OLS estimator can be similarly calculated under conditions 7.4, so it has the following form

$$\mathbf{f}(\mathbf{x})^\top \mathbf{M}_{OLS}^{-1}(\xi)\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \sigma_1^2\mathbf{f}_1(x_1)^\top \mathbf{M}_1^{-1}(\xi_1)\mathbf{f}_1(x_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \sigma_m^2\mathbf{f}_m(x_m)^\top \mathbf{M}_m^{-1}(\xi_m)\mathbf{f}_m(x_m) \end{pmatrix}$$

So the upper bounds of the trace, maximum eigenvalue, and the determinant for the covariance matrix for the prediction due to the product type design, as well as some of the considered efficiencies in chapter five can be calculated similarly to their counterparts there.

7.3 Examples, Simulations and Discussions

Remark 7.6 (The Restriction for Heterogeneous Correlation Structure). *If the correlation matrix has heterogeneous structure then theorem 7.1 is to be hold just locally, because there are different correlation terms.*

Example 7.1 (The D-optimality of the product designs). *The D-optimality of the product type design for the SUR model with the same, and different regression functions without intercepts for the marginals has been introduced in this example for three models in the same experiment regions $[0,1]$. The first model has the marginals $Y_{ij} = \beta_{j1}x_{ij} + \varepsilon_{ij}$ and the product type designs are D-optimal for intervals for ρ,s as given in the following tabular. Figure 7.1 illustrates the D-optimality of the product type design for the bivariate SUR model $Y_{ij} = \beta_{j1}x_{ij} + \varepsilon_{ij}$, $\varphi_D(\otimes_{j=1}^{2}\xi_0^*) = \frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{1-\rho^2}, x_2 = 1$.*

m	p	ξ^*	$\det \mathbf{C}^{1/p}$ -*I*	*I-L*	ρ^* - *I*	*I-L*
1	*1*	$\xi_0^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$	-	-	-	-
2	*2*	$\otimes_{j=1}^{2}\xi_0^*$	$[0.707, 1.000]$	0.293	$[-0.707, 0.707]$	1.414
3	*3*	$\otimes_{j=1}^{3}\xi_0^*$	$[0.666, 0.752]$	0.086	$[-0.390, 0.640]$	1.030
4	*4*	$\otimes_{j=1}^{4}\xi_0^*$	$[0.643, 0.779]$	0.136	$[-0.274, 0.607]$	0.881
5	*5*	$\otimes_{j=1}^{5}\xi_0^*$	$[0.628, 0.800]$	0.172	$[-0.212, 0.587]$	0.799
6	*6*	$\otimes_{j=1}^{6}\xi_0^*$	$[0.615, 0.813]$	0.198	$[-0.174, 0.574]$	0.748
7	*7*	$\otimes_{j=1}^{7}\xi_0^*$	$[0.606, 0.823]$	0.217	$[-0.148, 0.564]$	0.712
8	*8*	$\otimes_{j=1}^{8}\xi_0^*$	$[0.599, 0.837]$	0.238	$[-0.128, 0.556]$	0.684
9	*9*	$\otimes_{j=1}^{9}\xi_0^*$	$[0.593, 0.848]$	0.255	$[-0.113, 0.550]$	0.663
10	*10*	$\otimes_{j=1}^{10}\xi_0^*$	$[0.587, 0.858]$	0.271	$[-0.101, 0.546]$	0.647
11	*11*	$\otimes_{j=1}^{11}\xi_0^*$	$[0.582, 0.861]$	0.279	$[-0.092, 0.542]$	0.634
12	*12*	$\otimes_{j=1}^{12}\xi_0^*$	$[0.579, 0.869]$	0.290	$[-0.084, 0.538]$	0.622

The second model has the marginals $Y_{ij} = \beta_{j1}x_{ij} + \beta_{j2}x_{ij}^2 + \varepsilon_{ij}$ and the product type designs are D-optimal

m	p	ξ^*	$\det \mathbf{C}^{1/p}$ -*L*	*I-L*	ρ^* -*I*	*I-L*
1	*2*	$\xi_0^* = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$	-	-	-	-
2	*4*	$\otimes_{j=1}^{2}\xi_0^*$	$[0.760, 1.000]$	0.240	$[-0.816, 0.816]$	1.632
3	*6*	$\otimes_{j=1}^{3}\xi_0^*$	$[0.718, 0.805]$	0.087	$[-0.434, 0.767]$	1.201
4	*8*	$\otimes_{j=1}^{4}\xi_0^*$	$[0.696, 0.833]$	0.137	$[-0.298, 0.743]$	1.041
5	*10*	$\otimes_{j=1}^{5}\xi_0^*$	$[0.681, 0.851]$	0.170	$[-0.212, 0.728]$	0.940

The third model consists of different marginals, where the marginals have the form $Y_{ij} = \beta_{j1}x_{ij} + \beta_{j2}x_{ij}^2 + \varepsilon_{ij}$ for odd j and $Y_{ij} = \beta_{j1}x_{ij} + \varepsilon_{ij}$ for even j and for odd m the information matrices for product type design singular. Thus for $\xi_0^* = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, the product type designs are D-optimal for intervals for ρ,s as follos

m	p	ξ^*	$\det \mathbf{C}^{1/p}$ -*I*	*I-L*	ρ^* -*I*	*I-L*
2	*3*	ξ_0^*	$[0.578, 1.000]$	0.422	$[-0.816, 0.816]$	1.632
4	*6*	$\otimes_{j=1}^{2}\xi_0^*$	$[0.745, 0.847]$	0.102	$[-0.274, 0.607]$	0.881
6	*9*	$\otimes_{j=1}^{4}\xi_0^*$	$[0.723, 0.871]$	0.148	$[-0.174, 0.574]$	0.748

By comparing the results of the implemented examples and simulations it can be said, that m the number of components of the SUR model plays a mean role by the length of the ρ^* -interval, also, the intervals *length is monotonically decreasing in m. The three different SUR models illustrate, that the forms of the marginal regression functions play a role too,*

but these simulations have illustrated that the number of parameters p for the SUR model do not have the mean influence on the ρ^* -intervals length.
A possible argument for the role of m is the different dimensions and structure of the correlation matrix for grown m and m is inversely proportional to the determinants value of the correlation matrix as we have cleared in remark (7.1). Another weak argument could be developed and reinforced also, from the third part of the corollary (7.1)

$$\det \mathbf{C} \leq pm \frac{\sum_{j=1}^{m} trace\, \mathbf{M}_j(\xi_j^*)}{\max_{\mathbf{x}_j \in \mathcal{X}_j} \sum_{j=1}^{m} trace\, \mathbf{M}_j(x_j)}$$

For $m = 2, p = 2$ in the desgin region $[0,1]$, we have from the first tabular and the remark (7.1) $\det \mathbf{C}_{q\times q} \leq ...\mathbf{C}_{3\times 3} \leq \det \mathbf{C}_{2\times 2} \leq 4^{\frac{2}{2}} = 4$, *i.e.* $1-\rho^2 \leq 4$, *and it is known, that the* $\det \mathbf{C} \in [0,1]$ *and we have equality just for $p = m = 1$ and $\rho = 0$, which is the weakness of this argument.*
By investigating the influence of m on the p-square root for the determinant values of the correlation matrix for the different implemented examples, we can remark, that the upper and lower bounds of the $\det \mathbf{C}^{1/p}$ *-interval, for $m > 2$ because of symmetry by $m = 2$, are ordered monotonically increasing and decreasing in m and therewith the length of the* $\det \mathbf{C}^{1/p}$ *-intervals are monotonically increasing in m and that illustrates the inequality (7.25) of the determinant for the correlation matrix by corollary (7.1).*

Example 7.2 (When the product designs are not D-optimal)**.** *By this example we will consider the model with the marginals $Y_{ij} = \beta_{j1}x_{ij} + \varepsilon_{ij}$ for different numbers of components $m = 2,3,4,5$, also, for $m = 2$ the D-optimal design*

$$\xi_D^* = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 1-2.w^* & w^* & w^* \end{pmatrix}, \; w^* = \frac{1-2\rho^2}{1-4\rho^2}, \; \mid \rho \mid > \frac{\sqrt{2}}{2} = 0.707$$

For $\rho = \pm 1$, $w^* = 1/3$, we get the same D-optimal design for the spring weighing model for two objects

$$\xi_2^* = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \tag{7.42}$$

For $m = 3$ the D-optimal design is

$$\xi_D^* = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \\ 1-3.w^* & w^* & w^* & w^* \end{pmatrix}, \; w^* = \frac{1+\rho-4\rho^2}{1-9\rho^2}, \; \rho > 0.64 \;\&\; \rho < -0.39$$

Or

$$\xi_D^* = \begin{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \\ \frac{1-w^*}{3} & \frac{1-w^*}{3} & \frac{1-w^*}{3} & \frac{w^*}{3} & \frac{w^*}{3} & \frac{w^*}{3} \end{pmatrix} \tag{7.43}$$

where $w^* = \frac{1+2\rho+\rho^2}{2\rho^2-\rho-1}$, $\rho = -0.56$ *i.e.* $\frac{1-w^*}{3} = 0$, $\frac{w^*}{3} = \frac{1}{3}$ *or the design, where the last three points have weights equal to* $1/3$ *for* $\rho = -0.56$*, i.e. the D-optimal design for spring weighing model for three objects*

$$\xi_3^* = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \tag{7.44}$$

For $m = 4$ *the D-optimal designs are*

$$\xi_D^* = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \\ 1-4.w^* & w^* & w^* & w^* & w^* \end{pmatrix}, \ w^* = \frac{1+2\rho-6\rho^2}{1-16\rho^2}, \ \rho > 0.607 \ \& \ \rho < -0.274$$

Or

$$\xi_D^* = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ \frac{1-w^*}{4} & \frac{1-w^*}{4} & \frac{1-w^*}{4} & \frac{1-w^*}{4} & \frac{w^*}{6} & \frac{w^*}{6} & \frac{w^*}{6} & \frac{w^*}{6} & \frac{w^*}{6} & \frac{w^*}{6} \end{pmatrix} \tag{7.45}$$

Where $w^* = \frac{3(4\rho^2-8\rho-3)}{20\rho^2-4\rho-3}$, $\rho = -\frac{1}{3}$ *i.e.* $\frac{1-w^*}{4} = \frac{w^*}{6} = \frac{1}{10}$ *or the same design with the weights equal to* $1/10$ *for* $\rho = -\frac{1}{3}$ *, i.e. the D-optimal design for the spring weighing model for four objects*

$$\xi_4^* = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} \end{pmatrix} \tag{7.46}$$

When $m = 5$ *the optimal designs are*

$$\xi_D^* = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \\ 1-5.w^* & w^* & w^* & w^* & w^* & w^* \end{pmatrix}$$

Where $w^* = \frac{1+3\rho-8\rho^2}{1-25\rho^2}$, $\rho > 0.587$ & $\rho < -0.212$ *Or the 15-point design*

$$\xi_D^* = \left(\begin{pmatrix} 1\\1\\1\\1\\0 \end{pmatrix} \cdots \begin{pmatrix} 0\\1\\1\\1\\1 \end{pmatrix} \begin{pmatrix} 1\\1\\1\\0\\0 \end{pmatrix} \cdots \begin{pmatrix} 0\\0\\1\\1\\1 \end{pmatrix} \atop \frac{1-w^*}{5} \quad \cdots \quad \frac{1-w^*}{5} \quad \frac{w^*}{10} \quad \cdots \quad \frac{w^*}{10} \right) \tag{7.47}$$

where the weight $\frac{1-w^*}{5}$ *is related to the design points with four ones and one zero, and the weight* $\frac{w^*}{10}$ *is related to the design points with three ones and two zeros.* $w^* = \frac{2(9\rho^2-15\rho-4)}{27\rho^2-3\rho-2}$, $\rho = -\frac{1}{4}$ *i.e.* $\frac{1-w^*}{5} = 0$ *and* $\frac{w^*}{10} = \frac{1}{10}$ *or the design where the last ten points have weights equal to* $1/10$ *for* $\rho = -\frac{1}{4}$, *i.e. the D-optimal design for spring weighing model for five objects*

$$\xi_5^* = \left(\begin{pmatrix} 1\\1\\1\\0\\0 \end{pmatrix} \begin{pmatrix} 1\\1\\0\\1\\0 \end{pmatrix} \begin{pmatrix} 1\\0\\1\\1\\0 \end{pmatrix} \begin{pmatrix} 0\\1\\1\\1\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\1\\1\\1 \end{pmatrix} \begin{pmatrix} 1\\1\\0\\0\\1 \end{pmatrix} \begin{pmatrix} 0\\1\\0\\1\\1 \end{pmatrix} \begin{pmatrix} 1\\0\\1\\0\\1 \end{pmatrix} \begin{pmatrix} 0\\1\\1\\0\\1 \end{pmatrix} \begin{pmatrix} 1\\0\\0\\1\\1 \end{pmatrix} \atop \frac{1}{10} \quad \frac{1}{10} \quad \frac{1}{10} \quad \frac{1}{10} \quad \frac{1}{10} \quad \frac{1}{10} \quad \frac{1}{10} \quad \frac{1}{10} \quad \frac{1}{10} \quad \frac{1}{10} \right) \tag{7.48}$$

Figure 7.2 illustrates the D-optimality of design (7.42) for the bivariate SUR model $Y_{ij} = \beta_{j1}x_{ij} + \varepsilon_{ij}$, $\varphi_D(\xi_\rho^*) = 2(x_1^2 - x_1x_2 + x_2^2)$.

Example 7.3 (Approximative D-optimality for the univariate spring weighing models)**.** *We can remark, that the D-optimal designs (7.42), (7.43), (7.45) and (7.47) for the considered SUR models in example(7.3) are respectively D-optimal designs for the univariate spring weighing models for two, three, four and five objects, i.e. for*

$$\begin{aligned} Y_i &= \beta_1x_{i1} + \beta_2x_{i2} + \varepsilon_i \\ Y_i &= \beta_1x_{i1} + \beta_2x_{i2} + \beta_3x_{i3} + \varepsilon_i \\ Y_i &= \beta_1x_{i1} + \beta_2x_{i2} + \beta_3x_{i3} + \beta_4x_{i4} + \varepsilon_i \\ Y_i &= \beta_1x_{i1} + \beta_2x_{i2} + \beta_3x_{i3} + \beta_4x_{i4} + \beta_5x_{i5} + \varepsilon_i \end{aligned}$$

Where for all models $E(\varepsilon_i) = 0$, $Cov(\varepsilon_i, \varepsilon_k) = 0$, $Var(\varepsilon_i) = \sigma^2, i, k = 1, ..., n$ *and* n *is the number of weighing. To prove that, we just have to check the satisfaction of the conditions of the equivalence theorem in the univariate case for D-optimality for more information see* (Schwabe (1996)*) or* (Huda and Mukerjee (1988)*).*

To justify the equality of the local D-optimal designs for the SUR models and the D-optimal designs for their counterparts by the univariate spring weighing models, we can remark, that this equality occurs at most by the correlation terms ρ's, which make the correlation matrix non-positiver definite, i.e. it is not more a correlation matrix and the information matrices for the considered SUR models singular, i.e. the determinant of it is equal to zero, thus statistically we can see, when we can not get any information with respect to D-optimality because of unreality of the variance covariance matrix of the error variables, then *we have no difference between the obtained informations from the multivariate case and the counterpart in the univariate case, i.e. by the SUR model and in the counterpart by the univariate spring weighing models. Mathematically, we can check the information matrices for the both models, for example for the SUR model with five components their forms* $Y_{ij} = \beta_{j1}x_{ij} + \varepsilon_{ij}$ $j = 1, ..., 5$ *and these marginal can formulate a spring weighing regression for five objects, also, the univariate counterpart the spring weighing regression for five objects of this SUR model is* $Y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \varepsilon_i$*, also by* $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2 = 1$ *we have not lost the generality, because they are just scaled terms, thus*

$$\mathbf{M}_{SUR}(\mathbf{X}) = \frac{1}{1+3\rho-4\rho^2}\begin{pmatrix} (1+3\rho)x_1^2 & -\rho x_1x_2 & -\rho x_1x_3 & -\rho x_1x_4 & -\rho x_1x_5 \\ -\rho x_1x_2 & (1+3\rho)x_2^2 & -\rho x_2x_3 & -\rho x_2x_4 & -\rho x_2x_5 \\ -\rho x_1x_3 & -\rho x_2x_3 & (1+3\rho)x_3^2 & -\rho x_3x_4 & -\rho x_3x_5 \\ -\rho x_1x_4 & -\rho x_2x_4 & -\rho x_3x_4 & (1+3\rho)x_4^2 & -\rho x_4x_5 \\ -\rho x_1x_5 & -\rho x_2x_5 & -\rho x_3x_5 & -\rho x_4x_5 & (1+3\rho)x_5^2 \end{pmatrix} =$$

$$\frac{1}{1+3\rho-4\rho^2}\begin{pmatrix} x_1^2 & x_1x_2 & x_1x_3 & x_1x_4 & x_1x_5 \\ x_1x_2 & x_2^2 & x_2x_3 & x_2x_4 & x_2x_5 \\ x_1x_3 & x_2x_3 & x_3^2 & x_3x_4 & x_3x_5 \\ x_1x_4 & x_2x_4 & x_3x_4 & x_4^2 & x_4x_5 \\ x_1x_5 & x_2x_5 & x_3x_5 & x_4x_5 & x_5^2 \end{pmatrix}_{univariate} +$$

$$\frac{1}{1+3\rho-4\rho^2}\begin{pmatrix} 3\rho x_1^2 & -(1+\rho)x_1x_2 & -(1+\rho)x_1x_3 & -(1+\rho)x_1x_4 & -(1+\rho)x_1x_5 \\ -(1+\rho)x_1x_2 & 3\rho x_2^2 & -(1+\rho)x_2x_3 & -(1+\rho)x_2x_4 & -(1+\rho)x_2x_5 \\ -(1+\rho)x_1x_3 & -(1+\rho)x_2x_3 & 3\rho x_3^2 & -(1+\rho)x_3x_4 & -(1+\rho)x_3x_5 \\ -(1+\rho)x_1x_4 & -(1+\rho)x_2x_4 & -(1+\rho)x_3x_4 & 3\rho x_4^2 & -(1+\rho)x_4x_5 \\ -(1+\rho)x_1x_5 & -(1+\rho)x_2x_5 & -(1+\rho)x_3x_5 & -(1+\rho)x_4x_5 & 3\rho x_5^2 \end{pmatrix}_{residue}$$

i.e. $\mathbf{M}_{univariate}(\mathbf{X}) = (1+3\rho-4\rho^2)(\mathbf{M}_{SUR} - \mathbf{M}_{residue})$

And because of $\det \mathbf{M}_{univariate} = 0$*, then* $(1+3\rho-4\rho^2)^5 \det(\mathbf{M}_{SUR}(\mathbf{X}) - \mathbf{M}_{residue}(\mathbf{X})) = 0$

The solutions of this equation with respect to ρ *for* $x_1 = x_2 = x_3 = x_4 = x_5 = 1$

are $\rho = 1, \rho = -1/4$

also, for $\rho = -1/4$ *was the design 7.47 local D-optimal for the considered SUR model.*

We have identity for the D-optimal designs for the considered SUR models and respectively the formulated univariate spring weighing models from the same components of the SUR models for $m = 2, m = 4$*, and* $m = 5$ *respectively by* $\rho = \pm 1$*,* $\rho = -1/3$ *and* $\rho = -1/4$*, and they are the solutions with respect to the determinants. But for* $m = 3$ *the identity occurs for* $\rho = -0.56$ *the variance covariance matrix of the error variables is not positive definite but regular, and therewith the corresponding information is not positive definite , but that is not the obtained solution with respect to the determinant by checking the identity of the D-optimal design. Because of this contradiction, we can not take the relationship between the information matrices for both models with respect to their determinants calculation as an acceptable argument for all cases, but it clears at least the relationship between both models, which are formulated from the same marginal models. Also, by comparing the sensitivity functions with respect to D-optimality for the SUR model with three components and the counterpart univariate spring weighing model for three objects, for the design (7.43) with* $w^* = \frac{1+2\rho+\rho^2}{2\rho^2-\rho-1}$*,* $\rho = -0.56$ *i.e.* $\frac{1-w^*}{3} = 0$*,* $\frac{w^*}{3} = \frac{1}{3}$ *or* *the design with the last three points by weights equal to* $1/3$ *for* $\rho = -0.56$*, we get the same sensitivity functions, which are equal to* $2(x_1^2 + x_2^2 + x_3^2) - (x_1x_2 + x_2x_3 + x_1x_3)$*. The same justification can be ensued for the other models, i.e. for other values of* m*. Thus for* $m = 2$*,* $m = 4$*, and* $m = 5$ *the identity for the D-optimal designs for the SUR models and their counterpart by the spring weighing models, i.e. the both models have the same marginal model, occurs ordered by* $\rho = \pm 1$*,* $\rho = -1/3$ *and* $\rho = -1/4$*, also, when the information matrix for the SUR models are singular, i.e. the the determinant of them are equal to zero, and the correlation matrix indefinite. For* $m = 3$ *occurs the identity for the D-optimal designs for the SUR model and their counterpart by the spring weighing model with three objects by* $\rho = -0.56$*, also, when the correlation matrix indefinite resp. the information matrices for the considered SUR model indefinite, but regular.*

Remark 7.7. *We can generalize the discussion in the last example for arbitrary regression functions* $\mathbf{f}_j(\mathbf{x}_j)$ *by the components of the SUR model resp. by their corresponding multi-factor univariate model, which regression function has the form* $\cap_j^m \mathbf{f}_j(\mathbf{x}_j)$*, the information matrix for the SUR model has with respect to the product type design because of (7.15) the following form*

$$\mathbf{M}_{SUR} = \mathbf{M}_{univariate} + \begin{pmatrix} (\sigma^{(11)} - 1)\mathbf{M}_1 & (\sigma^{(12)} - 1)\mathbf{m}_1 \cdot \mathbf{m}_2^\top \\ (\sigma^{(12)} - 1)\mathbf{m}_2 \cdot \mathbf{m}_1^\top & (\sigma^{(22)} - 1)\mathbf{M}_2 \end{pmatrix} \tag{7.49}$$

Thus by calculating the sensitivity functions for both models and solving the equality of them for the same design with respect to the correlation, then we can get the value of the correlation, which makes the D-optimal design for the univariate model local D-optimal for the corresponding SUR model.

Example 7.4 (Approximate D-optimality for the multivariate spring weighing models)**.** *Another application of theorem (7.1) can be the determination of the D-optimal designs for the multivariate spring weighing models, where the multivariate word is used in this work for describing correlated components, and not correlated observations, as it is usually presented in the literature. Thus we can imagine the modeled problem through the algebraical balance in figure 7.3. We can also imagine, that some experiments in pharmaceutical industry can be modeled by the multivariate spring weighing models at least in the nesting form. Another problems such as Fitness studies or calculating calories in meals can be described with these models, see (* Eisenhauer (2003)*). It may also be useful, to study the optimal weighing designs in the multivariate case, because there is a relationship between it and the balance incomplete block designs, see (* Ceranka and Katulska (1987)*).*

In this example, we will present at first different regression functions for the bivariate components, i.e. the number of components is constant for all considered models $m = 2$*, and spring weighing models with equal and unequal numbers of objects for each components, i.e. we have models with* $p_1 = p_2$*,* $p_1 \neq p_2$ *and* $p = p_1 + p_2$*. Thus for* $\xi_{11}^* = (1, 1); w = 1$ *and* $\xi_{qj}^* = \xi_j^* \otimes \xi_q^*$ *for* $q, j = 2, 3, 4, 5$ *the product type designs are D-optimal for the considered bivariate spring weighing models for the given intervals for* ρ*,s as follows*

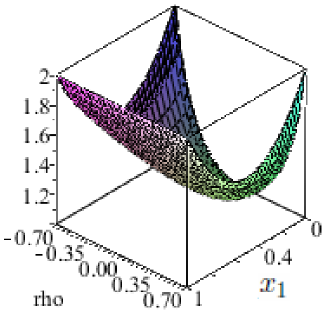

Figure 7.1: $\varphi_D(\otimes_{j=1}^{2}\xi_0^*)$ for SUR model without intercept

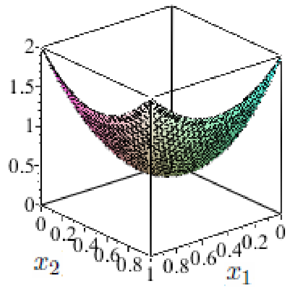

Figure 7.2: $\varphi_D(\xi_\rho^*)$ for SUR model without intercept

Figure 7.3: Four-Pan Algebra Balance

p_1	p_2	p	ξ^*	$\det \mathbf{C}^{1/p}$	$\det \mathbf{C}^{1/p}$ *-I*	ρ^* *-I*	*I-L*
1	*1*	*2*	ξ_{11}^*	$[0.670, 1.000]$	0.330	$[-0.707, 0.707]$	1.414
1	*2*	*3*	ξ_{12}^*	$[0.630, 1.000]$	0.370	$[-0.866, 0.866]$	1.732
1	*3*	*4*	ξ_{13}^*	$[0.867, 1.000]$	0.133	$[-0.660, 0.660]$	1.320
1	*4*	*5*	ξ_{14}^*	$[0.898, 1.000]$	0.102	$[-0.645, 0.645]$	1.290
1	*5*	*6*	ξ_{15}^*	$[0.900, 1.000]$	0.100	$[-0.683, 0.683]$	1.366
2	*2*	*4*	ξ_{22}^*	$[0.000, 1.000]$	1.000	$[0]$	0.000
2	*3*	*5*	ξ_{23}^*	$[0.848, 1.000]$	0.152	$[-0.750, 0.750]$	1.500
2	*4*	*6*	ξ_{24}^*	$[0.850, 1.000]$	0.150	$[-0.790, 0.790]$	1.580
3	*3*	*6*	ξ_{33}^*	$[0.859, 1.000]$	0.141	$[-0.774, 0.774]$	1.548
3	*4*	*7*	ξ_{34}^*	$[0.870, 1.000]$	0.130	$[-0.79, 0.79]$	1.580
4	*4*	*8*	ξ_{44}^*	$[0.000, 1.000]$	1.000	$[0]$	0.000

this simulation illustrates, that for a fixed number of components $m = 2$, the total number of parameters for the SUR model p do not play any roll by the length of the ρ^* -Interval resp. $\det \mathbf{C}^{1/p}$ -Interval, which have been illustrated by the first simulation of the first example for grown m. However we can remark, that for a fixed number of parameters for the first marginal model p_1 and grown number of parameters for the second marginal model p_2, we get longer ρ^* -Intervals resp. shorter $\det \mathbf{C}^{1/p}$ -Intervals. We can remark too, that the product type designs are D-optimal only for the bivariate weighing models with two or four objects, if $\rho = 0$, which means that the inequality $\max_{\mathbf{x}\in\mathcal{X}} \varphi_D(\mathbf{x}, \xi^*, \rho^*) \leq p$ by theorem (7.1) is valid just for one value, $\rho^* = 0$.

In the next example we will consider the multivariate spring weighing model for three objects by each marginal model, also, the product type designs are D-optimal for intervals for ρ's as follows

m	p	ξ^*	$\det \mathbf{C}^{1/p}$ *-I*	*I-L*	ρ^* *-I*	*I-L*
2	*6*	$\otimes_{j=1}^{2}\xi_3^*$	$[0.859, 1.000]$	0.141	$[-0.774, 0.774]$	1.548
3	*9*	$\otimes_{j=1}^{3}\xi_3^*$	$[0.802, 0.865]$	0.063	$[-0.434, 0.767]$	1.201
4	*12*	$\otimes_{j=1}^{4}\xi_3^*$	$[0.770, 0.877]$	0.107	$[-0.302, 0.763]$	1.065

Thus this simulation gives the same result from the first example, also, for grown m we have to have *shorter* ρ^* *-Intervals and respectively longer* $\det \mathbf{C}^{1/p}$ *-Intervals for the D-optimality of the product type designs ab* $m > 2$*.*

Example 7.5 (Approximate multivariate D-optimal chemical balance designs). *It is important to remark, that in the first tabula all intervals contains the value zero for* ρ*, also the product type designs are always D-optimal for* $\rho = 0$*. When we implemented the example in the experiments region* $[-1, 1]$*, then the product type designs*

$$\otimes_{j=1}^{m} \xi_0^* = \otimes_{j=1}^{m} \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{7.50}$$

are A-, D- and IMSE-optimal for the SUR models with the marginals $Y_{ij} = \beta_{j1} x_{ij} + \varepsilon_{ij}$ $j = 1, ..., m$*, as it has been mentioned in theorem (7.1).*
The same result is to be held for example (7.4) in the experiments region $[-1, 1]$*. Thus by considering the bivariate model*

$$\begin{aligned} Y_{i1} &= \beta_{11} x_{i11} + \beta_{12} x_{i12} + \varepsilon_{i1} \\ Y_{i2} &= \beta_{21} x_{i21} + \beta_{22} x_{i22} + \beta_{23} x_{i23} + \varepsilon_{i2} \end{aligned} \tag{7.51}$$

Then the product type designs of the form (7.50), also $\otimes_{j=1}^{2} \xi_0^*$ *and* $\otimes_{j=1}^{2} \xi_0^*$ *are ordered A-, D-, and IMSE-optimal for the marginal models in (7.51), thus the product type design* $\otimes_{j=1}^{5} \xi_0^*$ *is A-, D-, and IMSE-optimal for the SUR model resp. chemical balance model with the marginal models in (7.51). Figure 7.4 illustrates the D-optimality of the product type design* $\otimes_{j=1}^{5} \xi_0^*$ *resp. the graphic of the sensitivity function*

$$\varphi_D(\mathbf{x}; \otimes_{j=1}^{5} \xi_0^*) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2$$

for $x_3 = x_4 = x_5 = 1$*. Figure 7.5 illustrates the A-optimality of the product type design* $\otimes_{j=1}^{5} \xi_0^*$ *resp. the graphic of the function*

$$\frac{\varphi_A(\mathbf{x}; \xi_A^*)}{\operatorname{trace}\left(\mathbf{M}(\xi_A^*)^{-1}\right)} = \frac{\sigma_1^2 x_1^2 + \sigma_1^2 x_2^2 + \sigma_2^2 x_3^2 + \sigma_2^2 x_4^2 + \sigma_2^2 x_5^2}{2\sigma_1^2 + 3\sigma_2^2}, \ x_3 = x_4 = x_5 = 1, \sigma_1 = \sigma_2 = 1$$

Figure 7.6 illustrates the IMSE-optimality of the product type design $\otimes_{j=1}^{5} \xi_0^*$ *resp. the graphic of the function*

$$\frac{\varphi_{IMSE}(\mathbf{x}; \xi_{IMSE}^*)}{\operatorname{trace}\left(\mathbf{L}\ \mathbf{M}(\xi_{IMSE}^*)^{-1}\right)} = \frac{\sigma_1^2 x_1^2 + \sigma_1^2 x_2^2 + \sigma_2^2 x_3^2 + \sigma_2^2 x_4^2 + \sigma_2^2 x_5^2}{2\sigma_1^2 + 3\sigma_2^2}, \ x_3 = x_4 = x_5 = 1, \sigma_1 = 10, \sigma_2 = 0.1$$

Example 7.6 (The linear optimal design for general information matrix). *We will consider the simplest bivariate model with the marginals* $Y_{ij} = \beta_{j1} x_{ij} + \varepsilon_{ij}$ *for* $j = 1, 2$*, also, the product type design*

$$\xi = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ 1 \end{pmatrix}$$

is not A- nor IMSE-optimal for $\rho \neq 0$*, even locally. The following A- and IMSE-optimal designs for the corresponding multi-factor (additively) model or the spring weighing model for two objects are not A- and IMSE-optimal designs for the considered bivariate SUR model.*

$$\xi = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 1 - 2.w & w & w \end{pmatrix}, \ w_A \simeq 0.422, \ w_{IMSE} \simeq 0.264,$$

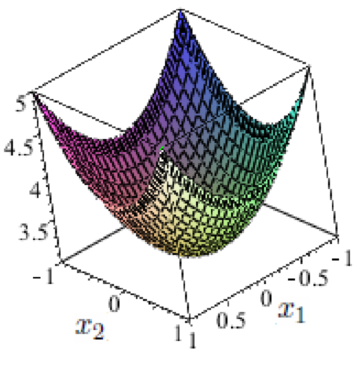

Figure 7.4: D-optimality for SUR model as Chemical balance model

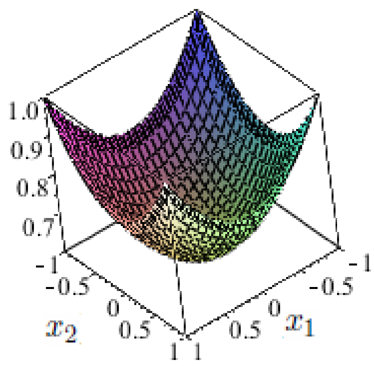

Figure 7.5: A-optimality for SUR model as Chemical balance model

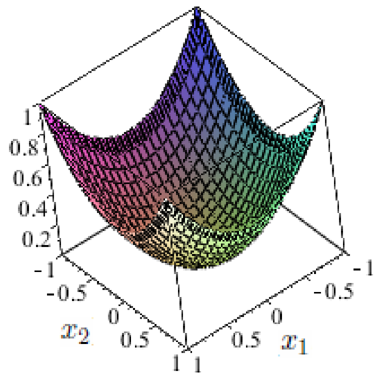

Figure 7.6: IMSE-optimality for SUR model as Chemical balance model

8 Discussion and Future Research

Finding optimal designs for multivariate linear models analytically is not easy but possible, and this work can be a positive signal by that, where algorithms are not the only method to overcome the complexity, which appear because of the different covariance resp. correlation structures, see for example *(Wijesinha and Khuri (1987).*

8.1 Conclusion

D- and linear optimal in 4.1 designs are determined for SUR models with different structures for the one- and multi-factor-marginal components, where the product type designs are D- and linear optimal in 4.1 designs for SUR models with intercepts by the one-factor and multiplicative marginals with different nesting forms, and by additive marginals with different nesting forms except for block diagonal information matrices. For SUR models without intercepts by the marginal components the product type designs are D- and linear optimal in 4.1 for the block diagonal form of the information matrices, as practical example for such models are the multivariate chemical balance regression models. These results can be held for a known covariance matrix for the error variables due to the Gauß Markov estimator and asymptotically for an unknown variance covariance matrix of the error variables, when the error variable is normally distributed. The D-optimality for the product type designs is restricted to the general form of the information matrix for SUR models without intercepts by the marginals and depended on the value of the correlation term, so for intervals, which include zero and their sides less than one in absolute value the product type designs can be D-optimal and these intervals will be closer for grown components numbers m, where the multivariate spring balance regression models can be a good example or a special case for such SUR models.
G-optimal criteria for the multivariate case were discussed and their upper bounds due to the D-optimal design resp. weighted G-optimal design, were determined by some inequalities from the Matrix theory for the product of the positive definite matrices, for general multivariate linear models, for MANOVA-models by the MANOVA-design, and SUR models due to the Gauß and OLS estimators by the product type designs. Efficiencies for the OLS estimator vs the Gauß Markov estimator for SUR models with respect to the product type designs, and due to D- and linear optimal criteria in 4.1 were determined, as well as the efficiencies for the MANOVA-design vs the product type designs due to the Gauß and OLS estimators. It is shown, that the MANOVA-designs are more efficient than the product type designs with respect to the OLS estimator due to the D-optimality, and the MANOVA-designs, and the product type designs, that have the same efficiency with respect to the OLS-estimator, and due to the linear optimality in 4.1. Integrated mean square error in the multivariate case as well as the IMSE-criterion were defined and derived, which is a type of the considered linear optimal criteria in 4.1. By implementing the IMSE-optimal criterion on the MANOVA-model, it is shown, that the reduction of IMSE-optimality for the MANOVA-model to the corresponding univariate problem of the marginal models as in *(Kurotschka and Schwabe (1996)* is possible.

8.1.1 Extensions

Through this work, it can be concluded, that the research of optimal designs for multivariate linear models are analytically not very restricted and can be reduced to the corresponding univariate case in many cases, where by an extension of the results for this work is possible. SUR models with fixed block effects can be a special case of the considered model by the dissertation of my college Jesus Alonso Cabrera, where optimal designs for models with random and fixed effects are explored.
When the sensitivity functions in the multivariate case due to D-optimality in general are independent on the variance terms and dependent only on the correlation term, then theorem 7.1 is valid and the D-optimal designs can be similarly restricted by some intervals of the correlation terms for different multivariate linear models.
The orthogonality conditions for the SUR models without intercepts by the corresponding marginal components, which makes the non block diagonal information matrix block diagonal, can be relieved for the nearest neighbor correlation structure.
Exploring the optimal designs for multivariate linear models through a more complex covariance structure, such as the Kronecker product covariance structure resp. when not only the components but the individual resp. observations are correlated, so the variance covariance matrix of the error variables for the multivariate model has the form $\mathbf{V} = \mathbf{\Sigma}_1 \otimes \mathbf{\Sigma}_2$, so the product type designs are no longer optimal for the simplest example for the considered SUR models, and the information matrix is no longer the sum of the information matrices for the individual. For the MANOVA-model the reduction of the optimal design problem to the corresponding univariate problem is impossible in this structure of $\mathbf{V}$ as in the work of *(Kurotschka and Schwabe (1996)*. However it can be reduced to the problem for finding optimal designs for univariate models with correlated observations, because the variance covariance matrix of the Gauß estimator has instead of (3.16) the following form

$$\begin{aligned}
&\mathbf{Cov} = \left(\mathbf{F}^\top \mathbf{V}^{-1}\mathbf{F}\right)^{-1} = \\
&\left((\mathbf{I}_{m\times m}\otimes \mathbf{f}_0(x_1),...,\mathbf{I}_{m\times m}\otimes \mathbf{f}_0(x_n))(\mathbf{\Sigma}_1^{-1}\otimes \mathbf{\Sigma}_2^{-1})(\mathbf{I}_{m\times m}\otimes \mathbf{f}_0(x_1),...,\mathbf{I}_{m\times m}\otimes \mathbf{f}_0(x_n))^\top\right)^{-1} \\
&= \mathbf{\Sigma}_1 \otimes \left(\sum_{i=1}^{n} \mathbf{f}_0(x_i)\mathbf{\Sigma}_2^{-1}\mathbf{f}_0(x_i)^\top\right)^{-1}
\end{aligned} \tag{8.1}$$

Then the results due to optimal designs for models with correlated observations in the works of *Bischoff (1992), (1993), and (1995)* and others are valid for these models for known $\mathbf{\Sigma}_1$.
The reduction of the optimal design problem for Growth curve models in *(Reinsel and Velu (1998)* to the optimal design for models with correlated observations is possible, if we are interested in the optimization of the time points for the observations t and not in the optimization of the number of subjects n, because of the form of information matrix, which is the Kronecker product for two information matrices, the first one is dependent on the numbers of subjects n and the second one dependent on the time points t, i.e.

$$\mathbf{M} = \mathbf{F}_1(\mathbf{n})^\top \mathbf{F}_1(\mathbf{n}) \otimes \mathbf{F}_2(\mathbf{t})^\top \mathbf{\Sigma}^{-1}\mathbf{F}_2(\mathbf{t}) = \mathbf{M}_1(\mathbf{sn}) \otimes \mathbf{M}_2(\mathbf{t})$$

So in general the information matrix for time points $\mathbf{M}_2(\mathbf{t})$ has the structure of the information matrix for models with correlated observations, and the considered regression functions by *(Reinsel and Velu (1998)* are polynomial regression function, so the design matrix $\mathbf{F}_2(\mathbf{t})$ is quadratic for equal numbers of observations and polynomial degrees resp. numbers of parameters , so the reduction of the D-optimality problem for correlated observations in the regular univariate problem for known variance covariance matrix of the error variables in this case is possible, because of the following form

$$\det \mathbf{M}_2(\mathbf{t}) = \det\left(\mathbf{F}_2(\mathbf{t})^\top \mathbf{\Sigma}^{-1}\mathbf{F}_2(\mathbf{t})\right) = \det \mathbf{F}_2(\mathbf{t})^\top \det \mathbf{\Sigma}^{-1} \det \mathbf{F}_2(\mathbf{t})$$

$$\text{thus } \det \mathbf{M}_2(\mathbf{t}) = \det \mathbf{\Sigma}^{-1} \det\left(\mathbf{F}_2(\mathbf{t})^\top \mathbf{F}_2(\mathbf{t})\right)$$

8.2 Possible Research

Much research in this region can be done, so the following research will be considered in the future

- Exploring further optimality criteria for the considered SUR models as the C-, E-, D_s- and linear optimal criteria, by non block diagonal weight matrix $\mathbf{L}$.
- Exploring the optimal designs for the considered SUR models by limited information estimators as the the feasible General least square estimator, as the two or three stage GLS, OLS-or ML-estimators in some cases see *(Amemiya (1985)* or *(Anderson (2005).* The big problem by dealing with such estimators is the non-equality between the inverse of their variance covariance matrices and the information matrices and therewith it is to prove the convexity for the set of the variance covariance matrices for such estimators to build the convex theory due to them.
- Usage of the results of the analytical optimal designs for the considered SUR models to test the efficiency for the developed algorithms or other optimization methods as the positive semi-definite programming *(Atashgah and Seifi (2009)* by finding optimal designs for multivariate linear models by their application to the considered SUR models.
- Exploring the optimality of the product type designs and the restrictions on D-optimality for the considered SUR models without intercepts for some special correlation structures as the autoregressive and the nearest neighbor correlation structures.
- Determining the efficiency bounds for product designs by the considered SUR models, where their marginals are additive models or multi-factor models, which can be formed additively but without intercepts in hierarchically and other nesting forms, similar to the work of *(Schwabe and Wong (1999).*
- Exploring the optimal designs for multivariate linear models through a more complex covariance structure, as the Kronecker product covariance structure resp. when not only the components but the individual resp. observations are correlated so the variance covariance matrix of the error variables for the multivariate model has the form $\mathbf{V} = \mathbf{\Sigma}_1 \otimes \mathbf{\Sigma}_2$.
- Finding optimal designs for the considered growth curve models by *(Reinsel and Velu (1998)* for non quadratic design matrix $\mathbf{F}_2(\mathbf{t})$.
- Investigating the optimal designs for the considered SUR models and other multivariate linear models but by other structures for the marginal regression functions resp. the parameter.

9 Bibliography

Bibliography

ABADIR, K. M. AND MAGNUS, J. R. (2005): Matrix Algebra (Econometric Exercises) . *Cambridge University Press, Cambridge.*

AMEMIYA, T. (1985): Advanced Econometrics. *Cambridge, Mass: Harvard University Press.*

ANDERSON, T.W. (2005): Origins of the limited information maximum likelihood and two-stage least squares estimators. *Journal of Econometrics, 127, 1-16.*

ATASHGAH , A. B. AND SEIFI, A. (2009): Optimal design of multi-response experiments using semi-definite programming. *Optimization and Engineering, 10 , 75-90.*

BALTAGI, B. (2011): Econometrics. *Springer, Berlin, Eighth edition.*

BANDEMER, H. ET. AL.(1977): Theorie und Anwendung der optimalen Versuchsplanung. I. Handbuch zur Theorie. *Akademie-Verlag, Berlin.*

BANDEMER, H. AND NÄTHER, W. (1980): Theorie und Anwendung der optimalen Versuchsplanung. II. Handbuch zur Anwendung. *Akademie-Verlag, Berlin.*

BANDEMER, H. AND BELLMANN, A. (1994) : Statistische Versuchsplanung. *Teubner Verlag; Leipzig, Auflage: 3.*

BERTRAND, J. AND MENTRé, F. (2008): Mathematical Expressions of the Pharmacokinetic and Pharmacodynamic Models implemented in the Monolix software. *http://www.lixoft.eu/wp-content/resources/docs/PKPDlibrary.pdf* (Last Access 01.06.2013).

BISCHOFF, W. (1992): On exact D-optimal designs for regression models with correlated observations. *Annals of the Institute of Statistical Mathematics, 44, 229-238 .*

BISCHOFF, W. (1993): On D-optimal designs for linear models under correlated observations with an application to a linear model with multiple response. *Journal of Statistical Planning and Inference, 37, 69-80.*

BISCHOFF, W. (1995): Lower bounds for the efficiency of designs with respect to the D-criterion when the observations are correlated. *Statistics, 27, 27-44 .*

BROWN, L. D. AND WONG, W. K. (2000): An algorithmic construction of optimal minimax designs for heteroscedastic linear models. *Journal of Statistical Planning and Inference, 85, 103-114.*

CASELLA, G. (1983): Leverage and Regression through the Origin. *American Statistician, 37, 147-52.*

CASELLA, G. AND LEHMANN, E. L. (1998) : Theory of Point Estimation (Springer Texts in Statistics). *Springer, New York, 3 edition.*

CERANKA, B. AND KATULSKA, K. (1987): A relation between BIB designs and chemical balance weighing designs. *Statistics & Probability Letters, 5, 339-341.*

CHANG, S.I. (1994): Some Properties of multiresponse D-optimal designs. *J. Math. Anal. Appl., 184, 256-262.*

CHANG, F.-C. AND HEILIGERS, B. (1996): E-optimal designs for polynomial regression without intercept. *J. Statist. Plann. Inference 55, 371-387.*

CHANGA, F.-C. , HUANGA, M.-N. L. , LIN, D. K. J. AND YANGA, H.-C. (2001): Optimal designs for dual response polynomial regression models. *Journal of Statistical Planning and Inference, 93, 309- 322 .*

CHEN, R.-B. , WONG, W. K., LI, K.-Y. (2008): Optimal minimax designs over a prespecified interval in a heteroscedastic polynomial model. *Statistics & Probability Letters, 78 (13) 1914-1921.*

CHRISTENSEN, R. (2001): Advanced linear modeling: multivariate, time series, and spatial data. *Springer New York.*

DETTE, H. (1990): A Generalization of D and D1-optimal design in polynomial regression. *Ann. Statist. 18, 1784-1804.*

DURRETT, R.(2010) : Probability: Theory and Examples. *Cambridge Series in Statistical and Probabilistic Mathematics, Fourth Edition, Cambridge.*

EISENHAUER, JOSEPH G. (2003): Regression through the Origin. *Teaching Statistics, 25, 76-80.*

FEDOROV V.V. (1972): Theory of optimal Experiments. *Academic Press, New York.*

GAFFKE, N. (2009/ 2010) : Maß- und Integrationstheorie (Maßtheorie) Kurzskript. `http://www.math.uni-magdeburg.de/institute/imst/ag_gaffke/` (Last Access 01.06.2013)

HACKL, P. (2008) : Einführung in die Ökonometrie. *Pearson Studium, München.*

HATZINGER, R. (1991): Quasi-Likelihood Methoden zur Analyse von unabhängigen und abhängigen Beobachtungen. *Forschungsberichte / Institut für Statistik, 13. Department of Statistics and Mathematics, WU Vienna University of Economics and Business, Vienna.*

HUANG, M.-N., CHANG, F.-C. AND WONG, W.-K. (1995): D-optimal designs for polynomial regression without an intercept. *Statist. Sinica 5, 441- 458.*

HUDA, S. AND MUKERJEE, R. (1988): Optimal weighing designs: Approximate theory. *statistics 19, 513-517.*

KIEFER, J. AND WOLFOWITZ, J. (1959): Optimum experimental designs. *J. R. Statist. Soc., 21, 272-304.*

KIEFER, J. (1961): Optimum designs in regression problems, . *II. Ann. Math. Statist. 32, 298-325.*

KIEFER, J. (1974): General equivalence theory for optimum designs(approximate theory). *Ann. Statist. 2, 849-879.*

KIEFER, J (1975): Variation in structure and performance under change of criterion. *Biometrika, 62, 277-288.*

KIEFER, J. AND WOLFOWITZ, J. (1959): Optimum designs in regression problems. *Ann. Math. Statist. 30, 271-294.*

KIEFER, J. AND WOLFOWITZ, J. (1960): The equivalence of two extremum problems. *Can. J. Math. 12, 363-366.*

KING, J. , WONG, W. K. (1998): Optimal minimax designs for prediction in heteroscedastic models. *Journal of Statistical Planning and Inference, 69, 371-383 .*

KRAFFT, O. AND SCHÄFER, M. (1992): D-optimal designs for a multivariate linear Model. *J. Multivar. Anal., 42, 130-140.*

KUROTSCHKA, V.G. AND SCHWABE, R.(1996): The reduction of design problems for multivariate experiments to univariate possibilities and their limitations . *In: E. Brunner and M. Denker (Eds.), Research Developments in Probability and Statistics. Festschrift in Honor of Madan L. Puri on the Occasion of his 65th Birthday. VSP, Utrecht 193-204.*

MAGNUS, J. R. (1978): Maximum likelihood estimation of the GLS model with unknown parameters in the disturbance covariance matrix. *Journal of Econometrics, 7, 281-312 .*

MYERS, H. M. , MONTGOMERY, D. C. , VINING,G. G. AND ROBINSON, T. J. (2010) : Generalized Linear Models: with Applications in Engineering and the Sciences. *Wiley, 2 edition, Hoboken.*

MULLER, K. E. AND STEWART, P. W. (2006): Linear Model Theory: Univariate, Multivariate, and Mixed Models. *Wiley, New York*

PÁZMAN, A. (1986) : Foundation of Optimum Experimental Design. *Reidel, Dordrecht.*

PENG, J., ZHU, J., BERGAMASCHI, A., HAN, W., NOH, D.-Y. , POLLACK, J. R., AND WANG, P. (2010): Regularized multivariate regression for identifying master predictors with application to integrative genomics study of breast cancer. *Ann. Appl. Stat., 4, 53-77.*

PETERSON, K.,B. AND PEDERSON, M., S. (2008): The Matrix Cookbook. *http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=3274* (Last Access 01.06.2013).

PUKELSHEIM, F. (1993): Optimal Experimental Design. *Wiley, New York.*

RAO, C. R. (1973): Linear Statistical Inference and Its Applications. *Wiley, New York, 2 edition.*

RAO, C.R., TOUTENBURG, H., SHALABH, HEUMANN, C. (2007) : Linear Models and Generalizations: Least Squares and Alternatives. *Springer, New York.*

REINSEL, C. G. AND VELU, P. R. (1998): Multivariate reduced-rank regression: theory and applications. *Springer, New York.*

SCHWABE, R. AND WONG, W., K. (1999): Efficiency bounds for product designs in linear models. *Ann. of the Institute of Stat. Math. 51, 723-730 .*

SCHWABE, R. (1996)A: Optimum Designs for Multi-factor Models. Lecture Notes in Statistics 113. *Springer, New York.*

SCHWABE, R. (1996)B: Optimal design of experiments for multivariate response in two-factor linear models. . *Invited paper presented at the sixth Eugene Lukacs Symposium on Multidimensional Statistical Analysis and Theory of Random Matrices. Proceedings of the Sixth Eugene Lukacs Symposium, Bowling Green, Ohio, 235-242.*

SCHWABE, R. (2008): Optimal Design for linear and non-linear Models (Tutorial). *Isaac Newton Inst. for Math. Sience,s DOE Workshop . http://www.newton.ac.uk/webseminars/pg+ws/2008/doe/doew01/0721/schwabe/frames.html* (Last Access 01.06.2013)

SILVEY, S. D. (1980): Optimal Design . *Chapman and Hall, London .*

SMITH, K. (1918): On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance they give towards a proper choice of the distribution of observations. *Biometrika 12, 1-85.*

SRIVASTAVA, V. K. AND GILES, D. E. A. (1987): Seemingly unrelated regression equations models: estimation and inference. *Marcel Dekker, New York.*

TURKINGTON, D. A. (2002) : Matrix Calculus and Zero-One Matrices Statistical and Econometric Applications. *Cambridge University Press, Cambridge.*

WEI, PENG AND PAN, WEI (2008): Incorporating Gene Functions into Regression Analysis of DNA-Protein Binding Data and Gene Expression Data to Construct Transcriptional Networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics. 5, 401-415.*

WIJESINHA, M.C. AND KHURI, A.I. (1987): The Sequential Generation of Multire-sponse D-optimal Designs when the Variance-Covariance Matrix is not Known. *Commun. Statist. Simula. Computa., 16, 239-259.*

WIJESINHA, MMC (1984): Design of experiments for multiresponse models. *Unpublished Ph.D. thesis, Dep. of Statistics, University of Florida, Gainesville.*

WONG, W. K. (1995): On the equivalence of D and G-optimal designs in heteroscedastic models. *Statistics & Probability Letters, 25, 317-321.*

WONG, W. K. (1993): Minimal number of support points for mini-max optimal designs. *Statistics & Probability Letters, 17, 405-409.*

WONG, W. K. (1994): G-optimal designs for multi-factor experiments with heteroscedastic errors. *Journal of Statistical Planning and Inference, 40, 127-133.*

YANG, X. (2000): A Matrix Trace Inequality. *J. Math. Anal. Appl. 250, 372-374 .*

YANG,Z. P. AND FENG, X. X. (2002) : A Note on the Trace Inequality for Products of Hermitian Matrix. *Journal of Inequalities in Pure and Applied Mathematics, 3, Article 78 .*

ZELLNER, A. (1962): An Efficient Method of Estimating Seemingly Unrelated Regression Equations and Tests for Aggregation Bias. *Journal of the American Statistical Association, 57, 348-368.*

Zhang F. (1999) : Matrix Theory: Basic Results and Thechniques. *Springer, New York.*

List of Figures