

D-Effizienzen von Versuchsplänen im einfachen linearen Regressionsmodell bei korrelierten Fehlern

Dissertation

zur Erlangung des Grades
eines Doktors der Naturwissenschaften
der Technischen Universität Dortmund

Der Fakultät Statistik
der Technischen Universität Dortmund

vorgelegt von

Adrian Wilk

Dortmund im Dezember 2012

1. Gutachter: Prof. Dr. Joachim Kunert

2. Gutachterin: Prof. Dr. Christine Müller

Kommissionsvorsitz: Prof. Dr. Jörg Rahnenführer

Tag der mündlichen Prüfung: 19. Februar 2013

Inhaltsverzeichnis

1. Einleitung	1
2. Modell und Problemstellung	3
3. Parameterschätzung in linearen Regressionsmodellen	7
3.1. Eine Schranke zur Beurteilung von Schätzern	7
3.2. Verallgemeinerter Kleinste-Quadrate-Schätzer	12
3.3. Maximum-Likelihood-Schätzer	13
3.4. Erwartungstreue und Kovarianzmatrix linearer Schätzer unter $\lambda = 1$	14
4. Das D-Kriterium zur Beurteilung von Versuchsplänen	21
5. D-Effizienzen von Versuchsplänen bei bekannten Störparametern	23
5.1. D-optimale Versuchspläne	24
5.2. Untersuchung der D-Effizienz	25
5.3. Vorüberlegungen zu den Abschnitten 5.1 und 5.2	33
5.3.1. Konstruktion des D-Kriteriums	33
5.3.2. Stetigkeit des D-Kriteriums	38
5.3.3. Grenzfall $\lambda \rightarrow 1$	42
5.3.4. Abschätzungen	44
5.3.5. Schranken für $\tilde{D}_\lambda(d)$	54
5.3.6. Relative D-Effizienzen	61
6. D-Effizienzen von Versuchsplänen bei unbekanntem Störparametern	75
6.1. Zweistufige Schätzung	75
6.2. Simulationsstudie	82
7. Untersuchungen zum D-Kriterium für den Parametervektor $[\beta_0, \beta_1, \lambda]^T$	97
7.1. Konstruktion des D-Kriteriums	98
7.2. Vergleich der Versuchspläne mit $d \rightarrow 0$ und $d = 1/2$	100
8. Relative D-Effizienzen bei verallgemeinerten Abhängigkeitsstrukturen	103
8.1. Ein Modell mit Nugget-Effekt	103
8.2. Verallgemeinerte Abhängigkeitsstrukturen	104
8.3. Vergleich von Versuchsplänen	107

9. Zusammenfassung	113
A. Lemmata	115
B. Tabellen	149
C. Programmcode	151
Symbolverzeichnis	157
Literatur	159

1. Einleitung

Die vorliegende Arbeit greift eine Problemstellung aus dem Artikel „*Exact optimal designs for weighted least squares analysis with correlated errors*“ von Dette *et al.* (2008) auf. In dem Artikel wird unter anderem der Frage nachgegangen, wie D-optimale Versuchspläne im einfachen linearen Regressionsmodell bei autoregressiver Fehlerstruktur auszusehen haben, wenn die Regressionsparameter mittels der verallgemeinerten Kleinste-Quadrate-Methode geschätzt werden. Die Autoren konnten zeigen, dass zwei Versuchspunkte eines D-optimalen Versuchsplans mit den beiden Randpunkten des Versuchsraums übereinstimmen. Ein weiteres Resultat lautet, dass im Grenzfall hochkorrelierter Störgrößen der D-optimale Versuchsplan gegen einen Versuchsplan mit äquidistanten Versuchspunkten konvergiert. Weitere Resultate bezüglich des einfachen linearen Regressionsmodells sind jedoch nur numerische Ergebnisse. Unter anderem wird für die Klasse von Versuchsplänen, die drei Versuchspunkte aufweisen, beschrieben, dass der D-optimale Versuchsplan von der Höhe der Korrelation abhängt. Es existieren somit lediglich sogenannte lokal D-optimale Versuchspläne. Die vorliegende Arbeit setzt an diesem Punkt an. Es stellt sich unter anderem die Frage, ob gewisse Versuchspläne, unabhängig von der Höhe der Korrelation, Vorteile beziehungsweise Nachteile gegenüber anderen Versuchsplänen besitzen. Ein weiterer offener Punkt ist die Konsequenz aus der Tatsache, dass in dem betrachteten Modell ein D-optimaler Versuchsplan nur lokal optimal ist. Dies bedeutet für die Angabe des optimalen Versuchsplans nämlich, dass gewisse Parameter bekannt sein müssen. Speziell hier ist für die Angabe eines D-optimalen Versuchsplans erforderlich, dass der Korrelationsparameter bekannt ist. Es stellt sich somit die Frage, wie damit umzugehen ist, falls der Korrelationsparameter unbekannt ist.

Die Arbeit ist folgendermaßen gegliedert. Modell und Problemstellung werden in Kapitel 2 ausführlich vorgestellt. Wie die Parameter des vorliegenden Modells geschätzt werden können, ist Gegenstand von Kapitel 3. Im Anschluss wird in Kapitel 4 das sogenannte D-Kriterium vorgestellt. Anhand dessen können Versuchspläne bewertet und miteinander verglichen werden. Damit besitzen die Kapitel 2, 3 und 4 einen einleitenden Charakter. Die Kapitel 5 bis 7 unterscheiden sich im Großen und Ganzen bezüglich zweier Aspekte. Der erste Aspekt ist der, dass unterschiedliche Anzahlen an Parametern, die in das D-Kriterium einfließen, betrachtet werden. Darüber hinaus gibt es Unterschiede bezüglich der Annahme, ob die Parameter, die die Verteilung der Störgrößen beschreiben, bekannt sind. Entweder sind sie wie in Dette *et al.* (2008) bekannt oder sie sind allesamt unbekannt

oder zum Teil unbekannt. In Kapitel 8 werden im Vergleich zu den vorherigen Kapiteln zwei verallgemeinerte Modellannahmen betrachtet, was dazu führt, dass sich die Anzahl der Modellparameter um eins erhöht. Die zentralen Kapitel 5 bis 8 weisen die Gemeinsamkeit auf, dass in ihnen Versuchspläne anhand von Effizienzen miteinander verglichen werden. Die Bestimmung dieser Effizienzen stellt den Kern der Arbeit dar. Abschließend erfolgt eine Zusammenfassung der Arbeit.

Als Hilfsmittel wurden die Programme R (R Development Core Team, 2011) und Maxima (Maxima.sourceforge.net, 2007) verwendet. Ersteres wurde zur Erstellung der Grafiken und im Rahmen einer Simulationsstudie herangezogen. Mit letzterem Programm wurden algebraische Umformungen durchgeführt.

2. Modell und Problemstellung

Die Grundlage der vorliegenden Arbeit ist die Annahme, dass eine Zielgröße das Ergebnis eines Experiments ist. Dabei hängt diese Zielgröße linear von einer unabhängigen Größe ab. Der lineare Zusammenhang ist unbekannt und wird additiv durch eine Störgröße überlagert. Das zentrale Ziel ist es, diesen Zusammenhang zu ermitteln, wobei die Werte der unabhängigen Größe nicht vorab festgelegt sind, sondern in einem bestimmten Rahmen frei gewählt werden können. Die Festlegung auf gewisse Werte geschieht durch die Wahl sogenannter Versuchspunkte. Da die Wahl der Versuchspunkte jedoch die Güte der Schätzung beeinflusst, ist es erstrebenswert, die Versuchspunkte möglichst geeignet im Sinne eines noch zu definierenden Gütekriteriums zu wählen. Die Menge der gewählten Versuchspunkte wird schließlich in dem sogenannten Versuchsplan zusammengefasst. Eine weitere Annahme in dieser Arbeit ist die, dass die Zugehörigkeit der Versuchspunkte zu Versuchseinheiten von Bedeutung ist. Denn Versuchspunkte einer Versuchseinheit stehen in der Beziehung, dass zwischen den entsprechenden Störgrößen ein Zusammenhang vorliegt. Dieser Zusammenhang äußert sich in der Art, dass die realisierten Werte der Störgrößen umso ähnlicher sind, je geringer der Abstand der zugehörigen Versuchspunkte ist. Es besteht somit ein räumlicher Zusammenhang. Dagegen wird angenommen, dass zwischen Versuchspunkten verschiedener Versuchseinheiten ein derartiger Zusammenhang nicht besteht. Die Anzahl der Versuchseinheiten innerhalb einer Versuchsreihe wird im Weiteren durch die Größe n gekennzeichnet. Bezüglich der Versuchseinheiten wird schließlich noch die Annahme getroffen, dass aus organisatorischen Gründen jeder Versuchseinheit die gleichen Versuchspunkte zugewiesen werden. Diese Situation liegt beispielsweise auch in der Arbeit von Potthoff und Roy (1964) vor.

Die im vorherigen Absatz beschriebene Situation wird nun formalisiert. Im Verlauf dieser Arbeit wird unter anderem der Fall untersucht, dass ein Versuchsplan ξ nur die Versuche einer einzigen Versuchseinheit enthält. Dies entspricht dem Fall, dass $n = 1$ ist. Dieser Fall wird nun vorgestellt und dann anschließend auf die Fälle $n > 1$ übertragen. Das zu untersuchende Modell lautet für $n = 1$

$$y_j = \beta_0 + \beta_1 x_j + e_j, \quad j \in \{1, 2, \dots, q\}. \quad (2.1)$$

Dabei bezeichnet y_j die j -te Messung einer Zielgröße, die vom j -ten Versuchspunkt x_j , den beiden Parametern β_0 und β_1 und dem Fehlerterm e_j abhängt.

Eine alternative Schreibweise für das Modell (2.1) ist

$$y_j = f(x_j)^T \beta + e_j = \sum_{k=0}^1 \beta_k f_k(x_j) + e_j, \quad (2.2)$$

wobei $f(x_j)^T = [f_0(x_j), f_1(x_j)] = [1, x_j]$ und $\beta = [\beta_0, \beta_1]^T \in \mathbb{R}^2$. Die Messungen einer Versuchseinheit können nun in einem Vektor y zusammengefasst werden. Dies ermöglicht die Matrixschreibweise

$$y = X_\xi \beta + e. \quad (2.3)$$

Dabei ist $y = [y_1, \dots, y_q]^T$, $X_\xi = [f(x_1)^T, \dots, f(x_q)^T]^T$ und $e = [e_1, \dots, e_q]^T$. Der Vektor e folgt einer multivariaten Normalverteilung mit Erwartungswert 0. Die Komponenten e_j , $1 \leq j \leq q$, haben jeweils die Varianz σ^2 mit $\sigma^2 > 0$. Zudem sind sie untereinander korreliert, wobei die Korrelation zwischen zwei Komponenten e_j und $e_{j'}$ gleich $\lambda^{|x_j - x_{j'}|}$ ist. Die Korrelationen hängen also von einem Parameter $\lambda \in [0, 1]$ und vom Abstand der zugehörigen Versuchspunkte ab. Diese Modellierung basiert auf der Annahme, dass zwei Fehler eine umso höhere Korrelation aufweisen, je näher ihre zugehörigen Versuchspunkte beieinander liegen. Im Weiteren soll für die Versuchspunkte gelten, dass $x_j \leq x_{j'}$ für $j < j'$. Diese Situation ist ohne Beschränkung der Allgemeinheit durch eine Ordnung der Versuchspunkte konstruierbar. Daraufhin resultiert die Korrelationsmatrix

$$\Sigma(\lambda) = \begin{bmatrix} 1 & \lambda^{x_2 - x_1} & \lambda^{x_3 - x_1} & \dots & \lambda^{x_q - x_1} \\ \lambda^{x_2 - x_1} & 1 & \lambda^{x_3 - x_2} & \dots & \lambda^{x_q - x_2} \\ \lambda^{x_3 - x_1} & \lambda^{x_3 - x_2} & \ddots & & \vdots \\ \vdots & \vdots & & 1 & \lambda^{x_q - x_{q-1}} \\ \lambda^{x_q - x_1} & \lambda^{x_q - x_2} & \dots & \lambda^{x_q - x_{q-1}} & 1 \end{bmatrix}.$$

Die Kovarianzmatrix lautet dann $\sigma^2 \Sigma(\lambda) =: \Sigma(\sigma^2, \lambda)$. In Kurzschreibweise gilt damit schließlich, dass

$$e \sim IN(0, \Sigma(\sigma^2, \lambda)) \quad (2.4)$$

und

$$y \sim IN(X\beta, \Sigma(\sigma^2, \lambda)). \quad (2.5)$$

Nun erfolgt die Übertragung auf die Fälle $n > 1$. Der Vektor der Beobachtungen lautet daraufhin $y = [y_{11}, \dots, y_{1q}, y_{21}, \dots, y_{nq}]^T$, wobei y_{ij} die j -te Beobachtung der i -ten Versuchseinheit bezeichnet. Entsprechend lautet der Fehlervektor $e = [e_{11}, \dots, e_{1q}, e_{21}, \dots, e_{nq}]^T$. Da zum einen die Fehler verschiedener Versuchseinheiten als unabhängig angesehen werden und zum anderen jede Versuchseinheit aus organisatorischen Gründen die gleichen

Versuchspunkte aufweist, lautet die zugehörige Kovarianzmatrix $I_n \otimes \Sigma(\sigma^2, \lambda) =: \Gamma(\sigma^2, \lambda)$. Damit hat die Korrelationsmatrix der Fehler die Form $I_n \otimes \Sigma(\lambda) =: \Gamma(\lambda)$. Schließlich hat die Designmatrix die Gestalt $1_n \otimes X_\xi$. Die Struktur dieser Matrix ist ebenso das Resultat der organisatorischen Gründe.

Innerhalb dieser Arbeit wird überwiegend der Spezialfall betrachtet, dass jede Versuchseinheit $q = 3$ Versuchspunkte aufweist, wobei zwei der drei Punkte durch $x_1 = 0$ und $x_3 = 1$ gegeben sind. Der dritte Punkt liegt zwischen den beiden bereits feststehenden Punkten oder stimmt mit einem der beiden überein. Eine Ausnahme stellt Kapitel 5 dar. Dort wird kurz auf Versuchspläne eingegangen, bei denen die Versuchseinheiten $q \geq 2$ Versuchspunkte aufweisen.

Bezüglich der Modellidentifikation interessieren in den Kapiteln 5, 6 und 8 in erster Linie die Parameter β_0 und β_1 , da diese den deterministischen Teil der Zielgröße y definieren, während die Parameter σ^2 und λ lediglich die Struktur des Fehlervektors festlegen und damit als Störparameter bezeichnet werden können. Lediglich in Kapitel 7 hat die Identifikation des Korrelationsparameters eine ebenso wichtige Bedeutung. Das Ziel ist es daher, die Parameter β_0 , β_1 und λ möglichst gut zu schätzen. Da sinnvolle Schätzer von der Designmatrix abhängen, stellt sich die Frage nach einer geeigneten Wahl der Versuchspunkte. Dies ist die zentrale Problemstellung der vorliegenden Arbeit. Dabei werden im Verlauf der Arbeit zwei verschiedene Konstellationen betrachtet. In den Kapiteln 5, 7 und 8 werden die Parameter, die nicht von primärem Interesse sind, wie in Dette *et al.* (2008) als bekannt vorausgesetzt. In den Kapiteln 5 und 8 sind dies die beiden Störparameter, in Kapitel 7 ist dies nur der Parameter für die Fehlervarianz. Dagegen sind die Störparameter in Kapitel 6 unbekannt und müssen ebenfalls geschätzt werden. Weiter unterscheiden sich die Inhalte in den Kapiteln 5 bis 8 im Detail wie folgt. Auf Grund der Tatsache, dass bei autoregressiven Störgrößen für den Spezialfall mit drei Versuchspunkten lediglich lokal D-optimale Versuchspläne existieren, werden in Kapitel 5 D-Effizienzen beziehungsweise Schranken für die D-Effizienzen bestimmt. Außerdem werden Schranken für relative D-Effizienzen angegeben, bei denen jeweils zwei Versuchspläne einem direkten Vergleich unterzogen werden. In Kapitel 6 wird eine Simulationsstudie durchgeführt. Dabei stimmt die Problemstellung bis auf die Annahme bezüglich der Störparameter mit der aus Kapitel 5 überein. Das siebte Kapitel unterscheidet sich von Kapitel 5 nur dadurch, dass dort der Korrelationsparameter ebenfalls für die Designwahl von Bedeutung ist. Schließlich werden im achten Kapitel abweichend zu den Kapiteln 5 bis 7 zwei verallgemeinerte Abhängigkeitsstrukturen des Fehlervektors betrachtet. Das Vorgehen in den Kapiteln 7

und 8 sieht so aus, dass verschiedene Versuchspläne wie unter anderem auch in Kapitel 5 mittels ihrer relativen D-Effizienzen miteinander verglichen werden.

Zur Formalisierung werden nun noch abschließend zwei Symbole definiert. Die Versuchspunkte eines Versuchsplans können aus einer vorgegebenen Menge gewählt werden, aus dem sogenannten Versuchsraum \mathcal{X} . Die Menge der möglichen Versuchspläne wird durch die Menge Ξ zusammengefasst.

3. Parameterschätzung in linearen Regressionsmodellen

In diesem Kapitel werden zwei Schätzer vorgestellt, mit denen die Parameter aus dem linearen Regressionsmodell des Kapitels 2 geschätzt werden können. Anhand der sogenannten Cramér-Rao-Schranke erfolgt eine Motivation des häufigen Einsatzes der beiden Schätzer. Der Satz von Cramér-Rao wird in Abschnitt 3.1 formuliert. Daraufhin werden in den Abschnitten 3.2 und 3.3 die beiden Schätzer vorgestellt. Abschließend wird in Abschnitt 3.4 der Frage nachgegangen, wie die Regressionsparameter sinnvollerweise unter $\lambda = 1$ geschätzt werden können.

3.1. Eine Schranke zur Beurteilung von Schätzern

Neben dem Wunsch der Erwartungstreue ist es gleichzeitig erstrebenswert, dass Schätzer eine möglichst geringe Streuung aufweisen. Es bietet sich also an, Schätzer anhand ihrer Varianzen beziehungsweise Kovarianzmatrizen miteinander zu vergleichen.

Definition 3.1. Sei θ ein Parametervektor.

Falls θ eindimensional ist, so heißt ein erwartungstreuer Schätzer gleichmäßig bester erwartungstreuer Schätzer, kurz $\hat{\theta}_{UMVU}$ (uniformly minimum variance unbiased), falls für jeden erwartungstreuen Schätzer $\hat{\theta}$ gilt, dass

$$\text{Var}(\hat{\theta}_{UMVU}) \leq \text{Var}(\hat{\theta}).$$

Falls θ mehrdimensional ist, so heißt ein erwartungstreuer Schätzer gleichmäßig bester erwartungstreuer Schätzer, kurz $\hat{\theta}_{UMVU}$, falls für jeden erwartungstreuen Schätzer $\hat{\theta}$ gilt, dass

$$\text{Cov}(\hat{\theta}_{UMVU}) \leq \text{Cov}(\hat{\theta}),$$

wobei die Ungleichung im Sinne der Löwner-Ordnung gilt.

Für den Nachweis, dass ein Schätzer der gleichmäßig beste erwartungstreue Schätzer ist, kann die sogenannte Cramér-Rao-Schranke hilfreich sein, da die Varianz oder Kovarianzmatrix eines Kandidaten nur mit einer Schranke zu vergleichen ist, siehe Satz 3.3. Ein Bestandteil von Satz 3.3 ist die sogenannte Fisher-Informationsmatrix, die zuvor definiert wird.

Definition 3.2. Sei $\theta \in \Theta$ ein k -dimensionaler Parametervektor. Der Zufallsvektor Y habe Dichte $f(Y|\theta)$. Dann ist der Eintrag $[i, j]$ der Fisher-Informationsmatrix $\mathcal{I}(\theta)$ definiert durch

$$\mathcal{I}(\theta)_{i,j} = E\left[\left(\frac{\partial}{\partial\theta_i}\ln(f(Y|\theta))\right)\left(\frac{\partial}{\partial\theta_j}\ln(f(Y|\theta))\right)\right].$$

Satz 3.3. Unter den Regularitätsbedingungen

- (1) die Menge $\Theta \subset \mathbb{R}^k$ ist offen,
- (2) die Menge $\{y : f(y|\theta) > 0\} =: A$ ist unabhängig von θ ,
- (3) für beliebiges $y \in A$ und $\theta \in \Theta$ existiert $\frac{\partial}{\partial\theta_i}f(y|\theta)$ mit $\frac{\partial}{\partial\theta_i}f(y|\theta) < \infty$, $1 \leq i \leq k$,
- (4) es gilt $E\left(\frac{\partial}{\partial\theta_i}\ln(f(Y|\theta))\right) = 0$, $1 \leq i \leq k$,

gilt für erwartungstreue Schätzer $\hat{\theta}$ im Sinne der Löwner-Ordnung, dass

$$\text{Cov}(\hat{\theta}) \geq \mathcal{I}(\theta)^{-1}.$$

Beweis. Siehe Härdle und Simar (2012), S. 189. □

Unter der Bedingung

$$\frac{\partial^2}{\partial\theta_i\partial\theta_j}f(y|\theta) < \infty, 1 \leq i, j \leq k,$$

können die Einträge der Fisher-Informationsmatrix alternativ mittels

$$\mathcal{I}(\theta)_{i,j} = -E\left[\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln(f(Y|\theta))\right)\right]$$

bestimmt werden. Dieser Weg vereinfacht in vielen Situationen die Angabe der Einträge.

Eine Eigenschaft der Fisher-Informationsmatrix lautet, dass für unabhängig identisch verteilte Zufallsvektoren Y_1, \dots, Y_n die Fisher-Informationsmatrix gerade der Summe der einzelnen Informationsmatrizen entspricht. Es gilt also, dass

$$\mathcal{I}_n(\theta) = n\mathcal{I}(\theta), \tag{3.1}$$

siehe Härdle und Simar (2012), S. 189.

Es kann der Fall eintreten, dass die Schätzer $\hat{\theta}_n$ aus einer Familie von Schätzern zwar weder erwartungstreu sind noch ihre Varianzen beziehungsweise Kovarianzmatrizen die untere Schranke annehmen, aber die Folge ihrer Verteilungen gegen eine Normalverteilung konvergiert mit

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} IN(0, \mathcal{I}(\theta)^{-1}).$$

In dieser Situation heißt eine Schätzerfamilie asymptotisch effizient.

In Kapitel 2 wurden verschiedene Konstellationen vorgestellt, bei denen unterschiedlich viele Parameter zu schätzen sind. Nun wird zunächst die Informationsmatrix für den Fall, dass die vier Parameter β_0 , β_1 , σ^2 und λ zu schätzen sind, bestimmt. In dieser Situation lautet die Dichte des vorliegenden Zufallsvektors

$$f(y|\beta_0, \beta_1, \sigma^2, \lambda) = \frac{1}{(2\pi)^{\frac{qn}{2}} (\det(\Gamma(\sigma^2, \lambda)))^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}(y - (1_n \otimes X)\beta)^T \Gamma^{-1}(\sigma^2, \lambda)(y - (1_n \otimes X)\beta)\right). \quad (3.2)$$

Die logarithmierte Dichte lautet daraufhin

$$\ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) = -\frac{qn}{2} \ln(2\pi) - \frac{1}{2} \ln(\det(\Gamma(\sigma^2, \lambda))) - \frac{1}{2}(y - (1_n \otimes X)\beta)^T \Gamma^{-1}(\sigma^2, \lambda)(y - (1_n \otimes X)\beta). \quad (3.3)$$

Es sei kurz angemerkt, dass die Dichten in (3.2) und (3.3) nur existieren, falls die Kovarianzmatrix des Fehlervektors invertierbar ist. Da insbesondere für $\lambda = 1$ die Invertierbarkeit nicht gegeben ist, wird später in Abschnitt 3.4 dieser Fall für sich untersucht.

Die aus (3.3) resultierende Fisher-Informationsmatrix wird jetzt in Behauptung 3.4 für den Fall $n = 1$ formuliert.

Behauptung 3.4. *Gegeben sei die Dichte gemäß (3.2). Dann weist die Fisher-Informationsmatrix für $n = 1$ folgende Einträge auf*

$$\mathcal{I}([\beta_0, \beta_1, \sigma^2, \lambda]^T) = \begin{pmatrix} X^T \Sigma^{-1}(\sigma^2, \lambda) X & [0, 0]^T & [0, 0]^T \\ [0, 0] & \frac{q}{2(\sigma^2)^2} & \frac{1}{2\sigma^2} \text{tr}(\Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda)) \\ [0, 0] & \frac{1}{2\sigma^2} \text{tr}(\Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda)) & \frac{1}{2} \text{tr}(\Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda) \Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda)) \end{pmatrix}. \quad (3.4)$$

Dabei ist $\dot{\Sigma}(\lambda) = \frac{\partial}{\partial \lambda} \Sigma(\lambda)$.

Beweis. Die ersten Ableitungen der Dichte lauten

$$\begin{aligned}\frac{\partial}{\partial \beta_0} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= (y - X\beta)^T \Sigma^{-1}(\sigma^2, \lambda) [1, \dots, 1]^T \\ \frac{\partial}{\partial \beta_1} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= (y - X\beta)^T \Sigma^{-1}(\sigma^2, \lambda) [x_1, \dots, x_q]^T \\ \frac{\partial}{\partial \sigma^2} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= -\frac{q}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} (y - X\beta)^T \Sigma^{-1}(\lambda) (y - X\beta) \\ \frac{\partial}{\partial \lambda} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= \frac{1}{2\sigma^2} (y - X\beta)^T \Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda) \Sigma^{-1}(\lambda) (y - X\beta) \\ &\quad - \frac{1}{2} \text{tr}(\Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda)).\end{aligned}$$

Für die Ableitung nach λ wird zum einen ausgenutzt, dass

$$\frac{\partial}{\partial \lambda} \Sigma^{-1}(\lambda) = -\Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda) \Sigma^{-1}(\lambda),$$

siehe Magnus und Neudecker (1999), S. 151. Darüber hinaus kann auf Grund des Vollrangs von $\Sigma(\lambda)$ ausgenutzt werden, dass

$$\frac{\partial}{\partial \lambda} \det(\Sigma(\lambda)) = \det(\Sigma(\lambda)) \text{tr}(\Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda)),$$

siehe Magnus und Neudecker (1999), S. 149.

Die zweiten Ableitungen der logarithmierten Dichte lauten daraufhin

$$\begin{aligned}\frac{\partial^2}{\partial \beta_0 \partial \beta_0} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= -[1, \dots, 1] \Sigma^{-1}(\sigma^2, \lambda) [1, \dots, 1]^T \\ \frac{\partial^2}{\partial \beta_0 \partial \beta_1} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= -[x_1, \dots, x_q] \Sigma^{-1}(\sigma^2, \lambda) [1, \dots, 1]^T \\ \frac{\partial^2}{\partial \beta_0 \partial \sigma^2} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= -\frac{1}{(\sigma^2)^2} (y - X\beta)^T \Sigma^{-1}(\lambda) [1, \dots, 1]^T \\ \frac{\partial^2}{\partial \beta_0 \partial \lambda} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= -\frac{1}{\sigma^2} (y - X\beta)^T \Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda) \Sigma^{-1}(\lambda) [1, \dots, 1]^T \\ \frac{\partial^2}{\partial \beta_1 \partial \beta_1} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= -\frac{1}{\sigma^2} [x_1, \dots, x_q] \Sigma^{-1}(\lambda) [x_1, \dots, x_q]^T \\ \frac{\partial^2}{\partial \beta_1 \partial \sigma^2} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= -\frac{1}{(\sigma^2)^2} (y - X\beta)^T \Sigma^{-1}(\lambda) [x_1, \dots, x_q]^T \\ \frac{\partial^2}{\partial \beta_1 \partial \lambda} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= -\frac{1}{\sigma^2} (y - X\beta)^T \Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda) \Sigma^{-1}(\lambda) [x_1, \dots, x_q]^T \\ \frac{\partial^2}{\partial \sigma^2 \partial \sigma^2} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= \frac{q}{2(\sigma^2)^2} - \frac{1}{(\sigma^2)^3} (y - X\beta)^T \Sigma^{-1}(\lambda) (y - X\beta)\end{aligned}$$

$$\begin{aligned} \frac{\partial^2}{\partial \sigma^2 \partial \lambda} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= -\frac{1}{2(\sigma^2)^2} (y - X\beta)^T \Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda) \Sigma^{-1}(\lambda) (y - X\beta) \\ \frac{\partial^2}{\partial \lambda \partial \lambda} \ln(f(y|\beta_0, \beta_1, \sigma^2, \lambda)) &= -\frac{1}{\sigma^2} (y - X\beta)^T \Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda) \Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda) \Sigma^{-1}(\lambda) (y - X\beta) \\ &\quad + \frac{1}{2\sigma^2} (y - X\beta)^T \Sigma^{-1}(\lambda) \ddot{\Sigma}(\lambda) \Sigma^{-1}(\lambda) (y - X\beta) \\ &\quad - \frac{1}{2} \text{tr}(\Sigma^{-1}(\lambda) \ddot{\Sigma}(\lambda)) + \frac{1}{2} \text{tr}(\Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda) \Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda)) \end{aligned}$$

mit $\frac{\partial}{\partial \lambda} \Sigma(\lambda) =: \dot{\Sigma}(\lambda)$ und $\frac{\partial^2}{\partial \lambda^2} \Sigma(\lambda) =: \ddot{\Sigma}(\lambda)$.

Für die Erwartungswerte der zweiten Ableitungen wird jetzt unter anderem ausgenutzt, dass $E(y - X\beta) = 0$. Außerdem wird folgende Regel benötigt. Allgemein gilt für den Erwartungswert der quadratischen Form $z^T A z$, dass $E(z^T A z) = \text{tr}(A \text{Cov}(z)) + (E(z))^T A E(z)$, wenn die Matrix A symmetrisch und deterministisch ist, siehe Schmidt und Trenkler (2006), S. 177. Die Umkehrung der Vorzeichen der Erwartungswerte liefert abschließend die Einträge der Fisher-Informationsmatrix. \square

Die Informationsmatrix hat die Besonderheit, dass sie eine diagonale Blockstruktur aufweist. Die zugehörige Inverse lässt sich daher besonders einfach berechnen. Sie lautet für $n = 1$

$$\mathcal{I}([\beta_0, \beta_1, \sigma^2, \lambda]^T)^{-1} = \begin{pmatrix} (X^T \Sigma^{-1}(\sigma^2, \lambda) X)^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix} \quad (3.5)$$

mit

$$B = \begin{pmatrix} \frac{q}{2(\sigma^2)^2} & \frac{1}{2\sigma^2} \text{tr}(\Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda)) \\ \frac{1}{2\sigma^2} \text{tr}(\Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda)) & \frac{1}{2} \text{tr}(\Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda) \Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda)) \end{pmatrix}. \quad (3.6)$$

Nun wird die Informationsmatrix sowie die zugehörige Inverse für den Fall bestimmt, bei dem die Störparameter als bekannt angenommen werden. Die Informationsmatrix für den Parametervektor $[\beta_0, \beta_1]^T$ ist per Definition eine Teilmatrix von $\mathcal{I}([\beta_0, \beta_1, \sigma^2, \lambda]^T)$. Sie lautet

$$\mathcal{I}([\beta_0, \beta_1]^T) = X^T \Sigma^{-1}(\sigma^2, \lambda) X. \quad (3.7)$$

Die Inverse hat damit die Gestalt

$$\mathcal{I}([\beta_0, \beta_1]^T)^{-1} = (X^T \Sigma^{-1}(\sigma^2, \lambda) X)^{-1}. \quad (3.8)$$

Diese stimmt somit mit dem ersten Block der Matrix $\mathcal{I}([\beta_0, \beta_1, \sigma^2, \lambda]^T)^{-1}$ überein, was die Folge der speziellen Blockstruktur von $\mathcal{I}([\beta_0, \beta_1, \sigma^2, \lambda]^T)$ ist.

Es stellt sich nun die Frage, ob Satz 3.3 für das Modell aus Kapitel 2 angewendet werden kann. Für den Fall, dass der Parametervektor $[\beta_0, \beta_1, \sigma^2, \lambda]^T$ zu schätzen ist, so ist die

erste Regularitätsbedingung nur erfüllt, wenn der Parameterraum für λ auf das Intervall $(0, 1)$ eingeschränkt wird. Die vierte Bedingung kann leicht mit Hilfe der Rechenregeln aus dem Beweis von Behauptung 3.4 nachgewiesen werden. Dass die weiteren Bedingungen gelten, ist offensichtlich.

Für den Fall, dass der Parametervektor $[\beta_0, \beta_1]^T$ zu schätzen ist, ist bezüglich der ersten Regularitätsbedingung keine Einschränkung vorzunehmen, da der Parameter λ als bekannt vorausgesetzt wird. Die weiteren Bedingungen können analog zum ersten Fall nachgewiesen werden.

3.2. Verallgemeinerter Kleinste-Quadrate-Schätzer

In diesem Abschnitt wird der verallgemeinerte Kleinste-Quadrate-Schätzer vorgestellt, sowie dessen Eigenschaften beschrieben.

Definition 3.5. *Unter den Annahmen*

1. $E(e) = 0$,
2. $Cov(e) = \sigma^2 W$, wobei W eine positiv definite bekannte Matrix ist,
3. die Design-Matrix X hat vollen Spaltenrang, also $rg(X) = p$,

heißt das Modell

$$y = X\beta + e \tag{3.9}$$

allgemeines lineares Regressionsmodell.

Satz 3.6. *Im allgemeinen linearen Regressionsmodell ist der verallgemeinerte Kleinste-Quadrate-Schätzer*

$$\hat{\beta}_{GLS} = (X^T W^{-1} X)^{-1} X^T W^{-1} y$$

der beste lineare unverzerrte Schätzer mit

$$Cov(\hat{\beta}_{GLS}) = \sigma^2 (X^T W^{-1} X)^{-1}.$$

Gilt zusätzlich, dass $e \sim IN(0, \sigma^2 W)$, so ist $\hat{\beta}_{GLS}$ sogar der gleichmäßig beste unverzerrte Schätzer.

Beweis. Der Beweis, dass $\hat{\beta}_{GLS}$ der beste lineare unverzerrte Schätzer ist, kann beispielsweise Rao *et al.* (2008), S. 152, entnommen werden.

Unter Normalverteilung liefert (3.8) eine untere Schranke für die Kovarianzmatrizen aller erwartungstreuen Schätzer. Da die Kovarianzmatrix von $\hat{\beta}_{GLS}$ diese Schranke annimmt, ist $\hat{\beta}_{GLS}$ der gleichmäßig beste unverzerrte Schätzer. \square

Im Modell aus Kapitel 2 kann der Parametervektor β für den Fall, dass die Störparameter σ^2 und λ bekannt sind, mittels der verallgemeinerten Kleinste-Quadrate-Methode geschätzt werden. Dabei ist zu beachten, dass die Korrelationsmatrix $\Sigma(\lambda)$ invertierbar sein muss. Da $\Sigma(\lambda)$ insbesondere für $\lambda = 1$ nicht invertierbar ist, wird dieser Fall in Abschnitt 3.4 gesondert untersucht.

Zusammenfassend kann festgehalten werden, dass der verallgemeinerte Kleinste-Quadrate-Schätzer wünschenswerte Eigenschaften besitzt. Dabei ist jedoch vorauszusetzen, dass die Störparameter σ^2 und λ bekannt sind. Für den Fall, dass diese unbekannt sind, kann der sogenannte Maximum-Likelihood-Schätzer herangezogen werden, welcher im folgenden Abschnitt vorgestellt wird.

3.3. Maximum-Likelihood-Schätzer

In vielen Situationen können Parameter anhand der Maximum-Likelihood-Methode geschätzt werden. Dabei wird angenommen, dass vorliegende Stichprobenwerte die Realisierungen von Zufallsvariablen sind, wobei die Verteilungen der Zufallsvariablen von einem oder mehreren unbekanntem Parametern abhängen. Werden die Stichprobenwerte in die Dichte eingesetzt und diese Funktion dann als Funktion des Parametervektors θ aufgefasst, so liegt die sogenannte Likelihood-Funktion $L_y(\theta)$ vor. Diejenigen Parameterwerte, die diese Funktion maximieren, liefern dann den Maximum-Likelihood-Schätzer $\hat{\theta}_{ML}$ für den Parameter bzw. Parametervektor θ . Die Motivation für dieses Vorgehen ist die, dass die Maximalstelle und damit die entsprechende Dichte eine besonders plausible Stichprobe liefert, da die Wahrscheinlichkeit für die Realisierung vergleichsweise groß ist.

Der Maximum-Likelihood-Schätzer $\hat{\theta}_{ML}$ hat die Eigenschaft, dass er asymptotisch normalverteilt ist. Es gilt nämlich

$$\sqrt{n}(\hat{\theta}_{ML} - \theta) \xrightarrow{d} IN(0, \mathcal{I}(\theta)^{-1}),$$

siehe Lehmann (1983), S. 430. Da die Inverse der Fisher-Informationsmatrix aber gleichzeitig auch die Schranke der Cramér-Rao-Ungleichung darstellt, ist der Maximum-Likelihood-Schätzer asymptotisch effizient und somit eine beliebte Wahl.

Für das Modell, das in dieser Arbeit Gegenstand der Untersuchungen ist, hat die Maximum-Likelihood-Methode gegenüber dem verallgemeinerten Kleinste-Quadrate-Schätzer den Vorteil, dass die Störparameter nicht als bekannt vorausgesetzt werden müssen. Es liegt nun folgende Situation vor. In Abschnitt 3.1 wurde gezeigt, dass die Inverse der Fisher-Informationsmatrix $\mathcal{I}([\beta_0, \beta_1]^T)$ eine Teilmatrix der Inversen von $\mathcal{I}([\beta_0, \beta_1, \sigma^2, \lambda]^T)$ ist. Es liegt also die Besonderheit vor, dass die Verallgemeinerung des Modells durch die unbekanntes Störparameter σ^2 und λ keine Auswirkung für die Maximum-Likelihood-Schätzer der Parameter β_0 und β_1 bezüglich ihrer asymptotischen Kovarianzmatrix hat.

Die Inverse der Fisher-Informationsmatrix (3.4) liefert nun die asymptotische Kovarianzmatrix des Zufallsvektors $\sqrt{n}\hat{\theta}_{ML}$. Somit gilt für große Stichprobenumfänge, dass

$$\text{Cov}(\hat{\theta}_{ML}) \approx \frac{1}{n} \mathcal{I}([\beta_0, \beta_1, \sigma^2, \lambda]^T)^{-1}.$$

Selbiges gilt auch für jeden anderen asymptotisch effizienten Schätzer. Für ausgewählte Stichprobenumfänge wird ein weiterer asymptotisch effizienter Schätzer in Kapitel 6 innerhalb einer Simulationsstudie herangezogen.

3.4. Erwartungstreue und Kovarianzmatrix linearer Schätzer unter $\lambda = 1$

In diesem Abschnitt liegt der Spezialfall $\lambda = 1$ vor. Dabei wird der Fall betrachtet, dass $n = 1$ ist. Damit liegt nur eine Versuchseinheit vor. Die folgenden Behauptungen zeigen in Abhängigkeit von der Größe q , die die Anzahl der Versuche pro Versuchseinheit beschreibt, auf, unter welchen Voraussetzungen lineare erwartungstreue Schätzer für die Regressionsparameter existieren. Daraufhin wird die Gestalt der Kovarianzmatrix dieser Schätzer bestimmt.

Behauptung 3.7. *Gegeben sei das Modell aus Kapitel 2 mit $q = 1$ und Versuchspunkt x_1 . Dann gibt es keinen linearen erwartungstreuen Schätzer für β .*

Beweis. Damit ein linearer Schätzer Uy erwartungstreu für β ist, muss gelten, dass

$$E(Uy) = UE(y) = UX\beta \stackrel{!}{=} \beta.$$

Dies bedeutet, dass

$$UX = I \tag{3.10}$$

gelten muss. Mit

$$U = \begin{bmatrix} u_{11} \\ u_{21} \end{bmatrix} \quad \text{und} \quad X = \begin{bmatrix} 1 & x_1 \end{bmatrix}$$

müssen jetzt wegen (3.10) die folgenden Gleichungen erfüllt sein:

$$\begin{aligned} u_{11} &= 1 \\ u_{11}x_1 &= 0 \\ u_{21} &= 0 \\ u_{21}x_1 &= 1. \end{aligned}$$

Die letzten beiden Gleichungen können nicht simultan erfüllt sein. Somit existiert für $q = 1$ kein linearer erwartungstreuer Schätzer für β . \square

Nun wird in Behauptung 3.8 die Situation betrachtet, dass der Versuchsplan ξ sich aus $q \geq 2$ Versuchspunkten zusammensetzt.

Behauptung 3.8. *Gegeben sei das Modell aus Kapitel 2 mit $q \geq 2$. Gegeben sei ein beliebiger Versuchsplan ξ , wobei die x_i nicht allesamt identisch sind. Sei $\lambda = 1$. Dann gilt für jeden linearen erwartungstreuen Schätzer $\tilde{\beta}_{\lambda=1} = Uy$, dass*

$$\text{Cov}(\tilde{\beta}_{\lambda=1}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \quad (3.11)$$

Beweis. Zunächst wird untersucht, wie in der vorliegenden Situation ein linearer Schätzer $\tilde{\beta} = [\tilde{\beta}_0, \tilde{\beta}_1]^T = Uy$ auszusehen hat, damit dieser erwartungstreu für β ist. Es soll also gelten, dass

$$E(Uy) = UE(y) = UX\beta \stackrel{!}{=} \beta.$$

Dies bedeutet wie in Behauptung 3.7, dass

$$UX = I \quad (3.12)$$

gelten muss. Mit

$$U = \begin{bmatrix} u_1^T \\ u_2^T \end{bmatrix} = \begin{bmatrix} u_{11} & \dots & u_{1q} \\ u_{21} & \dots & u_{2q} \end{bmatrix} \quad \text{und} \quad X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_q \end{bmatrix}$$

müssen jetzt wegen (3.12) für die Erwartungstreue von Uy die folgenden Gleichungen erfüllt sein:

$$\sum_{i=1}^q u_{1i} = 1 \quad (3.13)$$

$$\sum_{i=1}^q u_{1i}x_i = 0 \quad (3.14)$$

$$\sum_{i=1}^q u_{2i} = 0 \quad (3.15)$$

$$\sum_{i=1}^q u_{2i}x_i = 1. \quad (3.16)$$

Wenn nun der Fall vorliegen sollte, dass $x_1 = \dots = x_q$, so können die Gleichungen (3.16) und (3.15) nicht simultan erfüllt sein. Dies erklärt die Voraussetzung in der Behauptung, dass die x_i nicht allesamt identisch sein dürfen. Falls sich andererseits mindestens zwei Versuchspunkte voneinander unterscheiden, so hat die Designmatrix den Rang 2. In diesem Fall ist die Matrix $X^T X$ invertierbar, was dazu führt, dass die Bedingung (3.12) mit der Matrix $U = (X^T X)^{-1} X^T$ erfüllt ist.

Nun wird gezeigt, wie sich die Restriktionen (3.13) und (3.15) auf die Kovarianzmatrix des Schätzvektors auswirken. Es gilt

$$\begin{aligned} \text{Var}(\tilde{\beta}_0) &= \text{Var}(u_1^T y) \\ &= u_1^T \text{Cov}(y) u_1 \\ &= u_1^T \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & & 1 \\ \vdots & & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix} u_1 \\ &= u_1^T \mathbf{1}_q \mathbf{1}_q^T u_1 \\ &= \sum_{i=1}^q u_{1i} \sum_{i=1}^q u_{1i} \\ &= 1, \end{aligned} \quad (3.17)$$

$$\begin{aligned}
\text{Var}(\tilde{\beta}_1) &= \text{Var}(u_2^T y) \\
&= u_2^T \text{Cov}(y) u_2 \\
&= u_2^T \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & & 1 \\ \vdots & & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix} u_2 \\
&= u_2^T \mathbf{1}_q \mathbf{1}_q^T u_2 \\
&= \sum_{i=1}^q u_{2i} \sum_{i=1}^q u_{2i} \\
&= 0.
\end{aligned} \tag{3.18}$$

Weiter gilt auf Grund der Cauchy-Schwarz-Ungleichung, dass

$$\text{Var}(\tilde{\beta}_0) \text{Var}(\tilde{\beta}_1) \geq \text{Kov}(\tilde{\beta}_0, \tilde{\beta}_1)^2.$$

Äquivalent dazu ist

$$\text{Var}(\tilde{\beta}_0) \text{Var}(\tilde{\beta}_1) - \text{Kov}(\tilde{\beta}_0, \tilde{\beta}_1)^2 \geq 0. \tag{3.19}$$

Wenn nun eine der beiden Varianzen aus (3.19) beschränkt und die Zweite Null ist, muss wegen (3.19) zwangsläufig die Kovarianz von $\tilde{\beta}_0$ und $\tilde{\beta}_1$ ebenfalls Null sein. Mit $\text{Var}(\tilde{\beta}_0) = 1$ und $\text{Var}(\tilde{\beta}_1) = 0$ folgt, dass

$$\text{Kov}(\tilde{\beta}_0, \tilde{\beta}_1) = 0. \tag{3.20}$$

□

Behauptung 3.8 zeigt für $q \geq 2$, dass für den Spezialfall $\lambda = 1$ unabhängig von der Anzahl der Versuchspunkte jeder lineare erwartungstreue Schätzer für β die gleiche Kovarianzmatrix aufweist wie ein linearer erwartungstreuer Schätzer in der Situation, dass der Versuchsplan sich lediglich aus zwei Versuchspunkten zusammensetzt. Das hat zum Beispiel die Konsequenz, dass im Vergleich zum Versuchsplan mit $x_1 = 0$ und $x_2 = 1$ ein zusätzlicher Versuchspunkt zu keinem Informationsgewinn führt. Dieses Resultat wird im Weiteren mehrfach Anwendung finden.

Falls $q = 2$ ist, so liegt abweichend zu $q \geq 3$ außerdem die Besonderheit vor, dass die Wahl der Matrix U eindeutig ist, siehe Behauptung 3.9 und Behauptung 3.10.

Behauptung 3.9. *Gegeben sei das Modell aus Kapitel 2 mit $q = 2$. Gegeben seien die Versuchspunkte x_1 und x_2 , wobei $x_1 \neq x_2$ ist. Sei $\lambda = 1$. Dann gilt: Der Schätzvektor*

$$\hat{\beta}_{\lambda=1} = X^{-1}y = \frac{1}{x_2 - x_1} \begin{bmatrix} x_2 & -x_1 \\ -1 & 1 \end{bmatrix} y$$

ist der einzige beste lineare erwartungstreue Schätzer für β .

Beweis. Offensichtlich ist $\hat{\beta}_{\lambda=1}$ linear. Wegen

$$\begin{aligned} E(\hat{\beta}_{\lambda=1}) &= E(X^{-1}y) \\ &= X^{-1}X\beta = \beta \end{aligned}$$

ist $\hat{\beta}_{\lambda=1}$ erwartungstreu für β .

Sei nun $\tilde{\beta}_{\lambda=1} = Uy$ ein beliebiger erwartungstreuer Schätzer für β . Daher gilt:

$$\begin{aligned} E(\tilde{\beta}_{\lambda=1}) &= E(Uy) \\ &= UE(y) \\ &= UX\beta. \end{aligned}$$

Damit $\tilde{\beta}_{\lambda=1}$ erwartungstreu für β ist, muss gelten, dass $UX = I$, wobei I die Einheitsmatrix ist. Da X eine quadratische Matrix ist, bedeutet dies, dass U durch

$$X^{-1} = \frac{1}{x_2 - x_1} \begin{bmatrix} x_2 & -x_1 \\ -1 & 1 \end{bmatrix} =: U \quad (3.21)$$

eindeutig festgelegt ist. Dies liefert aber gerade den Schätzvektor $\hat{\beta}_{\lambda=1}$. Insgesamt heißt dies nun, dass es für $q = 2$ nur einen linearen Schätzer gibt, der erwartungstreu ist. Damit ist dieser auch gleichzeitig der Beste. \square

Behauptung 3.10. *In der Situation aus Behauptung 3.9 sei nun $q \geq 3$. Die Versuchspunkte x_i seien nicht allesamt identisch. Sei Uy erwartungstreu für β . Dann gilt: Es existieren unendlich viele Möglichkeiten für die Wahl von U . Allerdings gilt, dass die verschiedenen Möglichkeiten identische Schätzwerte liefern.*

Beweis. Ein erwartungstreuer Schätzer erfüllt die Gleichungen (3.13) bis (3.16). Es ist leicht zu sehen, dass die Gleichungen linear unabhängig sind. Darüber hinaus sind die Gleichungen unter der Voraussetzung, dass die x_i nicht allesamt identisch sind, simultan lösbar. Schließlich ist für $q \geq 3$ die Anzahl der Unbekannten stets größer als die Anzahl

der Gleichungen. Somit gibt es für den Vektor $[u_{11}, \dots, u_{1q}, u_{21}, \dots, u_{2q}]^T$ unendlich viele Lösungen. Damit existieren unendlich viele Möglichkeiten für die Wahl von U .

Es bleibt nur noch zu zeigen, dass die Schätzwerte von der Wahl der Matrix U unabhängig sind. Sei $\tilde{y} = X\beta + \tilde{e}$ eine beliebige Realisation des Zufallsvektors y . Dabei ist $\tilde{e} = [\tilde{e}_1, \dots, \tilde{e}_q]^T$ die Realisation des Zufallsvektors e . Wegen $\lambda = 1$ sind die Einträge von \tilde{e} mit Wahrscheinlichkeit 1 allesamt identisch. Es gilt also, dass $\tilde{e}_1 = \dots = \tilde{e}_q$. Daraufhin lauten die Schätzwerte

$$\begin{aligned} U\tilde{y} &= U(X\beta + \tilde{e}) \\ &= UX\beta + U\tilde{e} \\ &= \beta + \begin{bmatrix} \tilde{e}_1 \sum_{i=1}^q u_{1i} \\ \tilde{e}_1 \sum_{i=1}^q u_{2i} \end{bmatrix} && , \text{ wegen (3.12) und } \tilde{e}_1 = \dots = \tilde{e}_q, \\ &= \beta + \begin{bmatrix} \tilde{e}_1 \\ 0 \end{bmatrix} && , \text{ wegen (3.13) und (3.15).} \end{aligned}$$

Dies bedeutet, dass die Schätzwerte des Schätzers Uy lediglich von der Realisation des Fehlervektors abhängen aber nicht von der Wahl der Matrix U . \square

4. Das D-Kriterium zur Beurteilung von Versuchsplänen

In vielen Anwendungen sind die Daten gegeben, so dass die Wahl des Schätzers die zentrale Entscheidung darstellt. In manchen Situationen ist es jedoch so, dass die Daten erst noch erhoben werden müssen. Beispielsweise für das Modell aus Kapitel 2 hat dies die folgende Konsequenz. Zunächst gilt, dass die Varianzen und Kovarianzen innerhalb der Kovarianzmatrix des Kleinste-Quadrate-Schätzers von den Beobachtungen und diese wiederum von den Versuchspunkten abhängen. Daher könnte in Anlehnung an die Definition 3.1 das Ziel formuliert werden, dass innerhalb einer Menge von möglichen Versuchspunkten die Einträge der Kovarianzmatrix im Sinne der Löwner-Ordnung zu optimieren sind. Dies ist im Allgemeinen nicht möglich, wie folgendes Beispiel zeigt.

Beispiel 4.1. *Gegeben seien unkorrelierte Fehler mit Varianz $\sigma^2 = 1$ und zwei Versuchspläne ξ_1 und ξ_2 mit jeweils $q = 4$ Versuchspunkten. Der Versuchsraum \mathcal{X} sei das Intervall $[0, 1]$. Der Plan ξ_1 besitze die Versuchspunkte $0, 0, 1$ und 1 . Die Kovarianzmatrix des Kleinste-Quadrate-Schätzers für $[\beta_0, \beta_1]^T$ lautet daraufhin*

$$\begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1 \end{bmatrix}. \quad (4.1)$$

Somit gilt, dass $\text{Var}(\hat{\beta}_0) = 1/2$ und $\text{Var}(\hat{\beta}_1) = 1$. Es kann gezeigt werden, dass nur dieser Versuchsplan die Varianz des Schätzers für β_1 minimiert. Der Plan ξ_2 dagegen bestehe aus den Punkten $0, 0, 0$ und 1 . Die Kovarianzmatrix des Kleinste-Quadrate-Schätzers für $[\beta_0, \beta_1]^T$ lautet in diesem Fall

$$\begin{bmatrix} 1/3 & -1/3 \\ -1/3 & 4/3 \end{bmatrix}. \quad (4.2)$$

Es ist zu erkennen, dass hier die Varianz des Schätzers für β_1 vergleichsweise größer ist. Die Varianz des Schätzers für β_0 ist jedoch im zweiten Fall kleiner. Somit kann festgehalten werden, dass es keinen Plan gibt, der simultan alle Varianzen minimiert.

Einen Ausweg liefern jedoch sogenannte Optimalitätskriterien, mit denen verschiedene Versuchspläne anhand einer univariaten Kennzahl miteinander verglichen werden können.

Eine Kennzahl, mittels der ein Versuchsplan bewertet werden kann, ist das sogenannte D-Kriterium. Es lässt sich aus der Informationsmatrix ableiten. Die Definition des D-Kriteriums lautet in allgemeiner Form

$$D(\xi) = (\det(\mathcal{I}_\xi(\theta)^{-1}))^{\frac{1}{k}}, \quad (4.3)$$

wobei k der Rang der Informationsmatrix ist. Somit hängt dieses Kriterium von der Designmatrix X_ξ ab und damit insbesondere von den Versuchspunkten des Versuchsplans ξ . Ein D-optimaler Versuchsplan ξ^* enthält die Versuchspunkte, die das D-Kriterium minimieren. Da ein D-optimaler Versuchsplan für festes k ebenso den Term

$$\det(\mathcal{I}_\xi(\theta)^{-1}) =: \tilde{D}(\xi)$$

minimiert, kann bezüglich der Suche nach einem optimalen Versuchsplan der Einfachheit halber auch dieser Ausdruck analysiert werden. Eine Interpretation des D-Kriteriums ist die Folgende. Ein Konfidenzellipsoid zum Konfidenzniveau $1 - \alpha$ für den Vektor β aus Kapitel 2 ist bei normalverteilten Fehlern gegeben durch

$$A = \{\beta \in \mathbb{R}^2 : (\hat{\beta} - \beta)^T (X^T \Gamma^{-1}(\lambda) X) (\hat{\beta} - \beta) \leq 2\gamma \hat{\sigma}^2\}, \quad (4.4)$$

wobei γ das $(1 - \alpha)$ -Quantil der F-Verteilung mit 2 und $nq - 2$ Freiheitsgraden bezeichnet und $\hat{\sigma}^2 = \frac{1}{nq-2} \hat{e}^T \Gamma^{-1}(\lambda) \hat{e}$, siehe Schach und Schäfer (1978), S. 100. Dabei bezeichnet \hat{e} den Vektor der Residuen. Wird in (4.4) $\hat{\sigma}^2$ durch σ^2 ersetzt, so ist das Volumen des Konfidenzellipsoids proportional zu der Größe $(\det((X^T \Gamma^{-1}(\lambda) X)^{-1}))^{1/2}$, siehe Krafft (1978), S. 31 ff. Damit ist das Volumen genau dann minimal, wenn das D-Kriterium minimal ist.

Beispiel 4.2. Gegeben seien die Versuchspläne aus Beispiel 4.1. Dann gilt: $D(\xi_1) = (1/4)^{1/2}$ und $D(\xi_2) = (1/3)^{1/2}$. Die zugehörigen Konfidenzellipsoide zum Niveau 95% sind in Abbildung 4.1 veranschaulicht. Auf Grund des Zusammenhangs zwischen dem D-Kriterium und dem Volumen des Konfidenzellipsoids ist das Konfidenzellipsoid zum Versuchsplan ξ_1 kleiner.

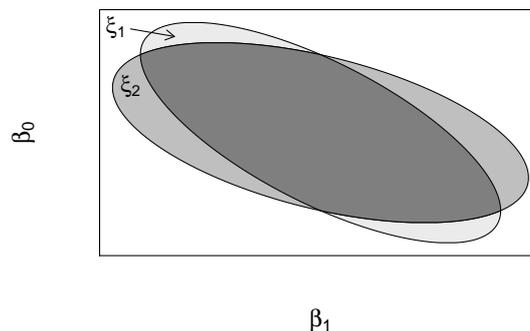


Abbildung 4.1: Konfidenzellipsoide für $[\beta_0, \beta_1]^T$. Die Ellipsoide resultieren aus den beiden Versuchsplänen ξ_1 und ξ_2 aus Beispiel 4.1. Es ist zu erkennen, dass das Ellipsoid zum Versuchsplan ξ_1 ein kleineres Volumen besitzt. Der Plan ξ_1 weist somit ein kleineres D-Kriterium auf.

5. D-Effizienzen von Versuchsplänen bei bekannten Störparametern

In diesem Kapitel ist es das Ziel, Versuchspläne zu finden, mit denen die Parameter β_0 und β_1 des Modells aus Kapitel 2 möglichst optimal im Sinne des D-Kriteriums aus Kapitel 4 geschätzt werden können. Dabei wird vorausgesetzt, dass die Parameter σ^2 und λ bekannt sind. Es werden überwiegend Versuchspläne mit $q = 3$ Versuchspunkten untersucht, wobei vorgegeben ist, dass die Versuchspläne die Versuchspunkte 0 und 1 enthalten. Der dritte Versuchspunkt d liegt entweder zwischen den beiden vorgegebenen Versuchspunkten oder entspricht einem der beiden Punkte. Die Begründung für die Vorgabe, dass zwei der drei Versuchspunkte durch die beiden Randpunkte des Versuchsraums gegeben sind, ist die, dass der D-optimale Versuchsplan diese beiden Versuchspunkte enthalten muss, siehe Dette *et al.* (2008). Darüber hinaus wird kurz auf einen allgemeineren Fall eingegangen, in dem die Versuchspläne mindestens zwei Versuchspunkte aufweisen. Dabei wird angenommen, dass die Fehler perfekt positiv korreliert sind. Bezüglich der Anzahl der Versuchseinheiten wird in allen Ausführungen dieses Kapitels nur der Fall betrachtet, dass $n = 1$ ist. Der Grund dafür ist der, dass die Ergebnisse für diesen Spezialfall auf die Fälle $n > 1$ direkt übertragen werden können.

Das Kapitel gliedert sich wie folgt. In Abschnitt 5.1 wird auf die Problematik eingegangen, dass der D-optimale Versuchsplan vom Korrelationsparameter abhängt, siehe Dette *et al.* (2008). Diese Problematik liefert eine Motivation für das Vorgehen in Abschnitt 5.2, in dem die sogenannte D-Effizienz von Versuchsplänen Gegenstand der Untersuchungen ist. Außerdem werden relative D-Effizienzen bestimmt, bei denen jeweils zwei Versuchspläne direkt miteinander verglichen werden. Da die Aussagen innerhalb des Abschnitts 5.2 auf umfangreichen Herleitungen basieren, werden die Herleitungen gesondert in Abschnitt 5.3 zusammengetragen, so dass die relevanten Aussagen kompakt in Abschnitt 5.2 vorliegen. In Abschnitt 5.3 wird zudem das D-Kriterium, das bereits in Abschnitt 5.1 betrachtet wird, konstruiert. Damit besitzt der Abschnitt 5.3 lediglich einen vorbereitenden Charakter. Falls der Leser nur an den Ergebnissen dieses Kapitels interessiert ist, so ermöglicht die Gliederung ein Überspringen des langen dritten Abschnitts.

5.1. D-optimale Versuchspläne

Ein numerisches Resultat aus Dette *et al.* (2008) lautet, dass lediglich lokal D-optimale Versuchspläne bestimmt werden können, da der optimale Versuchsplan mit Versuchspunkt d^* vom zugrunde liegenden Korrelationsparameter λ abhängt. Dieses Resultat wird in Abbildung 5.1 veranschaulicht. Es ist zu erkennen, dass für $\lambda \geq 0.0007798$ der Plan mit $d^* = 1/2$ D-optimal ist. Dagegen kann für $\lambda < 0.0007798$ festgehalten werden, dass der optimale Versuchspunkt d^* immer kleiner wird, je kleiner das gegebene λ ist. Dieses Resultat ist für die Bestimmung eines geeigneten Versuchsplans unbefriedigend. Einen möglichen Ausweg liefert der Ansatz aus Abschnitt 5.2. Dort werden die D-Effizienzen verschiedener Versuchspläne untersucht.

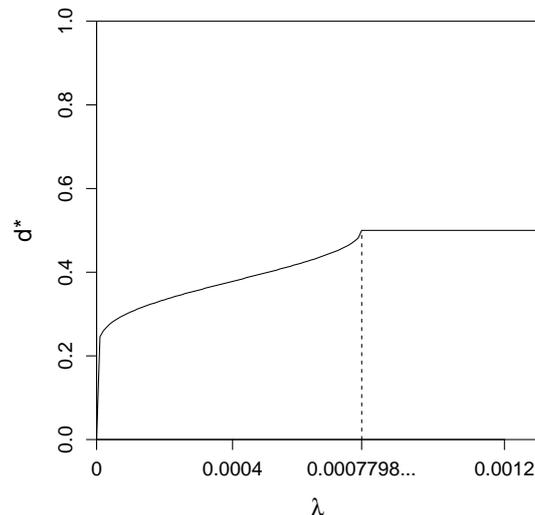


Abbildung 5.1: Lokal D-optimale Versuchspläne für $[\beta_0, \beta_1]^T$, $q=3$. Für $\lambda \geq 0.0007798$ ist der Versuchsplan mit $d^* = 1/2$ D-optimal. Für $\lambda < 0.0007798$ wird der optimale Versuchspunkt zunehmend kleiner, je kleiner die Korrelation ist.

In der vorliegenden Situation ist beim D-Kriterium die Wurzelfunktion die äußerste Funktion, was aus der Definition des D-Kriteriums für $k = 2$ folgt, siehe (4.3). Es bietet sich in den weiteren Ausführungen daher an, statt des D-Kriteriums das quadrierte D-Kriterium

$$D(d)^2 =: \tilde{D}(d) \quad (5.1)$$

zu analysieren. Dies hat den Vorteil, dass zukünftig in allen Rechenschritten auf den Exponenten $1/2$ verzichtet werden kann. Da das D-Kriterium und auch der optimale

Versuchsplan vom Parameter λ abhängen, werden das D-Kriterium und das quadrierte D-Kriterium im Folgenden mit $D_\lambda(d)$ und $\tilde{D}_\lambda(d)$ bezeichnet. Es sei kurz angemerkt, dass das D-Kriterium wegen $D_\lambda(d) = D_\lambda(1-d)$ symmetrisch um $d = 1/2$ ist. Das hat zur Folge, dass eine Aussage, die für den Versuchsplan mit Versuchspunkt d gilt, aus Symmetriegründen auch für den Versuchsplan mit Versuchspunkt $1-d$ gilt. Im Weiteren wird nicht immer explizit darauf hingewiesen.

Zur weiteren Veranschaulichung enthält Tabelle 5.1 Funktionswerte von $\tilde{D}_\lambda(d)$ für ausgewählte λ -Werte und ausgewählte Versuchspunkte. Zudem sind in der letzten Spalte die Funktionswerte der numerisch bestimmten lokal D-optimalen Versuchspläne mit Versuchspunkt d_λ^* für gegebene λ -Werte aufgeführt.

Tabelle 5.1: $\tilde{D}_\lambda(d)$ für ausgewählte Korrelationen und Versuchspläne. Die letzte Spalte enthält zum Vergleich $\tilde{D}_\lambda(d)$ für den lokal D-optimalen Versuchsplan.

λ	d									d_λ^*
	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1	
0	0.500	0.550	0.595	0.633	0.658	0.667	0.658	...	0.500	0.500 ($d_\lambda^* = 0$)
0.0001	1.000	0.763	0.684	0.668	0.672	0.676	0.672	...	1.000	0.668 (0.305)
0.001	1.000	0.816	0.734	0.703	0.695	0.694	0.695	...	1.000	0.694 (0.5)
0.01	1.000	0.881	0.812	0.774	0.756	0.751	0.756	...	1.000	0.751 (0.5)
0.1	0.990	0.948	0.918	0.898	0.887	0.883	0.887	...	0.990	0.883 (0.5)
0.25	0.938	0.924	0.913	0.906	0.901	0.900	0.901	...	0.938	0.900 (0.5)
0.5	0.750	0.748	0.747	0.745	0.745	0.745	0.745	...	0.750	0.745 (0.5)
0.9	0.190	0.190	0.190	0.190	0.190	0.190	0.190	...	0.190	0.190 (0.5)
1	0	0	0	0	0	0	0	...	0	0

5.2. Untersuchung der D-Effizienz

In Abschnitt 5.1 konnten lediglich lokal D-optimale Versuchspläne angegeben werden. Daher wird in diesem Abschnitt ein alternativer Ansatz herangezogen. Es handelt sich um die sogenannte D-Effizienz eines Versuchsplans. Berechnet wird die D-Effizienz eines Versuchsplans mit Versuchspunkt d mittels

$$\text{eff}_{D;\lambda_0}(d) = \frac{D_{\lambda_0}(d_{\lambda_0}^*)}{D_{\lambda_0}(d)} = \frac{\sqrt{\tilde{D}_{\lambda_0}(d_{\lambda_0}^*)}}{\sqrt{\tilde{D}_{\lambda_0}(d)}} = \sqrt{\frac{\tilde{D}_{\lambda_0}(d_{\lambda_0}^*)}{\tilde{D}_{\lambda_0}(d)}}. \quad (5.2)$$

Dabei bezeichnet $d_{\lambda_0}^*$ den Versuchspunkt, der bei gegebenem λ_0 zu einem D-optimalen Versuchsplan führt.

Ein Ziel könnte es sein, einen Versuchsplan zu bestimmen, für den bei beliebiger Variation von λ die minimale D-Effizienz maximal ist. Der Plan, der die minimale D-Effizienz maximiert, ist im Folgenden der Plan mit Versuchspunkt

$$d^{**} = \arg \max_d \min_{\lambda} \text{eff}_{D;\lambda}(d).$$

In Tabelle 5.2 sind einige Effizienzen zusammengetragen. Diese resultieren aus den Einträgen der Tabelle 5.1. Da Tabelle 5.1 überwiegend lediglich auf numerischen Untersuchungen basiert, sind die Einträge in Tabelle 5.2 ebenfalls überwiegend nur numerische Resultate. Die Untersuchungen in Abschnitt 5.3 liefern jedoch Erkenntnisse, die die Ergebnisse aus Tabelle 5.2 bestätigen und vertiefen. Die Erkenntnisse werden nun im Folgenden beschrieben und in Abbildung 5.3 veranschaulicht.

Tabelle 5.2: *D-Effizienzen für ausgewählte Korrelationen und Versuchspläne.*

λ	d								
	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1
0	1.000	0.954	0.917	0.889	0.872	0.866	0.872	...	1.000
0.0001	0.817	0.936	0.989	1.000	0.997	0.995	0.997	...	0.817
0.001	0.833	0.923	0.973	0.994	0.999	1.000	0.999	...	0.833
0.01	0.867	0.923	0.962	0.985	0.997	1.000	0.997	...	0.867
0.1	0.944	0.965	0.981	0.992	0.998	1.000	0.998	...	0.944
0.25	0.980	0.987	0.993	0.997	0.999	1.000	0.999	...	0.980
0.5	0.996	0.998	0.999	0.999	1.000	1.000	1.000	...	0.996
0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	...	1.000

Werden zunächst korrelierte Fehler außen vor gelassen, so folgt aus Behauptung 5.18 (Seite 54) direkt, dass die D-Effizienz für sämtliche Versuchspläne bei unkorrelierten Fehlern mindestens $(\tilde{D}_{\lambda=0}(0)/\tilde{D}_{\lambda=0}(1/2))^{1/2} = (3/4)^{1/2} \approx 0.866$ beträgt. Somit werden die entsprechenden Einträge in Tabelle 5.2 bestätigt.

Nun wird der allgemeine Fall betrachtet, der korrelierte Fehler beinhaltet. Die Einträge in Tabelle 5.2 zu $d = 0$ lassen die Vermutung zu, dass die D-Effizienz des Versuchsplans mit $d = 0$ kleiner wird, je kleiner die Korrelation ist, während für unkorrelierte Fehler gerade dieser Versuchsplan D-optimal ist, siehe auch Dette *et al.* (2008). Anhand der Untersuchungen, die in den Abschnitten 5.3.1 und 5.3.2 getätigt werden, kann sogar gefolgert werden, dass die D-Effizienz im Grenzfall $\lambda \rightarrow 0$ nur $(1/2)^{1/2} \approx 0.707$ beträgt. Denn aus den Grenzwerten

$$\lim_{\lambda \rightarrow 0} \tilde{D}_{\lambda}(0) = 1$$

und

$$\lim_{d \rightarrow 0} \lim_{\lambda \rightarrow 0} \tilde{D}_\lambda(d) = 1/2,$$

siehe (5.15) und Behauptung 5.4, folgt zunächst, dass die D-Effizienz des Versuchsplans mit $d = 0$ höchstens den Wert $(1/2)^{1/2}$ aufweist. Somit wird die in Tabelle 5.2 für $d = 0$ mit 0.817 kleinste aufgeführte D-Effizienz noch deutlich unterschritten. Die D-Effizienz von $(1/2)^{1/2}$ ist darüber hinaus sogar die kleinste D-Effizienz bezüglich aller Versuchspunkte $d \in [0, 1]$, was aus einer oberen Schranke und einer unteren Schranke des quadrierten D-Kriteriums folgt. Die Argumentation sieht wie folgt aus. Die obere Schranke ist durch die Funktion $L_0(\lambda) = 1 - \lambda^2$ gegeben. Die beiden Funktionen $L_1(\lambda) = 1/2 + \lambda$ und $L_2(\lambda) = 81(1 - \lambda^2)/100$ liefern zusammen die untere Schranke. Die Schranken werden in Abbildung 5.2 veranschaulicht. Sie werden später in Abschnitt 5.3.5 hergeleitet. Auf Grund der beiden Schranken haben sämtliche Versuchspläne mit $d \in [0, 1]$ mindestens

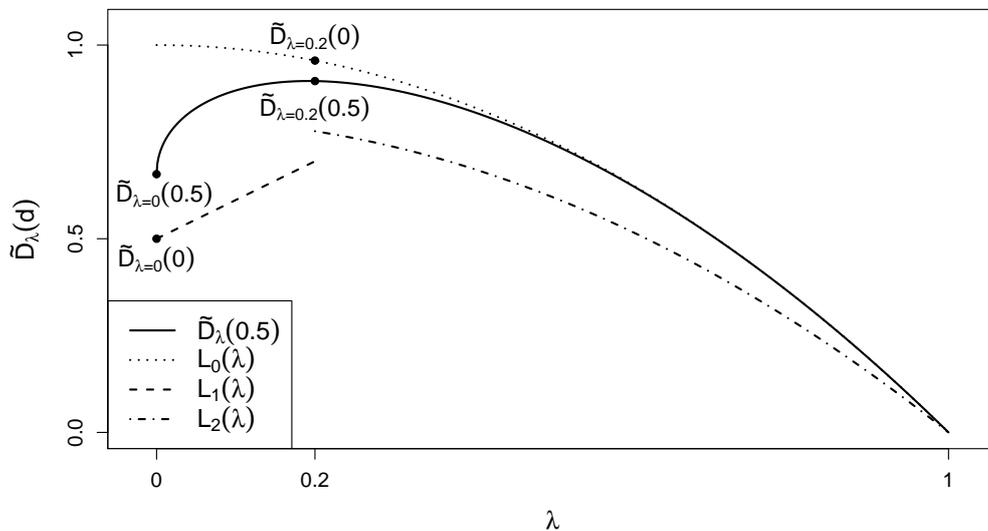


Abbildung 5.2: Eine obere und eine untere Schranke der Funktion $\tilde{D}_\lambda(d)$ für $\lambda \in [0, 1]$. Die Funktion $L_0(\lambda) = 1 - \lambda^2$ ist eine obere Schranke für $\tilde{D}_\lambda(d)$. Die Gerade $L_1(\lambda) = 1/2 + \lambda$ stellt zusammen mit der Funktion $L_2(\lambda) = 81(1 - \lambda^2)/100$ eine untere Schranke dar. Der Funktionsgraph von $\tilde{D}_\lambda(1/2)$ wird exemplarisch für alle weiteren Versuchspunkte eingezeichnet. Die Schranken sind Korollar 5.20, Behauptung 5.22 und Behauptung 5.23 aus Abschnitt 5.3.5 entnommen.

D-Effizienzen von

$$\begin{aligned} & (L_1(\lambda_0)/L_0(\lambda_0))^{1/2}, \text{ falls } \lambda_0 \in [0, 1/5), \\ & (L_2(\lambda_0)/L_0(\lambda_0))^{1/2}, \text{ falls } \lambda_0 \in [1/5, 1]. \end{aligned} \quad (5.3)$$

Nun ist die Funktion $L_1(\lambda)/L_0(\lambda) = (1/2 + \lambda)/(1 - \lambda^2)$ monoton steigend in λ . Sie nimmt somit für $\lambda = 0$ ein Minimum an. Dies führt zu der Mindesteffizienz $(1/2)^{1/2}$ für $\lambda < 1/5$. Da die Funktion $L_2(\lambda)/L_0(\lambda)$ stets den Funktionswert $81/100$ aufweist, weisen die D-Effizienzen für $\lambda \geq 1/5$ mindestens den Wert $(81/100)^{1/2} = 9/10$ auf. Die globale Mindesteffizienz beträgt demnach $(1/2)^{1/2} \approx 0.707$. Insgesamt liegt nun die Situation vor,

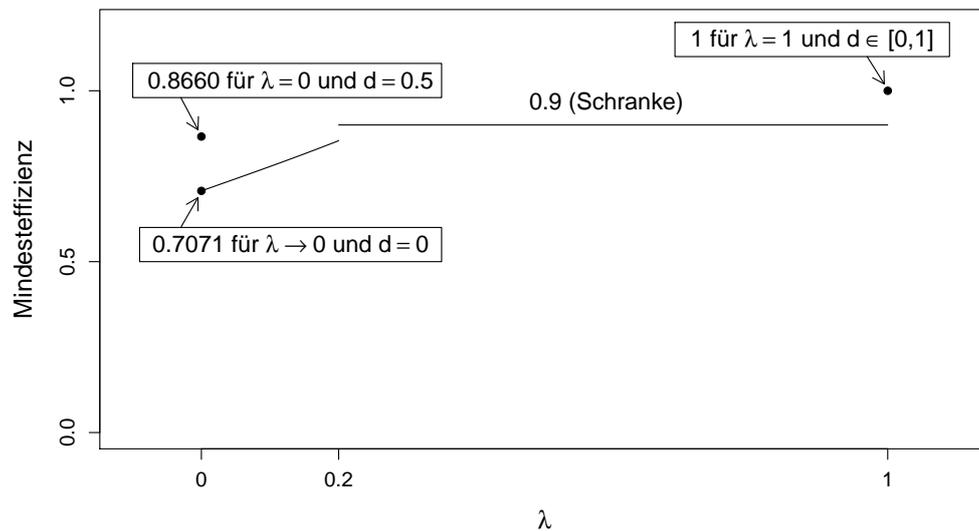


Abbildung 5.3: Mindesteffizienzen für $\lambda \in [0, 1]$ bezüglich sämtlicher Versuchspläne mit Versuchspunkten $d \in [0, 1]$. Die Geraden stellen eine untere Schranke dar, welche durch (5.3) gegeben ist. Zudem sind die Mindesteffizienzen der Korrelationen $\lambda = 0$ und $\lambda = 1$ durch die Punkte mit den Koordinaten $(0, 0.866)$ und $(1, 1)$ eingezeichnet. Erstere resultiert aus Behauptung 5.18, letztere resultiert aus (5.18). Schließlich ist durch den Punkt mit den Koordinaten $(0, 0.707)$ eine obere Schranke für die D-Effizienz des Versuchsplans mit $d = 0$ unter der Korrelation $\lambda \rightarrow 0$ aufgeführt, welche das Minimum der unteren Schranke annimmt. Damit ist der Versuchsplan mit $d = 0$ der Plan, der bezüglich $\lambda \in [0, 1]$ im ungünstigsten Fall die kleinste D-Effizienz aller Versuchspläne aufweist.

dass die obere Schranke der D-Effizienzen von $d = 0$ mit der globalen Mindesteffizienz übereinstimmt. Damit weist der Plan mit $d = 0$ die kleinste Mindesteffizienz bezüglich aller Pläne auf, wobei diese für $\lambda \rightarrow 0$ realisiert wird. Dagegen hat in der Situation $\lambda \rightarrow 0$ der Versuchsplan mit $d = 1/2$ wegen

$$\lim_{\lambda \rightarrow 0} \tilde{D}_\lambda(1/2) = 2/3$$

und der Schranke

$$\lim_{\lambda \rightarrow 0} L_1(\lambda) = 1/2$$

eine Effizienz von $(3/4)^{1/2} \approx 0.866$. Es fällt auf, dass für $\lambda \rightarrow 0$ der äquidistante Plan dem Plan mit $d = 0$ überlegen ist, während für $\lambda = 0$ wie bereits beschrieben der Fall eintritt, dass der Plan mit $d = 0$ den optimalen Versuchsplan liefert. Es gilt sogar für $\lambda \in [0, 1]$, dass der Versuchsplan mit $d = 1/2$ eine größere Mindesteffizienz als der Plan mit $d = 0$ aufweist. Dies folgt aus obigem Resultat, dass für $\lambda \in [1/5, 1]$ alle Pläne mindestens Effizienzen von $9/10$ aufweisen, und der Schranke für die Mindesteffizienz, die in Behauptung 5.1 formuliert wird.

Behauptung 5.1. *Sei $\lambda \in [0, 1/5]$. Die Zahl $(16/25)^{1/2} = 0.8$ stellt eine untere Schranke für die D-Effizienzen des Versuchsplans mit $d = 1/2$ dar.*

Beweis. Es ist zu zeigen, dass

$$\frac{L_1(\lambda)}{\tilde{D}_\lambda(1/2)} = \frac{(\lambda - 4\lambda^{1/2} + 3)(2\lambda + 1)}{4(1 - \lambda)^2} \geq \frac{16}{25}.$$

Äquivalenzumformungen liefern die Ungleichung

$$(1 - \lambda^{1/2})(14\lambda^{3/2} + 214\lambda - 89\lambda^{1/2} + 11) \geq 0.$$

Wegen

$$\begin{aligned} (1 - \lambda^{1/2})(14\lambda^{3/2} + 214\lambda - 89\lambda^{1/2} + 11) &\geq (1 - \lambda^{1/2})\left(\frac{7921}{44}\lambda - 89\lambda^{1/2} + 11\right) \\ &= \frac{7921}{44}(1 - \lambda^{1/2})\left(\lambda^{1/2} - \frac{22}{89}\right)^2 \\ &\geq 0 \end{aligned}$$

sind die obigen Ungleichungen gültig. □

Als weiteres Resultat kann festgehalten werden, dass die in Tabelle 5.2 für stärkere Korrelationen größer werdenden Effizienzen wegen Behauptung 5.7 aus Abschnitt 5.3.3 plausibel sind.

Nach den bisherigen Untersuchungen drängt sich die Frage auf, welcher Versuchsplan die minimale Effizienz maximiert. Für die ausgewählten Werte aus Tabelle 5.2 hat beispielsweise der Versuchsplan mit $d = 0.1$ eine minimale Effizienz von $\text{eff}_{D;0.001}(0.1) = 0.923$. Es ist zu erkennen, dass die weiteren aufgeführten Versuchspunkte eine kleinere minimale Effizienz aufweisen. Innerhalb dieser Tabelle maximiert somit der Versuchspunkt $d = 0.1$ die minimale Effizienz. Eine numerische Untersuchung auf einem feineren Gitter kommt zu dem Ergebnis, dass der Versuchsplan mit $d^{**} = 0.138$ die minimale Effizienz mit einem Wert von 0.939 maximiert. In Abbildung 5.4 werden die D-Effizienzen aller Versuchspläne veranschaulicht. Dabei werden unterschiedliche Skalierungen für λ gewählt, damit vor allem die D-Effizienzen für kleine Korrelationen zu erkennen sind.

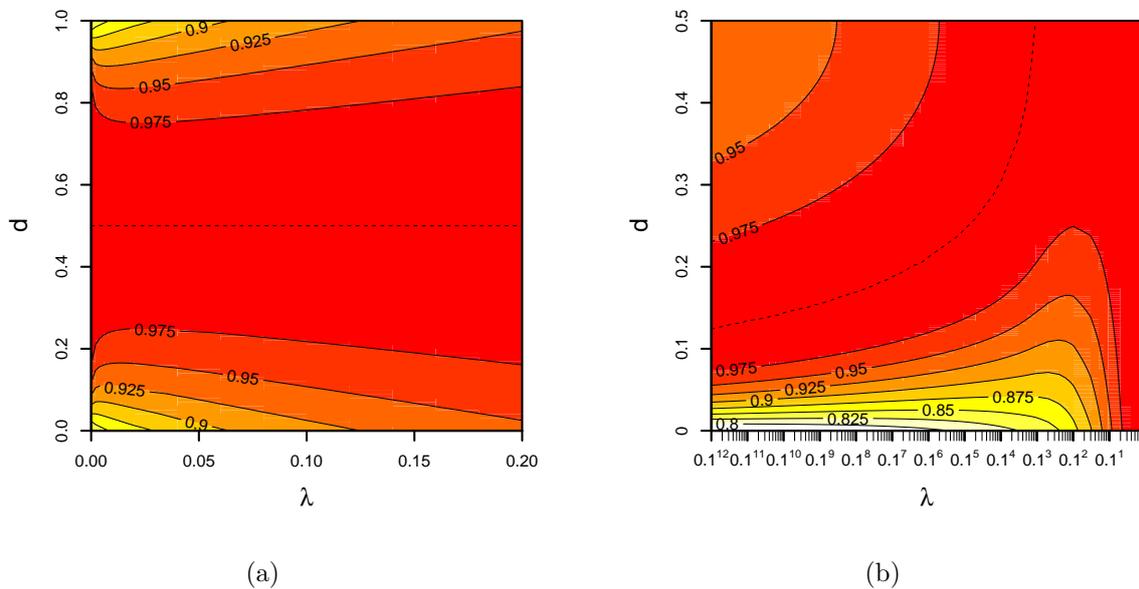


Abbildung 5.4: D-Effizienzen ($\text{eff}_{D;\lambda_0}(d)$) für $d \in [0, 1]$ in Abhängigkeit vom Korrelationsparameter λ . Die lokal D-optimalen Versuchspläne sind durch (---) gekennzeichnet ((a): λ aus $[0, 0.2]$, (b): λ aus $[0.1^{12}, 0.9]$ mit logarithmischer Skala.)

Es ist festzuhalten, dass die Beweise bezüglich der Schranken, die in Abschnitt 5.3.5 und zum Teil auch im Anhang formuliert sind, recht aufwendig und umfangreich sind. Ein Nachweis, dass ein Versuchsplan mit $d \approx 0.138$ die minimale D-Effizienz maximiert, erscheint noch weitaus schwerer zu sein. Aus diesem Grund werden nun anhand der sogenannten relativen D-Effizienz Aussagen präsentiert, die die Vermutung, dass ein Versuchsplan mit $d \approx 0.138$ die minimale D-Effizienz maximiert, bestärken.

Die relativen D-Effizienzen eines Versuchsplans mit Versuchspunkt d_1 bezüglich eines Versuchsplans mit Versuchspunkt d_2 sind durch

$$\frac{D_\lambda(d_2)}{D_\lambda(d_1)} =: q_{d_1, d_2}(\lambda)$$

definiert. Abweichend zur D-Effizienz wird also bei der relativen D-Effizienz statt des D-optimalen Versuchsplans ein beliebiger Versuchsplan mit Versuchspunkt d_2 als Referenz herangezogen. Ein Versuchsplan mit Versuchspunkt d_1 ist nun bezüglich des Plans mit Versuchspunkt d_2 umso effizienter, je größer seine relativen D-Effizienzen sind. Es sei angemerkt, dass hier abweichend zur D-Effizienz auch Werte größer 1 auftreten können.

Im Weiteren wird anstelle des bisher betrachteten Versuchsraums $\mathcal{X} = [0, 1]$ der Teilraum $\{i/8 : i = 0, 1, 2, \dots, 8\}$ herangezogen. Wegen $q_{d_1, d_2}(\lambda) = q_{d_1, 1-d_2}(\lambda) = q_{1-d_1, d_2}(\lambda) = q_{1-d_1, 1-d_2}(\lambda)$ kann zunächst festgehalten werden, dass die relativen D-Effizienzen bestimmter Versuchspläne mit bestimmten Referenzen aus Symmetriegründen identisch sind. Somit genügt es bezüglich des Teilraums $\{i/8 : i = 0, 1, 2, \dots, 8\}$ die Versuchspläne mit den Punkten $d = 0, d = 1/8, d = 1/4, d = 3/8, d = 1/2$ zu betrachten.

Die folgenden Ausführungen vergleichen die Versuchspläne mit den genannten Versuchspunkten bezüglich ihrer relativen D-Effizienzen miteinander. Dabei wird wegen $D_{\lambda=1}(d) = 0$ der Fall $\lambda = 1$ außen vor gelassen. Das Ziel der Ausführungen ist es nun, einen Plan zu finden, dessen kleinstes Minimum seiner relativen D-Effizienzen maximal ist.

In Tabelle 5.3 sind untere Schranken und Minima der relativen D-Effizienzen, die für den Versuchsraum $\{i/8 : i = 0, 1, 2, \dots, 8\}$ relevant sind, angegeben. Die Herleitungen der Schranken und Minima können in Abschnitt 5.3.6 nachvollzogen werden.

Tabelle 5.3: *Minima beziehungsweise untere Schranken der relativen D-Effizienzen der Versuchspläne mit Versuchspunkt d_1 bezüglich der Versuchspläne mit Versuchspunkt d_2 . Die unteren Schranken können an den Ungleichheitszeichen ausgemacht werden.*

		d_2				
		0	1/8	1/4	3/8	1/2
d_1	0		0.749	0.785	0.808	0.817
	1/8	0.944		≥ 0.949	≥ 0.922	≥ 0.922
	1/4	0.901	0.955		≥ 0.901	≥ 0.975
	3/8	0.875	0.927	0.971		≥ 0.922
	1/2	0.866	0.918	0.961	0.990	

Anhand der Einträge in Tabelle 5.3 ist zu erkennen, dass der Versuchsplan mit $d = 0$ im

Vergleich zu den Plänen mit $d = 1/8$, $d = 1/4$, $d = 3/8$ bzw. $d = 1/2$ einen größeren Verlust bezüglich des D-Kriteriums erleiden kann als umgekehrt. Wegen

$$\min q_{0,1/2}(\lambda) > \min q_{0,3/8}(\lambda) > \min q_{0,1/4}(\lambda) > \min q_{0,1/8}(\lambda)$$

ist der maximal mögliche Verlust gegenüber dem Versuchsplan mit $d = 1/8$ am größten. Somit kann festgehalten werden, dass der Versuchsplan mit $d = 0$ bezüglich der relativen D-Effizienz vergleichsweise schlecht abschneidet. Andererseits ist zu erkennen, dass der Plan mit $d = 1/8$ bezüglich der untersuchten Versuchspläne den geringsten maximal möglichen Verlust aufweist, denn im ungünstigsten Fall beträgt die relative D-Effizienz des Versuchsplans mit $d = 1/8$ zumindest 0.922.

Die Resultate bezüglich des Versuchsraums $\{i/8 : i = 0, 1, 2, \dots, 8\}$, dass der Versuchsplan mit $d = 0$ als vergleichsweise schlecht zu beurteilen ist und dagegen der Plan mit $d = 1/8$ als relativ gut, decken sich mit den bisherigen Ergebnissen dieses Abschnitts, denn es wurde für den Versuchsraum $\mathcal{X} = [0, 1]$ gezeigt, dass der Plan mit $d = 0$ die Mindesteffizienzen sämtlicher Versuchspläne minimiert. Weiter wurde auf Grund numerischer Untersuchungen beschrieben, dass ein Plan mit $d \approx 0.138$ die Mindesteffizienzen mit 0.939 maximiert. Letzteres Ergebnis ist vergleichsweise in der Nähe der unteren Schranke für die Mindesteffizienz des Versuchsplans mit $d = 0.125$. Dabei betrug die Schranke 0.922. Die Ergebnisse in Tabelle 5.3 lassen jedoch noch die Vermutung zu, dass die Mindesteffizienzen der Versuchspläne mit $d \in (0, 1/2]$ umso größer werden, je kleiner der Versuchspunkt gewählt wird. Zunächst gilt jedoch beispielsweise, dass $q_{1/16,1/2}(\lambda = 1/1000) \approx 0.894$ ist. Dies bedeutet, dass die Mindesteffizienz des Versuchsplans mit $d = 1/16$ auf jeden Fall kleiner als die kleinste untere Schranke der relativen D-Effizienzen, die den Versuchspunkt $d = 1/8$ in Tabelle 5.3 betreffen, ist. Falls der Versuchsraum $\{i/8 : i = 0, 1, 2, \dots, 8\}$ um den Versuchspunkt $d = 1/16$ ergänzt wird, so führt dies für den Versuchsplan mit $d = 1/8$ jedoch nicht dazu, dass dessen bisherige Schranke der relativen D-Effizienzen von 0.922 unterboten wird, siehe Behauptung 5.2.

Behauptung 5.2. *Gegeben seien die beiden Versuchspläne mit $d = 1/16$ beziehungsweise $d = 1/8$. Dann gilt:*

$$q_{1/8,1/16}(\lambda) \geq (17/20)^{1/2} \approx 0.922.$$

Beweis. Die Ungleichung

$$(q_{1/8,1/16}(\lambda))^2 \geq \frac{17}{20}$$

ist äquivalent zu der Ungleichung

$$\begin{aligned}
& 305\lambda^3 + 272\lambda^{47/16} - 1200\lambda^{23/8} + 895\lambda^{11/4} - 272\lambda^{43/16} - 4480\lambda^{17/8} \\
& \quad + 4080\lambda^{33/16} + 4943\lambda^2 - 4560\lambda^{15/8} - 4080\lambda^{29/16} + 4097\lambda^{7/4} \\
& \quad + 255\lambda^{5/4} - 272\lambda^{19/16} - 560\lambda^{9/8} + 945\lambda + 272\lambda^{15/16} - 640\lambda^{7/8} \\
& \quad - 4080\lambda^{5/16} + 8577\lambda^{1/4} - 9040\lambda^{1/8} + 4080\lambda^{1/16} + 463 \geq 0. \quad (5.4)
\end{aligned}$$

Eine Substitution mittels $\lambda = x^{16}$ liefert für die linke Seite von (5.4) den Term

$$(1-x)^3(x+1)^2m(x)$$

mit

$$\begin{aligned}
m(x) = & -305x^{43} - 577x^{42} + 13x^{41} - 531x^{40} + 59x^{39} - 485x^{38} + 105x^{37} - 439x^{36} \\
& + 151x^{35} - 393x^{34} + 197x^{33} - 347x^{32} + 243x^{31} - 301x^{30} + 4769x^{29} + 145x^{28} \\
& + 4752x^{27} - 3952x^{26} + 4752x^{25} - 3952x^{24} + 4752x^{23} - 3952x^{22} + 4752x^{21} \\
& - 3952x^{20} + 4752x^{19} - 3952x^{18} + 4752x^{17} - 3952x^{16} + 4497x^{15} - 3935x^{14} \\
& + 4819x^{13} - 3341x^{12} + 4773x^{11} - 3387x^{10} + 4727x^9 - 3433x^8 + 4681x^7 \\
& - 3479x^6 + 4635x^5 - 3525x^4 + 4589x^3 - 3571x^2 + 4543x + 463.
\end{aligned}$$

Mit Behauptung 5.13 aus Abschnitt 5.3.4 (Seite 50) kann gezeigt werden, dass $m(x)$ positiv ist. Somit gilt die Behauptung. \square

Somit kann festgehalten werden, dass das Minimum der relativen D-Effizienzen eines Versuchspunkts d_1 , der im Vergleich zu einem anderen Versuchspunkt d_2 näher am Rand des Versuchsraums liegt, im Allgemeinen nicht kleiner ist als das Minimum der relativen D-Effizienzen des Versuchspunkts d_2 . Damit deuten die Ergebnisse in Tabelle 5.3 umso mehr darauf hin, dass ein Versuchsplan mit $d \approx 0.138$ die Mindesteffizienz maximiert.

5.3. Vorüberlegungen zu den Abschnitten 5.1 und 5.2

5.3.1. Konstruktion des D-Kriteriums

Das D-Kriterium, das in den Abschnitten 5.1 und 5.2 herangezogen wird, wird nun im Folgenden konstruiert, wobei abweichend zu Dette *et al.* (2008) hier zusätzlich die Fälle $d = 0$, $d = 1$ und $\lambda = 1$ betrachtet werden. Die Konstruktion des D-Kriteriums geschieht

durch eine Fallunterscheidung mit drei Fällen. Die Begründungen für die Fallunterscheidung können den entsprechenden Ausführungen entnommen werden. Zunächst wird die Gestalt des D-Kriteriums für den Fall $\lambda \in (0, 1)$, wobei $d \in (0, 1)$, sowie für den Fall $\lambda = 0$ bestimmt. Die Designmatrix des Versuchsplans ξ hat in der vorliegenden Situation die Gestalt

$$X_{(1)} = \begin{bmatrix} 1 & 0 \\ 1 & d \\ 1 & 1 \end{bmatrix}. \quad (5.5)$$

Die Kovarianzmatrix des Fehlervektors lautet

$$\Sigma_{(1)}(\sigma^2, \lambda) = \sigma^2 \begin{bmatrix} 1 & \lambda^d & \lambda \\ \lambda^d & 1 & \lambda^{1-d} \\ \lambda & \lambda^{1-d} & 1 \end{bmatrix}. \quad (5.6)$$

Da die Fehlervarianz jedoch lediglich als Vorfaktor in das D-Kriterium eingeht, siehe (3.8), hat sie bezüglich des Versuchsdesigns keine Relevanz. Also kann im Weiteren aus Gründen der Vereinfachung die Korrelationsmatrix

$$\Sigma_{(1)}(\lambda) = \begin{bmatrix} 1 & \lambda^d & \lambda \\ \lambda^d & 1 & \lambda^{1-d} \\ \lambda & \lambda^{1-d} & 1 \end{bmatrix} \quad (5.7)$$

statt der Kovarianzmatrix $\Sigma_{(1)}(\sigma^2, \lambda)$ betrachtet werden. Analog wird auch in Kürze in den weiteren Fällen vorgegangen. Da für die Bestimmung der Informationsmatrix unter anderem $\Sigma_{(1)}^{-1}(\lambda)$ herangezogen wird, ist vorab zu klären, für welche Fälle $\Sigma_{(1)}(\lambda)$ invertierbar ist. Als Kriterium für die Invertierbarkeit kann die Determinante von $\Sigma_{(1)}(\lambda)$ herangezogen werden. Ist diese gleich Null so ist $\Sigma_{(1)}(\lambda)$ nicht invertierbar, andernfalls ist $\Sigma_{(1)}(\lambda)$ invertierbar, siehe Schmidt und Trenkler (2006), S. 72. Nun ist

$$\det(\Sigma_{(1)}(\lambda)) = (1 - \lambda^{2(1-d)})(1 - \lambda^{2d}).$$

Die beiden Faktoren $1 - \lambda^{2(1-d)}$ und $1 - \lambda^{2d}$ sind für $\lambda = 1$ stets Null. Für $\lambda \in (0, 1)$, wobei $d \in \{0, 1\}$ ist, ist zumindest ein Faktor Null. Damit ist in diesen Fällen $\Sigma_{(1)}(\lambda)$ nicht invertierbar. Dagegen ist die Determinante für alle weiteren Fälle stets ungleich Null. Demnach existiert für $\lambda = 0$ und in den Fällen $\lambda \in (0, 1)$, wobei $d \in (0, 1)$ ist, die

Inverse von $\Sigma_{(1)}(\lambda)$. Sie lautet

$$\Sigma_{(1)}^{-1}(\lambda) = \begin{bmatrix} \frac{1}{1-\lambda^{2d}} & -\frac{\lambda^d}{1-\lambda^{2d}} & 0 \\ -\frac{\lambda^d}{1-\lambda^{2d}} & \frac{\lambda^2-1}{\lambda^{2d}+\lambda^{2-2d}-\lambda^2-1} & -\frac{\lambda^{1-d}}{1-\lambda^{2-2d}} \\ 0 & -\frac{\lambda^{1-d}}{1-\lambda^{2-2d}} & \frac{1}{1-\lambda^{2-2d}} \end{bmatrix}. \quad (5.8)$$

Daraufhin resultiert die Informationsmatrix

$$\mathcal{I}_{(1)}([\beta_0, \beta_1]^T) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad (5.9)$$

mit

$$\begin{aligned} a_{11} &= \frac{\lambda^d + \lambda^{1-d} - \lambda + 3}{\lambda^d + \lambda^{1-d} + \lambda + 1}, \\ a_{12} = a_{21} &= \frac{(1-\lambda)d + \lambda^d + 1}{\lambda^d + \lambda^{1-d} + \lambda + 1}, \\ a_{22} &= \frac{(1-\lambda^2)d^2 + (2\lambda^{d+1} - 2\lambda^{1-d})d - \lambda^{2d} + 1}{-\lambda^{2d} - \lambda^{2-2d} + \lambda^2 + 1}. \end{aligned}$$

In der vorliegenden Situation ist $k = 2$, so dass für $\lambda = 0$ und für $\lambda \in (0, 1)$, wobei $d \in (0, 1)$, das D-Kriterium wie folgt lautet

$$\begin{aligned} D_\lambda(d) &= (\det(\mathcal{I}_{(1)}([\beta_0, \beta_1]^T)^{-1}))^{1/2} \\ &= \left(\frac{(1-\lambda^{2d})(1-\lambda^{2(1-d)})}{2((1-(1-d)\lambda^d)(1-d\lambda^{1-d})-d(1-d))} \right)^{1/2}. \end{aligned} \quad (5.10)$$

Speziell für $\lambda = 0$ vereinfacht sich der Ausdruck (5.10) zu

$$D_\lambda(d) = \left(\frac{1}{2(1-d+d^2)} \right)^{1/2}. \quad (5.11)$$

Als Nächstes wird die Gestalt des D-Kriteriums für den Fall $\lambda \in (0, 1)$, wobei $d \in \{0, 1\}$ ist, bestimmt. Vorab sei daran erinnert, dass das D-Kriterium eine univariate Kennzahl darstellt, die die Varianzen mehrerer Schätzer sowie deren Kovarianzen zusammenfassend beschreibt. Dabei wird in dem zuvor beschriebenen Fall die Kovarianzmatrix des verallgemeinerten Kleinste-Quadrate-Schätzers bewertet, wobei dessen Optimalitätseigenschaft der Grund für seine Wahl ist. Für $d \in \{0, 1\}$ existiert der verallgemeinerte Kleinste-Quadrate-Schätzer jedoch nicht, so dass zu untersuchen ist, ob es in dieser Situation

dennoch einen Schätzer gibt, der die wünschenswerten Eigenschaften des verallgemeinerten Kleinste-Quadrate-Schätzers aufweist. Dafür wird zunächst die Situation betrachtet, dass lediglich zwei Versuchspunkte vorliegen, wobei diese an den Rand des Versuchsraums gesetzt werden. Die Designmatrix lautet demnach

$$X_{(0)} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \quad (5.12)$$

Die zugehörige Kovarianzmatrix hat die Gestalt

$$\Sigma_{(0)}(\lambda) = \begin{bmatrix} 1 & \lambda \\ \lambda & 1 \end{bmatrix}. \quad (5.13)$$

Daraufhin resultiert die Informationsmatrix

$$\mathcal{I}_{(0)}([\beta_0, \beta_1]^T) = X_{(0)}^T \Sigma_{(0)}^{-1}(\lambda) X_{(0)}. \quad (5.14)$$

Damit lautet die Kovarianzmatrix des Schätzers $\hat{\beta}_{GLS}$ für den betrachteten Spezialfall

$$(X_{(0)}^T \Sigma_{(0)}^{-1}(\lambda) X_{(0)})^{-1} = \begin{bmatrix} 1 & -1 + \lambda \\ -1 + \lambda & 2 - 2\lambda \end{bmatrix}.$$

Wird nun in den Versuchsplan zusätzlich zu den Versuchspunkten $x_1 = 0$ und $x_2 = 1$ ein dritter Versuchspunkt x_3 an der Stelle $d = 1$ aufgenommen, so hat die neu hinzugekommene Zufallsvariable y_3 den gleichen Erwartungswert wie die Zufallsvariable y_2 . Wegen $\lambda^{x_3 - x_2} = \lambda^0 = 1$ weisen die Zufallsvariablen y_2 und y_3 eine Korrelation von 1 auf, so dass die Realisierungen von y_2 und y_3 stets identisch sind. Dies bedeutet aber, dass der Versuchspunkt x_3 keine zusätzlichen Informationen für die Schätzung der Parameter β_0 und β_1 liefert. Für die Suche nach einem optimalen Schätzer bedeutet dies wiederum, dass der Schätzer für die Situation mit drei Versuchspunkten mit dem Schätzer für die Situation mit zwei Versuchspunkten übereinstimmt. Selbiges gilt für den Fall, bei dem der dritte Versuchspunkt an der Stelle $d = 0$ platziert wird. Somit kann letztendlich für $\lambda \in (0, 1)$, wobei $d \in \{0, 1\}$ ist, die Kovarianzmatrix des verallgemeinerten Kleinste-Quadrate-Schätzers aus dem Fall, bei dem lediglich zwei Versuchspunkte vorliegen, übernommen werden. Damit lautet das D-Kriterium in diesem Fall

$$D_\lambda(d) = (\det(X_{(0)}^T \Sigma_{(0)}^{-1}(\lambda) X_{(0)})^{-1})^{1/2} = (1 - \lambda^2)^{1/2}. \quad (5.15)$$

Nun ist abschließend noch die Gestalt des D-Kriteriums für $\lambda = 1$ zu klären. Auch hier existiert der verallgemeinerte Kleinste-Quadrate-Schätzer nicht, so dass erneut

ein Schätzer zu bestimmen ist, der die Eigenschaften des verallgemeinerten Kleinst-Quadrat-Schätzers aufweist. Das heißt analog zu $\lambda < 1$ wird nun der Schätzvektor $\hat{\beta}_{\lambda=1}$ gesucht, für dessen Kovarianzmatrix im Vergleich zu einem beliebigen anderen linearen erwartungstreuen Schätzvektor $\tilde{\beta}_{\lambda=1}$ gilt, dass $\text{Cov}(\tilde{\beta}_{\lambda=1}) \geq \text{Cov}(\hat{\beta}_{\lambda=1})$. Der beste lineare erwartungstreue Schätzer für $\lambda = 1$ und $q = 3$ resultiert aus Behauptung 3.8. Für die Gestalt des D-Kriteriums ist nun entscheidend, dass dieser Schätzer wegen (3.11) die Kovarianzmatrix

$$\text{Cov}(\tilde{\beta}_{\lambda=1}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad (5.16)$$

aufweist. Es gilt für den besten linearen erwartungstreuen Schätzer somit stets, dass

$$D_{\lambda}(d) = (\det(\text{Cov}(\tilde{\beta}_{\lambda=1}))^{1/2}) = 0. \quad (5.17)$$

Da in der Behauptung 3.8 der Fall $q \geq 2$ betrachtet wird, gilt sogar für $q \geq 2$, dass der beste lineare erwartungstreue Schätzer die Kovarianzmatrix (5.16) aufweist. Für das D-Kriterium folgt dann, dass

$$D_{\lambda=1}(\xi) = (\det(\text{Cov}(\tilde{\beta}_{\lambda=1, \xi}))^{1/2}) = 0.$$

Dies bedeutet zum einen, dass für eine vorgegebene Anzahl von Versuchspunkten der D-optimale Versuchsplan nicht von der Wahl der Versuchspunkte abhängt. Darüber hinaus ist ein Plan mit zwei Versuchspunkten bezüglich des D-Kriteriums genauso gut wie ein Plan mit beliebig vielen Versuchspunkten. Somit liegt für $\lambda = 1$ die Besonderheit vor, dass sämtliche Versuchspläne ein D-Kriterium von 0 aufweisen, unabhängig von der Anzahl und der Wahl der Versuchspunkte. Die einzige Bedingung ist die, dass sich in allen Fällen mindestens zwei Versuchspunkte voneinander unterscheiden müssen, siehe Beweis von Behauptung 3.8.

Wegen (5.10), (5.11), (5.15) und (5.17) kann das D-Kriterium für $q = 3$ und $n = 1$ insgesamt wie folgt zusammengefasst werden:

$$D_{\lambda}(d) = \begin{cases} \left(\frac{1}{2(1-d+d^2)} \right)^{1/2} & \text{falls } \lambda = 0, \\ 0 & \text{falls } \lambda = 1, \\ (1-\lambda^2)^{1/2} & \text{falls } \lambda \in (0, 1), d \in \{0, 1\}, \\ \left(\frac{(1-\lambda^{2d})(1-\lambda^{2(1-d)})}{2((1-(1-d)\lambda^d)(1-d\lambda^{1-d})-d(1-d))} \right)^{1/2} & \text{sonst.} \end{cases} \quad (5.18)$$

Im Verlauf dieser Arbeit wird statt des D-Kriteriums der Einfachheit halber die Funktion

$$D_\lambda(d)^2 =: \tilde{D}_\lambda(d) \quad (5.19)$$

analysiert. Das hat den Vorteil, dass in allen Rechenschritten auf einen Exponenten verzichtet werden kann.

Abschließend wird auf die Fälle $n > 1$ eingegangen. Zwischen dem D-Kriterium für die Fälle $n > 1$ und dem D-Kriterium für $n = 1$ besteht der Zusammenhang

$$D_{\lambda, n > 1}(d) = \frac{1}{n} D_\lambda(d) \quad (5.20)$$

mit $D_\lambda(d)$ aus (5.18). Da die Größe n in (5.20) als multiplikativer Faktor einfließt, ist der D-optimale Versuchsplan unabhängig von der Größe n . Ausschlaggebend dafür ist zunächst die vorliegende Annahme, dass die Fehler verschiedener Versuchseinheiten voneinander unabhängig sind, so dass bezüglich der Versuchseinheiten unabhängige Zufallsvektoren vorliegen. Dies führt nun dazu, dass die Informationsmatrix für die Fälle $n > 1$ leicht aus dem Fall $n = 1$ berechnet werden kann, siehe (3.1).

5.3.2. Stetigkeit des D-Kriteriums

In diesem Abschnitt wird die Stetigkeit von (5.18) untersucht. Dies geschieht, indem die Stetigkeit des Ausdrucks $\tilde{D}_\lambda(d)$ analysiert wird. Da $D_\lambda(d)$ gerade die Wurzel von $\tilde{D}_\lambda(d)$ ist, übertragen sich alle Unstetigkeitsstellen von $\tilde{D}_\lambda(d)$ direkt auf $D_\lambda(d)$.

Die Suche nach Unstetigkeitsstellen geschieht nun, indem verschiedene Fälle betrachtet werden. Die Fälle lauten:

1. $\lambda = 0, d \in [0, 1]$,
2. $\lambda \in (0, 1], d \in \{0, 1\}$,
3. $\lambda = 1, d \in (0, 1)$,
4. $\lambda \in (0, 1), d \in (0, 1)$.

Fall 1 ($\lambda = 0, d \in [0, 1]$):

Für $\lambda = 0$ gilt, dass

$$\tilde{D}_{\lambda=0}(d) = \frac{1}{2(1-d+d^2)}. \quad (5.21)$$

Dieser Ausdruck ist stetig in d .

Fall 2 ($\lambda \in (0, 1]$, $d \in \{0, 1\}$):

Für $\lambda \in (0, 1]$ und $d \in \{0, 1\}$ gilt auf Grund von (5.18), dass

$$\tilde{D}_\lambda(d) = 1 - \lambda^2. \quad (5.22)$$

Das D-Kriterium ist damit für $d = 0$ in $\lambda = 0$ unstetig, da

$$\lim_{\lambda \rightarrow 0} \tilde{D}_\lambda(0) = 1 \neq 1/2 = \tilde{D}_{\lambda=0}(0).$$

Aus (5.22) ist dagegen ersichtlich, dass in allen weiteren $\lambda \in (0, 1]$ Stetigkeit in λ vorliegt.

Fall 3 ($\lambda = 1$, $d \in (0, 1)$):

Für $\lambda = 1$ und $d \in (0, 1)$ gilt auf Grund von (5.18), dass

$$\tilde{D}_{\lambda=1}(d) = 0. \quad (5.23)$$

Mit (5.18) gilt sogar für $d \in [0, 1]$, dass

$$\tilde{D}_{\lambda=1}(0) = 0.$$

Somit liegt für $\lambda = 1$ Stetigkeit in d vor.

Fall 4 ($\lambda \in (0, 1)$, $d \in (0, 1)$):

Behauptung 5.3. Sei $d \in (0, 1)$. Dann gilt:

$$\lim_{\lambda \rightarrow 1} \tilde{D}_\lambda(d) = 0. \quad (5.24)$$

Beweis. Mit der Regel von l'Hospital (*), siehe Hildebrandt (2006), S. 348, gilt, dass

$$\lim_{\lambda \rightarrow 1} \text{Var}(\hat{\beta}_1) = \lim_{\lambda \rightarrow 1} \frac{a_d(\lambda)}{b_d(\lambda)} \stackrel{*}{=} \lim_{\lambda \rightarrow 1} \frac{a_d^{(1)}(\lambda)}{b_d^{(1)}(\lambda)} = 0,$$

wobei

$$\begin{aligned} a_d(\lambda) &= \lambda^{4d} + (2 - 2\lambda)\lambda^{3d} + \lambda^2 + \lambda^{2d}(\lambda^2 - 3) + \lambda^d(2\lambda - 2\lambda^2), \\ b_d(\lambda) &= \lambda^{2d}(2\lambda - 2)d^2 + (-2\lambda^{3d} + (2 - 2\lambda)\lambda^{2d} + 2\lambda^{d+1})d + 2\lambda^{3d} - 2\lambda^{2d}, \\ a_d^{(1)}(\lambda) &= \frac{\partial}{\partial \lambda} a_d(\lambda) \\ &= 2\lambda^{d-1}(-3\lambda^{2d+1} + 2\lambda^{3d} + \lambda^{d+2} + 3\lambda^{2d} - 3\lambda^d - \lambda^2 + \lambda)d \\ &\quad - 2(\lambda^d - 1)(\lambda^{2d} - \lambda^{d+1} + \lambda^d + \lambda), \\ b_d^{(1)}(\lambda) &= \frac{\partial}{\partial \lambda} b_d(\lambda) \\ &= 2\lambda^{d-1}d((2\lambda^{d+1} - 2\lambda^d)d^2 + (2\lambda^d - 3\lambda^{2d} - \lambda^{d+1} + \lambda)d + 3\lambda^{2d} - \lambda^{d+1} - 2\lambda^d + \lambda) \end{aligned}$$

und

$$\begin{aligned} \lim_{\lambda \rightarrow 1} \text{Var}(\hat{\beta}_0) &= \lim_{\lambda \rightarrow 1} \frac{\lambda^d(\lambda^2 - 1)d^2 + (2\lambda - 2\lambda^{2d+1})d + \lambda^{3d} - \lambda^d}{\lambda^d(2\lambda - 2)d^2 + (-2\lambda^{2d} + (2 - 2\lambda)\lambda^d + 2\lambda)d + 2\lambda^{2d} - 2\lambda^d} \\ &\stackrel{*}{=} \lim_{\lambda \rightarrow 1} \frac{(\lambda^{d+1} - \lambda^{d-1})d^3 - (4\lambda^{2d} + 2\lambda^{d+1})d^2 + (3\lambda^{3d-1} - 2\lambda^{2d} - \lambda^{d-1} + 2)d}{(2\lambda^d - 2\lambda^{d-1})d^3 + (2\lambda^{d-1} - 4\lambda^{2d-1})d^2 + (4\lambda^{2d-1} - 2\lambda^d - 2\lambda^{d-1} + 2)d} \\ &= 1. \end{aligned}$$

Wegen (3.19) gilt, dass

$$\lim_{\lambda \rightarrow 1} \text{Kov}(\hat{\beta}_0, \hat{\beta}_1) = 0.$$

Daher folgt insgesamt, dass das D-Kriterium für $\lambda \rightarrow 1$ den Wert 0 hat. \square

Aus (5.23) und (5.24) folgt für $d \in (0, 1)$, dass

$$\lim_{\lambda \rightarrow 1} \tilde{D}_\lambda(d) = \tilde{D}_{\lambda=1}(d).$$

Somit ist das D-Kriterium für $d \in (0, 1)$ an der Stelle $\lambda = 1$ stetig in λ .

Behauptung 5.4. Sei $\lambda \in (0, 1)$. Für beliebiges $d \in (0, 1)$ gilt:

$$\lim_{\lambda \rightarrow 0} \tilde{D}_\lambda(d) = \frac{1}{2(1 - d + d^2)}. \quad (5.25)$$

Beweis.

$$\begin{aligned} \lim_{\lambda \rightarrow 0} \tilde{D}_\lambda(d) &= \lim_{\lambda \rightarrow 0} \frac{(1 - \lambda^{2d})(1 - \lambda^{2(1-d)})}{2((1 - (1 - d)\lambda^d)(1 - d\lambda^{1-d}) - d(1 - d))} \\ &= \frac{1}{2(1 - d(1 - d))} \\ &= \frac{1}{2(1 - d + d^2)}. \end{aligned}$$

\square

Dies bedeutet, dass für beliebiges $d \in (0, 1)$ das D-Kriterium an der Stelle $\lambda = 0$ stetig in λ ist, da $\tilde{D}_\lambda(d)$ für $\lambda \rightarrow 0$ gerade (5.21) entspricht.

Behauptung 5.5. Sei $d \in (0, 1)$. Für beliebiges $\lambda \in (0, 1)$ gilt:

$$\lim_{d \rightarrow 0} \tilde{D}_\lambda(d) = 1 - \lambda^2. \quad (5.26)$$

Beweis. Es gilt:

$$\tilde{D}_\lambda(d) = \frac{(1 - \lambda^{2d})(1 - \lambda^{2(1-d)})}{2((1 - (1-d)\lambda^d)(1 - d\lambda^{1-d}) - d(1-d))}.$$

Zu bestimmen ist der Grenzwert $\lim_{d \rightarrow 0} \tilde{D}_\lambda(d)$. Sowohl Zähler als auch Nenner haben für $d \rightarrow 0$ den Grenzwert Null, da

$$\lim_{d \rightarrow 0} (1 - \lambda^{2d})(1 - \lambda^{2(1-d)}) = \lim_{d \rightarrow 0} 1 - \lambda^{2d} - \lambda^{2-2d} + \lambda^2 = 1 - \lambda^0 - \lambda^2 + \lambda^2 = 0$$

und

$$\lim_{d \rightarrow 0} 2(1 - \lambda^{1-d}d - \lambda^d + \lambda d + \lambda^d d - \lambda d^2 - d + d^2) = 2(1 - 1) = 0.$$

Mit der Regel von l'Hospital (*) folgt:

$$\begin{aligned} \lim_{d \rightarrow 0} \tilde{D}_\lambda(d) &= \lim_{d \rightarrow 0} \frac{(1 - \lambda^{2d})(1 - \lambda^{2(1-d)})}{2((1 - (1-d)\lambda^d)(1 - d\lambda^{1-d}) - d(1-d))} \\ &\stackrel{*}{=} \lim_{d \rightarrow 0} \frac{2\ln(\lambda)(\lambda^{2-2d} - \lambda^{2d})}{2((\lambda^d \ln(\lambda) + \lambda^{1-d} \ln(\lambda) - 2\lambda + 2)d + \lambda^d - \lambda^{1-d} - \lambda^d \ln(\lambda) + \lambda - 1)} \\ &= \frac{2\ln(\lambda)(\lambda^2 - 1)}{-2\ln(\lambda)} \\ &= 1 - \lambda^2. \end{aligned}$$

□

Es ist zu erkennen, dass der Grenzwert aus Behauptung 5.5 mit dem Funktionswert (5.22) übereinstimmt. Folglich ist das D-Kriterium für $\lambda \in (0, 1)$ an den Stellen $d = 0$ bzw. $d = 1$ stetig in d . Da das D-Kriterium für $d \in (0, 1)$ und $\lambda \in (0, 1)$ eine Komposition stetiger Funktionen ist, kann insgesamt festgehalten werden, dass das D-Kriterium bis auf die eine Ausnahme stetig ist. Eine anschauliche Erklärung für dieses Resultat ist die Folgende. Wenn einerseits zwei Versuchspunkte identisch sind und der Korrelationsparameter λ größer Null ist, so hat die Korrelation der beiden Zufallsexperimente den Wert $\lambda^0 = 1$. Dies hat dann die Konsequenz, dass die beiden Ergebnisse der Experimente identisch sind. Folglich liegen nur zwei an Stelle von drei Versuchen vor. Wenn dagegen die Korrelation gleich Null ist, so führen zwei Versuche an der gleichen Stelle zu verschiedenen Beobachtungen. Dies führt letztendlich zur Unstetigkeit des D-Kriteriums. Die hier gewonnenen Erkenntnisse sind in Abschnitt 5.2 von Bedeutung.

5.3.3. Grenzfall $\lambda \rightarrow 1$

In Dette *et al.* (2008) wird gezeigt, dass für $\lambda \rightarrow 1$ der D-optimale Versuchsplan gegen den äquidistanten Versuchsplan konvergiert, also gegen den Plan mit $d = 1/2$. Behauptung 5.7 zeigt ergänzend, dass für $\lambda \rightarrow 1$ jeder Versuchsplan mit $d \in [0, 1]$ eine Effizienz von 1 aufweist. Zuvor wird Behauptung 5.6 formuliert.

Behauptung 5.6. *Für beliebiges $d \in (0, 1)$ gilt:*

$$\lim_{\lambda \rightarrow 1} \frac{\tilde{D}_\lambda(d)}{\tilde{D}_\lambda(0)} = 1.$$

Beweis.

$$\begin{aligned} \lim_{\lambda \rightarrow 1} \frac{\tilde{D}_\lambda(d)}{\tilde{D}_\lambda(0)} &= \lim_{\lambda \rightarrow 1} \frac{(1 - \lambda^{2d})(1 - \lambda^{2(1-d)})}{2(1 - \lambda^2)((1 - (1 - d)\lambda^d)(1 - d\lambda^{1-d}) - d(1 - d))} \\ &= \lim_{\lambda \rightarrow 1} \frac{a_d(\lambda)}{b_d(\lambda)}. \end{aligned}$$

Dabei ist

$$a_d(\lambda) = (1 - \lambda^{2d})(1 - \lambda^{2(1-d)})$$

und

$$b_d(\lambda) = 2(1 - \lambda^2)((1 - (1 - d)\lambda^d)(1 - d\lambda^{1-d}) - d(1 - d)).$$

Wegen $\lim_{\lambda \rightarrow 1} a_d(\lambda) = \lim_{\lambda \rightarrow 1} b_d(\lambda) = 0$ gilt mit der Regel von l'Hospital:

$$\lim_{\lambda \rightarrow 1} \frac{a_d(\lambda)}{b_d(\lambda)} = \lim_{\lambda \rightarrow 1} \frac{a_d^{(1)}(\lambda)}{b_d^{(1)}(\lambda)},$$

falls der Grenzwert auf der rechten Seite der Gleichung existiert. Dabei ist

$$\begin{aligned} a_d^{(1)}(\lambda) &= \frac{\partial}{\partial \lambda} a_d(\lambda) \\ &= -2\lambda^{-2d-1}(\lambda^{4d}d - \lambda^{2d}d - \lambda^{2d+2} + \lambda^2) \end{aligned}$$

und

$$\begin{aligned} b_d^{(1)}(\lambda) &= \frac{\partial}{\partial \lambda} b_d(\lambda) \\ &= -\lambda^{-d-1}((2\lambda^3 + \lambda^d(-6\lambda^3 + 4\lambda^2 + 2\lambda) + \lambda^{2d}(2\lambda^2 - 2) - 2\lambda)d^2 \\ &\quad + (\lambda^d(6\lambda^3 - 4\lambda^2 - 2\lambda) - 6\lambda^3 + \lambda^{2d}(2\lambda^2 + 2) + 2\lambda)d - 4\lambda^{2d+2} + 4\lambda^{d+2}). \end{aligned}$$

Nun ist $\lim_{\lambda \rightarrow 1} a_d^{(1)}(\lambda) = \lim_{\lambda \rightarrow 1} b_d^{(1)}(\lambda) = 0$, so dass mit der Regel von l'Hospital erneut gilt:

$$\lim_{\lambda \rightarrow 1} \frac{a_d^{(1)}(\lambda)}{b_d^{(1)}(\lambda)} = \lim_{\lambda \rightarrow 1} \frac{a_d^{(2)}(\lambda)}{b_d^{(2)}(\lambda)}.$$

Dabei ist

$$\begin{aligned} a_d^{(2)}(\lambda) &= \frac{\partial^2}{\partial \lambda^2} a_d(\lambda) \\ &= -2\lambda^{-2d-2}(2\lambda^{4d}d^2 + 2\lambda^2d^2 - \lambda^{4d}d - 3\lambda^2d - \lambda^{2d+2} + \lambda^2) \end{aligned}$$

und

$$\begin{aligned} b_d^{(2)}(\lambda) &= \frac{\partial^2}{\partial \lambda^2} b_d(\lambda) \\ &= -2\lambda^{-d-2}((-\lambda^3 + \lambda^{2d}(\lambda^2 - 1) + \lambda)d^3 + (5\lambda^3 + 2\lambda^d(\lambda^2 - 3\lambda^3) \\ &\quad + 2\lambda^{2d}(\lambda^2 + 1) - \lambda)d^2 + (2\lambda^d(3\lambda^3 - \lambda^2) - 6\lambda^3 + \lambda^{2d}(-\lambda^2 - 1))d \\ &\quad - 2\lambda^{2d+2} + 2\lambda^{d+2}). \end{aligned}$$

Wegen $\lim_{\lambda \rightarrow 1} a_d^{(2)}(\lambda) = \lim_{\lambda \rightarrow 1} b_d^{(2)}(\lambda) = -8d^2 + 8d$ gilt schließlich:

$$\lim_{\lambda \rightarrow 1} \frac{\tilde{D}_\lambda(d)}{\tilde{D}_\lambda(0)} = \lim_{\lambda \rightarrow 1} \frac{a_d(\lambda)}{b_d(\lambda)} = \lim_{\lambda \rightarrow 1} \frac{a_d^{(1)}(\lambda)}{b_d^{(1)}(\lambda)} = \lim_{\lambda \rightarrow 1} \frac{a_d^{(2)}(\lambda)}{b_d^{(2)}(\lambda)} = \frac{-8d^2 + 8d}{-8d^2 + 8d} = 1.$$

□

Behauptung 5.7. Für beliebiges $d \in [0, 1]$ gilt:

$$\lim_{\lambda \rightarrow 1} \frac{\tilde{D}_\lambda(d^*)}{\tilde{D}_\lambda(d)} = 1. \quad (5.27)$$

Beweis. **Fall 1** ($d \in (0, 1)$):

Mit Behauptung 5.6 gilt, dass

$$\begin{aligned} 1 &= \frac{\lim_{\lambda \rightarrow 1} \left(\frac{\tilde{D}_\lambda(d^*)}{\tilde{D}_\lambda(0)} \right)}{\lim_{\lambda \rightarrow 1} \left(\frac{\tilde{D}_\lambda(d)}{\tilde{D}_\lambda(0)} \right)} \\ &= \left(\lim_{\lambda \rightarrow 1} \left(\frac{\tilde{D}_\lambda(d^*)}{\tilde{D}_\lambda(0)} \right) \right) \left(\lim_{\lambda \rightarrow 1} \left(\frac{\tilde{D}_\lambda(0)}{\tilde{D}_\lambda(d)} \right) \right) \\ &= \lim_{\lambda \rightarrow 1} \left(\frac{\tilde{D}_\lambda(d^*)}{\tilde{D}_\lambda(0)} \cdot \frac{\tilde{D}_\lambda(0)}{\tilde{D}_\lambda(d)} \right) \\ &= \lim_{\lambda \rightarrow 1} \left(\frac{\tilde{D}_\lambda(d^*)}{\tilde{D}_\lambda(d)} \right). \end{aligned}$$

Fall 2 ($d \in \{0, 1\}$):

Für $d^* \in \{0, 1\}$ gilt (5.27) direkt. Für $d^* \in (0, 1)$ gilt (5.27) auf Grund von Behauptung 5.6. \square

5.3.4. Abschätzungen

Für die Beweise der Behauptungen in Abschnitt 5.3.5 werden zahlreiche Abschätzungen herangezogen. Vorgehen, die mehrfach Anwendung finden, werden nun vorgestellt, so dass die eigentlichen Beweise abgekürzt werden können. Darüber hinaus enthält dieser Abschnitt Abschätzungen der Terme $\ln(2)$ und $\ln(5)$, welche im weiteren Verlauf ebenfalls mehrfach herangezogen werden.

Vorgehen 1:

Vorgehen 1 verwendet Behauptung 5.8.

Behauptung 5.8. *Das Krümmungsverhalten der Funktion $f : [x_1, x_2] \rightarrow \mathbb{R}$ sei in $[x_1, x_2]$ bekannt. Dann gilt für alle $x \in [x_1, x_2]$:*

a)

$$f(x) \geq 0, \text{ falls } f(x_1) \geq 0, f(x_2) \geq 0 \text{ sowie } f(x) \text{ in } [x_1, x_2] \text{ konkav,} \quad (5.28)$$

b)

$$f(x) \leq 0, \text{ falls } f(x_1) \leq 0, f(x_2) \leq 0 \text{ sowie } f(x) \text{ in } [x_1, x_2] \text{ konvex.} \quad (5.29)$$

Beweis. Folgt direkt aus der Definition der Konkavität bzw. Konvexität. \square

Vorgehen 2:

Vorgehen 2 zieht das Resultat aus Behauptung 5.9 heran. Der Teil a) von Behauptung 5.9 wird in Abbildung 5.5 veranschaulicht.

Behauptung 5.9. *Die Funktion $f : [x_1, x_2] \rightarrow \mathbb{R}$ sei konvex. Sei $a \geq 0$. Dann gilt für alle $x \in [x_1, x_2]$:*

a)

$$f(x) \leq f(x_1) + a - x_1 \frac{f(x_2) - f(x_1) - a}{x_2 - x_1} + x \frac{f(x_2) - f(x_1) - a}{x_2 - x_1} =: g_1(x), \quad (5.30)$$

b)

$$f(x) \leq f(x_2) + a - x_2 \frac{f(x_2) + a - f(x_1)}{x_2 - x_1} + x \frac{f(x_2) + a - f(x_1)}{x_2 - x_1} =: g_2(x). \quad (5.31)$$

Sei $f : [x_1, x_2] \rightarrow \mathbb{R}$ konkav. Dann gilt für alle $x \in [x_1, x_2]$:

c)

$$f(x) \geq f(x_1) - a - x_1 \frac{f(x_2) - f(x_1) + a}{x_2 - x_1} + x \frac{f(x_2) - f(x_1) + a}{x_2 - x_1} =: g_3(x), \quad (5.32)$$

d)

$$f(x) \geq f(x_2) - a - x_2 \frac{f(x_2) - a - f(x_1)}{x_2 - x_1} + x \frac{f(x_2) - a - f(x_1)}{x_2 - x_1} =: g_4(x). \quad (5.33)$$

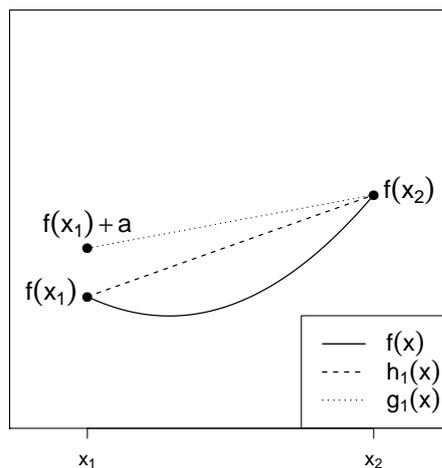


Abbildung 5.5: Die Gerade $h_1(x)$ schneidet die Funktion $f(x)$ an den Stellen x_1 und x_2 . Auf Grund der Konkavität von $f(x)$ stellt die Gerade $h_1(x)$ eine obere Schranke für $f(x)$ dar. Die Gerade $g_1(x)$ schneidet die Funktionen $f(x)$ und $h_1(x)$ an der Stelle x_2 und stellt im Vergleich zu $h_1(x)$ eine konservativere Schranke dar, siehe Behauptung 5.9.

Beweis. Zu a):

Die Gerade $h_1(x) := f(x_1) - x_1 \frac{f(x_2) - f(x_1)}{x_2 - x_1} + x \frac{f(x_2) - f(x_1)}{x_2 - x_1}$ schneidet die Funktion $f(x)$ an den Stellen x_1 und x_2 . Somit gilt für $x \in [x_1, x_2]$ auf Grund der Konkavität von $f(x)$, dass

$$h_1(x) \geq f(x). \quad (5.34)$$

Wegen $g_1(x) - h_1(x) = a(x_2 - x)/(x_2 - x_1)$ gilt für $x \leq x_2$, dass

$$g_1(x) \geq h_1(x). \quad (5.35)$$

Somit folgt für $x \in [x_1, x_2]$ wegen (5.34) und (5.35) insgesamt, dass

$$g_1(x) \geq f(x).$$

Zu b), c) und d):

Analog zum Beweis von a). □

Vorgehen 3:

Vorgehen 3 verwendet Behauptung 5.10. Die Idee wird in Abbildung 5.6 veranschaulicht.

Behauptung 5.10. *Die Funktion $f : [x_1, x_2] \rightarrow \mathbb{R}$ sei konkav. Definiere $\frac{\partial}{\partial x}f(x) =: f^{(1)}(x)$. Dann gilt für alle $x \in [x_1, x_2]$:*

a)

$$f(x) \leq f(x_1) - x_1 f^{(1)}(x_1) + x f^{(1)}(x_1) =: g_5(x), \quad (5.36)$$

b)

$$f(x) \geq f(x_2) - x_2 f^{(1)}(x_1) + x f^{(1)}(x_1) =: g_6(x), \quad (5.37)$$

c)

$$f(x) \leq f(x_2) - x_2 f^{(1)}(x_2) + x f^{(1)}(x_2) =: g_7(x), \quad (5.38)$$

d)

$$f(x) \geq f(x_1) - x_1 f^{(1)}(x_2) + x f^{(1)}(x_2) =: g_8(x). \quad (5.39)$$

Die Funktion $f : [x_1, x_2] \rightarrow \mathbb{R}$ sei konvex. Dann gilt für alle $x \in [x_1, x_2]$:

e)

$$f(x) \geq f(x_1) - x_1 f^{(1)}(x_1) + x f^{(1)}(x_1) =: g_9(x), \quad (5.40)$$

f)

$$f(x) \leq f(x_2) - x_2 f^{(1)}(x_1) + x f^{(1)}(x_1) =: g_{10}(x), \quad (5.41)$$

g)

$$f(x) \geq f(x_2) - x_2 f^{(1)}(x_2) + x f^{(1)}(x_2) =: g_{11}(x), \quad (5.42)$$

h)

$$f(x) \leq f(x_1) - x_1 f^{(1)}(x_2) + x f^{(1)}(x_2) =: g_{12}(x). \quad (5.43)$$

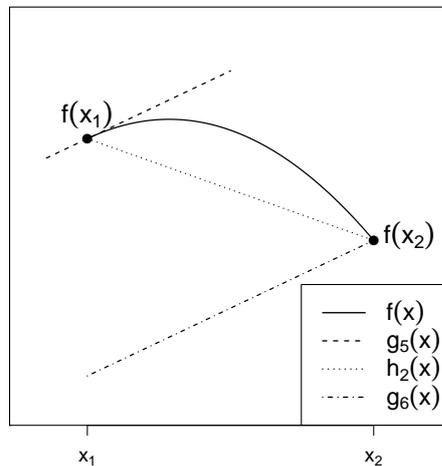


Abbildung 5.6: Die Gerade $g_5(x)$ stellt die Tangente von $f(x)$ an der Stelle x_1 dar. Auf Grund der Konkavität von $f(x)$ stellt $h_2(x)$ eine untere Schranke von $f(x)$ dar. Die Gerade $g_6(x)$ weist die Steigung von $g_5(x)$ auf und schneidet $h_2(x)$ an der Stelle x_2 . Da $g_6(x)$ eine untere Schranke von $h_2(x)$ ist, ist $g_6(x)$ eine weitere untere Schranke von $f(x)$, siehe Behauptung 5.10.

Beweis. Zu a):

Die Gerade $g_5(x)$ schneidet $f(x)$ an der Stelle x_1 . Da die Steigung von $g_5(x)$ mit der ersten Ableitung von $f(x)$ an der Stelle x_1 übereinstimmt, gilt auf Grund der Konkavität von $f(x)$, dass

$$g_5(x) \geq f(x).$$

Zu b):

Definiere

$$h_2(x) := f(x_2) - x_2 \frac{f(x_2) - f(x_1)}{x_2 - x_1} + x \frac{f(x_2) - f(x_1)}{x_2 - x_1}. \quad (5.44)$$

Die Gerade $h_2(x)$ schneidet $f(x)$ an den Stellen x_1 und x_2 . Wegen der beiden Schnittstellen und der Konkavität gilt, dass

$$f(x) \geq h_2(x). \quad (5.45)$$

Wegen der Konkavität von $f(x)$ in $[x_1, x_2]$ gilt außerdem, dass

$$\max_x f^{(1)}(x) = f^{(1)}(x_1). \quad (5.46)$$

Nun gilt auf Grund des Mittelwertsatzes der Differentialrechnung, siehe Hildebrandt (2006), S. 204:

$$\exists x_0 \in [x_1, x_2] \text{ mit } f^{(1)}(x_0) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}. \quad (5.47)$$

Wegen (5.46), (5.47) und $x_2 \geq x$ kann jetzt gefolgert werden, dass

$$h_2(x) - g_6(x) = (x_2 - x)(f^{(1)}(x_1) - \frac{f(x_2) - f(x_1)}{x_2 - x_1}) \geq 0. \quad (5.48)$$

Abschließend gilt wegen (5.45) und (5.48), dass

$$f(x) \geq g_6(x).$$

Zu c), e) und g):

Analog zum Beweis von a).

Zu d), f) und h):

Analog zum Beweis von b). □

Vorgehen 4:

Vorgehen 4 verwendet Behauptung 5.11. Abbildung 5.7 liefert eine Veranschaulichung.

Behauptung 5.11. *Sind $f : [x_1, x_2] \rightarrow \mathbb{R}$ und $g_{13} : [x_1, x_2] \rightarrow \mathbb{R}$ stetig mit $f(x_1) \geq g_{13}(x_1)$, $f(x_2) \geq g_{13}(x_2)$ und $\frac{\partial^2}{\partial x^2} f(x) \leq \frac{\partial^2}{\partial x^2} g_{13}(x)$ für $x \in [x_1, x_2]$. Dann gilt für alle $x \in [x_1, x_2]$, dass $f(x) \geq g_{13}(x)$.*

Beweis. Definiere $h_3(x) := f(x) - g_{13}(x)$. Wegen

$$\frac{\partial^2}{\partial x^2} f(x) \leq \frac{\partial^2}{\partial x^2} g_{13}(x)$$

gilt

$$\frac{\partial^2}{\partial x^2} h_3(x) = \frac{\partial^2}{\partial x^2} f(x) - \frac{\partial^2}{\partial x^2} g_{13}(x) \leq 0.$$

Somit ist $h_3(x)$ konkav. Wegen $h_3(x_1) \geq 0$, $h_3(x_2) \geq 0$ und der Konkavität von $h_3(x)$ gilt schließlich, dass $h_3(x) \geq 0$. Daraus folgt schließlich die Gültigkeit der Behauptung. □

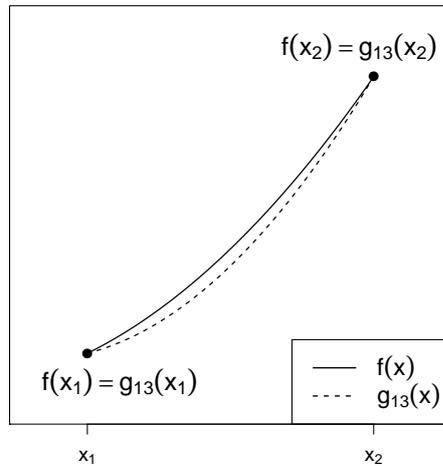


Abbildung 5.7: Die Funktion $f(x)$ wird durch $g_{13}(x)$ nach unten abgeschätzt. Beide Funktionen haben an den Stellen x_1 und x_2 identische Funktionswerte. Anhand der zweiten Ableitungen kann letztendlich gefolgert werden, dass $f(x)$ für $x \in [x_1, x_2]$ größere oder gleich große Funktionswerte liefert, siehe Behauptung 5.11.

Vorgehen 5:

Das fünfte Vorgehen bedient sich Behauptung 5.12.

Behauptung 5.12. Sind $f : [x_1, x_2] \rightarrow \mathbb{R}$ und $g_{14} : [x_1, x_2] \rightarrow \mathbb{R}$ differenzierbar mit $f(x_2) \leq g_{14}(x_2)$ und $\frac{\partial}{\partial x} f(x) \geq \frac{\partial}{\partial x} g_{14}(x)$ für $x \in [x_1, x_2]$. Dann gilt für alle $x \in [x_1, x_2]$, dass $f(x) \leq g_{14}(x)$.

Beweis. Sei $x_0 \in [x_1, x_2]$ beliebig. Es gilt:

$$\begin{aligned} g_{14}(x_0) - f(x_0) &\geq f(x_2) - g_{14}(x_2) - (f(x_0) - g_{14}(x_0)) \\ &= \int_{x_0}^{x_2} \underbrace{\left(\frac{\partial}{\partial x} f(x) - \frac{\partial}{\partial x} g_{14}(x) \right)}_{\geq 0} dx \\ &\geq 0. \end{aligned}$$

□

In anstehenden Beweisen wiederholt sich ein weiteres Vorgehen mehrfach. Daher wird es in der folgenden Behauptung einmal allgemein vorgestellt.

Vorgehen 6:

Behauptung 5.13. Sei $x \in [0, 1]$. Gegeben sei das Polynom

$$\sum_{i=0}^n a_i x^i =: R. \quad (5.49)$$

Falls für alle $j \in \{1, \dots, n\}$ mit $a_j < 0$ gilt, dass $\sum_{i=0}^j a_i \geq 0$, so folgt, dass $R \geq 0$.

Beweis. Sei $i < i'$. Dann gilt $x^i \geq x^{i'}$. Die Menge $A_1 = \{a_0, \dots, a_n\}$ setzt sich aus den Mengen $P = \{a_i : a_i \geq 0\}$ und $N = \{a_i : a_i < 0\}$ zusammen. Falls $a_n, a_{n-1}, \dots, a_{n-q}$ nicht negativ sind, so gilt, dass

$$R \geq \sum_{i=0}^{n-q-1} a_i x^i =: S.$$

Definiere $n - q - 1 =: m$. Es gilt $a_0 \geq 0$ und $a_m < 0$. Die Menge $\{a_0, \dots, a_m\} =: A_2$ wird nun in die Teilmengen

$$\begin{aligned} P_1 &:= \{a_0, \dots, a_{n_1}\} \\ N_1 &:= \{a_{n_1+1}, \dots, a_{n_2}\} \\ P_2 &:= \{a_{n_2+1}, \dots, a_{n_3}\} \\ &\vdots \\ P_z &:= \{a_{n_{2z-2}+1}, \dots, a_{n_{2z-1}}\} \\ N_z &:= \{a_{n_{2z-1}+1}, \dots, a_{n_{2z}}\} \end{aligned}$$

zerlegt. Dabei seien die Mengen P_i und N_i , $1 \leq i \leq z$, so, dass die Elemente der Mengen P_i allesamt größergleich 0 sind, während die Elemente der Mengen N_i negativ sind. Für die jeweiligen Elemente der Mengen P_i und N_i gelte dabei außerdem, dass sie bereits in der Menge A_2 benachbart sind. Wegen $a_0 \geq 0$ und $a_m < 0$ stimmt die Anzahl an Mengen mit Elementen positiver Vorzeichen mit der Anzahl an Mengen negativer Vorzeichen überein.

Es gilt:

$$S = S_1 + S_N + S_P$$

mit

$$\begin{aligned} S_1 &= \sum_{i=0}^{n_1} a_i x^i \\ S_N &= \sum_{k=1}^z \sum_{i=n_{2k-1}+1}^{n_{2k}} a_i x^i \\ S_P &= \sum_{k=2}^z \sum_{i=n_{2k-2}+1}^{n_{2k-1}} a_i x^i. \end{aligned}$$

Es gilt, dass S_1 positiv ist. Da außerdem alle Summanden von S_P positiv und die von S_N negativ sind, gilt

$$\begin{aligned} S &\geq (x^{n_1} \sum_{i=0}^{n_1} a_i) + \left(\sum_{k=1}^z x^{n_{2k-1}} \sum_{i=n_{2k-1}+1}^{n_{2k}} a_i \right) + \left(\sum_{k=2}^z x^{n_{2k-1}} \sum_{i=n_{2k-2}+1}^{n_{2k-1}} a_i \right) \\ &= (x^{n_1} \sum_{i=0}^{n_2} a_i) + \left(\sum_{k=2}^z x^{n_{2k-1}} \sum_{i=n_{2k-2}+1}^{n_{2k}} a_i \right) \\ &=: T. \end{aligned}$$

Definiere für $2 \leq m \leq z$ die Partialsummen

$$T_m := (x^{n_1} \sum_{i=0}^{n_2} a_i) + \left(\sum_{k=2}^m x^{n_{2k-1}} \sum_{i=n_{2k-2}+1}^{n_{2k}} a_i \right).$$

Für $m = 2$ gilt:

$$\begin{aligned} T_2 &= (x^{n_1} \sum_{i=0}^{n_2} a_i) + (x^{n_3} \sum_{i=n_2+1}^{n_4} a_i) \\ &\geq x^{n_3} \sum_{i=0}^{n_4} a_i \\ &\geq 0. \end{aligned}$$

Nun gilt induktiv, dass sämtliche Partialsummen T_m positiv sind. Wegen $T_z = T$ gilt die Behauptung. \square

Abschätzung 1:

Lemma 5.14. *Es gelten die Abschätzungen*

$$\frac{8}{5} < \ln(5) < \frac{21}{13}. \quad (5.50)$$

Beweis. Zwischen dem Areatangens Hyperbolicus und dem natürlichen Logarithmus besteht der Zusammenhang

$$\operatorname{artanh}(x) = \frac{1}{2} \ln \left(\frac{1+x}{1-x} \right) \text{ für } |x| < 1,$$

siehe Stöcker (2009), S. 237. Zudem gilt die Reihendarstellung

$$\operatorname{artanh}(x) = \sum_{k=0}^{\infty} \frac{1}{2k+1} x^{2k+1} \text{ für } |x| < 1,$$

siehe Stöcker (2009), S. 239. Somit gilt auch

$$\ln \left(\frac{1+x}{1-x} \right) = \sum_{k=0}^{\infty} \frac{2}{2k+1} x^{2k+1} \text{ für } |x| < 1.$$

Die Substitution $(1+x)/(1-x) =: z$ liefert die Darstellung

$$\ln(z) = \sum_{k=0}^{\infty} \frac{2}{2k+1} \left(\frac{z-1}{z+1} \right)^{2k+1} \text{ für } z > 0. \quad (5.51)$$

Nun ist es möglich den Term $\ln(5)$ nach unten abzuschätzen. Wegen

$$\ln(5) - \frac{8}{5} = \sum_{k=0}^{\infty} \frac{2}{2k+1} \left(\frac{2}{3} \right)^{2k+1} - \frac{8}{5} > \sum_{k=0}^3 \frac{2}{2k+1} \left(\frac{2}{3} \right)^{2k+1} - \frac{8}{5} = \frac{4}{15309} > 0$$

gilt

$$\ln(5) > 8/5.$$

Außerdem gilt

$$e^{21/13} = \sum_{k=0}^{\infty} \frac{(21/13)^k}{k!} > \sum_{k=0}^6 \frac{(21/13)^k}{k!} = 5 \frac{8777701}{386144720} > 5.$$

Daraufhin kann gefolgert werden, dass

$$\ln(5) < 21/13.$$

□

Abschätzung 2:

Lemma 5.15. *Es gelten die Abschätzungen*

$$\frac{9}{13} < \ln(2) < \frac{25}{36}. \quad (5.52)$$

Beweis. Für den Beweis können wie im Beweis von Lemma 5.14 die Reihendarstellung (5.51) und die Reihendarstellung der Exponentialfunktion herangezogen werden.

Zunächst wird der Term $\ln(2)$ nach unten abgeschätzt. Wegen

$$\ln(2) - \frac{9}{13} = \sum_{k=0}^{\infty} \frac{2}{2k+1} \left(\frac{1}{3}\right)^{2k+1} - \frac{9}{13} > \sum_{k=0}^2 \frac{2}{2k+1} \left(\frac{1}{3}\right)^{2k+1} - \frac{9}{13} = \frac{11}{15795} > 0$$

gilt

$$\ln(2) > 9/13.$$

Außerdem gilt

$$e^{25/36} = \sum_{k=0}^{\infty} \frac{(25/36)^k}{k!} > \sum_{k=0}^4 \frac{(25/36)^k}{k!} = 2 \frac{43441}{40310784} > 2.$$

Daraufhin kann gefolgert werden, dass

$$\ln(2) < 25/36.$$

□

In manchen Beweisen, die noch folgen, bietet es sich aus Gründen der Vereinfachung an, im Vergleich zu Lemma 5.14 und Lemma 5.15 konservativere Abschätzungen heranzuziehen. Diese werden nun formuliert.

Abschätzung 3:

Lemma 5.16. *Es gelten die Abschätzungen*

$$\frac{3}{2} < \ln(5) < \frac{13}{7}. \quad (5.53)$$

Beweis. Folgt aus Lemma 5.14, da $3/2 < 8/5$ und $21/13 < 13/7$. □

Abschätzung 4:

Lemma 5.17. *Es gelten die Abschätzungen*

$$\frac{11}{16} < \ln(2) < \frac{7}{10}. \quad (5.54)$$

Beweis. Folgt aus Lemma 5.15, da $11/16 < 9/13$ und $25/36 < 7/10$. □

5.3.5. Schranken für $\tilde{D}_\lambda(d)$

In diesem Abschnitt werden eine obere sowie eine untere Schranke für die Funktion $\tilde{D}_\lambda(d) = (D_\lambda(d))^2$ mit $D_\lambda(d)$ aus (5.18) bestimmt. Die Formulierung der Schranken erfolgt letztendlich in Korollar 5.20, Behauptung 5.22 und Behauptung 5.23. Die Abbildung 5.2 auf Seite 27 liefert eine Veranschaulichung der Schranken. Sie sind für Abschnitt 5.2 relevant.

Eine obere Schranke in Abhängigkeit von λ

Behauptung 5.18. *Die Versuchspläne mit $d = 0$ und $d = 1$ minimieren $\tilde{D}_{\lambda=0}(d)$, wobei $\tilde{D}_{\lambda=0}(0) = \tilde{D}_{\lambda=0}(1) = 1/2$. Der Versuchsplan mit $d = 1/2$ maximiert $\tilde{D}_{\lambda=0}(d)$. Dabei ist $\tilde{D}_{\lambda=0}(1/2) = 2/3$.*

Beweis. Die Ableitung des Nenners von $\tilde{D}_{\lambda=0}(d)$ lautet $-2 + 4d$, so dass an der Stelle $d = 1/2$ die einzige lokale Extremstelle zu finden ist. Der Rand und die lokale Extremstelle sind somit die absoluten Extrema mit $\tilde{D}_{\lambda=0}(0) = \tilde{D}_{\lambda=0}(1) = 1/2$ und $\tilde{D}_{\lambda=0}(1/2) = 2/3$. \square

Behauptung 5.19. *Seien $\lambda \in (0, 1]$ und $d \in (0, 1)$. Dann gilt:*

$$\tilde{D}_\lambda(0) \geq \tilde{D}_\lambda(d). \quad (5.55)$$

Beweis. Fall 1 ($\lambda \in (0, 1)$):

Es besteht der Zusammenhang

$$\det(\mathcal{I}_{(1)}([\beta_0, \beta_1]^T)) = \det(\mathcal{I}_{(0)}([\beta_0, \beta_1]^T)) \left[1 + \frac{1}{1 - s^T \Sigma_{(0)}^{-1}(\lambda) s} p^T (\mathcal{I}_{(0)}([\beta_0, \beta_1]^T))^{-1} p \right] \quad (5.56)$$

mit $s = [\lambda^d, \lambda^{1-d}]^T$ und $p = [1, d]^T - X_{(0)}^T \Sigma_{(0)}^{-1}(\lambda) s$, siehe Fedorov (1996), S. 524. Für die Matrix $(\mathcal{I}_{(0)}([\beta_0, \beta_1]^T))^{-1}$ existiert auf Grund ihrer Symmetrie eine Darstellung der Form $R^T R$. Wegen

$$\frac{1}{1 - s^T \Sigma_{(0)}^{-1}(\lambda) s} = \frac{(1 - \lambda) \lambda^{2d} (\lambda + 1)}{(1 - \lambda^d)(\lambda^d + 1)(\lambda^d - \lambda)(\lambda^d + \lambda)} \geq 0$$

und

$$p^T (\mathcal{I}_{(0)}([\beta_0, \beta_1]^T))^{-1} p = (Rp)^T (Rp) \geq 0$$

kann anhand der Gleichung (5.56) gefolgert werden, dass

$$\det(\mathcal{I}_{(1)}([\beta_0, \beta_1]^T)) \geq \det(\mathcal{I}_{(0)}([\beta_0, \beta_1]^T)).$$

Auf Grund der Regel

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

gilt schließlich, dass

$$\tilde{D}_\lambda(d) = \det((\mathcal{I}_{(1)}([\beta_0, \beta_1]^T))^{-1}) \leq \det((\mathcal{I}_{(0)}([\beta_0, \beta_1]^T))^{-1}) = \tilde{D}_\lambda(0).$$

Fall 2 ($\lambda = 1$):

Es gilt, dass $\tilde{D}_\lambda(0) = \tilde{D}_\lambda(d) = 0$, siehe (5.18). \square

Das Resultat von Behauptung 5.19 ist plausibel, da ein Versuchsplan, der sich lediglich aus zwei Versuchspunkten zusammensetzt, zu größeren Schätzervarianzen führen sollte als ein Plan, der neben den beiden Versuchspunkten des konkurrierenden Plans einen weiteren dritten Versuchspunkt aufweist. Vergleiche dazu die Konstruktion des D-Kriteriums in Abschnitt 5.3.1.

Eine Konsequenz von Behauptung 5.19 ist die, dass der Versuchsplan mit $d = 0$ den Plänen mit $d \in (0, 1)$ für $\lambda \in (0, 1]$ unterlegen ist. Für $\lambda = 0$ jedoch ist der Plan $d = 0$ D-optimal, siehe Behauptung 5.18.

Das folgende Korollar zeigt, dass die Funktion $\tilde{D}_\lambda(d)$ für jedes λ durch $1 - \lambda^2$ nach oben beschränkt ist. Und damit insgesamt durch 1, da $1 - \lambda^2 \leq 1$.

Korollar 5.20. *Seien $\lambda \in [0, 1]$ und $d \in (0, 1)$. Dann gilt:*

$$\tilde{D}_\lambda(d) \leq 1 - \lambda^2. \quad (5.57)$$

Beweis. Folgt aus Behauptung 5.18 und Behauptung 5.19. \square

Eine untere Schranke in Abhängigkeit von λ

Für die Bestimmung einer unteren Schranke für $\tilde{D}_\lambda(d)$ wird eine Fallunterscheidung vorgenommen. Die Fallunterscheidung gliedert sich in die Fälle $\lambda \in [0, 1/5)$ und $\lambda \in [1/5, 1]$, siehe Behauptung 5.22 und Behauptung 5.23. Um die Beweise der beiden Behauptungen übersichtlicher zu gestalten, werden einige Lemmata formuliert. Die rein technischen Lemmata sind in Anhang A zu finden. Das Ineinandergreifen der Lemmata und Behauptungen wird in Abbildung 5.8 veranschaulicht.

Zunächst wird in Behauptung 5.22 die Schranke für den Fall $\lambda \in [0, 1/5)$ formuliert. Für den Nachweis wird Lemma 5.21 benötigt.

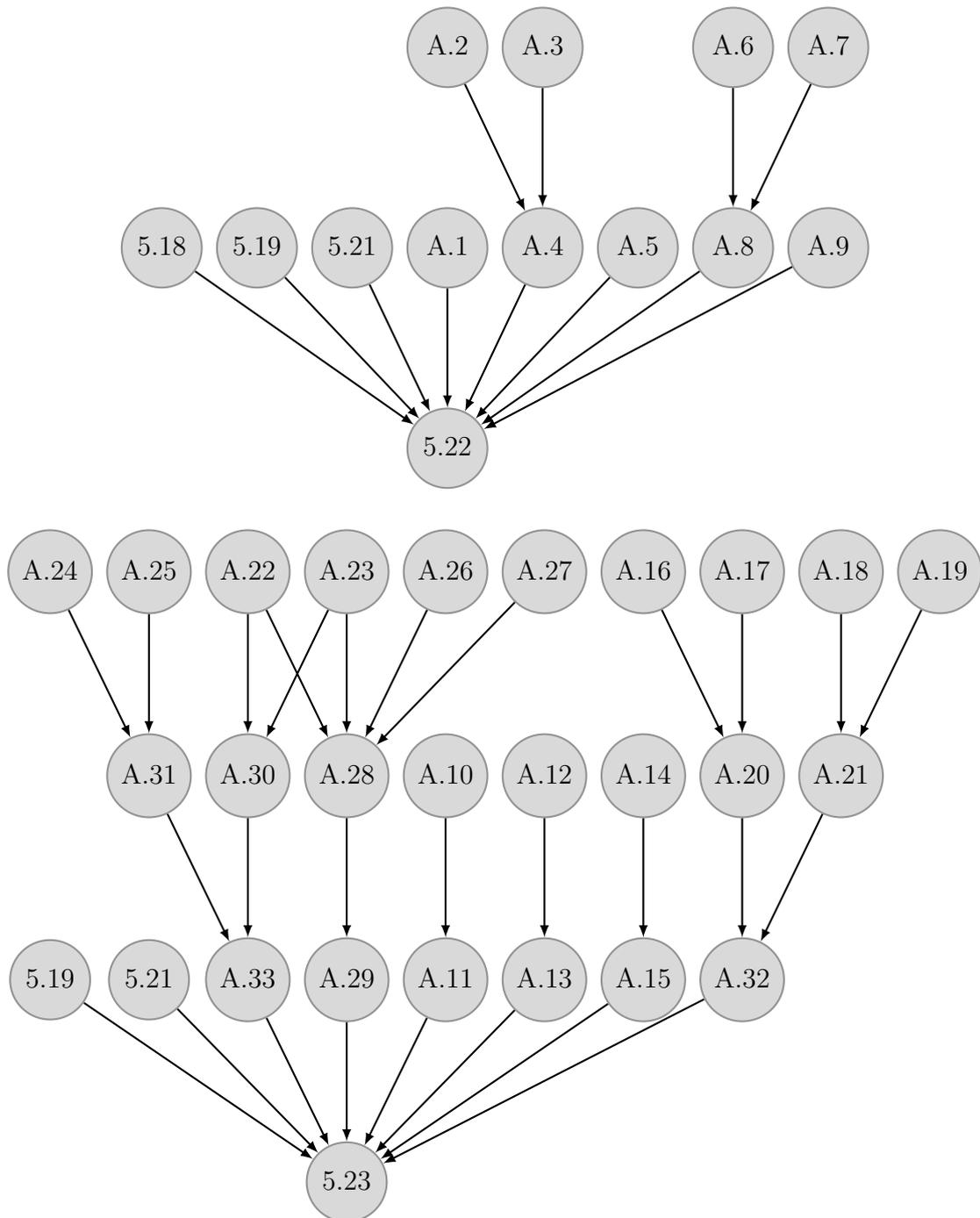


Abbildung 5.8: Zwei Bäume, die die Bewisaufbauten zu den Behauptungen 5.22 und 5.23 veranschaulichen. Die beiden Behauptungen ziehen die Lemmata A.1 bis A.33 aus Anhang A, die Behauptungen 5.18 und 5.19 sowie Lemma 5.21 heran. In die aufgeführten Lemmata selbst fließen wiederum die Behauptungen 5.8 bis 5.13 und die Lemmata 5.14 bis 5.17 ein. Die letztgenannten Abhängigkeiten sind im Hinblick auf die Übersichtlichkeit nicht Bestandteil der Bäume.

Lemma 5.21. *Seien $\lambda \in (0, 1)$ und $d \in (0, 1)$. Betrachte für $\tilde{D}_\lambda(d)$, die Notation*

$$\tilde{D}_\lambda(d) = e_d(\lambda)/f_d(\lambda)$$

mit

$$e_d(\lambda) = (1 - \lambda^{2d})(1 - \lambda^{2(1-d)}),$$

$$f_d(\lambda) = 2((1 - (1 - d)\lambda^d)(1 - d\lambda^{1-d}) - d(1 - d)).$$

Dann gilt:

$$\tilde{D}_\lambda(d) > 0, \tag{5.58}$$

$$e_d(\lambda) > 0, \tag{5.59}$$

$$f_d(\lambda) > 0. \tag{5.60}$$

Beweis. Im Weiteren bezeichnet die Schreibweise $\text{rg}(A)$ den Rang einer Matrix A . Für eine beliebige Matrix $A \in \mathbb{R}^{m \times n}$ gilt

$$A^T A \text{ ist positiv definit} \Leftrightarrow \text{rg}(A) = n. \tag{5.61}$$

Darüber hinaus gilt

$$A^T A \text{ ist positiv definit} \Rightarrow \det(A^T A) > 0, \tag{5.62}$$

siehe Schmidt und Trenkler (2006), S. 99.

Nun zum eigentlichen Beweis. Die Matrix $\Sigma_{(1)}^{-1}(\lambda)$, siehe (5.8), ist das Ergebnis des Matrizenprodukts $V^T V$ mit

$$V = \begin{bmatrix} 1 & 0 & 0 \\ \frac{\lambda^d}{\sqrt{1 - \lambda^{2d}}} & \frac{1}{\sqrt{1 - \lambda^{2d}}} & 0 \\ 0 & -\frac{\lambda^{1-d}}{\sqrt{1 - \lambda^{2(1-d)}}} & \frac{1}{\sqrt{1 - \lambda^{2(1-d)}}} \end{bmatrix},$$

siehe Dette *et al.* (2008). Dies ermöglicht für die Matrix $\mathcal{I}_{(1)}([\beta_0, \beta_1]^T)$, siehe (5.9), die Schreibweise

$$\mathcal{I}_{(1)}([\beta_0, \beta_1]^T) = X_{(1)}^T V^T V X_{(1)} = M^T M \tag{5.63}$$

mit $VX_{(1)} =: M$. Die Matrix M hat dabei die Gestalt

$$M = \begin{bmatrix} 1 & 0 \\ \frac{1}{\sqrt{1-\lambda^{2d}}} - \frac{\lambda^d}{\sqrt{1-\lambda^{2d}}} & \frac{d}{\sqrt{1-\lambda^{2d}}} \\ \frac{1}{\sqrt{1-\lambda^{2(1-d)}}} - \frac{\lambda^{1-d}}{\sqrt{1-\lambda^{2(1-d)}}} & \frac{1}{\sqrt{1-\lambda^{2(1-d)}}} - \frac{d\lambda^{1-d}}{\sqrt{1-\lambda^{2(1-d)}}} \end{bmatrix}.$$

Im nächsten Schritt wird jetzt gezeigt, dass $\text{rg}(M) = 2$ ist. Es ist zu erkennen, dass der zweite Eintrag der zweiten Zeile von M unter den Voraussetzungen $d \in (0, 1)$ und $\lambda \in (0, 1)$ stets größer Null ist. Das hat die Konsequenz, dass die ersten beiden Zeilen von M linear unabhängig sind. Somit gilt, dass $\text{rg}(M) \geq 2$. Da andererseits $M \in \mathbb{R}^{3 \times 2}$, gilt für die Matrix M , dass $\text{rg}(M) \leq 2$. Insgesamt folgt also, dass $\text{rg}(M) = 2$. Das bedeutet nun, dass $\mathcal{I}_{(1)}([\beta_0, \beta_1]^T)$ wegen (5.61) und (5.63) positiv definit ist. Auf Grund der Tatsache, dass $\mathcal{I}_{(1)}([\beta_0, \beta_1]^T)$ positiv definit ist, folgt, dass $\mathcal{I}_{(1)}([\beta_0, \beta_1]^T)^{-1}$ ebenfalls positiv definit ist, siehe Schmidt und Trenkler (2006), S. 99. Dies hat wiederum zur Folge, dass $\det((\mathcal{I}_{(1)}([\beta_0, \beta_1]^T))^{-1})$ stets größer Null ist. Damit ist (5.58) gezeigt.

Die Aussage (5.59) gilt auf Grund der Voraussetzungen $\lambda \in (0, 1)$ und $d \in (0, 1)$ unmittelbar. Schließlich gilt (5.60) auf Grund von (5.58) und (5.59). \square

Behauptung 5.22. Sei $\lambda \in [0, 1/5)$. Dann ist die Funktion $L_1(\lambda) = 1/2 + \lambda$ eine untere Schranke für $\tilde{D}_\lambda(d)$.

Beweis. **Fall 1** ($\lambda = 0$):

Folgt aus Behauptung 5.18.

Fall 2 ($\lambda \in (0, 1/5)$):

Wegen Behauptung 5.19 genügt es den Fall $d \in (0, 1)$ zu betrachten. Im Weiteren wird die Schreibweise $\tilde{D}_\lambda(d) = e_d(\lambda)/f_d(\lambda)$ mit

$$e_d(\lambda) = (1 - \lambda^{2d})(1 - \lambda^{2(1-d)})$$

und

$$f_d(\lambda) = 2((1 - (1 - d)\lambda^d)(1 - d\lambda^{1-d}) - d(1 - d))$$

verwendet. Es ist zu zeigen, dass

$$\tilde{D}_\lambda(d) \geq \frac{1}{2} + \lambda.$$

Diese Ungleichung ist wegen (5.60) äquivalent zu der Ungleichung

$$e_d(\lambda) - \frac{1}{2}f_d(\lambda) - \lambda f_d(\lambda) \geq 0. \quad (5.64)$$

Definiere nun

$$e_d(\lambda) - \frac{1}{2}f_d(\lambda) - \lambda f_d(\lambda) = s_1(\lambda, d) + s_2(\lambda, d) + s_3(\lambda, d) + s_4(\lambda, d) =: s(\lambda, d)$$

mit

$$\begin{aligned} s_1(\lambda, d) &= d^2(2\lambda^2 - \lambda - 1) + d(-2\lambda^2 + \lambda + 1), \\ s_2(\lambda, d) &= -2\lambda^{d+1}d + 2\lambda^{2-d}d + 2\lambda^{d+1} - 2\lambda, \\ s_3(\lambda, d) &= -\lambda^d d - \lambda^{2d} + \lambda^d + \lambda^2, \\ s_4(\lambda, d) &= \lambda^{1-d}d - \lambda^{2-2d}. \end{aligned}$$

Wegen Lemma A.1 und Lemma A.4 gilt, dass

$$s(\lambda, d) = s_1(\lambda, d) + s_2(\lambda, d) + s_3(\lambda, d) + s_4(\lambda, d) \geq s_5(\lambda, d) + s_2(\lambda, d) =: t_1(\lambda, d). \quad (5.65)$$

Eine Zentrierung des Versuchsraums $[0, 1]$ auf $[-1/2, 1/2]$ liefert:

$$\begin{aligned} t_1(\lambda, d) &= t_1(\lambda, c + 1/2) \\ &= -\lambda c^2 - \frac{23}{25}c^2 - 2\lambda^{c+3/2}c + 2\lambda^{3/2-c}c + \lambda^{c+3/2} + \lambda^{3/2-c} - \frac{7}{4}\lambda + \frac{23}{100} \\ &=: t_2(\lambda, c). \end{aligned}$$

Die Funktion $t_2(\lambda, c)$ ist wegen Lemma A.5 konvex in λ . Wegen Lemma A.8 ist $t_2(\lambda, c)$ an der Stelle $\lambda = 1/5$ monoton nicht steigend in λ . Da schließlich wegen Lemma A.9 $t_2(1/5, c) \geq 0$, gilt $t_2(\lambda, c) \geq 0$. Somit gilt wegen (5.65) die Ungleichung (5.64). Damit gilt insbesondere auch $\tilde{D}_\lambda(d) \geq 1/2 + \lambda$. \square

Die Funktion $L_1(\lambda)$ aus Behauptung 5.22 liefert also eine untere Schranke für den Fall $\lambda \in [0, 1/5)$. Nun wird eine Funktion präsentiert, die eine untere Schranke für den Fall $\lambda \in [1/5, 1]$ darstellt. Diese Funktion wird in Behauptung 5.23 formuliert.

Behauptung 5.23. *Sei $\lambda \in [1/5, 1]$. Dann ist die Funktion $L_2(\lambda) = 81(1 - \lambda^2)/100$ eine untere Schranke für $\tilde{D}_\lambda(d)$.*

Beweis. Fall 1 ($\lambda \in [1/2, 1)$):

Wegen Behauptung 5.19 genügt es den Fall $d \in (0, 1)$ zu betrachten. Erneut wird die Schreibweise $\tilde{D}_\lambda(d) = e_d(\lambda)/f_d(\lambda)$ mit

$$e_d(\lambda) = (1 - \lambda^{2d})(1 - \lambda^{2(1-d)})$$

und

$$f_d(\lambda) = 2((1 - (1 - d)\lambda^d)(1 - d\lambda^{1-d}) - d(1 - d))$$

verwendet. Es ist zu zeigen, dass

$$\tilde{D}_\lambda(d) \geq \frac{81}{100}(1 - \lambda^2). \quad (5.66)$$

Diese Ungleichung ist wegen (5.60) äquivalent zu der Ungleichung

$$100e_d(\lambda) + 81f_d(\lambda)\lambda^2 - 81f_d(\lambda) \geq 0. \quad (5.67)$$

Nun wird (5.67) bewiesen. Definiere

$$100e_d(\lambda) + 81f_d(\lambda)\lambda^2 - 81f_d(\lambda) =: s(\lambda, d).$$

Eine Zentrierung des Versuchsraums $[0, 1]$ auf $[-1/2, 1/2]$ liefert:

$$\begin{aligned} s(\lambda, d) &= s(\lambda, c + 1/2) \\ &= s_1(\lambda, c) + s_2(\lambda, c) + s_3(\lambda, c) + s_4(\lambda, c) + s_5(\lambda, c) \\ &=: z_1(\lambda, c) \end{aligned}$$

mit

$$\begin{aligned} s_1(\lambda, c) &= -162(\lambda - 1)^2(\lambda + 1)c^2, \\ s_2(\lambda, c) &= 81\lambda^{c+5/2}(-1 + 2c) + 81\lambda^{5/2-c}(-1 - 2c), \\ s_3(\lambda, c) &= 81\lambda^{c+1/2}(1 - 2c) + 81\lambda^{1/2-c}(1 + 2c), \\ s_4(\lambda, c) &= -100\lambda^{2c+1} - 100\lambda^{1-2c}, \\ s_5(\lambda, c) &= \frac{81}{2}\lambda^3 + \frac{443}{2}\lambda^2 - \frac{81}{2}\lambda - \frac{43}{2}. \end{aligned}$$

Wegen Lemma A.11, Lemma A.13 und Lemma A.15 gilt für $c \in [0, 1/2]$, dass

$$\begin{aligned} z_1(\lambda, c) &= s_1(\lambda, c) + s_2(\lambda, c) + s_3(\lambda, c) + s_4(\lambda, c) + s_5(\lambda, c) \\ &\geq s_1(\lambda, c) + m_2(\lambda, c) + m_3(\lambda, c) + m_4(\lambda, c) + s_5(\lambda, c) =: z_2(\lambda, c). \end{aligned} \quad (5.68)$$

Nun ist $z_2(\lambda, c)$ monoton fallend in c , siehe Lemma A.32. Da außerdem $z_2(\lambda, 1/2) = 0$ ist, folgt zunächst, dass $z_2(\lambda, c) \geq 0$ ist. Wegen $s(\lambda, c + 1/2) \geq z_2(\lambda, c)$ gilt (5.67) und damit auch (5.66). Aus Symmetriegründen gilt (5.66) auch für $c \in [-1/2, 0)$.

Fall 2 ($\lambda \in [1/5, 1/2)$):

Anfangs analog zum ersten Fall. Statt (5.68) gilt hier abweichend für $c \in [0, 1/2]$, dass

$$\begin{aligned} z_1(\lambda, c) &= s_1(\lambda, c) + s_2(\lambda, c) + s_6(\lambda, c) + s_5(\lambda, c) \\ &\geq s_1(\lambda, c) + m_2(\lambda, c) + m_6(\lambda, c) + s_5(\lambda, c) =: z_3(\lambda, c) \end{aligned} \quad (5.69)$$

mit $s_6(\lambda, c)$ und $m_6(\lambda, c)$ aus Lemma A.29.

Die Funktion $z_3(\lambda, c)$ ist monoton fallend in c , siehe Lemma A.33. Wegen $z_3(\lambda, 0) = 0$ folgt, dass $z_3(\lambda, c)$ für $c \in [0, 1/2]$ positiv ist. Die weiteren Beweisschritte entsprechen denen aus dem ersten Fall.

Fall 3 ($\lambda = 1$):

Folgt aus (5.18). □

Die hergeleiteten Schranken aus Korollar 5.20, Behauptung 5.22 und Behauptung 5.23 sind für Abschnitt 5.2 relevant.

5.3.6. Relative D-Effizienzen

In diesem Abschnitt wird der Quotient

$$\frac{D_\lambda(d_2)}{D_\lambda(d_1)} =: q_{d_1, d_2}(\lambda)$$

betrachtet. Genauer gesagt werden bezüglich dieses Quotienten untere Schranken für ausgewählte Versuchspunkte d_1 und d_2 bestimmt. Die Ergebnisse sind für Abschnitt 5.2 relevant. Die Quotienten werden als relative D-Effizienz bezeichnet.

Effizienzvergleich der Pläne mit $d = 0$ und $d = 1/2$

Als Erstes werden die relativen D-Effizienzen des Versuchsplans mit $d = 0$ bezüglich des Versuchsplans mit $d = 1/2$ untersucht, wobei zunächst der Fall $\lambda \in (0, 1)$ betrachtet wird. Sie lauten

$$q_{0, 1/2}(\lambda) = \frac{D_\lambda(1/2)}{D_\lambda(0)} = \left(\frac{2(1-\lambda)}{(\lambda+1)(\lambda-4\lambda^{1/2}+3)} \right)^{1/2}.$$

Die erste Ableitung der Funktion $2(1-\lambda)/((\lambda+1)(\lambda-4\lambda^{1/2}+3))$ ist gegeben durch

$$\frac{2(1-\lambda^{1/2})^4(\lambda^{1/2}+2)\lambda^{-1/2}}{((\lambda+1)(\lambda-4\lambda^{1/2}+3))^2}.$$

Diese ist für $\lambda \in (0, 1)$ stets positiv. Somit sind die relativen D-Effizienzen $q_{0,1/2}(\lambda)$ monoton wachsend in λ . Das Minimum von $q_{0,1/2}(\lambda)$ wird folglich für den Grenzfall $\lambda \rightarrow 0$ realisiert. Das Maximum wird dagegen für $\lambda \rightarrow 1$ angenommen. Es gilt

$$\lim_{\lambda \rightarrow 0} q_{0,1/2}(\lambda) = (2/3)^{1/2} \approx 0.817.$$

Wegen Abschnitt 5.3.3 gilt, dass

$$\lim_{\lambda \rightarrow 1} q_{0,1/2}(\lambda) = 1.$$

Somit nimmt die Überlegenheit des äquidistanten Versuchsplans mit wachsender Korrelation ab. Dagegen ist speziell für $\lambda = 0$ der Plan mit Versuchspunkt $d = 0$ dem Plan mit $d = 1/2$ überlegen. In diesem Fall beträgt die relative D-Effizienz des Plans mit $d = 1/2$ bezüglich des Plans mit $d = 0$

$$q_{1/2,0}(0) = \frac{D_{\lambda=0}(0)}{D_{\lambda=0}(1/2)} = (3/4)^{1/2} \approx 0.866.$$

Effizienzvergleich der Pläne mit $d = 0$ und $d = 3/8$

Nun werden die relativen D-Effizienzen des Versuchsplans mit $d = 0$ bezüglich des Versuchsplans mit $d = 3/8$ untersucht. Sie lauten für $\lambda \in (0, 1)$

$$q_{0,3/8}(\lambda) = \frac{D_{\lambda}(3/8)}{D_{\lambda}(0)} = \left(\frac{32(\lambda^2 - \lambda^{5/4} - \lambda^{3/4} + 1)}{(1 - \lambda)(\lambda + 1)(15\lambda - 24\lambda^{5/8} - 40\lambda^{3/8} + 49)} \right)^{1/2}.$$

Die erste Ableitung von $32(\lambda^2 - \lambda^{5/4} - \lambda^{3/4} + 1)/[(1 - \lambda)(\lambda + 1)(15\lambda - 24\lambda^{5/8} - 40\lambda^{3/8} + 49)]$ lautet

$$\frac{m_1(\lambda)}{((1 - \lambda)(\lambda + 1)(15\lambda - 24\lambda^{5/8} - 40\lambda^{3/8} + 49))^2}$$

mit

$$\begin{aligned} m_1(\lambda) &= m_1(x^8) \\ &= 8(1 - x)^6(x + 1)^2(60x^{29} + 240x^{28} + 720x^{27} + 1620x^{26} + 3240x^{25} + 5700x^{24} \\ &\quad + 9255x^{23} + 13860x^{22} + 19560x^{21} + 26112x^{20} + 33183x^{19} + 40344x^{18} \\ &\quad + 46896x^{17} + 52416x^{16} + 55884x^{15} + 57188x^{14} + 55532x^{13} + 51660x^{12} \\ &\quad + 45300x^{11} + 37968x^{10} + 29832x^9 + 22456x^8 + 15709x^7 + 10464x^6 \\ &\quad + 6336x^5 + 3612x^4 + 1773x^3 + 780x^2 + 240x + 60). \end{aligned}$$

Auf Grund der Substitution $\lambda = x^8$ ist leicht zu erkennen, dass die erste Ableitung für $\lambda \in (0, 1)$ stets positiv ist. Damit sind hier die relativen D-Effizienzen erneut monoton wachsend in λ . Es gilt:

$$\lim_{\lambda \rightarrow 0} q_{0,3/8}(\lambda) = (32/49)^{1/2} \approx 0.808.$$

Natürlich gilt hier auch, dass

$$\lim_{\lambda \rightarrow 1} q_{0,3/8}(\lambda) = 1.$$

Für $\lambda = 0$ gilt hier, dass

$$q_{3/8,0}(0) = (49/64)^{1/2} = 0.875.$$

Effizienzvergleich der Pläne mit $d = 0$ und $d = 1/4$

Als Nächstes werden die relativen D-Effizienzen des Versuchsplans mit $d = 0$ bezüglich des Versuchsplans mit $d = 1/4$ untersucht. Sie lauten für $\lambda \in (0, 1)$

$$q_{0,1/4}(\lambda) = \frac{D_\lambda(1/4)}{D_\lambda(0)} = \left(\frac{8(1 - \lambda^{1/2})^2(\lambda + \lambda^{1/2} + 1)}{(1 - \lambda)(\lambda + 1)(3\lambda - 4\lambda^{3/4} - 12\lambda^{1/4} + 13)} \right)^{1/2}.$$

Die erste Ableitung von $8(1 - \lambda^{1/2})^2(\lambda + \lambda^{1/2} + 1)/[(1 - \lambda)(\lambda + 1)(3\lambda - 4\lambda^{3/4} - 12\lambda^{1/4} + 13)]$ lautet dann

$$\frac{m_2(\lambda)}{((1 - \lambda)(\lambda + 1)(3\lambda - 4\lambda^{3/4} - 12\lambda^{1/4} + 13))^2}$$

mit

$$m_2(\lambda) = 4(\lambda^{1/4} - 1)^6(\lambda^{1/4} + 1)^2(6\lambda^{11/4} + 18\lambda^{5/2} + 39\lambda^{9/4} + 64\lambda^2 + 88\lambda^{7/4} + 114\lambda^{3/2} + 114\lambda^{5/4} + 102\lambda + 54\lambda^{3/4} + 32\lambda^{1/2} + 11\lambda^{1/4} + 6)\lambda^{-3/4}.$$

Es ist leicht zu erkennen, dass dieser Ausdruck ebenfalls für $\lambda \in (0, 1)$ stets positiv ist. Damit sind hier die relativen D-Effizienzen erneut monoton wachsend in λ . Es gilt:

$$\lim_{\lambda \rightarrow 0} q_{0,1/4}(\lambda) = (8/13)^{1/2} \approx 0.785.$$

Natürlich gilt hier auch, dass

$$\lim_{\lambda \rightarrow 1} q_{0,1/4}(\lambda) = 1.$$

Für $\lambda = 0$ gilt hier, dass

$$q_{1/4,0}(0) = (13/16)^{1/2} \approx 0.901.$$

Effizienzvergleich der Pläne mit $d = 0$ und $d = 1/8$

Analog zum bisherigen Vorgehen lautet der Zähler der ersten Ableitung hier

$$\begin{aligned} m_3(\lambda) = & 8(\lambda^{1/4} - 1)^6(\lambda^{1/4} + 1)^2(28\lambda^{31/4} + 84\lambda^{15/2} + 189\lambda^{29/4} + 344\lambda^7 + 564\lambda^{27/4} \\ & + 852\lambda^{13/2} + 1218\lambda^{25/4} + 1640\lambda^6 + 2096\lambda^{23/4} + 2596\lambda^{11/2} + 3086\lambda^{21/4} + 3608\lambda^5 \\ & + 4076\lambda^{19/4} + 4564\lambda^{9/2} + 4877\lambda^{17/4} + 5128\lambda^4 + 5132\lambda^{15/4} + 5032\lambda^{7/2} \\ & + 4623\lambda^{13/4} + 4076\lambda^3 + 3164\lambda^{11/4} + 2504\lambda^{5/2} + 1866\lambda^{9/4} + 1436\lambda^2 + 1016\lambda^{7/4} \\ & + 760\lambda^{3/2} + 502\lambda^{5/4} + 364\lambda + 212\lambda^{3/4} + 136\lambda^{1/2} + 55\lambda^{1/4} + 28)\lambda^{-7/4}. \end{aligned}$$

Dieser ist erneut stets positiv. Es gilt

$$\lim_{\lambda \rightarrow 0} q_{0,1/8}(\lambda) = (32/57)^{1/2} \approx 0.749.$$

Für $\lambda = 0$ gilt, dass

$$q_{1/8,0}(0) = (57/64)^{1/2} \approx 0.944.$$

Effizienzvergleich der Pläne mit $d = 1/4$ und $d = 1/2$

Es gilt, dass der Versuchsplan mit $d = 1/4$ gegenüber dem Versuchsplan mit $d = 1/2$ im ungünstigsten Fall circa 2.5 Prozent an relativer D-Effizienz verliert. Dagegen verliert andersherum der Versuchsplan mit $d = 1/2$ gegenüber dem Versuchsplan mit $d = 1/4$ im ungünstigsten Fall etwa 3.9 Prozent, siehe Behauptung 5.24.

Behauptung 5.24. *Gegeben seien die beiden Versuchspläne mit $d = 1/4$ beziehungsweise $d = 1/2$. Dann gilt:*

a)

$$q_{1/4,1/2}(\lambda) \geq (19/20)^{1/2} \approx 0.975,$$

b)

$$\min_{\lambda} q_{1/2,1/4}(\lambda) = (12/13)^{1/2} \approx 0.961.$$

Beweis. Zu a):

Für $\lambda \in [0, 1)$ gilt:

$$q_{1/4,1/2}(\lambda) = \left(\frac{((\lambda - 1)^2(3\lambda - 4\lambda^{3/4} - 12\lambda^{1/4} + 13))}{4(\lambda^{1/2} - 1)^2(\lambda - 4\lambda^{1/2} + 3)(\lambda + \lambda^{1/2} + 1)} \right)^{1/2}.$$

Die Behauptung gilt, falls

$$(q_{1/4,1/2}(\lambda))^2 = \frac{((\lambda - 1)^2(3\lambda - 4\lambda^{3/4} - 12\lambda^{1/4} + 13))}{4(\lambda^{1/2} - 1)^2(\lambda - 4\lambda^{1/2} + 3)(\lambda + \lambda^{1/2} + 1)} \geq \frac{19}{20}.$$

Äquivalenzumformungen liefern die Ungleichung

$$\begin{aligned} -16\lambda^3 - 80\lambda^{11/4} + 380\lambda^{5/2} - 240\lambda^{9/4} - 392\lambda^2 + 160\lambda^{7/4} + 304\lambda^{3/2} \\ + 480\lambda^{5/4} - 840\lambda - 80\lambda^{3/4} + 532\lambda^{1/2} - 240\lambda^{1/4} + 32 \geq 0. \end{aligned} \quad (5.70)$$

Eine Substitution mittels $\lambda = x^4$ liefert für die linke Seite der Ungleichung (5.70) den Term

$$4(1 - x)^3(x + 1)^2(4x^7 + 24x^6 - 63x^5 + 37x^4 - 43x^3 + 97x^2 - 52x + 8).$$

Es bleibt zu zeigen, dass der Term

$$4x^7 + 24x^6 - 63x^5 + 37x^4 - 43x^3 + 97x^2 - 52x + 8 \quad (5.71)$$

für $x \in [0, 1)$ positiv ist.

Fall 1 ($x \in [0, 1/2]$):

Die Substitution $x = y/2$ in (5.71) liefert den Term

$$\frac{1}{32}(y^7 + 12y^6 - 63y^5 + 74y^4 - 172y^3 + 776y^2 - 832y + 256).$$

Nun ist für $y \in [0, 1]$ zu zeigen, dass

$$y^7 + 12y^6 - 63y^5 + 74y^4 - 172y^3 + 776y^2 - 832y + 256 =: p_1(y)$$

positiv ist.

Es gilt

$$\begin{aligned} p_1(y) &\geq -2y^7 + 12y^6 - 63y^5 + 74y^4 - 172y^3 + 776y^2 - 832y + 256 \\ &= (2 - y)(2y^6 - 8y^5 + 47y^4 + 20y^3 + 212y^2 - 352y + 128). \end{aligned}$$

Nun ist zu zeigen, dass

$$2y^6 - 8y^5 + 47y^4 + 20y^3 + 212y^2 - 352y + 128 =: p_2(y)$$

positiv ist. Die Funktion $p_2(y)$ ist wegen

$$\begin{aligned} \frac{\partial^2}{\partial y^2} p_2(y) &= 60y^4 - 160y^3 + 564y^2 + 120y + 424 \\ &\geq 60y^4 - 160y^3 + 564y^3 + 120y + 424 \\ &= 60y^4 + 404y^3 + 120y + 424 \\ &\geq 0 \end{aligned}$$

konvex.

Die Gerade

$$s_1(y) = \frac{33499903}{6400000} - \frac{4134721}{800000}y$$

stellt die Tangente der Funktion $p_2(y)$ an der Stelle $y = 13/20$ dar. Da die Tangente am Rand die Funktionswerte $s_1(0) = 33499903/6400000$ und $s_1(1) = 84427/1280000$ aufweist, sind die Funktionswerte für $y \in [0, 1]$ stets positiv. Auf Grund der Tatsache,

dass die Tangente eine untere Schranke der konvexen Funktion $p_2(y)$ darstellt, ist die Funktion $p_2(y)$ ebenfalls positiv.

Fall 2 ($x \in (1/2, 1)$):

Die Substitution $x = 1/2 + y/2$ in (5.71) liefert den Term

$$\frac{1}{32}(y^7 + 19y^6 + 30y^5 - 26y^4 - 231y^3 + 275y^2 + 264y + 52).$$

Nun ist für $y \in (0, 1]$ zu zeigen, dass

$$y^7 + 19y^6 + 30y^5 - 26y^4 - 231y^3 + 275y^2 + 264y + 52 =: p_3(y)$$

positiv ist. Dies gilt mit Behauptung 5.13.

Zu b):

Das Minimum der Funktion $q_{1/2,1/4}(\lambda)$ stimmt mit dem Minimum des Terms $(q_{1/2,1/4}(\lambda))^2$ überein. Es gilt

$$(q_{1/2,1/4}(\lambda))^2 = \frac{m_4(\lambda)}{m_5(\lambda)} =: r_1(\lambda)$$

mit

$$m_4(\lambda) = 4(\lambda^{1/2} - 1)^2(\lambda - 4\lambda^{1/2} + 3)(\lambda + \lambda^{1/2} + 1)$$

und

$$m_5(\lambda) = (\lambda - 1)^2(3\lambda - 4\lambda^{3/4} - 12\lambda^{1/4} + 13).$$

Der Zähler der ersten Ableitung von $r_1(\lambda)$ lautet

$$\begin{aligned} & -4\lambda^{19/4} + 30\lambda^{9/2} - 56\lambda^{17/4} - 56\lambda^4 + 184\lambda^{15/4} + 2\lambda^{7/2} + 52\lambda^{13/4} \\ & -304\lambda^3 - 584\lambda^{11/4} + 956\lambda^{5/2} + 592\lambda^{9/4} - 736\lambda^2 - 272\lambda^{7/4} - 172\lambda^{3/2} - 56\lambda^{5/4} \\ & + 560\lambda + 556\lambda^{3/4} - 634\lambda^{1/2} - 568\lambda^{1/4} + 120\lambda^{-1/4} - 182\lambda^{-1/2} + 36\lambda^{-3/4} + 536. \end{aligned}$$

Die Substitution $\lambda = x^4$ liefert den Term $2(x-1)^8(x+1)^4(x^2+1)p_4(x)x^{-3}$ mit

$$p_4(x) = -2x^8 + 7x^7 + 6x^6 - 9x^5 - 48x^4 - 99x^3 - 70x^2 - 19x + 18.$$

Wegen $p_4(0) = 18$ und $p_4(1) = -216$ hat die Funktion $p_4(x)$ auf Grund des Zwischenwertsatzes mindestens eine Nullstelle in dem Intervall $[0, 1]$. Andererseits existiert wegen

$$\begin{aligned} \frac{\partial}{\partial x} p_4(x) &= -16x^7 + 49x^6 + 36x^5 - 45x^4 - 192x^3 - 297x^2 - 140x - 19 \\ &\leq -16x^7 + 49x^6 + 36x^5 - 45x^5 - 192x^6 - 297x^2 - 140x - 19 \\ &= -16x^7 - 143x^6 - 9x^5 - 297x^2 - 140x - 19 < 0 \end{aligned}$$

höchstens eine Nullstelle. Somit existiert genau eine lokale Extremstelle. Zudem kann auf Grund des Vorzeichens der Ableitung gefolgert werden, dass $q_{1/2,1/4}(\lambda)$ bis zur Extremstelle monoton wächst und danach monoton fällt. Dies bedeutet, dass das globale Minimum am Rand zu finden ist. Wegen

$$\lim_{\lambda \rightarrow 1} q_{1/2,1/4}(\lambda) = 1$$

und

$$q_{1/2,1/4}(0) = (12/13)^{1/2} < 1$$

ist das globale Minimum an der Stelle $\lambda = 0$ zu finden. Die Minimalstelle hat den Funktionswert $(12/13)^{1/2} \approx 0.961$. \square

Effizienzvergleich der Pläne mit $d = 1/8$ und $d = 1/2$

Behauptung 5.25. *Gegeben seien die beiden Versuchspläne mit $d = 1/8$ beziehungsweise $d = 1/2$. Dann gilt:*

a)

$$q_{1/8,1/2}(\lambda) \geq (17/20)^{1/2} \approx 0.922,$$

b)

$$\min_{\lambda} q_{1/2,1/8}(\lambda) = (16/19)^{1/2} \approx 0.918.$$

Beweis. Zu a):

Für $\lambda \in [0, 1)$ gilt:

$$q_{1/8,1/2}(\lambda) = \left(\frac{(\lambda - 1)^2(7\lambda - 8\lambda^{7/8} - 56\lambda^{1/8} + 57)}{16(\lambda - 4\lambda^{1/2} + 3)(\lambda^2 - \lambda^{7/4} - \lambda^{1/4} + 1)} \right)^{1/2}.$$

Die Behauptung gilt, falls

$$(q_{1/8,1/2}(\lambda))^2 = \frac{(\lambda - 1)^2(7\lambda - 8\lambda^{7/8} - 56\lambda^{1/8} + 57)}{16(\lambda - 4\lambda^{1/2} + 3)(\lambda^2 - \lambda^{7/4} - \lambda^{1/4} + 1)} \geq \frac{17}{20}.$$

Äquivalenzumformungen liefern die Ungleichung

$$\begin{aligned} -\frac{33}{16}\lambda^3 - \frac{5}{2}\lambda^{23/8} + \frac{17}{4}\lambda^{11/4} + 17\lambda^{5/2} - 17\lambda^{9/4} - \frac{35}{2}\lambda^{17/8} + \frac{11}{16}\lambda^2 + 5\lambda^{15/8} + \frac{51}{4}\lambda^{7/4} \\ + \frac{17}{4}\lambda^{5/4} + 35\lambda^{9/8} - \frac{603}{16}\lambda - \frac{5}{2}\lambda^{7/8} - 17\lambda^{3/4} + 17\lambda^{1/2} + \frac{51}{4}\lambda^{1/4} - \frac{35}{2}\lambda^{1/8} + \frac{81}{16} \geq 0. \end{aligned}$$

Eine Substitution mittels $\lambda = x^8$ liefert für die linke Seite der Ungleichung den Term

$$\frac{1}{16}(1-x)^3(x+1)^2(x^2+1)m_6(x)$$

mit

$$\begin{aligned} m_6(x) = & 33x^{17} + 73x^{16} + 38x^{15} + 78x^{14} - 196x^{13} - 116x^{12} - 186x^{11} + 174x^{10} - 146x^9 \\ & + 174x^8 - 146x^7 + 454x^6 - 116x^5 + 444x^4 - 194x^3 + 86x^2 - 199x + 81. \end{aligned} \quad (5.72)$$

Es bleibt zu zeigen, dass der Term $m_6(x)$ für $x \in [0, 1)$ positiv ist.

Fall 1 ($x \in [0, 1/2]$):

Die Substitution $x = y/2$ in (5.72) liefert den Term $(p_5(y) + p_6(y))/131072$ mit

$$\begin{aligned} p_5(y) = & 33y^{17} + 146y^{16} + 152y^{15} + 624y^{14} - 3136y^{13} - 3712y^{12} - 11904y^{11} + 22272y^{10} \\ & - 37376y^9 + 89088y^8 - 149504y^7 + 929792y^6 \end{aligned}$$

und

$$p_6(y) = -475136y^5 + 3637248y^4 - 3178496y^3 + 2818048y^2 - 13041664y + 10616832.$$

Wegen Behauptung 5.13 gilt, dass $p_5(y)$ positiv ist. Es bleibt somit zu zeigen, dass für $p_6(y)$ selbiges gilt. Es gilt

$$\begin{aligned} p_6(y) & \geq -475136y^5 + 3637248y^4 - 3178496y^3 + 2818048y^2 - 13041664y + 10616832 \\ & =: p_7(y). \end{aligned}$$

Weiter gilt

$$\begin{aligned} p_7(y) & \geq \frac{1}{100000}(-5y^5 + 36y^4 - 32y^3 + 28y^2 - 131y + 106) \\ & \geq \frac{1}{100000}(31y^4 - 32y^3 + 28y^2 - 131y + 106). \end{aligned}$$

Definiere $p_8(y) := 31y^4 - 32y^3 + 28y^2 - 131y + 106$. Wegen

$$\begin{aligned} \frac{\partial}{\partial y} p_8(y) & = 124y^3 - 96y^2 + 56y - 131 \\ & \leq 124y^3 - 96y^3 + 56y - 56y - 75y^3 \\ & = -47y^3 \\ & < 0 \end{aligned}$$

und $p_8(1) = 2$ folgt, dass $p_8(y) \geq 0$. Dies bedeutet, dass $m_6(x)$ positiv ist.

Fall 2 ($x \in (1/2, 1)$):

Die Substitution $x = 1/2 + y/2$ in (5.72) liefert den Term $p_9(y)/131072$ mit

$$\begin{aligned} p_9(y) = & 33y^{17} + 707y^{16} + 6976y^{15} + 42864y^{14} + 181860y^{13} + 551388y^{12} + 1179696y^{11} \\ & + 1641472y^{10} + 954222y^9 - 1549446y^8 - 4927584y^7 - 6183440y^6 - 2071820y^5 \\ & + 9348076y^4 + 21227120y^3 + 22263584y^2 + 161177y + 1213307. \end{aligned}$$

Mit Behauptung 5.13 kann gezeigt werden, dass $p_9(y)$ positiv ist. Damit ist $m_6(x)$ ebenfalls positiv.

Zu b):

Das Minimum der Funktion $q_{1/2,1/8}(\lambda)$ stimmt mit dem Minimum des Terms $(q_{1/2,1/8}(\lambda))^2$ überein. Es gilt

$$(q_{1/2,1/8}(\lambda))^2 = \frac{m_7(\lambda)}{m_8(\lambda)} =: r_2(\lambda)$$

mit

$$m_7(\lambda) = 16(\lambda^{1/4} - 1)^2(\lambda - 4\lambda^{1/2} + 3)(\lambda^{3/2} + \lambda^{5/4} + \lambda + \lambda^{3/4} + \lambda^{1/2} + \lambda^{1/4} + 1)$$

und

$$m_8(\lambda) = (\lambda - 1)^2(7\lambda - 8\lambda^{7/8} - 56\lambda^{1/8} + 57).$$

Zunächst wird gezeigt, dass $r_2(\lambda) \geq 16/19$. Diese Ungleichung ist äquivalent zu

$$\begin{aligned} 3\lambda^3 + 2\lambda^{23/8} - \frac{19}{4}\lambda^{11/4} - 19\lambda^{5/2} + 19\lambda^{9/4} + 14\lambda^{17/8} + \frac{7}{2}\lambda^2 - 4\lambda^{15/8} - \frac{57}{4}\lambda^{7/4} \\ - \frac{19}{4}\lambda^{5/4} - 28\lambda^{9/8} + \frac{63}{2}\lambda + 2\lambda^{7/8} + 19\lambda^{3/4} - 19\lambda^{1/2} - \frac{57}{4}\lambda^{1/4} + 14\lambda^{1/8} \geq 0. \end{aligned} \quad (5.73)$$

Eine Substitution mittels $\lambda = x^8$ liefert für die linke Seite der Ungleichung den Term

$$\frac{1}{4}(1-x)^3x(x+1)^2(x^2+1)(m_9(x) + m_{10}(x))$$

mit

$$\begin{aligned} m_9(x) = & -12x^{16} - 20x^{15} - 13x^{14} - 21x^{13} + 50x^{12} + 34x^{11} + 48x^{10} - 24x^9 \\ & + 40x^8 - 24x^7 + 25x^6 \end{aligned}$$

und

$$m_{10}(x) = 15x^6 - 80x^5 + 34x^4 - 78x^3 + 55x^2 - x + 56.$$

Mit Behauptung 5.13 kann gezeigt werden, dass $m_9(x)$ positiv ist. Somit bleibt zu zeigen, dass die Funktion $m_{10}(x)$ für $x \in [0, 1)$ positiv ist. Es gilt

$$m_{10}(x) = (1 - x)m_{11}(x)$$

mit

$$m_{11}(x) = -14x^5 + 66x^4 + 32x^3 + 110x^2 + 55x + 56.$$

Wieder kann Behauptung 5.13 herangezogen werden um nachzuweisen, dass

$$m_{11}(x) \geq 0.$$

Somit ist die Ungleichung (5.73) gezeigt. Wegen $q_{1/2,1/8}(0) = (16/19)^{1/2}$ wird die Schranke angenommen. \square

Effizienzvergleich der Pläne mit $d = 1/8$ und $d = 1/4$

Behauptung 5.26. *Gegeben seien die beiden Versuchspläne mit $d = 1/8$ beziehungsweise $d = 1/4$. Dann gilt:*

a)

$$q_{1/8,1/4}(\lambda) \geq (9/10)^{1/2} \approx 0.949,$$

b)

$$\min_{\lambda} q_{1/4,1/8}(\lambda) = (52/57)^{1/2} \approx 0.955.$$

Beweis. Zu a):

Die Ungleichung

$$(q_{1/8,1/4}(\lambda))^2 \geq \frac{9}{10}$$

ist äquivalent zu der Ungleichung

$$\begin{aligned} & \frac{1093}{16}\lambda^3 - 80\lambda^{23/8} + \frac{63}{16}\lambda^{11/4} - \frac{289}{4}\lambda^{5/2} + 80\lambda^{19/8} + \frac{27}{4}\lambda^{9/4} - 560\lambda^{17/8} \\ & + \frac{8895}{16}\lambda^2 + \frac{117}{16}\lambda^{7/4} + 560\lambda^{13/8} - 640\lambda^{3/2} + 80\lambda^{11/8} + \frac{27}{16}\lambda^{5/4} + \frac{1057}{16}\lambda \\ & - 80\lambda^{7/8} + \frac{9}{4}\lambda^{3/4} + 560\lambda^{5/8} - \frac{2307}{4}\lambda^{1/2} + \frac{225}{16}\lambda^{1/4} - 560\lambda^{1/8} + \frac{9003}{16} \geq 0. \end{aligned} \quad (5.74)$$

Eine Substitution mittels $\lambda = x^8$ liefert für die linke Seite von (5.74) den Term

$$\frac{1}{16}(1 - x)^3(x + 1)^2m_{12}(x)$$

mit

$$\begin{aligned} m_{12}(x) = & -1093x^{19} + 187x^{18} - 2062x^{17} + 498x^{16} - 1751x^{15} + 809x^{14} - 1548x^{13} \\ & + 9972x^{12} - 1388x^{11} + 19092x^{10} - 1388x^9 + 19092x^8 - 268x^7 \\ & + 18932x^6 + 665x^5 + 18585x^4 + 354x^3 + 18274x^2 + 43x + 9003. \end{aligned}$$

Mit Behauptung 5.13 kann gezeigt werden, dass $m_{12}(x)$ positiv ist. Somit gilt die Behauptung.

Zu b):

Die Ungleichung

$$(q_{1/4,1/8}(\lambda))^2 \geq \frac{52}{57}$$

ist äquivalent zu der Ungleichung

$$\begin{aligned} \frac{15}{16}\lambda^3 - \frac{35}{16}\lambda^{11/4} + \frac{5}{4}\lambda^{5/2} - \frac{15}{4}\lambda^{9/4} + \frac{125}{16}\lambda^2 - \frac{65}{16}\lambda^{7/4} \\ - \frac{15}{16}\lambda^{5/4} + \frac{35}{16}\lambda - \frac{5}{4}\lambda^{3/4} + \frac{15}{4}\lambda^{1/2} - \frac{125}{16}\lambda^{1/4} + \frac{65}{16} \geq 0. \end{aligned} \quad (5.75)$$

Die Substitution $\lambda = x^4$ führt auf der linken Seite von (5.75) zu dem Term

$$\frac{5}{16}(1-x)^3(-3x^3+x^2+x+13)(x^6+x^5+x^4+x^3+x^2+x+1).$$

Der Term $-3x^3 + x^2 + x + 13$ ist wegen Behauptung 5.13 positiv. Da die weiteren Terme ebenfalls positiv sind, folgt, dass der gesamte Ausdruck positiv ist. Somit gilt die Behauptung. \square

Effizienzvergleich der Pläne mit $d = 1/8$ und $d = 3/8$

Behauptung 5.27. *Gegeben seien die beiden Versuchspläne mit $d = 1/8$ beziehungsweise $d = 3/8$. Dann gilt:*

a)

$$q_{1/8,3/8}(\lambda) \geq (17/20)^{1/2} \approx 0.922,$$

b)

$$\min_{\lambda} q_{3/8,1/8}(\lambda) = (49/57)^{1/2} \approx 0.927.$$

Beweis. Zu a):

Die ersten Überlegungen können analog zum Beweis von Behauptung 5.26 getätigt werden. Statt $\frac{1}{16}(1-x)^3(x+1)^2m_{12}(x) \geq 0$ ist hier dann aber zu zeigen, dass $(1-x)m_{13}(x)$ mit

$$\begin{aligned} m_{13}(x) = & 115x^{19} + 275x^{18} + 250x^{17} + 162x^{16} - 3x^{15} - 611x^{14} - 776x^{13} - 264x^{12} \\ & - 736x^{11} + 896x^{10} - 576x^9 + 2016x^8 - 456x^7 + 1976x^6 - 611x^5 \\ & + 1661x^4 - 926x^3 + 634x^2 - 813x + 307 \end{aligned} \quad (5.76)$$

positiv ist. Es gilt mit Behauptung 5.13

$$m_{13}(x) \geq -611x^5 + 1661x^4 - 926x^3 + 634x^2 - 813x + 307 =: m_{14}(x).$$

Der Term $m_{14}(x)$ ist konvex, denn die Minimalstelle der zweiten Ableitung von $m_{14}(x)$ hat den Funktionswert $515383/1928$. Analog zu den Überlegungen in Behauptung 5.24 kann nun gezeigt werden, dass die Tangente an der Stelle $121/200$ positiv ist. Da diese eine untere Schranke liefert, ist die Behauptung gezeigt.

Zu b):

Der Beweis kann zu Beginn wieder analog zum Beweis von Behauptung 5.26 geführt werden. Hier ist zu zeigen, dass $\frac{1}{64}(1-x)^3x(x+1)^2m_{15}(x)$ mit

$$\begin{aligned} m_{15}(x) = & -512x^{18} - 904x^{17} - 1073x^{16} - 489x^{15} - 315x^{14} + 2157x^{13} + 2331x^{12} \\ & + 2059x^{11} + 2233x^{10} - 783x^9 + 1841x^8 - 3527x^7 + 1547x^6 - 3429x^5 \\ & + 2157x^4 - 2427x^3 + 3159x^2 - 49x + 2744 \end{aligned}$$

für $x \in [0, 1)$ positiv ist. Für den Nachweis genügt es, Behauptung 5.13 heranzuziehen. \square

Effizienzvergleich der Pläne mit $d = 1/4$ und $d = 3/8$

Behauptung 5.28. *Gegeben seien die beiden Versuchspläne mit $d = 1/4$ beziehungsweise $d = 3/8$. Dann gilt:*

a)

$$q_{1/4,3/8}(\lambda) \geq (49/52)^{1/2} \approx 0.901,$$

b)

$$\min_{\lambda} q_{3/8,1/4}(\lambda) = (49/52)^{1/2} \approx 0.971.$$

Beweis. Zu a):

Erneut können die ersten Überlegungen analog zu den bisherigen Beweisen geführt werden. Nach Äquivalenzumformungen und der Substitution $\lambda = x^8$ ist hier zu zeigen, dass

$$\frac{1}{32}(1-x)^3(x+1)^2(x^2-x+1)(x^2+x+1)m_{16}(x) \geq 0$$

mit

$$\begin{aligned} m_{16} = & 3x^{15} + 3x^{14} + 262x^{13} - 50x^{12} + 14x^{11} - 818x^{10} + 209x^9 - 311x^8 \\ & + 521x^7 + 209x^6 + 1038x^5 + 206x^4 + 142x^3 - 378x^2 + 195x + 195. \end{aligned}$$

Die Funktion $m_{16}(x)$ ist wegen Behauptung 5.13 positiv.

Zu b):

Auch Teil b) kann analog gezeigt werden. Äquivalenzumformungen und die Substitution $\lambda = x^8$ liefern die Ungleichung

$$\frac{1}{16}(1-x)^3x^2(x+1)^2(x^2-x+1)(x^2+x+1)(m_{17}(x) + m_{18}(x)) \geq 0$$

mit

$$m_{17}(x) = -48x^{13} - 48x^{12} - 292x^{11} + 20x^{10} - 29x^9 + 397x^8$$

und

$$\begin{aligned} m_{18} = & 406x^8 - 29x^7 + 491x^6 - 341x^5 - 29x^4 - 813x^3 + 19x^2 + 68x + 588 \\ = & x(6x^7 + 42x^6 + (20 - 20x)^2x^5 + 28x^5 + (27 - 27x)^2x^4 + 37x^4 + (39 - 39x)^2x^3 \\ & + x^3 + x(p_{10}(x)) + (44 - 44x)^2x^2 + 14x^2 + 40x \\ & + (10 - 10x)^3 + 20) + (2 - x)^2(3 - x)(7 - x)^2. \end{aligned}$$

Dabei ist $p_{10}(x) = 3300x^2 - 5600x + 2400$. Die quadratische Funktion $p_{10}(x)$ hat ihr Minimum an der Stelle $x = 28/33$. Daraufhin gilt wegen $p_{10}(28/33) = 800/33 > 0$, dass $p_{10}(x)$ nur positive Funktionswerte besitzt. Somit ist zu erkennen, dass $m_{18}(x)$ positiv ist. Es bleibt zu zeigen, dass $m_{17}(x)$ positiv ist. Dies kann mit Behauptung 5.13 gezeigt werden. \square

Effizienzvergleich der Pläne mit $d = 3/8$ und $d = 1/2$

Behauptung 5.29. Gegeben seien die beiden Versuchspläne mit $d = 3/8$ beziehungsweise $d = 1/2$. Dann gilt:

a)

$$q_{3/8,1/2}(\lambda) \geq (17/20)^{1/2} \approx 0.922,$$

b)

$$\min_{\lambda} q_{1/2,3/8}(\lambda) = (48/49)^{1/2} \approx 0.990.$$

Beweis. Zu a):

Der Beweis kann erneut analog durchgeführt werden. Nach Äquivalenzumformungen und der Substitution $\lambda = x^8$ ist dann hier zu zeigen, dass

$$\frac{1}{8}(1-x)^3(x+1)^2(x^2+1)m_{19}(x) \geq 0$$

mit

$$\begin{aligned} m_{19}(x) = & -7x^{17} - 7x^{16} - 14x^{15} + 106x^{14} - 180x^{13} + 140x^{12} - 214x^{11} + 226x^{10} - 298x^9 \\ & + 342x^8 - 46x^7 + 474x^6 - 84x^5 + 236x^4 - 118x^3 + 82x^2 + 41x + 41. \end{aligned}$$

Die Funktion $m_{19}(x)$ ist wegen Behauptung 5.13 positiv.

Zu b):

Hier liefern Äquivalenzumformungen und die Substitution $\lambda = x^8$ die Ungleichung

$$\frac{1}{2}(1-x)^3x^3(x+1)^2(x^2+1)(m_{20}(x) + m_{21}(x)) \geq 0$$

mit

$$m_{20}(x) = -4x^{14} - 4x^{13} - 8x^{12} - 80x^{11} + 108x^{10} - 84x^9 + 89x^8$$

und

$$m_{21}(x) = 64x^8 - 111x^7 + 228x^6 - 156x^5 + 85x^4 - 227x^3 + 116x^2 - 76x + 120.$$

Wegen Behauptung 5.13 ist $m_{20}(x)$ positiv. Es bleibt zu zeigen, dass $m_{21}(x)$ ebenfalls positiv ist. Nun

$$\begin{aligned} m_{21}(x) = & x^7 + (8-8x)^2x^6 + (4-4x)^2x^5 + 24x^5 + (14-14x)^2x^4 + (14-14x)^2x^3 + 18x^3 \\ & + (5-5x)^2x^2 + (16-16x)^2x^2 + (11-11x)^2x + (120-77x)(1-x). \end{aligned}$$

Anhand dieser Darstellung ist zu erkennen, dass $m_{21}(x)$ positiv ist. Damit gilt die Behauptung. \square

6. D-Effizienzen von Versuchsplänen bei unbekanntem Störparametern

In diesem Kapitel wird vorausgesetzt, dass die vier Parameter β_0 , β_1 , σ^2 und λ unbekannt sind. Abweichend zu Kapitel 5 wird also der Fall betrachtet, dass λ und σ^2 zusätzlich unbekannt sind und somit ebenfalls geschätzt werden müssen. Bezüglich der Versuchsdesigns wird in diesem Kapitel die Einschränkung vorgenommen, dass jede Versuchseinheit drei Versuchspunkte aufweist. Dabei stehen nur Versuchspläne zur Auswahl, bei denen innerhalb einer Versuchseinheit zwei Versuchspunkte mit $x_1 = 0$ und $x_3 = 1$ bereits festgelegt sind, so dass nur noch die Wahl des dritten Versuchspunkts zu untersuchen ist. Es stellt sich nun die Frage, welcher Versuchsplan in der vorliegenden Situation eine gute Schätzung der Regressionskoeffizienten ermöglicht. Da die Restriktionen bezüglich der Versuchspläne auch im vorherigen Kapitel Gegenstand der Untersuchungen war, ist es das Ziel, die jeweiligen Resultate miteinander zu vergleichen.

In Abschnitt 6.1 wird zunächst ein Verfahren vorgestellt, mit dem es möglich ist, die Parameter β_0 und β_1 bei unbekanntem Störparametern zu schätzen. In Abschnitt 6.2 wird daraufhin eine Simulationsstudie durchgeführt, in der Versuchspläne bezüglich der Schätzung des Vektors $[\beta_0, \beta_1]^T$ miteinander verglichen werden. Da in dem Verfahren aus Abschnitt 6.1 auch eine Schätzung des Störparameters λ erfolgt, wird ebenfalls in Abschnitt 6.2 untersucht, wie sich die Wahl des Versuchsplans auf die Schätzungen dieses Parameters auswirkt. Die Simulationsstudie in Abschnitt 6.2 erfolgt für vier verschiedene Anzahlen an Versuchseinheiten, so dass die Anzahl der Versuchseinheiten als weiterer Aspekt in die Untersuchungen eingeht.

6.1. Zweistufige Schätzung

In diesem Abschnitt wird ein zweistufiges Verfahren vorgestellt, das dem Vorgehen in Fahrmeir *et al.* (2007), S. 142 f., angelehnt ist. In der zitierten Arbeit werden lineare Modelle betrachtet, in denen die Fehlerterme einem autoregressiven Prozess erster Ordnung folgen. Hier dagegen hängt die Abhängigkeit nicht von der Ordnung der Beobachtungen ab, sondern von der Distanz der Versuchspunkte.

Ein Bestandteil des Vorgehens hier ist es, zwei Parameter mittels der Maximum-Likelihood-Methode zu schätzen. Für die Bestimmung der Likelihood-Funktion ist jedoch

eine Fallunterscheidung erforderlich. Der Grund ist der, dass für $\lambda = 0$ und für $\lambda \in (0, 1)$, wobei $d \in (0, 1)$ ist, die Matrix $\Sigma_{(1)}^{-1}(\lambda)$ herangezogen wird, diese jedoch für $\lambda \in (0, 1)$, wobei $d \in \{0, 1\}$ ist, nicht verwendet werden kann, siehe Abschnitt 5.3.1. Dabei zeigte sich in den Ausführungen in Abschnitt 5.3.1, dass der dritte Versuchspunkt in den Fällen $\lambda \in (0, 1)$, wobei $d \in \{0, 1\}$, keine zusätzlichen Informationen liefert. Da eine doppelte Beobachtung auch hier keinen Informationsgewinn liefert, wird in diesem Abschnitt analog vorgegangen. Somit setzt sich eine Versuchsreihe für den Spezialfall $\lambda \in (0, 1)$ mit $d \in \{0, 1\}$ lediglich aus den beiden Randpunkten zusammen.

Fall 1 ($\lambda = 0$):

Die Likelihood-Funktion lautet

$$L_y^{(1)}(\beta_0, \beta_1, \sigma^2, \lambda) = \frac{1}{(2\pi)^{3n/2} \det(I_n \otimes \sigma^2 I_3)^{1/2}} \exp\left(-\frac{1}{2\sigma^2} (y - X_{(1)}^{[n]}\beta)^T (y - X_{(1)}^{[n]}\beta)\right), \quad (6.1)$$

wobei $X_{(1)}^{[n]} := 1_n \otimes X_{(1)}$.

Fall 2 ($\lambda \in (0, 1)$):

Die Likelihood-Funktion lautet

$$L_y^{(1)}(\beta_0, \beta_1, \sigma^2, \lambda) = \begin{cases} \frac{1}{(2\pi)^n \det(I_n \otimes \Sigma_{(0)}(\sigma^2, \lambda))^{1/2}} \exp\left(-\frac{1}{2} (y - X_{(0)}^{[n]}\beta)^T \dots \right. \\ \quad \left. \dots (I_n \otimes \Sigma_{(0)}^{-1}(\sigma^2, \lambda))(y - X_{(0)}^{[n]}\beta)\right) & \text{falls } d \in \{0, 1\}, \\ \frac{1}{(2\pi)^{3n/2} \det(I_n \otimes \Sigma_{(1)}(\sigma^2, \lambda))^{1/2}} \exp\left(-\frac{1}{2} (y - X_{(1)}^{[n]}\beta)^T \dots \right. \\ \quad \left. \dots (I_n \otimes \Sigma_{(1)}^{-1}(\sigma^2, \lambda))(y - X_{(1)}^{[n]}\beta)\right) & \text{falls } d \in (0, 1), \end{cases} \quad (6.2)$$

wobei $X_{(0)}^{[n]} := 1_n \otimes X_{(0)}$.

Für die Schätzung der Parameter wird nun das zweistufige Verfahren herangezogen. In einem ersten Schritt werden die Parameter β_0 und β_1 mittels der gewöhnlichen Kleinst-Quadrate-Methode geschätzt. Diese Schätzwerte werden jedoch nicht für die Schätzung des Modells verwendet. Stattdessen werden sie für eine Schätzung der beiden anderen Parameter benutzt. Dies geschieht im zweiten Schritt wie folgt. Der aus der gewöhnlichen Kleinst-Quadrate-Methode resultierende Vektor der Residuen $\hat{e}^{(1)} = y - X(X^T X)^{-1} X^T y$ liefert eine Schätzung des Fehlervektors e . Wird nun in (6.1) beziehungsweise (6.2) der Vektor $y - X\beta$ durch $\hat{e}^{(1)}$ ersetzt, so erscheinen die Parameter β_0 und β_1 nicht mehr in

$L_y^{(1)}(\beta_0, \beta_1, \sigma^2, \lambda)$. Die neue Likelihood-Funktion lautet daraufhin für $\lambda = 0$

$$L_y^{(2)}(\sigma^2, \lambda) = \frac{1}{(2\pi)^{3n/2} \det(I_n \otimes \sigma^2 I_3)^{1/2}} \exp\left(-\frac{1}{2\sigma^2} (\hat{e}^{(1)})^T \hat{e}^{(1)}\right) \quad (6.3)$$

beziehungsweise für $\lambda \in (0, 1)$

$$L_y^{(2)}(\sigma^2, \lambda) = \begin{cases} \frac{1}{(2\pi)^n \det(I_n \otimes \Sigma_{(0)}(\sigma^2, \lambda))^{1/2}} \exp\left(-\frac{1}{2} (\hat{e}^{(1)})^T (I_n \otimes \Sigma_{(0)}^{-1}(\sigma^2, \lambda)) \hat{e}^{(1)}\right) & \text{falls } d \in \{0, 1\}, \\ \frac{1}{(2\pi)^{3n/2} \det(I_n \otimes \Sigma_{(1)}(\sigma^2, \lambda))^{1/2}} \exp\left(-\frac{1}{2} (\hat{e}^{(1)})^T (I_n \otimes \Sigma_{(1)}^{-1}(\sigma^2, \lambda)) \hat{e}^{(1)}\right) & \text{falls } d \in (0, 1). \end{cases} \quad (6.4)$$

Als Nächstes werden die Werte für die verbliebenen Parameter σ^2 und λ gesucht, welche $L_y^{(2)}(\sigma^2, \lambda)$ maximieren. An dieser Stelle ist die erste Stufe zunächst einmal abgeschlossen. In der zweiten Stufe werden die Regressionskoeffizienten β_0 und β_1 mittels der verallgemeinerten Kleinste-Quadrate-Methode geschätzt, wobei die Maximum-Likelihood-Schätzer für σ^2 und λ aus dem zweiten Schritt der ersten Stufe herangezogen werden. Das Vorgehen kann nun mehrfach wiederholt werden. Dabei werden in den Wiederholungen des zweiten Schritts in der ersten Stufe von nun an die Residuen, die aus der zweiten Stufe resultieren, herangezogen. Die Wiederholungen haben die Absicht, die Schätzungen der verschiedenen Parameter zu verbessern.

Das Vorgehen zur Bestimmung der finalen Schätzwerte für β sei nun noch einmal kurz zusammengefasst.

Stufe 1:

- (1) Schätze die Parameter β_0 und β_1 mittels der gewöhnlichen Kleinste-Quadrate-Methode. Bestimme daraufhin die Residuen.
- (2) Ersetze in der Likelihood-Funktion den Term $y - X\beta$ durch den Vektor der Residuen aus Schritt (1). Schätze daraufhin die Parameter σ^2 und λ mittels der Maximum-Likelihood-Methode.

Stufe 2:

- (3) Schätze die Parameter β_0 und β_1 mittels der verallgemeinerten Kleinste-Quadrate-Methode. Verwende dabei die Schätzwerte für σ^2 und λ aus Schritt (2).

Wiederholung des zweistufigen Verfahrens:

- (4) Wiederhole die Schritte (2) und (3) bis Konvergenz der Schätzwerte erreicht ist. Verwende in Schritt (2) dabei abweichend die Residuen aus Schritt (3) statt aus Schritt (1).

Die folgenden Ausführungen gehen auf den Maximum-Likelihood-Schätzer für σ^2 und λ ein, der im zweiten Schritt der ersten Stufe Anwendung findet. Die Überlegungen reduzieren die Komplexität der zu untersuchenden Log-Likelihood-Funktionen, welche in den Simulationsstudien in Abschnitt 6.2 herangezogen werden. Dies hat den Vorteil, dass der Rechenaufwand bedeutsam reduziert werden kann. Für $\lambda = 0$ lautet die Log-Likelihood-Funktion

$$\ln(L_y^{(2)}(\sigma^2, \lambda)) = -\frac{3n}{2}\ln(2\pi) - \frac{1}{2}\ln(\det(I_n \otimes \sigma^2 I_3)) - \frac{1}{2\sigma^2}(\hat{e}^{(2)})^T \hat{e}^{(2)} \quad (6.5)$$

beziehungsweise für $\lambda \in (0, 1)$

$$\ln(L_y^{(2)}(\sigma^2, \lambda)) = \begin{cases} -n\ln(2\pi) - \frac{1}{2}\ln(\det(I_n \otimes \Sigma_{(0)}(\sigma^2, \lambda))) \\ \quad - \frac{1}{2\sigma^2}(\hat{e}^{(2)})^T (I_n \otimes \Sigma_{(0)}^{-1}(\lambda)) \hat{e}^{(2)} & \text{falls } d \in \{0, 1\}, \\ -\frac{3n}{2}\ln(2\pi) - \frac{1}{2}\ln(\det(I_n \otimes \Sigma_{(1)}(\sigma^2, \lambda))) \\ \quad - \frac{1}{2\sigma^2}(\hat{e}^{(2)})^T (I_n \otimes \Sigma_{(1)}^{-1}(\lambda)) \hat{e}^{(2)} & \text{falls } d \in (0, 1). \end{cases} \quad (6.6)$$

Dabei stellt $\hat{e}^{(2)}$ den Vektor der Residuen dar, der aus der verallgemeinerten Kleinst-Quadrat-Methode resultiert. Die Log-Likelihood-Funktion wird nun als univariate Funktion in σ^2 aufgefasst, so dass eine Funktionenschar mit Parameter λ vorliegt. Definiere daher $\ln(L_y^{(2)}(\sigma^2, \lambda)) =: f_\lambda(\sigma^2)$.

Die erste Ableitung von $f_\lambda(\sigma^2)$ nach σ^2 lautet für $\lambda = 0$

$$\frac{\partial}{\partial \sigma^2} f_\lambda(\sigma^2) = -\frac{3n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}(\hat{e}^{(2)})^T \hat{e}^{(2)} \quad (6.7)$$

beziehungsweise für $\lambda \in (0, 1)$

$$\frac{\partial}{\partial \sigma^2} f_\lambda(\sigma^2) = \begin{cases} -\frac{n}{\sigma^2} + \frac{1}{2(\sigma^2)^2}(\hat{e}^{(2)})^T (I_n \otimes \Sigma_{(0)}^{-1}(\lambda)) \hat{e}^{(2)} & \text{falls } d \in \{0, 1\}, \\ -\frac{3n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}(\hat{e}^{(2)})^T (I_n \otimes \Sigma_{(1)}^{-1}(\lambda)) \hat{e}^{(2)} & \text{falls } d \in (0, 1). \end{cases} \quad (6.8)$$

Nullsetzen von (6.7) und (6.8) liefert für $\lambda = 0$

$$\sigma_{(1)}^2 = \frac{(\hat{e}^{(2)})^T \hat{e}^{(2)}}{3n} \quad (6.9)$$

beziehungsweise für $\lambda \in (0, 1)$

$$\sigma_{(1)}^2 = \begin{cases} \frac{(\hat{e}^{(2)})^T (I_n \otimes \Sigma_{(0)}^{-1}(\lambda)) \hat{e}^{(2)}}{2n} & \text{falls } d \in \{0, 1\}, \\ \frac{(\hat{e}^{(2)})^T (I_n \otimes \Sigma_{(1)}^{-1}(\lambda)) \hat{e}^{(2)}}{3n} & \text{falls } d \in (0, 1) \end{cases} \quad (6.10)$$

als jeweils einzige Lösung. Somit sind dies die alleinigen Kandidaten für lokale Extremstellen.

Die zweite Ableitung von $f_\lambda(\sigma^2)$ nach σ^2 lautet für $\lambda = 0$

$$\frac{\partial^2}{\partial(\sigma^2)^2} f_\lambda(\sigma^2) = \frac{3n}{2(\sigma^2)^2} - \frac{(\hat{e}^{(2)})^T \hat{e}^{(2)}}{(\sigma^2)^3} \quad (6.11)$$

beziehungsweise für $\lambda \in (0, 1)$

$$\frac{\partial^2}{\partial(\sigma^2)^2} f_\lambda(\sigma^2) = \begin{cases} \frac{n}{(\sigma^2)^2} - \frac{(\hat{e}^{(2)})^T (I_n \otimes \Sigma_{(0)}^{-1}(\lambda)) \hat{e}^{(2)}}{(\sigma^2)^3} & \text{falls } d \in \{0, 1\}, \\ \frac{3n}{2(\sigma^2)^2} - \frac{(\hat{e}^{(2)})^T (I_n \otimes \Sigma_{(1)}^{-1}(\lambda)) \hat{e}^{(2)}}{(\sigma^2)^3} & \text{falls } d \in (0, 1). \end{cases} \quad (6.12)$$

Einsetzen der Kandidaten aus (6.9) und (6.10) in die zweiten Ableitungen liefert für $\lambda = 0$

$$\frac{\partial^2}{\partial(\sigma^2)^2} f_\lambda(\sigma_{(1)}^2) = -\frac{27n^3}{2((\hat{e}^{(2)})^T \hat{e}^{(2)})^2} \quad (6.13)$$

beziehungsweise für $\lambda \in (0, 1)$

$$\frac{\partial^2}{\partial(\sigma^2)^2} f_\lambda(\sigma_{(1)}^2) = \begin{cases} -\frac{4n^3}{((\hat{e}^{(2)})^T (I_n \otimes \Sigma_{(0)}^{-1}(\lambda)) \hat{e}^{(2)})^2} & \text{falls } d \in \{0, 1\}, \\ -\frac{27n^3}{2((\hat{e}^{(2)})^T (I_n \otimes \Sigma_{(1)}^{-1}(\lambda)) \hat{e}^{(2)})^2} & \text{falls } d \in (0, 1). \end{cases} \quad (6.14)$$

Die Größe n ist stets positiv. Die Nenner können auf Grund der Quadrierung jeweils höchstens Null werden, falls der Vektor $\hat{e}^{(2)}$ den Nullvektor darstellt oder falls die Inversen der Korrelationsmatrizen keinen Vollrang aufweisen. Da der Nullvektor die Wahrscheinlichkeit Null besitzt, wird dieser Fall nicht weiter verfolgt. Zudem weisen die Matrizen

Vollrang auf, da es sich um Inversen handelt, deren Existenz gegeben ist. Somit liefern die Kandidaten $\sigma_{(1)}^2$ lokale Maximalstellen.

Nun ist noch zu klären, ob die lokalen Maximalstellen auch die globalen Maximalstellen darstellen. Da jedoch für jede Funktion der Funktionenschar nur genau eine lokale Extremstelle existiert und die Funktionen $f_\lambda(\sigma^2)$ in σ^2 stetig sind, ist die jeweilige lokale Extremstelle gleichzeitig auch das globale Maximum. Dieses Erkenntnis erlaubt es, dass die Log-Likelihood-Funktion bezüglich der Suche nach dem Maximum-Likelihood-Schätzer nicht mehr über den gesamten Parameterraum untersucht werden muss, sondern nur über den Teilraum, der bezüglich der Funktionenschar $f_\lambda(\sigma^2)$ zu den globalen Maximalstellen führt. Der Teilraum kann nun analysiert werden, indem die lokalen Extremstellen $\sigma_{(1)}^2$ in die Log-Likelihood-Funktion eingesetzt werden. Die neuen Log-Likelihood-Funktionen lauten für $\lambda = 0$

$$\ln(L_y^{(2)}(\sigma_{(1)}^2, \lambda)) = -\frac{3n}{2}\ln(2\pi) - \frac{3n}{2}\ln\left(\frac{(\hat{e}^{(2)})^T \hat{e}^{(2)}}{3n}\right) - \frac{3n}{2} =: g(0) \quad (6.15)$$

beziehungsweise für $\lambda \in (0, 1)$

$$\ln(L_y^{(2)}(\sigma_{(1)}^2, \lambda)) = \begin{cases} -n\ln(2\pi) - n\ln((\hat{e}^{(2)})^T (I_n \otimes \Sigma_{(0)}^{-1}(\lambda)) \hat{e}^{(2)}) + n\ln(2n) \\ \quad - \frac{n}{2}\ln(\det(\Sigma_{(0)}(\lambda))) - n & \text{falls } d \in \{0, 1\}, \\ -\frac{3n}{2}\ln(2\pi) - \frac{3n}{2}\ln((\hat{e}^{(2)})^T (I_n \otimes \Sigma_{(1)}^{-1}(\lambda)) \hat{e}^{(2)}) + \frac{3n}{2}\ln(3n) \\ \quad - \frac{n}{2}\ln(\det(\Sigma_{(1)}(\lambda))) - \frac{3n}{2} & \text{falls } d \in (0, 1) \end{cases} \quad (6.16)$$

$=: g(\lambda).$

Mit $g(\lambda)$ liegt jetzt eine Funktion vor, die nur noch vom Parameter λ abhängt. Die Maximalstelle dieser Funktion wird im Weiteren jedoch nur numerisch bestimmt, da sich das Auffinden einer geschlossenen Darstellung auf Grund der Komplexität der Funktion $g(\lambda)$ als schwierig erweist.

Der Schätzwert für λ kann nun in den verallgemeinerten Kleinste-Quadrate-Schätzer eingesetzt werden, so dass die Regressionskoeffizienten β_0 und β_1 geschätzt werden können.

In Oberhofer und Kmenta (1974) wird im Vergleich zu der hier vorliegenden Situation für allgemeinere Rahmenbedingungen gezeigt, dass die Schätzwerte aus den Wiederholungen des zweistufigen Verfahrens konvergieren, falls gewisse Annahmen erfüllt sind. In der

Arbeit wird daraufhin unter anderem die Anwendbarkeit des Verfahrens für das verallgemeinerte lineare Regressionsmodell demonstriert. Die ersten fünf der sechs Annahmen, die die Konvergenz garantieren, lauten in der hier vorliegenden Notation:

(A1) Der Fehlervektor e ist normalverteilt.

(A2) $E(e) = 0$.

(A3) $Cov(e) = W(\rho)$, wobei die Determinante von $W(\rho)$ mit $\rho = [\sigma^2, \lambda]^T$ ungleich Null ist.

(A4) Die Designmatrix $X \in \mathbb{R}^{nq \times p}$ ist nicht stochastisch und besitzt Vollrang mit $nq > p$.

(A5) Die Parameter des Vektors β sind unabhängig von denen, die in W vorkommen.

Für das Modell aus Kapitel 2 sind die erste, zweite und fünfte Annahme offensichtlich erfüllt. Damit die dritte Annahme erfüllt ist, ist der Parameterraum so einzuschränken, dass dieser die Fälle $\sigma^2 = 0$ und $\lambda = 1$ nicht enthält. Diese Einschränkung stellt kein Problem dar, da in den beiden Fällen sämtliche Realisierungen des Zufallsvektors y auf einer Geraden liegen, so dass alle Residuen Null sind, falls der Kleinste-Quadrate-Schätzer herangezogen wird. Damit ist die Varianz des Kleinste-Quadrate-Schätzers Null. Weiter ist die vierte Annahme erfüllt, wenn hinreichend viele Versuchspunkte vorliegen. Sie müssen zudem so gewählt werden, dass die Designmatrix Vollrang besitzt.

Schließlich wird in der sechsten Annahme gefordert, dass die verschiedenen Schätzwerte für den Parametervektor $\rho = [\sigma^2, \lambda]^T$, die in den Iterationen nach und nach vorliegen, allesamt in einem speziellen Parameterraum liegen. Dieser hat unter anderem die Eigenschaft, dass er abgeschlossen ist. Analog zum ersten Beispiel in Kapitel 4 in Oberhofer und Kmenta (1974) genügt es in der hier vorliegenden Situation lediglich zu fordern, dass die Schätzwerte für σ^2 größer Null und die für λ aus dem Intervall $[0, 1)$ sind. Anhand aller vorliegenden Schätzwerte können dann im Nachhinein abgeschlossene Intervalle konstruiert werden, so dass auch die sechste Annahme erfüllt ist.

Das zweistufige Vorgehen weist außerdem die Eigenschaft auf, dass es asymptotisch effizient ist, falls die Parameter des Vektors ρ konsistent geschätzt werden, siehe Greene (2012), S. 308. Da hier für die Schätzung der beiden Parameter σ^2 und λ der Maximum-Likelihood-Schätzer herangezogen wird, liegt ein konsistenter Schätzer vor. Damit stellt das iterative Vorgehen für den Fall $n \rightarrow \infty$ ein effizientes Verfahren dar. Im endlichen Fall hängt die Güte des zweistufigen Vorgehens insbesondere von der Gestalt der Kovari-

anzmatrix ab, siehe Greene (2012), S. 308. Daher wird für das vorliegende Modell im nun folgenden Abschnitt eine Simulationsstudie durchgeführt, die die zweistufige Schätzung aus diesem Abschnitt heranzieht.

6.2. Simulationsstudie

In diesem Abschnitt wird eine Simulationsstudie durchgeführt. Dabei wird das zweistufige Schätzverfahren aus dem vorherigen Abschnitt herangezogen. Im Detail sieht die Simulation, die mit der Software R in der Version 2.14.0, R Development Core Team (2011), durchgeführt wird, wie folgt aus. Bezüglich der Größe n , welche die Anzahl an Versuchseinheiten bezeichnet, werden vier verschiedene Fälle betrachtet. Diese lauten $n = 2$, $n = 5$, $n = 10$ und $n = 50$. Für jeden Fall ist es zunächst erforderlich, dass Datensätze erzeugt werden, anhand derer die Parameter geschätzt werden können. Zur Erzeugung der Daten werden Zufallszahlen aus multivariaten Normalverteilungen mittels der Funktion `rmvnorm` erzeugt, wobei nacheinander sämtliche Konstellationen der Parameterwerte und Versuchspunkte aus Tabelle 6.1 betrachtet werden.

Tabelle 6.1: *Parameterwerte und Versuchspunkte anhand derer die Daten erzeugt werden.*

λ	0, 0.1 ⁴ , 0.1 ³ , 0.1 ² , 0.1, 0.25, 0.5, 0.9
σ^2	1
β_0	0
β_1	0
d	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 1

Für alle möglichen Konstellationen wird nun die Güte des zweistufigen Schätzverfahrens untersucht. Dabei ist das Vorgehen innerhalb einer einzelnen Konstellation wie folgt. Für jeden Datensatz erfolgt eine Schätzung der Parameter. Die Schätzwerte in den Schritten (1) und (3) resultieren einfach aus der Multiplikation von Matrizen und Vektoren. In Schritt (2) dagegen ist der Schätzwert des Maximum-Likelihood-Schätzers numerisch zu bestimmen. Dies geschieht mittels der Funktion `nlminb`, die ein Quasi-Newton-Verfahren durchführt. Um zu verhindern, dass dabei auf Grund eines ungünstig gewählten Startwerts das globale Maximum nicht gefunden wird, werden $m = 9$ verschiedene Startwerte herangezogen. Sie lauten 0.1, 0.2, ..., 0.9. Jeder Startwert liefert nun jeweils eine Maximalstelle.

Daraufhin wird die Maximalstelle ausgewählt, die das größte Maximum liefert. Dabei beschränkt sich die Suche nach den Maximalstellen zunächst auf das Intervall $[0.1^6, 1 - 0.1^6]$. Im Anschluss erfolgt dann noch ein Abgleich mit dem Wert der Likelihood-Funktion an der Stelle $\lambda = 0$. Dieses Vorgehen ist erforderlich, da die Funktion `nllminb` die Maximalstellen nicht findet, falls das Intervall die Null enthält. Das zweistufige Verfahren wird nun weitere neun Mal wiederholt mit dem Ziel, dass die Schätzwerte konvergieren. Dabei zeigt sich für jeden Datensatz, dass ab der dritten Wiederholung die Schätzwerte für die Parameter β_0 und β_1 in mindestens sechs Nachkommastellen identisch sind. Die Anzahl der Wiederholungen ist damit ausreichend. Das Vorgehen bis hier hin wird jetzt $r = 10000$ Mal wiederholt, so dass für jede Konstellation 10000 Vektoren mit Schätzwerten vorliegen. Im nächsten Schritt wird die empirische Kovarianzmatrix der in einer Matrix zusammengefassten Vektoren bestimmt. Daraufhin wird das Maximum der absoluten Abweichungen zwischen den vier Einträgen der empirischen Kovarianzmatrix und den zugehörigen Einträgen der Inversen der Fisher-Informationsmatrix ermittelt. Letztere sind im Anhang in Tabelle B.5 für sämtliche Konstellationen aus Tabelle 6.1 aufgeführt. Damit liegt die erste Kennzahl vor, die für die Simulationsstudie relevant ist. Um eine weitere univariate Kennzahl zu erhalten, wird abschließend die Determinante der empirischen Kovarianzmatrix berechnet. Zum Vergleich wird daraufhin die Determinante der Inversen der Fisher-Informationsmatrix herangezogen.

In die Simulationsstudie fließt nun die folgende Idee ein. Falls die erste Kennzahl kleine Werte liefern sollte, so hat das die Konsequenz, dass die Differenz der beiden Determinanten ebenfalls klein wird. Dies folgt aus Feststellung 6.1, denn kleine absolute Abweichungen führen dazu, dass das Produkt $a_1 a_2$ einen ähnlichen Wert wie das Produkt $b_1 b_2$ aufweist. Gleiches gilt für die beiden verbleibenden Quadrate a_3^2 und b_3^2 .

Feststellung 6.1. Seien $A \in \mathbb{R}^{2 \times 2}$ und $B \in \mathbb{R}^{2 \times 2}$ symmetrische Matrizen mit

$$A = \begin{bmatrix} a_1 & a_3 \\ a_3 & a_2 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} b_1 & b_3 \\ b_3 & b_2 \end{bmatrix}.$$

Dann gilt:

$$\det(A) - \det(B) = a_1 a_2 - b_1 b_2 - (a_3^2 - b_3^2).$$

Eine besondere Situation liegt für die Versuchspläne mit den Versuchspunkten $d = 0$ oder $d = 1$ bei unkorrelierten Fehlern vor. In dieser Situation ist es nicht sinnvoll den Korrelationsparameter zu schätzen, da Schätzwerte größer Null nicht auftreten können. Dies liegt daran, dass die Residuen der beiden Beobachtungen, die an der selben Stelle erhoben

wurden, für $\lambda > 0$ auf Grund einer Korrelation von $\lambda^0 = 1$ identisch sein müssen. Unter Normalverteilung besitzt dieser Fall bei unkorrelierten Fehlern jedoch die Wahrscheinlichkeit Null. Das hat zur Folge, dass das zweistufige Verfahren hier nicht angebracht ist. Da die Fehlerterme unkorreliert sind, ist der gewöhnliche Kleinste-Quadrate-Schätzer die beste Wahl.

Die ersten Ergebnisse der Simulationsstudie können im Anhang in den Tabellen B.1 bis B.4 eingesehen werden. Dort werden für sämtliche Korrelationen und ausgewählte Versuchspunkte aus Tabelle 6.1 die maximalen absoluten Abweichungen zwischen den Einträgen der empirischen Kovarianzmatrizen und den entsprechenden Einträgen der Inversen der Fisher-Informationsmatrix aufgeführt, wobei die Einträge der empirischen Kovarianzmatrizen aus Normierungsgründen mit dem Faktor n multipliziert worden sind. In Abbildung 6.1 werden die in den Tabellen aufgeführten Abweichungen grafisch veranschaulicht. Es ist zu erkennen, dass die maximalen absoluten Abweichungen für $n = 2$ tendenziell am größten sind. Für die weiteren Fälle kann registriert werden, dass die Quartile recht ähnlich sind. Entscheidend ist aber, dass die absoluten Abweichungen zwischen den Einträgen der beiden Matrizen unter Beachtung ihrer Größenordnung und Variabilität klein ausfallen, was anhand der Einträge in Tabelle B.5 und dem Skalenniveau in Abbildung 6.1 ersichtlich ist. Diese Beobachtung zeigt, dass das zweistufige Schätzverfahren auf Grund der geringen absoluten Abweichungen bereits für eine kleine Anzahl an Versuchseinheiten ein effizientes Verfahren zur Schätzung der Regressionsparameter darstellt. Es sei angemerkt, dass die Versuchspläne mit den Versuchspunkten $d = 0$ und $d = 1$ in den Tabellen außen vor bleiben. Dies hat zwei Ursachen. Für unkorrelierte Fehler ist die Ursache die, dass, wie im vorherigen Absatz angesprochen, das zweistufige Schätzverfahren für derartige Fälle nicht angewendet wird. Bei korrelierten Fehlern liegt dagegen das Problem vor, dass bei jeder Versuchseinheit eine Beobachtung weniger in die Schätzung einfließt. Damit ist im Vergleich zu den anderen Versuchsplänen zu erwarten, dass das zweistufige Schätzverfahren zu vergleichsweise schlechteren Schätzungen führt. Wären die Versuchspunkte in die Untersuchungen aufgenommen worden, so wäre die Güte des zweistufigen Schätzverfahrens für die vorliegenden Rahmenbedingungen nicht so leicht zu erkennen gewesen.

In Tabelle 6.2 bis 6.5 (Seite 88 f.) können die zentralen Ergebnisse der Simulationsstudie eingesehen werden. Die Tabellen enthalten zum einen für sämtliche Konstellationen und Versuchspunkte aus Tabelle 6.1 die Determinanten der empirischen Kovarianzmatrizen. Diese Werte sind mit dem Faktor n^2 multipliziert worden. In der darunter liegenden Zeile

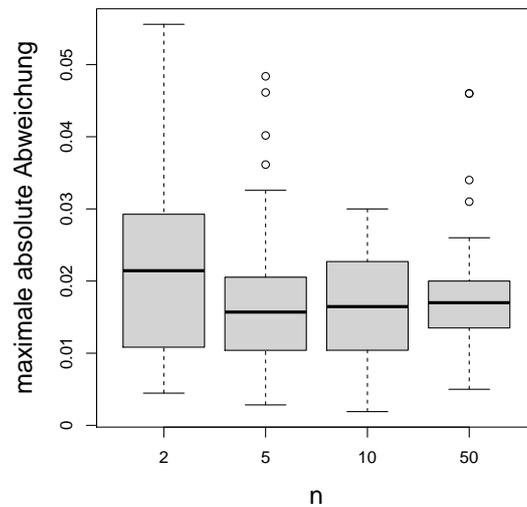


Abbildung 6.1: *Boxplots zur Illustration der Konsistenz des Schätzers für β aus dem zwei-stufigen Verfahren. Die Boxplots geben die maximalen absoluten Abweichungen zwischen den Einträgen der empirischen Kovarianzmatrix und den Einträgen der Inversen der Fisher-Informationsmatrix wider. Die empirischen Kovarianzmatrizen sind mit dem Faktor n multipliziert worden, damit für die verschiedenen Anzahlen an Versuchseinheiten ein Bezug zur selben Informationsmatrix hergestellt werden kann. In die Abbildung fließen die Ergebnisse aus den Tabellen B.1 bis B.4 ein (Seite 149 f. im Anhang).*

können jeweils die in Klammern gesetzten Determinanten der Inversen der Informationsmatrix für den Fall, dass eine Versuchseinheit vorliegt, abgelesen werden. Die Normierung hat den Vorteil, dass mit steigender Anzahl an Versuchseinheiten stets die gleichen Referenzwerte herangezogen werden, so dass eine mögliche Konvergenz der Determinanten erkannt werden kann. Es ist zu erkennen, dass für die gewählten Korrelationsparameter die normierten Determinanten der empirischen Kovarianzmatrizen der Art ähnliche Werte aufweisen wie die Determinanten der Inversen der Informationsmatrizen, dass die Funktionsverläufe im Großen und Ganzen miteinander übereinstimmen, siehe Abbildung 6.2. Dies gilt insbesondere auch für ein kleines n . Weiter ist in den Tabellen 6.2 bis 6.5 zu erkennen, dass mit steigender Anzahl an Versuchseinheiten sich die Differenzen der Determinanten tendenziell immer weiter der Null annähern, siehe Abbildung 6.3 (linke Grafik). Da die Determinanten der Inversen der Informationsmatrix in etwa der Hälfte aller

Fälle größere Werte als die normierten Determinanten der empirischen Kovarianzmatrizen aufweisen, ist zu vermuten, dass die Abweichungen auch teilweise auf Grund von zufälligen Fehlern zustandekommen, siehe Abbildung 6.3 (rechte Grafik). Dies hätte die Konsequenz, dass eine Erhöhung der Anzahl an Datensätzen zu geringeren Abweichungen führen könnte. Insgesamt kann nun bezüglich der Wahl des Versuchsdesigns der Schluss gezogen werden, dass wie in Abschnitt 5.1 nur lokal optimale Versuchspläne existieren. Insbesondere lassen die Ergebnisse für den Parameter $\lambda = 0.0001$ die Vermutung zu, dass für Korrelationen, die nahe bei Null sind, der optimale Versuchspunkt umso kleiner wird, je kleiner die Korrelation ist. Außerdem ist hier ebenso für hinreichend große Korrelationen der äquidistante Versuchsplan D-optimal. Schließlich kann festgehalten werden, dass sich auf Grund der guten Übereinstimmung der Determinanten die Überlegungen bezüglich der Effizienzen aus Abschnitt 5.2 auf die hier vorliegende Situation übertragen lassen.

Da in dem zweistufigen Schätzverfahren aus Abschnitt 6.1 auch der Korrelationsparameter λ geschätzt wird, wird nun abschließend beschrieben, wie sich die Wahl des Versuchsplans auf die Schätzung dieses Parameters auswirkt. Zudem wird untersucht, ob die Einträge der Inversen der Informationsmatrix bereits für eine kleine Anzahl an Versuchseinheiten als Näherung der Varianzen in Frage kommt. Die Tabellen 6.6 und 6.7 (Seite 93 f.) enthalten die empirischen Varianzen für den Maximum-Likelihood-Schätzer, der als Schätzer für den Parameter λ Bestandteil des zweistufigen Verfahrens ist. Jeweils direkt darunter sind eingeklammert die Einträge der Inversen der Informationsmatrix für den Fall, dass eine Versuchseinheit vorliegt, aufgeführt. Die empirischen Varianzen sind hier ebenfalls mit einem Faktor, diesmal mit dem Faktor n , multipliziert worden, so dass für die verschiedenen Anzahlen an Versuchseinheiten stets der Bezug zum gleichen Referenzwert gegeben ist. Es ist zu erkennen, dass für den Parameterwert $\lambda = 0$ keine Referenzwerte aufgeführt sind. Dies liegt daran, dass die Inverse der Informationsmatrix für den Vektor $[\beta_0, \beta_1, \sigma^2, \lambda]^T$ für $\lambda = 0$ nicht existiert, da drei der vier Einträge der Matrix (3.6) für $\lambda = 0$ Null sind. Die Werte in den Tabellen 6.6 und 6.7 zeigen nun für den Korrelationsparameter, dass die Einträge der Inversen der Informationsmatrix erst mit wachsendem n eine gute Näherung für die Varianzen des Schätzers darstellen. Zudem ist zu erkennen, dass die Höhe der Differenzen aus den empirischen Varianzen und den Referenzwerten auch vom wahren Wert des Korrelationsparameters abhängt. Außerdem weichen für die Fälle $n = 2$ und $n = 5$ die Quotienten aus den empirischen Varianzen und den Referenzwerten insbesondere für sehr schwache und sehr starke Korrelationen deutlich vom Sollwert 1 ab. Das Problem, dass die Näherung erst für großes n gut ist, wird in Abbildung 6.4 (linke

Grafik) veranschaulicht. Weiterhin fällt in den Tabellen 6.6 und 6.7 auf, dass die Referenzwerte von den empirischen Varianzen bei mittelstarken Korrelationen systematisch unterschritten werden. Ein positiver Aspekt ist jedoch der Folgende. In den Tabellen 6.6 und 6.7 ist zum einen zu erkennen, dass der Plan mit dem Versuchspunkt $d = 0.1$ die Einträge der Inversen der Informationsmatrix minimiert. Zum anderen ist zu vermuten, dass selbiges auch für die wahren Varianzen gilt, da die empirischen Varianzen tendenziell ebenfalls für $d = 0.1$ minimal sind. Ein genaues Urteil ist schwierig, da die empirischen Varianzen Zufallsschwankungen unterliegen. Numerische Untersuchungen für den Grenzfall $d \rightarrow 0$ zeigen, dass die Einträge der Inversen der Informationsmatrix umso kleiner sind, je näher der Punkt d an der Null platziert wird. An dieser Stelle sei daran erinnert, dass der D-optimale Versuchsplan bezüglich der Regressionsparameter β_0 und β_1 in den Fällen $\lambda \geq 0.0007798$ durch den äquidistanten Versuchsplan gegeben ist, siehe Abschnitt 5.1. Die Simulationsstudie zeigt selbige Tendenz, siehe Tabelle 6.2 bis 6.5. Somit kann festgehalten werden, dass bezüglich des Parametervektors $[\beta_0, \beta_1, \lambda]^T$ die Versuchspläne $d = 1/2$ und $d \rightarrow 0$ in vielen Fällen miteinander konkurrieren. Weitere Untersuchungen hierzu erfolgen in Kapitel 7.

Bisher wurden in diesem Abschnitt nur Varianzen betrachtet. Die Varianz eines Schätzers ist jedoch nur ein mögliches Gütekriterium. Weitere Gütekriterien sind die Erwartungstreue und die Mediantreue. Die Einträge in den Tabellen 6.8 bis 6.11 (Seite 95 f.) zeigen, dass der Maximum-Likelihood-Schätzer für den Korrelationsparameter diesen für kleine Korrelationen im Mittel eher überschätzt und für hohe Korrelationen systematisch unterschätzt. Die Verzerrungen sind umso kleiner, je größer die Anzahl der Versuchseinheiten ist, siehe Abbildung 6.4 (rechte Grafik). Bezüglich der Mediantreue ist anhand der Tabellen zu erkennen, dass der Korrelationsparameter überwiegend unterschätzt wird. Mit wachsender Anzahl an Versuchseinheiten verkleinern sich jedoch auch hier die Abweichungen zwischen den Medianen der Schätzwerte und den wahren Werten des Korrelationsparameters.

Tabelle 6.2: *Determinante der empirischen Kovarianzmatrix des Schätzers für β aus dem zweistufigen Schätzverfahren in Abhängigkeit vom Korrelationsparameter λ und vom Versuchspunkt d . In Klammern stehen die Determinanten der Inversen der Informationsmatrix für den Fall $n = 1$. Die Determinanten der empirischen Kovarianzmatrizen sind mit dem Faktor n^2 multipliziert worden, damit in den Tabellen 6.2 bis 6.5 ein Bezug zur selben Informationsmatrix hergestellt werden kann. Es liegt der Spezialfall $n = 2$ mit 10000 Wiederholungen vor.*

λ	d								
	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1
0		0.554 (0.550)	0.577 (0.595)	0.633 (0.633)	0.650 (0.658)	0.658 (0.667)	0.668 (0.658)	...	
0.0001	0.972 (1.000)	0.744 (0.763)	0.692 (0.684)	0.689 (0.668)	0.680 (0.672)	0.663 (0.676)	0.690 (0.672)	...	0.976 (1.000)
0.001	1.000 (1.000)	0.835 (0.816)	0.736 (0.734)	0.709 (0.703)	0.709 (0.695)	0.709 (0.694)	0.697 (0.695)	...	0.999 (1.000)
0.01	0.969 (1.000)	0.884 (0.881)	0.838 (0.812)	0.775 (0.774)	0.773 (0.756)	0.754 (0.751)	0.734 (0.756)	...	1.015 (1.000)
0.1	0.974 (0.990)	0.959 (0.948)	0.960 (0.918)	0.909 (0.898)	0.911 (0.887)	0.884 (0.883)	0.890 (0.887)	...	1.005 (0.990)
0.25	0.935 (0.938)	0.950 (0.924)	0.900 (0.913)	0.899 (0.906)	0.889 (0.901)	0.910 (0.900)	0.917 (0.901)	...	0.930 (0.938)
0.5	0.743 (0.750)	0.759 (0.748)	0.780 (0.747)	0.751 (0.745)	0.751 (0.745)	0.710 (0.745)	0.744 (0.745)	...	0.763 (0.750)
0.9	0.187 (0.190)	0.193 (0.190)	0.193 (0.190)	0.193 (0.190)	0.187 (0.190)	0.193 (0.190)	0.188 (0.190)	...	0.196 (0.190)

Tabelle 6.3: *Wie Tabelle 6.2. Hier liegt jedoch der Spezialfall $n = 5$ vor.*

λ	d								
	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1
0		0.534 (0.550)	0.604 (0.595)	0.643 (0.633)	0.648 (0.658)	0.677 (0.667)	0.658 (0.658)	...	
0.0001	0.980 (1.000)	0.776 (0.763)	0.697 (0.684)	0.686 (0.668)	0.668 (0.672)	0.661 (0.676)	0.683 (0.672)	...	0.983 (1.000)
0.001	1.008 (1.000)	0.835 (0.816)	0.718 (0.734)	0.718 (0.703)	0.696 (0.695)	0.706 (0.694)	0.706 (0.695)	...	0.977 (1.000)
0.01	0.978 (1.000)	0.877 (0.881)	0.818 (0.812)	0.780 (0.774)	0.757 (0.756)	0.772 (0.751)	0.778 (0.756)	...	1.003 (1.000)
0.1	0.977 (0.990)	0.955 (0.948)	0.927 (0.918)	0.954 (0.898)	0.933 (0.887)	0.903 (0.883)	0.911 (0.887)	...	0.996 (0.990)
0.25	0.917 (0.938)	0.896 (0.924)	0.922 (0.913)	0.901 (0.906)	0.906 (0.901)	0.932 (0.900)	0.924 (0.901)	...	0.937 (0.938)
0.5	0.747 (0.750)	0.742 (0.748)	0.770 (0.747)	0.753 (0.745)	0.730 (0.745)	0.736 (0.745)	0.728 (0.745)	...	0.768 (0.750)
0.9	0.190 (0.190)	0.186 (0.190)	0.193 (0.190)	0.192 (0.190)	0.194 (0.190)	0.189 (0.190)	0.190 (0.190)	...	0.192 (0.190)

Tabelle 6.4: Wie Tabelle 6.2. Hier liegt jedoch der Spezialfall $n = 10$ vor.

λ	d								
	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1
0		0.552 (0.550)	0.577 (0.595)	0.652 (0.633)	0.669 (0.658)	0.683 (0.667)	0.657 (0.658)
0.0001	0.989 (1.000)	0.769 (0.763)	0.682 (0.684)	0.641 (0.668)	0.679 (0.672)	0.682 (0.676)	0.673 (0.672)	...	1.004 (1.000)
0.001	1.010 (1.000)	0.823 (0.816)	0.757 (0.734)	0.694 (0.703)	0.685 (0.695)	0.697 (0.694)	0.685 (0.695)	...	0.997 (1.000)
0.01	0.994 (1.000)	0.871 (0.881)	0.806 (0.812)	0.760 (0.774)	0.779 (0.756)	0.763 (0.751)	0.753 (0.756)	...	1.025 (1.000)
0.1	1.024 (0.990)	0.954 (0.948)	0.893 (0.918)	0.930 (0.898)	0.903 (0.887)	0.908 (0.883)	0.888 (0.887)	...	0.966 (0.990)
0.25	0.933 (0.938)	0.931 (0.924)	0.913 (0.913)	0.868 (0.906)	0.913 (0.901)	0.911 (0.900)	0.903 (0.901)	...	0.950 (0.938)
0.5	0.748 (0.750)	0.754 (0.748)	0.768 (0.747)	0.732 (0.745)	0.733 (0.745)	0.747 (0.745)	0.731 (0.745)	...	0.772 (0.750)
0.9	0.193 (0.190)	0.185 (0.190)	0.190 (0.190)	0.191 (0.190)	0.190 (0.190)	0.187 (0.190)	0.187 (0.190)	...	0.188 (0.190)

Tabelle 6.5: Wie Tabelle 6.2. Hier liegt jedoch der Spezialfall $n = 50$ vor.

λ	d								
	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1
0		0.540 (0.550)	0.592 (0.595)	0.644 (0.633)	0.638 (0.658)	0.677 (0.667)	0.644 (0.658)
0.0001	1.006 (1.000)	0.785 (0.763)	0.668 (0.684)	0.668 (0.668)	0.685 (0.672)	0.674 (0.676)	0.682 (0.672)	...	0.993 (1.000)
0.001	1.002 (1.000)	0.802 (0.816)	0.723 (0.734)	0.708 (0.703)	0.707 (0.695)	0.690 (0.694)	0.708 (0.695)	...	0.992 (1.000)
0.01	1.014 (1.000)	0.866 (0.881)	0.815 (0.812)	0.789 (0.774)	0.756 (0.756)	0.773 (0.751)	0.740 (0.756)	...	1.006 (1.000)
0.1	0.996 (0.990)	0.931 (0.948)	0.921 (0.918)	0.883 (0.898)	0.889 (0.887)	0.899 (0.883)	0.880 (0.887)	...	0.983 (0.990)
0.25	0.943 (0.938)	0.911 (0.924)	0.944 (0.913)	0.889 (0.906)	0.908 (0.901)	0.902 (0.900)	0.890 (0.901)	...	0.930 (0.938)
0.5	0.752 (0.750)	0.735 (0.748)	0.749 (0.747)	0.754 (0.745)	0.747 (0.745)	0.756 (0.745)	0.743 (0.745)	...	0.744 (0.750)
0.9	0.193 (0.190)	0.188 (0.190)	0.192 (0.190)	0.189 (0.190)	0.193 (0.190)	0.188 (0.190)	0.189 (0.190)	...	0.191 (0.190)

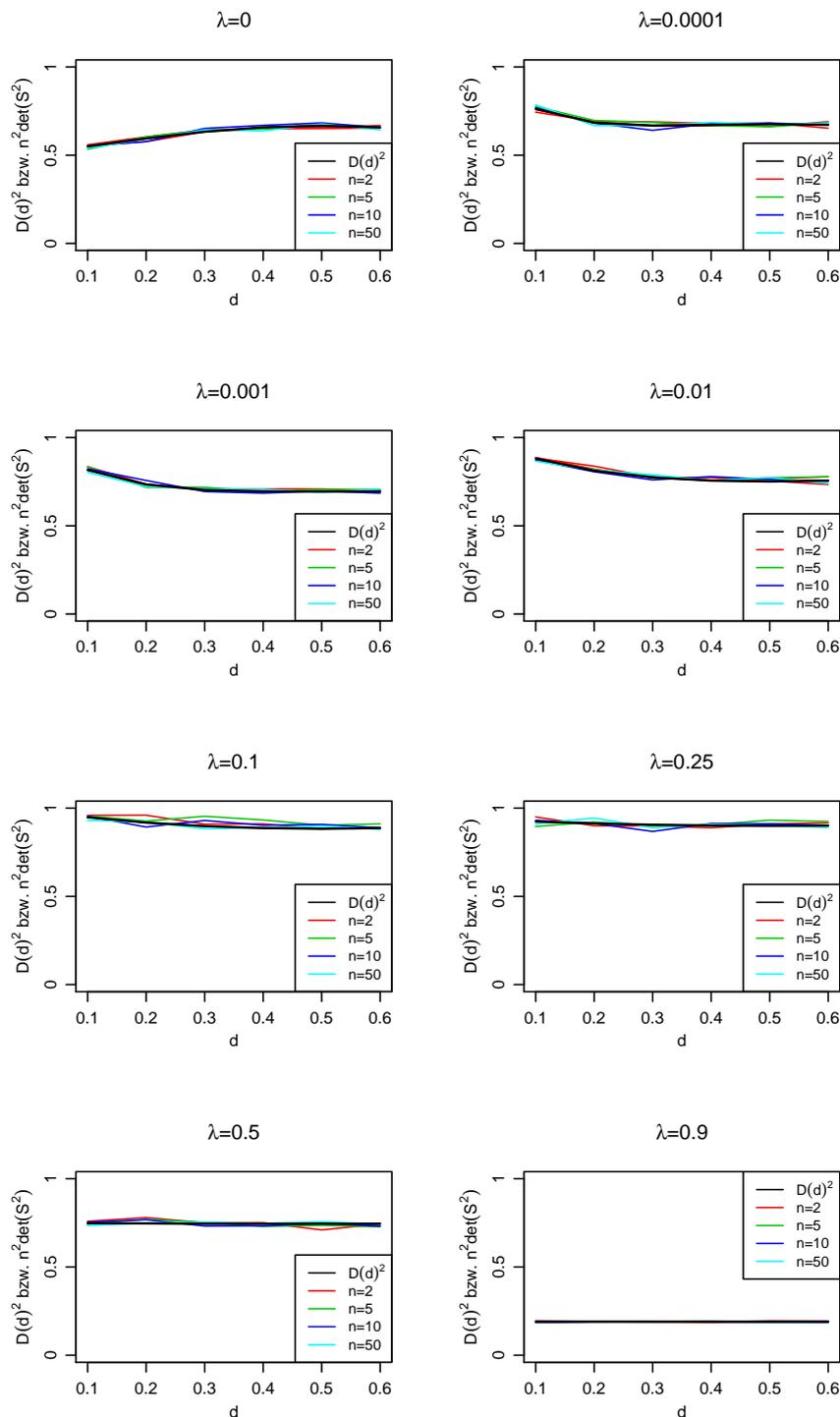


Abbildung 6.2: Determinanten der empirischen Kovarianzmatrizen und Determinanten der Inversen der Informationsmatrix für $d \in \{0.1, 0.2, \dots, 0.6\}$. Jeder Polygonzug resultiert aus den Werten einer Zeile der Tabellen 6.2 bis 6.5. Es ist zu erkennen, dass die Polygonzüge innerhalb eines Korrelationsparameterwerts ähnlich verlaufen.

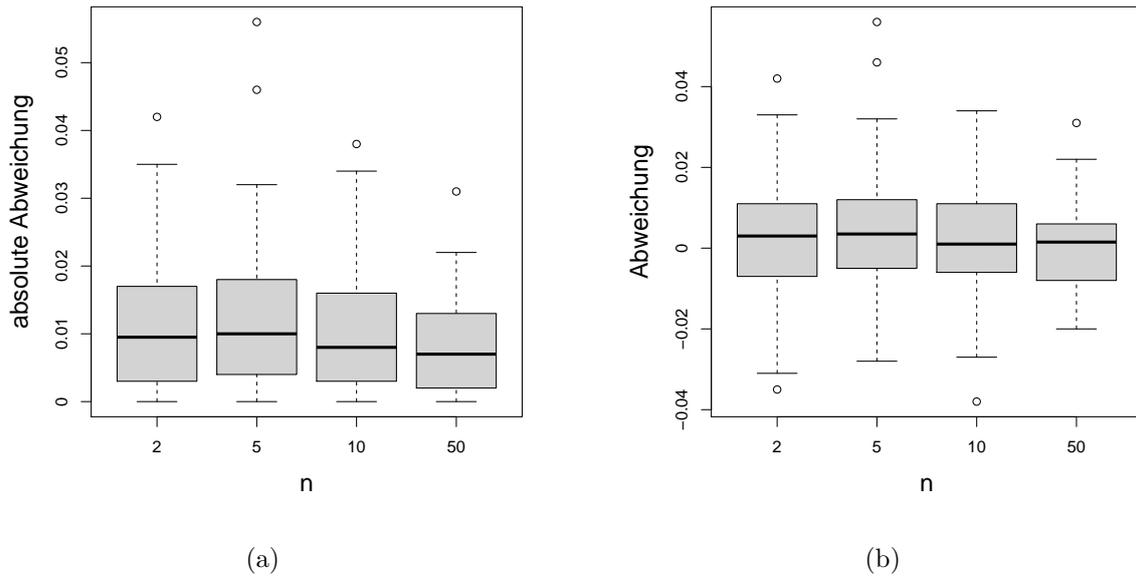


Abbildung 6.3: Boxplots zur Illustration der Konsistenz des Schätzers für β aus dem zwei-stufigen Verfahren. Die Datensätze resultieren aus den Tabellen 6.2 bis 6.5. (a): Ein Boxplot veranschaulicht die 62 absoluten Abweichungen zwischen der normierten Determinanten der empirischen Kovarianzmatrix und der Determinanten der Inversen der Informationsmatrix. Es ist zu erkennen, dass die Quartile der absoluten Abweichungen für $n = 2$ und $n = 5$ recht ähnlich sind. Die Quartile werden aber für $n = 10$ und insbesondere für $n = 50$ erkennbar kleiner. (b): Ein Boxplot veranschaulicht die 62 Abweichungen zwischen den Determinanten. Es ist für alle vier Fälle zu erkennen, dass nahezu die Hälfte der Abweichungen negativ ist. Somit wird die Determinante der Inversen der Informationsmatrix recht oft unterschritten.

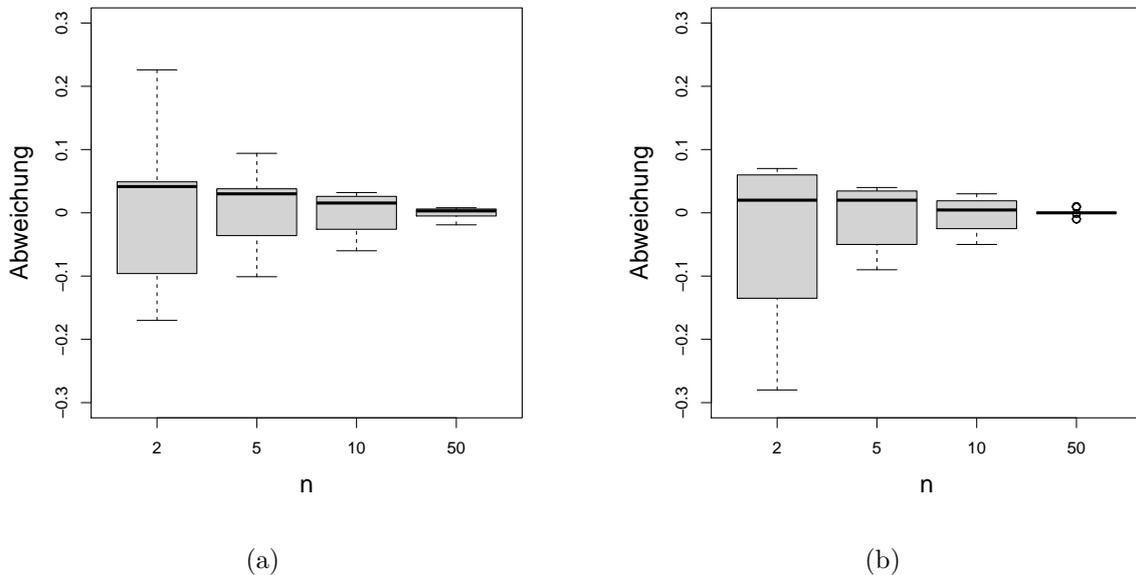


Abbildung 6.4: Boxplots zur Illustration der Konsistenz des Schätzers für λ aus dem zwei-stufigen Verfahren. (a): Die Datensätze resultieren aus den Tabellen 6.6 und 6.7. Ein Boxplot veranschaulicht 42 Abweichungen zwischen den empirischen Varianzen der Maximum-Likelihood-Schätzer für λ und den dazu passenden Einträgen aus den Inversen der Informationsmatrizen. Es ist zu erkennen, dass die Abweichungen mit wachsender Anzahl an Versuchseinheiten deutlich kleiner werden. (b): Die Datensätze resultieren aus den Tabellen 6.8 bis 6.11. Ein Boxplot veranschaulicht 42 Abweichungen zwischen den Mittelwerten der Maximum-Likelihood-Schätzer für λ und den wahren zugrunde liegenden Parameterwerten. Hier ist unter anderem für $n = 2$ und $n = 5$ zu sehen, dass die Parameterwerte mehrheitlich überschätzt werden. Die größten Abweichungen resultieren jedoch durch Unterschätzungen. In beiden Grafiken werden nur die Ergebnisse für die Versuchspunkte $d \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6\}$ herangezogen.

Tabelle 6.6: *Empirische Varianzen des Schätzers für λ aus dem zweistufigen Schätzverfahren. In Klammern stehen die Einträge der Inversen der Informationsmatrix für den Fall $n = 1$. Die empirischen Varianzen sind mit dem Faktor n multipliziert worden, damit in den Tabellen 6.6 und 6.7 ein Bezug zum gleichen Referenzwert hergestellt werden kann. Oben: Es liegt der Spezialfall $n = 2$ mit 10000 Wiederholungen vor. Unten: Es liegt der Spezialfall $n = 5$ mit 10000 Wiederholungen vor.*

λ	d								
	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1
0		0.019	0.030	0.042	0.047	0.048	0.046	...	
0.0001	0.296 (1.000)	0.032 (0.00000)	0.042 (0.00001)	0.043 (0.00003)	0.048 (0.00010)	0.049 (0.00020)	0.045 (0.00010)	...	0.301 (1.000)
0.001	0.300 (1.000)	0.041 (0.000)	0.042 (0.000)	0.050 (0.001)	0.053 (0.001)	0.053 (0.002)	0.054 (0.001)	...	0.303 (1.000)
0.01	0.296 (1.000)	0.052 (0.008)	0.059 (0.010)	0.056 (0.014)	0.060 (0.018)	0.061 (0.020)	0.060 (0.018)	...	0.298 (1.000)
0.1	0.312 (0.980)	0.091 (0.151)	0.094 (0.158)	0.092 (0.163)	0.098 (0.166)	0.094 (0.168)	0.092 (0.166)	...	0.314 (0.980)
0.25	0.336 (0.879)	0.139 (0.298)	0.138 (0.302)	0.135 (0.305)	0.141 (0.306)	0.138 (0.307)	0.143 (0.306)	...	0.331 (0.879)
0.5	0.354 (0.563)	0.197 (0.295)	0.198 (0.297)	0.203 (0.299)	0.204 (0.300)	0.199 (0.300)	0.204 (0.300)	...	0.355 (0.563)
0.9	0.364 (0.036)	0.243 (0.026)	0.247 (0.026)	0.249 (0.026)	0.251 (0.026)	0.250 (0.026)	0.252 (0.026)	...	0.358 (0.036)

λ	d								
	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1
0		0.005	0.016	0.025	0.033	0.037	0.033	...	
0.0001	0.327 (1.000)	0.019 (0.00000)	0.025 (0.00001)	0.032 (0.00003)	0.035 (0.00010)	0.037 (0.00020)	0.034 (0.00010)	...	0.324 (1.000)
0.001	0.324 (1.000)	0.029 (0.000)	0.031 (0.000)	0.037 (0.001)	0.039 (0.001)	0.043 (0.002)	0.038 (0.001)	...	0.335 (1.000)
0.01	0.328 (1.000)	0.047 (0.008)	0.052 (0.010)	0.054 (0.014)	0.055 (0.018)	0.056 (0.020)	0.053 (0.018)	...	0.341 (1.000)
0.1	0.365 (0.980)	0.122 (0.151)	0.121 (0.158)	0.123 (0.163)	0.130 (0.166)	0.122 (0.168)	0.124 (0.166)	...	0.374 (0.980)
0.25	0.439 (0.879)	0.200 (0.298)	0.205 (0.302)	0.204 (0.305)	0.205 (0.306)	0.208 (0.307)	0.210 (0.306)	...	0.430 (0.879)
0.5	0.473 (0.563)	0.275 (0.295)	0.283 (0.297)	0.284 (0.299)	0.292 (0.300)	0.294 (0.300)	0.288 (0.300)	...	0.474 (0.563)
0.9	0.204 (0.036)	0.108 (0.026)	0.110 (0.026)	0.112 (0.026)	0.115 (0.026)	0.120 (0.026)	0.109 (0.026)	...	0.192 (0.036)

Tabelle 6.7: Wie Tabelle 6.6. Oben: Hier liegt der Spezialfall $n = 10$ vor. Unten: Hier liegt der Spezialfall $n = 50$ vor.

λ	d								
	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1
0		0.000	0.006	0.012	0.019	0.022	0.020	...	
0.0001	0.334 (1.000)	0.007 (0.00000)	0.013 (0.00001)	0.016 (0.00003)	0.022 (0.00010)	0.023 (0.00020)	0.022 (0.00010)	...	0.329 (1.000)
0.001	0.333 (1.000)	0.014 (0.000)	0.018 (0.000)	0.022 (0.001)	0.027 (0.001)	0.028 (0.002)	0.026 (0.001)	...	0.332 (1.000)
0.01	0.340 (1.000)	0.036 (0.008)	0.041 (0.010)	0.041 (0.014)	0.043 (0.018)	0.044 (0.020)	0.044 (0.018)	...	0.351 (1.000)
0.1	0.435 (0.980)	0.128 (0.151)	0.132 (0.158)	0.134 (0.163)	0.139 (0.166)	0.137 (0.168)	0.133 (0.166)	...	0.432 (0.980)
0.25	0.537 (0.879)	0.244 (0.298)	0.242 (0.302)	0.249 (0.305)	0.255 (0.306)	0.250 (0.307)	0.251 (0.306)	...	0.539 (0.879)
0.5	0.562 (0.563)	0.297 (0.295)	0.303 (0.297)	0.312 (0.299)	0.315 (0.300)	0.308 (0.300)	0.302 (0.300)	...	0.560 (0.563)
0.9	0.083 (0.036)	0.055 (0.026)	0.056 (0.026)	0.057 (0.026)	0.058 (0.026)	0.057 (0.026)	0.056 (0.026)	...	0.081 (0.036)

λ	d								
	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1
0		0.000	0.000	0.001	0.004	0.006	0.003	...	
0.0001	0.340 (1.000)	0.000 (0.00000)	0.001 (0.00001)	0.002 (0.00003)	0.006 (0.00010)	0.007 (0.00020)	0.005 (0.00010)	...	0.339 (1.000)
0.001	0.343 (1.000)	0.002 (0.000)	0.003 (0.000)	0.006 (0.001)	0.008 (0.001)	0.009 (0.002)	0.009 (0.001)	...	0.343 (1.000)
0.01	0.367 (1.000)	0.015 (0.008)	0.018 (0.010)	0.021 (0.014)	0.026 (0.018)	0.026 (0.020)	0.025 (0.018)	...	0.366 (1.000)
0.1	0.621 (0.980)	0.142 (0.151)	0.150 (0.158)	0.150 (0.163)	0.162 (0.166)	0.159 (0.168)	0.153 (0.166)	...	0.612 (0.980)
0.25	0.813 (0.879)	0.281 (0.298)	0.287 (0.302)	0.286 (0.305)	0.301 (0.306)	0.293 (0.307)	0.287 (0.306)	...	0.833 (0.879)
0.5	0.583 (0.563)	0.303 (0.295)	0.302 (0.297)	0.301 (0.299)	0.302 (0.300)	0.302 (0.300)	0.298 (0.300)	...	0.583 (0.563)
0.9	0.040 (0.036)	0.029 (0.026)	0.029 (0.026)	0.030 (0.026)	0.029 (0.026)	0.030 (0.026)	0.030 (0.026)	...	0.041 (0.036)

Tabelle 6.8: Oben: Empirische Mittelwerte des Schätzers für λ aus dem zweistufigen Schätzverfahren. Unten: Empirische Mediane des Schätzers für λ aus dem zweistufigen Schätzverfahren. Es liegt der Spezialfall $n = 2$ mit 10000 Wiederholungen vor.

	d								
λ	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1
0		0.02	0.04	0.05	0.06	0.06	0.06	...	
0.0001	0.28	0.04	0.05	0.05	0.06	0.06	0.06	...	0.29
0.001	0.28	0.05	0.05	0.06	0.06	0.06	0.07	...	0.28
0.01	0.28	0.07	0.07	0.07	0.07	0.08	0.07	...	0.28
0.1	0.30	0.12	0.12	0.12	0.12	0.12	0.12	...	0.31
0.25	0.35	0.20	0.19	0.18	0.19	0.18	0.19	...	0.34
0.5	0.43	0.30	0.29	0.30	0.30	0.29	0.29	...	0.42
0.9	0.64	0.62	0.62	0.62	0.62	0.62	0.62	...	0.65

0	0	0	0	0	0	0	0	...	
0.0001	0.00	0	0	0	0	0	0	...	0.00
0.001	0.00	0	0	0	0	0	0	...	0.00
0.01	0.00	0.00	0	0	0	0	0	...	0.00
0.1	0.00	0.01	0.00	0.00	0.00	0	0.00	...	0.00
0.25	0.00	0.05	0.04	0.02	0.02	0.02	0.01	...	0.00
0.5	0.36	0.18	0.17	0.17	0.17	0.17	0.17	...	0.34
0.9	0.89	0.76	0.77	0.77	0.76	0.76	0.76	...	0.90

Tabelle 6.9: Wie Tabelle 6.8. Hier liegt jedoch der Spezialfall $n = 5$ vor. Oben: Empirische Mittelwerte. Unten: Empirische Mediane.

	d								
λ	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1
0		0.01	0.02	0.03	0.03	0.04	0.03	...	
0.0001	0.19	0.02	0.03	0.03	0.03	0.04	0.03	...	0.19
0.001	0.19	0.03	0.03	0.04	0.04	0.04	0.04	...	0.19
0.01	0.19	0.05	0.05	0.05	0.05	0.05	0.05	...	0.19
0.1	0.23	0.13	0.12	0.12	0.12	0.12	0.12	...	0.23
0.25	0.30	0.22	0.22	0.22	0.22	0.22	0.22	...	0.31
0.5	0.46	0.41	0.41	0.41	0.41	0.41	0.41	...	0.46
0.9	0.83	0.83	0.83	0.83	0.83	0.83	0.83	...	0.84

0	0	0	0	0	0	0	0	...	
0.0001	0.00	0	0	0	0	0	0	...	0.00
0.001	0.00	0	0	0.00	0	0	0	...	0.00
0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	...	0.00
0.1	0.10	0.06	0.06	0.05	0.05	0.05	0.05	...	0.09
0.25	0.25	0.18	0.17	0.18	0.18	0.18	0.18	...	0.26
0.5	0.50	0.42	0.42	0.42	0.42	0.42	0.43	...	0.50
0.9	0.90	0.88	0.88	0.88	0.88	0.88	0.88	...	0.90

Tabelle 6.10: Wie Tabelle 6.8. Hier liegt jedoch der Spezialfall $n = 10$ vor. Oben: Empirische Mittelwerte. Unten: Empirische Mediane.

	d									
λ	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1	
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	...		
0.0001	0.13	0.01	0.01	0.02	0.02	0.02	0.02	...	0.13	
0.001	0.13	0.01	0.02	0.02	0.02	0.02	0.02	...	0.13	
0.01	0.13	0.03	0.03	0.04	0.04	0.04	0.04	...	0.14	
0.1	0.18	0.11	0.11	0.11	0.11	0.11	0.11	...	0.18	
0.25	0.27	0.24	0.23	0.24	0.24	0.23	0.23	...	0.27	
0.5	0.46	0.46	0.46	0.45	0.45	0.45	0.45	...	0.47	
0.9	0.88	0.87	0.87	0.87	0.87	0.87	0.87	...	0.88	

0	0	0	0	0	0	0	0	...	
0.0001	0.00	0.00	0.00	0.00	0	0	0	...	0.00
0.001	0.00	0.00	0.00	0.00	0.00	0	0.00	...	0.00
0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.00	...	0.01
0.1	0.11	0.08	0.08	0.08	0.08	0.07	0.08	...	0.11
0.25	0.25	0.22	0.22	0.22	0.22	0.21	0.22	...	0.25
0.5	0.50	0.47	0.47	0.47	0.47	0.46	0.46	...	0.50
0.9	0.90	0.89	0.89	0.89	0.89	0.89	0.89	...	0.90

Tabelle 6.11: Wie Tabelle 6.8. Hier liegt jedoch der Spezialfall $n = 50$ vor. Oben: Empirische Mittelwerte. Unten: Empirische Mediane.

	d									
λ	0	0.1	0.2	0.3	0.4	0.5	0.6	...	1	
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	...		
0.0001	0.06	0.00	0.00	0.00	0.00	0.01	0.00	...	0.06	
0.001	0.06	0.00	0.00	0.00	0.01	0.01	0.01	...	0.06	
0.01	0.06	0.01	0.02	0.02	0.02	0.02	0.02	...	0.06	
0.1	0.12	0.10	0.10	0.10	0.10	0.10	0.10	...	0.12	
0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	...	0.25	
0.5	0.49	0.49	0.49	0.49	0.49	0.49	0.49	...	0.49	
0.9	0.90	0.90	0.90	0.90	0.90	0.90	0.90	...	0.90	

0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	...	
0.0001	0.00	0.00	0.00	0.00	0.00	0.00	0.00	...	0.00
0.001	0.00	0.00	0.00	0.00	0.00	0.00	0.00	...	0.00
0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	...	0.01
0.1	0.10	0.10	0.10	0.10	0.09	0.10	0.10	...	0.10
0.25	0.25	0.24	0.24	0.24	0.24	0.24	0.24	...	0.25
0.5	0.50	0.50	0.49	0.49	0.49	0.49	0.49	...	0.50
0.9	0.90	0.90	0.90	0.90	0.90	0.90	0.90	...	0.90

7. Untersuchungen zum D-Kriterium für den Parametervektor $[\beta_0, \beta_1, \lambda]^T$

In diesem Kapitel wird nach wie vor das Modell aus Kapitel 2 betrachtet. Zudem werden erneut Versuchspläne untersucht, bei denen jede Versuchseinheit drei Versuchspunkte aufweist. Zwei Versuchspunkte jeder Versuchseinheit sind durch die beiden Randpunkte weiterhin vorgegeben. Abweichend zu Kapitel 5 wird hier jedoch die Situation betrachtet, dass neben den beiden Parametern β_0 und β_1 auch der Korrelationsparameter λ möglichst effizient geschätzt werden soll. Eine Möglichkeit, geeignete Versuchspläne zu benennen, ist erneut durch das D-Kriterium gegeben. Zunächst wird in Abschnitt 7.1 das D-Kriterium für zwei verschiedene Fälle konstruiert. Daraufhin erfolgt in Abschnitt 7.2 ein Vergleich zwischen dem Versuchsplan mit Versuchspunkt $d = 1/2$ und dem Versuchsplan, bei dem der dritte Versuchspunkt gegen den Rand des Versuchsraums konvergiert. Das naheliegende Ziel D-optimale Versuchspläne anzugeben ist auf Grund der Komplexität der zu untersuchenden Funktionen nicht Bestandteil dieses Kapitels. Im Weiteren wird wie in Kapitel 5 wieder nur der Fall $n = 1$ untersucht, da die zugehörige Inverse der Informationsmatrix mit der asymptotischen Kovarianzmatrix des mit dem Faktor \sqrt{n} multiplizierten Maximum-Likelihood-Schätzers übereinstimmt.

Bevor in Abschnitt 7.1 das D-Kriterium zum Parametervektor $[\beta_0, \beta_1, \lambda]^T$ konstruiert wird, wird zuvor noch kurz auf eine Arbeit von Kiselák und Stehlík (2008) eingegangen, die eine Problemstellung enthält, welche der Problemstellung dieses Kapitels nah ist. Sie unterscheidet sich von der hier vorliegenden Situation aber wie folgt. In Kiselák und Stehlík (2008) werden lediglich äquidistante Versuchspläne untersucht. Außerdem enthält das betrachtete Modell nicht den Steigungsparameter β_1 . Unter diesen Bedingungen stellen die Autoren fest, dass der D-optimale Versuchsplan für den Fall, dass zwei oder drei Versuchspunkte vorliegen, "kollabiert". Damit ist gemeint, dass die Größe d , welche in ihrer Arbeit den Abstand zwischen den äquidistanten Versuchspunkten bezeichnet, gegen Null geht. Falls nur der Korrelationsparameter von Bedeutung ist, so stellen die Autoren fest, dass für jede beliebige Anzahl an Versuchspunkten der Plan mit $d \rightarrow 0$ den D-optimalen Versuchsplan darstellt.

7.1. Konstruktion des D-Kriteriums

Die Informationsmatrix für den Parametervektor $[\beta_0, \beta_1, \lambda]^T$ lautet

$$\mathcal{I}([\beta_0, \beta_1, \lambda]^T) = \begin{pmatrix} X^T \Sigma^{-1}(\sigma^2, \lambda) X & [0, 0]^T \\ [0, 0] & M(\lambda) \end{pmatrix} \quad (7.1)$$

mit $M(\lambda) = \frac{1}{2} \text{tr}(\Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda) \Sigma^{-1}(\lambda) \dot{\Sigma}(\lambda))$. Dies folgt aus den Einträgen der Matrix aus Behauptung 3.4. Die zugehörige Inverse lautet

$$\mathcal{I}([\beta_0, \beta_1, \lambda]^T)^{-1} = \begin{pmatrix} (X^T \Sigma^{-1}(\sigma^2, \lambda) X)^{-1} & 0 \\ 0 & M(\lambda)^{-1} \end{pmatrix}. \quad (7.2)$$

Auf Grund der Blockstruktur der Matrix (7.2) ist ihre Determinante das Produkt aus dem Term $M(\lambda)^{-1}$ und der Determinanten von $\mathcal{I}([\beta_0, \beta_1]^T)^{-1}$. Für den Fall, dass die Informationsmatrix den Rang 3 aufweist, lautet das D-Kriterium also

$$(\det(\mathcal{I}([\beta_0, \beta_1]^T)^{-1}) M(\lambda)^{-1})^{1/3} =: \Phi(d). \quad (7.3)$$

Statt $\Phi(d)$ wird im Folgenden aus Gründen, die der Übersichtlichkeit dienen, die Funktion

$$\Phi(d)^3 =: \tilde{\Phi}(d) \quad (7.4)$$

untersucht. Dies hat den Vorteil, dass der Exponent $1/3$ nicht stets in jedem Schritt mitgeschleppt werden muss.

Für das vorliegende Modell hängen die Terme $\det(\mathcal{I}([\beta_0, \beta_1]^T)^{-1})$ und $M(\lambda)^{-1}$ vom Korrelationsparameter und vom Versuchsplan ab. Da auch hier für $\lambda \in (0, 1)$ die Versuchspunkte $d = 0$ und $d = 1$ keine zusätzliche Information liefern, vergleiche Abschnitt 5.3.1, erfolgt erneut eine Fallunterscheidung. Wie in Kapitel 5 haben die Anzahl der Versuchseinheiten und der Parameter σ^2 keinen Einfluss auf die Designwahl, da sie innerhalb des D-Kriteriums lediglich als zusätzlicher Multiplikator auftreten. Damit genügt es den Spezialfall $n = 1$ und $\sigma^2 = 1$ zu untersuchen.

Fall 1 ($\lambda \in (0, 1)$, $d \in (0, 1)$):

Für die Funktion $\tilde{\Phi}(d)$ wird die Informationsmatrix benötigt. Diese lautet

$$\mathcal{I}_{(1)}([\beta_0, \beta_1, \lambda]^T) = \begin{pmatrix} \mathcal{I}_{(1)}([\beta_0, \beta_1]^T) & [0, 0]^T \\ [0, 0] & M_{(1)}(\lambda) \end{pmatrix} \quad (7.5)$$

mit $\mathcal{I}_{(1)}([\beta_0, \beta_1]^T)$ aus (5.9) und $M_{(1)}(\lambda) = m_1(\lambda, d)/n_1(\lambda, d)$. Dabei ist

$$\begin{aligned} m_1(\lambda, d) &= (\lambda^{8d-4} - \lambda^{6d-2} + \lambda^{6d-4} - 4\lambda^{4d-2} + \lambda^{2d-2} + 2\lambda^{4d} - \lambda^{2d} + 1)d^2 \\ &\quad + (-2\lambda^{6d-2} + 4\lambda^{4d-2} - 2\lambda^{2d-2} - 2\lambda^{4d} + 4\lambda^{2d} - 2)d \\ &\quad + \lambda^{6d-2} - 2\lambda^{4d-2} + \lambda^{2d-2} + \lambda^{4d} - 2\lambda^{2d} + 1, \\ n_1(\lambda, d) &= \lambda^{8d-2} - 2\lambda^{6d-2} + \lambda^{4d+2} + \lambda^{4d-2} - 2\lambda^{2d+2} - 2\lambda^{6d} + 4\lambda^{4d} - 2\lambda^{2d} + \lambda^2. \end{aligned}$$

Die Informationsmatrix $\mathcal{I}_{(1)}([\beta_0, \beta_1, \lambda]^T)$ weist den Rang 3 auf. Es gilt dann:

$$\tilde{\Phi}(d) = \det(\mathcal{I}_{(1)}([\beta_0, \beta_1]^T)^{-1})M_{(1)}(\lambda)^{-1}. \quad (7.6)$$

Fall 2 ($\lambda \in (0, 1)$, $d \in \{0, 1\}$):

In diesem Fall gehen die Matrix $\mathcal{I}_{(0)}([\beta_0, \beta_1]^T)$ aus (5.14) und der Term $M_{(0)}(\lambda) = m_0(\lambda)/n_0(\lambda)$ mit

$$\begin{aligned} m_0(\lambda) &= \lambda^2 + 1, \\ n_0(\lambda) &= (\lambda - 1)^2(\lambda + 1)^2 \end{aligned}$$

in die Informationsmatrix $\mathcal{I}_{(0)}([\beta_0, \beta_1, \lambda]^T)$ ein. Die Informationsmatrix $\mathcal{I}_{(0)}([\beta_0, \beta_1, \lambda]^T)$ weist den Rang 3 auf. Es gilt:

$$\tilde{\Phi}(d) = \det(\mathcal{I}_{(0)}([\beta_0, \beta_1]^T)^{-1})M_{(0)}(\lambda)^{-1}. \quad (7.7)$$

Fall 3 ($\lambda = 0$):

Es gilt, dass $M_{(1)}(0)^{-1} = 0$ ist. Somit ist der Eintrag in der Inversen der Informationsmatrix für alle Versuchspunkte Null. Damit ist dieser Fall bezüglich des Vergleichs von Versuchspunkten uninteressant.

Fall 4 ($\lambda = 1$):

Die Inverse der Korrelationsmatrix ist die Matrix (5.8). Für $\lambda = 1$ sind mehrere Einträge nicht definiert, so dass die Informationsmatrix nicht berechnet werden kann. Damit kann dieser Fall nicht weiter betrachtet werden.

Im nun folgenden Abschnitt werden für $\lambda \in (0, 1)$ die Versuchspläne mit $d \rightarrow 0$ und $d = 1/2$ bezüglich ihrer relativen D-Effizienz miteinander verglichen. Der Plan mit $d = 0$ bleibt außen vor, da dieser den beiden Konkurrenten $d \rightarrow 0$ und $d = 1/2$ auf Grund der auftretenden perfekten positiven Korrelation wie in Behauptung 5.19 stets unterlegen ist.

7.2. Vergleich der Versuchspläne mit $d \rightarrow 0$ und $d = 1/2$

In diesem Abschnitt sind der äquidistante Versuchsplan und der Versuchsplan, bei dem der dritte Versuchspunkt gegen Null konvergiert, Gegenstand der Untersuchungen. Da das D-Kriterium für $\lambda = 0$ und $\lambda = 1$ nicht betrachtet wird, siehe Abschnitt 7.1, wird im Weiteren nur der Fall $\lambda \in (0, 1)$ analysiert. Auf die Frage, welcher der beiden Versuchspläne zu einem kleineren D-Kriterium führt, gibt die folgende Behauptung eine Antwort.

Behauptung 7.1. *Es gilt für $\lambda \in (0, 1)$, dass*

$$\frac{\lim_{d \rightarrow 0} \tilde{\Phi}(d)}{\tilde{\Phi}(1/2)} \leq 1.$$

Beweis. Es gilt:

$$\tilde{\Phi}(1/2) = \frac{4(\lambda - 1)^4 \lambda}{(\lambda + 1)(\lambda - 4\lambda^{1/2} + 3)}.$$

Weiter gilt, dass

$$\lim_{d \rightarrow 0} \tilde{\Phi}(d) = (1 - \lambda^2) \lim_{d \rightarrow 0} M_{(1)}(\lambda)^{-1}.$$

Dabei wird ausgenutzt, dass $\lim_{d \rightarrow 0} \tilde{D}_\lambda(d) = 1 - \lambda^2$, siehe Behauptung 5.5.

Nun gilt mit der Regel von l'Hospital (*), dass

$$\begin{aligned} \lim_{d \rightarrow 0} M_{(1)}(\lambda)^{-1} &= \lim_{d \rightarrow 0} \frac{n_1(\lambda, d)}{m_1(\lambda, d)} \\ &= \frac{\lim_{d \rightarrow 0} n_1(\lambda, d)}{\lim_{d \rightarrow 0} m_1(\lambda, d)} \\ &\stackrel{*}{=} \frac{\lim_{d \rightarrow 0} n_2(\lambda, d)}{\lim_{d \rightarrow 0} m_2(\lambda, d)} \\ &\stackrel{*}{=} \frac{\lim_{d \rightarrow 0} n_3(\lambda, d)}{\lim_{d \rightarrow 0} m_3(\lambda, d)} \\ &= \frac{8(1 - \lambda)^2 \lambda^2 (\lambda + 1)^2 \ln(\lambda)^2}{8\lambda^4 \ln(\lambda)^2 + 8\lambda^2 \ln(\lambda)^2 + 4\lambda^4 - 8\lambda^2 + 4}, \end{aligned}$$

wobei

$$\begin{aligned} m_2(\lambda, d) &= \frac{\partial}{\partial d} m_1(\lambda, d), & n_2(\lambda, d) &= \frac{\partial}{\partial d} n_1(\lambda, d), \\ m_3(\lambda, d) &= \frac{\partial^2}{\partial d^2} m_1(\lambda, d), & n_3(\lambda, d) &= \frac{\partial^2}{\partial d^2} n_1(\lambda, d). \end{aligned}$$

Auf Grund der bisherigen Ausführungen gilt nun:

$$\frac{\lim_{d \rightarrow 0} \tilde{\Phi}(d)}{\tilde{\Phi}(1/2)} = \frac{f_1(\lambda)}{f_2(\lambda)} =: h_1(\lambda)$$

mit

$$\begin{aligned} f_1(\lambda) &= (1 - \lambda)^3 \lambda^3 (\lambda + 1)^4 (3 - \lambda^{1/2}) (1 - \lambda^{1/2}) \ln(\lambda)^2, \\ f_2(\lambda) &= 2(\lambda^{1/2} - 1)^2 (\lambda^{1/2} + 1)^2 (\lambda - \lambda^{1/2})^2 (\lambda + \lambda^{1/2})^2 (2\lambda^4 \ln(\lambda)^2 \\ &\quad + 2\lambda^2 \ln(\lambda)^2 + \lambda^4 - 2\lambda^2 + 1). \end{aligned}$$

Es ist für $\lambda \in (0, 1)$ zu zeigen, dass

$$h_1(\lambda) \leq 1. \quad (7.8)$$

Die Funktion $h_1(\lambda)$ ist positiv, da die Funktion $\tilde{\Phi}(d)$ positiv ist. Weiter ist leicht zu erkennen, dass die Funktion $f_1(\lambda)$ ebenfalls positiv ist. Dies hat die Konsequenz, dass $f_2(\lambda)$ ebenso positiv sein muss. Daraufhin ist die Ungleichung (7.8) äquivalent zu

$$f_2(\lambda) - f_1(\lambda) \geq 0. \quad (7.9)$$

Definiere $f_2(\lambda) - f_1(\lambda) =: h_2(\lambda)$. Die Substitution $\lambda = x^2$ liefert

$$h_2(x^2) = 2(x - 1)^4 x^4 (x + 1)^3 (\ln(x)^2 f_3(x) + f_4(x)) =: h_3(x) \quad (7.10)$$

mit

$$\begin{aligned} f_3(x) &= 2x^2(-(-x^9 + 3x^8 - 8x^7 + 8x^6 - 6x^5 + 18x^4 - 8x^3 + 8x^2 - x + 1) - 2), \\ f_4(x) &= x^9 + x^8 - 2x^5 - 2x^4 + x + 1. \end{aligned}$$

Jetzt ist zu zeigen, dass

$$h_3(x) \geq 0.$$

In (7.10) ist zu erkennen, dass $2(x - 1)^4 x^4 (x + 1)^3$ positiv ist. Es bleibt damit zu zeigen, dass

$$h_4(x) \geq 0 \quad (7.11)$$

mit $h_4(x) = \ln(x)^2 f_3(x) + f_4(x)$. Mit Behauptung 5.13 gilt, dass

$$-x^9 + 3x^8 - 8x^7 + 8x^6 - 6x^5 + 18x^4 - 8x^3 + 8x^2 - x + 1 =: g_1(x)$$

positiv ist. Wegen $f_3(x) = 2x^2(-g_1(x) - 2)$ ist $f_3(x)$ daraufhin negativ. Damit kann anstelle von (7.11) gezeigt werden, dass

$$\ln(x)^2 + \frac{f_4(x)}{f_3(x)} =: h_5(x)$$

negativ ist. Es gilt $h_5(1) = 0$. Um zu zeigen, dass $h_5(x)$ negativ ist, genügt es nun zu zeigen, dass $h_5(x)$ monoton wachsend ist. Dies gilt, da die Ableitung von $h_5(x)$ positiv ist, siehe Lemma A.35. \square

Die Behauptung 7.1 hat nun die Konsequenz, dass der Versuchsplan, bei dem der dritte Versuchspunkt gegen Null konvergiert, dem äquidistanten Versuchsplan bezüglich des D-Kriteriums in den Fällen $\lambda \in (0, 1)$ überlegen ist. Dies ist insofern interessant, da gerade dieser Versuchsplan für den Parametervektor $[\beta_0, \beta_1]^T$ der Versuchsplan ist, der innerhalb der Versuchspläne mit $d \in [0, 1]$ stets das größte D-Kriterium aufweist, siehe Behauptung 5.19.

8. Relative D-Effizienzen bei verallgemeinerten Abhängigkeitsstrukturen

Wie in Kapitel 5 zu sehen ist, liegt für das Modell aus Kapitel 2 die Besonderheit vor, dass im Grenzfall $d \rightarrow 0$ perfekte positive Korrelationen vorliegen, was zu überraschenden Ergebnissen führen kann, siehe Abschnitt 7.2. Unter anderem in Müller und Stehlik (2009) und Stehlik (2004) werden perfekte positive Korrelationen bei Abständen $d \rightarrow 0$ durch einen sogenannten Nugget-Effekt umgangen. Im Folgenden wird die Bedeutung des Nugget-Effekts sowie einer weiteren Modellverallgemeinerung hinsichtlich der Designwahl in der Situation, dass die Parameter β_0 und β_1 zu schätzen sind, untersucht.

Das Kapitel gliedert sich wie folgt. In Abschnitt 8.1 wird die bisherige Modellierung der Abhängigkeitsstruktur des Fehlervektors um den sogenannten Nugget-Effekt erweitert. Dieser Abschnitt stellt eine Motivation für eine Modellverallgemeinerung dar, die in Abschnitt 8.2 vorgestellt wird. Der Abschnitt 8.2 enthält zudem eine zweite Modellverallgemeinerung. Für beide Modellverallgemeinerungen wird daraufhin in Abschnitt 8.3 der Fall betrachtet, dass der Parametervektor $[\beta_0, \beta_1]^T$ zu schätzen ist. Dabei werden erneut Versuchspläne betrachtet, bei denen die Versuchsreihen die gleichen drei Versuchspunkte besitzen. Die beiden Randpunkte sind weiterhin fester Bestandteil des Versuchsdesigns, so dass nur der dritte Versuchspunkt festgelegt werden muss. Im Speziellen werden Versuchspläne mit den Versuchspunkten $d = 0$, $d \rightarrow 0$ und $d = 1/2$ anhand der relativen D-Effizienz miteinander verglichen. Dabei genügt es auch hier, den Fall $n = 1$ zu untersuchen, da die zugehörigen Ergebnisse auf die Fälle $n > 1$ übertragen werden können.

8.1. Ein Modell mit Nugget-Effekt

Bevor zwei Modellverallgemeinerungen in Abschnitt 8.2 vorgestellt werden, wird der sogenannte Nugget-Effekt beschrieben. Dieser ist mit einer der beiden Modellverallgemeinerungen eng verknüpft.

Der Nugget-Effekt kann im sogenannten Semi-Variogramm veranschaulicht werden. Es ist an das Variogramm angelehnt, welches die Beschreibung der Abhängigkeitsstrukturen eines stochastischen Prozesses $Z(x)$ ermöglicht, siehe Cressie (1993), S. 58 ff. Es ist wie folgt definiert

$$2\gamma(x, y) = \text{var}(Z(x) - Z(y)).$$

Dabei wird die Größe $\gamma(x, y)$ als Semi-Variogramm bezeichnet. Ist ein stochastischer Prozess stationär, so kann das Semi-Variogramm auch durch die Funktion

$$\gamma(h) = \gamma(0, 0 + h)$$

beschrieben werden, wobei $h = y - x$. Das Semi-Variogramm hat unter anderem die Eigenschaft, dass

$$\gamma(0) = 0.$$

Zwischen der Kovarianzmatrix eines stationären stochastischen Prozesses und dem Semi-Variogramm besteht der Zusammenhang

$$\gamma(h) = \frac{\text{Var}(x) + \text{Var}(x + h) - 2\text{Kov}(x, x + h)}{2}.$$

Falls das Semi-Variogramm an der Stelle $h = 0$ eine Sprungstelle aufweist, so wird die Höhe der Sprungstelle als Nugget-Effekt bezeichnet. Der Nugget-Effekt wird im Folgenden durch die Größe τ^2 beschrieben.

Für das Modell aus Kapitel 2 sieht das Semi-Variogramm folgendermaßen aus:

$$\gamma(h) = \sigma^2(1 - \lambda^h). \quad (8.1)$$

In Abbildung 8.1 wird das Semi-Variogramm aus (8.1) anhand des konkreten Beispiels $\sigma^2 = 1$ und $\lambda = 1/2$ veranschaulicht. Es ist zu erkennen, dass dieses Modell keinen Nugget-Effekt aufweist. Wird das Modell aus Kapitel 2 dagegen durch einen Nugget-Effekt erweitert, so muss das Semi-Variogramm (8.1) lediglich um eine Konstante additiv ergänzt werden. Das Semi-Variogramm dieses Prozesses lautet dann

$$\gamma(h) = \sigma^2(1 - \lambda^h) + \tau^2. \quad (8.2)$$

Das bisherige Beispiel wird in Abbildung 8.1 um die Größe $\tau^2 = 1/5$ erweitert.

8.2. Verallgemeinerte Abhängigkeitsstrukturen

Nun wird die erste Modellverallgemeinerung, die in Abschnitt 8.3 untersucht wird, im Detail vorgestellt. Sie ist unter anderem in Stehlík (2004) zu finden. Da die Fehlervarianz lediglich als multiplikativer Faktor Bestandteil des D-Kriteriums ist, wird im Folgenden die Einschränkung vorgenommen, dass die Fehlervarianz gleich 1 ist. Diese Annahme vereinfacht das weitere Vorgehen.

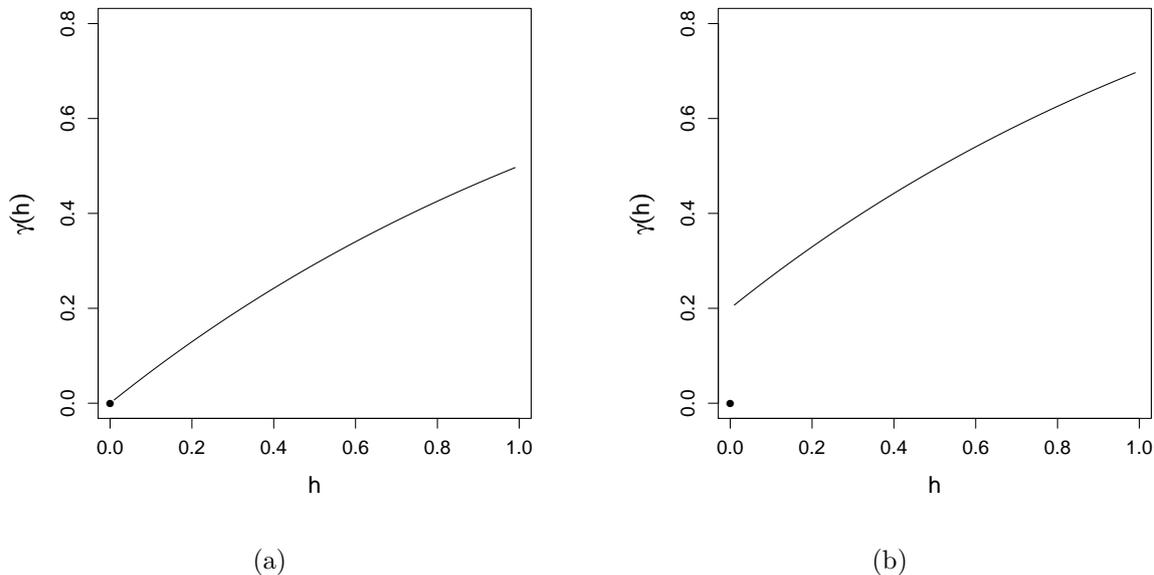


Abbildung 8.1: (a): Semi-Variogramm aus (8.1) mit $\sigma^2 = 1$ und $\lambda = 1/2$. Dies ist ein Modell ohne Nugget-Effekt. (b): Semi-Variogramm aus (8.2). Im Vergleich zum Semi-Variogramm aus (8.1) existiert hier an der Stelle $h = 0$ auf Grund des Nugget-Effekts $\tau^2 = 1/5$ eine Sprungstelle.

In Stehlík (2004) wird eine Modellannahme bezüglich der Korrelationsmatrix des Fehlervektors betrachtet, die der dieser Arbeit sehr nah ist. Und zwar besteht dort zwischen zwei Komponenten des Fehlervektors die Korrelation $\exp(-rd)$ mit Parameter $r \in [0, \infty)$, wenn der Abstand der zugehörigen Versuchspunkte d beträgt. Dieses Modell wird als exponentielles Modell bezeichnet. Erfolgt eine Substitution des Parameters r durch die Funktion $-\ln(\lambda)$ so resultiert wegen $\exp(-rd) = \lambda^d$ das Modell, das in dieser Arbeit Gegenstand der Untersuchungen ist. Da die Transformation bijektiv ist, bedeutet dies unter anderem, dass auch bei einem exponentiellen Modell bezüglich des Parametervektors $[\beta_0, \beta_1]^T$ lediglich lokal D-optimale Versuchspläne existieren. Die weiteren Erläuterungen zu den Modellannahmen erfolgen nun in der Parametrisierung, wie sie in Kapitel 2 gewählt ist. In Stehlík (2004) wird das Modell mit einem Nugget-Effekt $1 - \rho$ erweitert. Zudem wird für zwei Versuchspunkte, die identisch sind, angenommen, dass die zugehörigen Fehler perfekt positiv korreliert sind. Damit lautet die Korrelation der Fehler zweier Versuchspunkte x und y

$$\text{Korr}(x, y) = \begin{cases} \rho\lambda^{|x-y|} & \text{falls } x - y \neq 0, \\ 1 & \text{falls } x - y = 0. \end{cases} \quad (8.3)$$

Die Abhängigkeitsstruktur aus (8.3) stellt nun die erste Modellverallgemeinerung dar. Dabei gilt, dass $\rho \in [0, 1]$. Das zugehörige Semi-Variogramm lautet

$$\gamma(d) = \rho(1 - \lambda^d) + 1 - \rho,$$

wobei $d = |x - y|$. Es ist zu erkennen, dass für $\rho = 1$ die Abhängigkeitsstruktur, die in Kapitel 2 angenommen wird, vorliegt.

In der zweiten Modellverallgemeinerung werden die Korrelationen

$$Korr(x, y) = \begin{cases} \rho\lambda^{|x-y|} & \text{falls } x - y \neq 0, \\ \rho & \text{falls } x - y = 0, \end{cases} \quad (8.4)$$

angenommen. Es ist erneut zu erkennen, dass für $\rho = 1$ die Abhängigkeitsstruktur, die in Kapitel 2 angenommen wird, vorliegt. Die zweite Modellverallgemeinerung unterscheidet sich von der ersten dadurch, dass für $\rho < 1$ die Fehler zweier gleicher Versuchspunkte nicht perfekt positiv korreliert sind. Für den Spezialfall, dass der Versuchsplan sich wie in Kapitel 5 aus den drei Versuchspunkten $x_1 = 0$, $x_2 = d$ und $x_3 = 1$ zusammensetzt, führt die Abhängigkeitsstruktur (8.4) zu der Kovarianzmatrix

$$\Sigma_{(3)}(\sigma_{(1)}^2, \sigma_{(2)}^2, \lambda) = \sigma_{(1)}^2 \begin{bmatrix} 1 & \lambda^d & \lambda \\ \lambda^d & 1 & \lambda^{1-d} \\ \lambda & \lambda^{1-d} & 1 \end{bmatrix} + \sigma_{(2)}^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (8.5)$$

Dabei beträgt die Fehlervarianz $\sigma_{(1)}^2 + \sigma_{(2)}^2$. Somit setzt sich die Fehlervarianz aus zwei Komponenten zusammen. Der Parameter ρ aus der Korrelationsmatrix ergibt sich dabei durch $\rho = \sigma_{(1)}^2 / (\sigma_{(1)}^2 + \sigma_{(2)}^2)$ mit $\rho \in [0, 1]$. Falls $\rho = 1$ ist, so muss für die Komponente $\sigma_{(2)}^2$ der Fehlervarianz gelten, dass $\sigma_{(2)}^2 = 0$ ist. Falls andererseits $\rho = 0$ ist, so sind die Fehler unkorreliert.

Im nächsten Abschnitt wird statt der Kovarianzmatrix (8.5) die zugehörige Korrelationsmatrix herangezogen. Sie lautet

$$\Sigma_{(3)}(\rho, \lambda) = \begin{bmatrix} 1 & \rho\lambda^d & \rho\lambda \\ \rho\lambda^d & 1 & \rho\lambda^{1-d} \\ \rho\lambda & \rho\lambda^{1-d} & 1 \end{bmatrix}. \quad (8.6)$$

Falls der Versuchspunkt d nicht einem der Randpunkte entspricht, so kann die Korrelationsmatrix (8.6) auch für die erste Modellverallgemeinerung herangezogen werden. In dem

Fall $d = 0$ liegt bei der ersten Modellverallgemeinerung jedoch die Korrelationsmatrix

$$\Sigma_{(4)}(\rho, \lambda) = \begin{bmatrix} 1 & 1 & \rho\lambda \\ 1 & 1 & \rho\lambda \\ \rho\lambda & \rho\lambda & 1 \end{bmatrix} \quad (8.7)$$

vor. Für $d = 1$ weist die Korrelationsmatrix die gleichen Einträge wie die der Matrix (8.7) auf. Allein die Positionen der Einträge unterscheiden sich. Da für $d = 0$ und $d = 1$ der Versuchspunkt d keine zusätzlichen Informationen zur Schätzung der Regressionsparameter beiträgt, kann wie in Abschnitt 5.3.1 der Versuchsplan betrachtet werden, der lediglich die beiden Randpunkte enthält. Die Korrelationsmatrix des Fehlervektors lautet in dieser Situation

$$\Sigma_{(5)}(\rho, \lambda) = \begin{bmatrix} 1 & \rho\lambda \\ \rho\lambda & 1 \end{bmatrix}. \quad (8.8)$$

In dem nun folgenden Abschnitt wird, abweichend zu dem in den bisherigen Kapiteln betrachteten Modell, der Parameter ρ in die Untersuchungen einbezogen.

8.3. Vergleich von Versuchsplänen

In diesem Abschnitt werden die beiden Modellverallgemeinerungen aus Abschnitt 8.2 betrachtet. Im Speziellen werden drei verschiedene Versuchspläne anhand der relativen D-Effizienz miteinander verglichen. Es handelt sich um die Versuchspläne mit $d = 0$, $d \rightarrow 0$ und $d = 1/2$.

Zu Beginn wird das D-Kriterium für den Parametervektor $[\beta_0, \beta_1]^T$ formuliert. Dabei wird der Fall $\lambda = 1$ nicht betrachtet, da für $\rho = 1$ manche Einträge in der invertierten Kovarianzmatrix des Fehlervektors nicht definiert sind. Weil für den Fall $\lambda = 0$ im Allgemeinen unkorrelierte Fehler vorliegen und diese Situation bereits in Kapitel 5 Gegenstand der Untersuchungen ist, wird dieser Fall ebenfalls nicht untersucht. Somit werden im Weiteren nur die Fälle $\lambda \in (0, 1)$ betrachtet.

Erfolgt die Konstruktion des D-Kriteriums analog zum Vorgehen in Abschnitt 5.3.1, so lautet es für die erste Modellverallgemeinerung aus Abschnitt 8.2

$$D_{(1)}(d) = \begin{cases} (1 - \lambda^2 \rho^2)^{1/2} & \text{falls } d \in \{0, 1\}, \\ \left(\frac{2\rho^3 \lambda^2 + \rho^2(-\lambda^{2d} - \lambda^{2-2d} - \lambda^2) + 1}{2(\rho(-\lambda d^2 + d\lambda^d - d\lambda^{1-d} + \lambda d - \lambda^d) + d^2 - d + 1)} \right)^{1/2} & \text{falls } d \in (0, 1). \end{cases} \quad (8.9)$$

Dabei gehen die Matrizen (8.8), (5.12), (8.6) und (5.5) in die Konstruktion ein.

Für die zweite Modellverallgemeinerung aus Abschnitt 8.2 lautet das D-Kriterium

$$D_{(2)}(d) = \begin{cases} \left(\frac{1 + \rho - 2\lambda^2 \rho^2}{2} \right)^{1/2} & \text{falls } d \in \{0, 1\}, \\ \left(\frac{2\rho^3 \lambda^2 + \rho^2(-\lambda^{2d} - \lambda^{2-2d} - \lambda^2) + 1}{2(\rho(-\lambda d^2 + d\lambda^d - d\lambda^{1-d} + \lambda d - \lambda^d) + d^2 - d + 1)} \right)^{1/2} & \text{falls } d \in (0, 1). \end{cases} \quad (8.10)$$

Hier gehen die Matrizen (8.6) und (5.5) in die Konstruktion ein. Es ist zu erkennen, dass bezüglich des D-Kriteriums nur für den Versuchspunkt $d = 0$ ein Unterschied zwischen den beiden Modellverallgemeinerungen besteht. Dabei gilt wegen

$$(D_{(1)}(0))^2 - (D_{(2)}(0))^2 = \frac{1 - \rho}{2}, \quad (8.11)$$

dass die erste Modellverallgemeinerung für $\rho \in [0, 1)$ im Vergleich zur zweiten Verallgemeinerung stets zu einem größeren D-Kriterium führt. Dieses Resultat ist plausibel, da in dem Modell mit Nugget-Effekt eine zusätzliche Beobachtung an gleicher Stelle zu keinem Informationsgewinn führt. Für $\rho = 1$ dagegen stimmen die D-Kriterien überein.

Weiter gilt, dass das D-Kriterium bei der zweiten Verallgemeinerung wegen

$$\lim_{d \rightarrow 0} D_{(2)}(d) = D_{(2)}(0) \quad (8.12)$$

für $\lambda \in (0, 1)$ stetig in $d = 0$ ist. Dies bedeutet, dass die Versuchspläne mit $d = 0$ und $d \rightarrow 0$ das gleiche D-Kriterium aufweisen. Für die erste Verallgemeinerung liegt dagegen die Besonderheit vor, dass das D-Kriterium an der Stelle $d = 0$ größer ist als im Grenzfall $d \rightarrow 0$. Dies folgt aus (8.11) und (8.12).

Da in der zweiten Verallgemeinerung das D-Kriterium des Versuchsplans mit $d = 0$ mit dem D-Kriterium des Versuchsplans mit $d \rightarrow 0$ übereinstimmt, wird im Weiteren nur noch die erste Verallgemeinerung untersucht. Denn Ergebnisse für den Versuchsplan mit $d \rightarrow 0$ können direkt auf den Versuchsplan mit $d = 0$ übertragen werden.

In dem D-Kriterium der ersten Modellverallgemeinerung (8.9) resultieren für $\rho = 1$ die Ausdrücke (5.10) und (5.15). Dagegen resultiert für $\rho = 0$ der Ausdruck (5.11). Dies bedeutet, dass der D-optimale Versuchsplan nicht wie bisher nur vom Korrelationsparameter abhängt, sondern insbesondere auch von der Größe des Parameters ρ . Denn für $\rho = 0$ ist der Versuchsplan mit $d = 0$ D-optimal, siehe Behauptung 5.18. Dies gilt hier unabhängig vom Korrelationsparameter. In dem Fall $\rho = 1$ liegt die Situation vor, die ausführlich in Kapitel 5 analysiert wurde. Es existieren damit lediglich lokal D-optimale Versuchspläne.

Da die Hinzunahme des Parameters ρ die in Kapitel 5 vorliegende Komplexität hinsichtlich des Auffindens D-optimaler Versuchspläne weiter verschärft, werden nun lediglich für einen Spezialfall weitere Untersuchungen vorgenommen. Und zwar wird der Fall $\rho = 1/2$ angeschaut. Dieser hat die Eigenschaft, dass er nicht wie die im vorherigen Absatz angesprochenen Extremfälle ($\rho = 1$ und $\rho = 0$) zu Situationen führt, die in Kapitel 5 bereits untersucht worden sind. Dabei werden im Folgenden die drei Versuchspläne mit $d = 0$, $d \rightarrow 0$ und $d = 1/2$ untersucht. Zunächst werden wie in Abschnitt 7.2 die Versuchspläne mit $d \rightarrow 0$ und $d = 1/2$ miteinander verglichen. Es zeigt sich, dass keiner der beiden Versuchspläne für sämtliche Korrelationen $\lambda \in (0, 1)$ besser als der andere ist. Dies kann anhand der relativen D-Effizienz des Versuchsplans mit $d \rightarrow 0$ bezüglich des Versuchsplans mit $d = 1/2$ gezeigt werden, welche durch

$$q_{d \rightarrow 0, 1/2}(\lambda) = \frac{D_{(1)}(1/2)}{\lim_{d \rightarrow 0} D_{(2)}(d)} = \left(\frac{8(2 - \lambda)}{(\lambda - 4\lambda^{1/2} + 6)(3 - \lambda^2)} \right)^{1/2} =: z_1(\lambda)$$

gegeben ist. Denn zunächst gilt für die Funktion $z_1(\lambda)$, dass $\lim_{\lambda \rightarrow 0} z_1(\lambda) = (8/9)^{1/2}$ und $\lim_{\lambda \rightarrow 1} z_1(\lambda) = (4/3)^{1/2}$. Damit ist der Versuchsplan mit $d = 1/2$ dem Versuchsplan mit $d \rightarrow 0$ für $\lambda \rightarrow 1$ überlegen. Andererseits ist er für $\lambda \rightarrow 0$ dem Versuchsplan mit $d \rightarrow 0$ unterlegen. Es gilt sogar, dass die relative D-Effizienz des Versuchsplans mit $d \rightarrow 0$ bezüglich des Versuchsplans mit $d = 1/2$ in λ monoton wächst, siehe Behauptung 8.1.

Behauptung 8.1. *Sei $\rho = 1/2$. Gegeben seien die beiden Versuchspläne mit $d \rightarrow 0$ und $d = 1/2$. Dann gilt, dass die relative D-Effizienz des Plans mit $d \rightarrow 0$ bezüglich des Versuchsplans mit $d = 1/2$ in λ monoton wachsend ist.*

Beweis. Die quadrierte relative D-Effizienz des Versuchsplans mit $d \rightarrow 0$ bezüglich des Versuchsplans mit $d = 1/2$ ist für $\rho = 1/2$ gegeben durch

$$\frac{8(2 - \lambda)}{(\lambda - 4\lambda^{1/2} + 6)(3 - \lambda^2)}.$$

Die Substitution $\lambda = x^2$ liefert die Funktion

$$k(x) = \frac{8(2 - x^2)}{(x^2 - 4x + 6)(3 - x^4)}.$$

Es gilt

$$\frac{\partial}{\partial x} k(x) = \frac{32r_1(x)}{(x^2 - 4x + 6)^2(x^4 - 3)^2}$$

mit $r_1(x) = -x^7 + 3x^6 - 10x^4 + 12x^3 + 3x^2 - 12x + 6$. Nun gilt

$$r_1(x) \geq -2x^7 + 3x^6 - 10x^4 + 12x^3 + 3x^2 - 12x + 6 = (1 - x)r_2(x)$$

mit

$$r_2(x) = 2x^6 - x^5 - x^4 + 9x^3 - 3x^2 - 6x + 6.$$

Weiter gilt

$$r_2(x) \geq -4x^6 - x^5 - x^4 + 9x^3 - 3x^2 - 6x + 6 = (1-x)r_3(x)$$

mit

$$r_3(x) = 4x^5 + 5x^4 + 6x^3 - 3x^2 + 6.$$

Die Funktion $r_3(x)$ ist wegen Behauptung 5.13 positiv. Insgesamt folgt, dass die betrachtete relative D-Effizienz monoton wachsend in λ ist. \square

Die Behauptung 8.1 hat zusammen mit den relativen D-Effizienzen, die für $\lambda \rightarrow 0$ und $\lambda \rightarrow 1$ vorliegen, die Konsequenz, dass der Versuchsplan mit $d \rightarrow 0$ gegenüber dem Plan mit $d = 1/2$ bezüglich des D-Kriteriums im ungünstigsten Fall prozentual weniger verliert als umgekehrt.

Als Nächstes erfolgt ein Vergleich der Versuchspläne mit $d = 0$ und $d \rightarrow 0$. Die relative D-Effizienz des Versuchsplans mit $d = 0$ bezüglich des Versuchsplans mit $d \rightarrow 0$ lautet

$$q_{0,d \rightarrow 0}(\lambda) = \left(\frac{3 - \lambda^2}{4 - \lambda^2} \right)^{1/2} =: z_2(\lambda).$$

Die Funktion $z_2(\lambda)$ ist monoton fallend in λ , siehe Behauptung 8.2.

Behauptung 8.2. *Sei $\rho = 1/2$. Gegeben seien die beiden Versuchspläne mit $d = 0$ und $d \rightarrow 0$. Dann gilt, dass die relative D-Effizienz des Plans mit $d = 0$ bezüglich des Versuchsplans mit $d \rightarrow 0$ in λ monoton fallend ist.*

Beweis. Die quadrierte relative D-Effizienz des Versuchsplans mit $d = 0$ bezüglich des Versuchsplans mit $d \rightarrow 0$ ist für $\rho = 1/2$ gegeben durch

$$\frac{3 - \lambda^2}{4 - \lambda^2}.$$

Die erste Ableitung der quadrierten relativen D-Effizienz lautet

$$-\frac{2\lambda}{(\lambda - 2)^2(\lambda + 2)^2}.$$

Da dieser Ausdruck negativ ist, ist die relative D-Effizienz monoton fallend in λ . \square

Außerdem gilt $\lim_{\lambda \rightarrow 0} z_2(\lambda) = (3/4)^{1/2}$ und $\lim_{\lambda \rightarrow 1} z_2(\lambda) = (2/3)^{1/2}$. Damit ist der Versuchsplan mit $d = 0$ dem Versuchsplan mit $d \rightarrow 0$ für $\lambda \in (0, 1)$ bezüglich des D-Kriteriums auf Grund der Monotonie der relativen D-Effizienz unterlegen.

Abschließend werden nun noch die Versuchspläne mit $d = 0$ und $d = 1/2$ anhand ihrer relativen D-Effizienzen untersucht. Die relative D-Effizienz des Versuchsplans mit $d = 0$ bezüglich des Versuchsplans mit $d = 1/2$ lautet

$$q_{0,1/2}(\lambda) = \left(\frac{8}{(\lambda + 2)(\lambda - 4\lambda^{1/2} + 6)} \right)^{1/2} =: z_3(\lambda).$$

Erneut ist der Quotient monoton wachsend in λ , siehe Behauptung 8.3.

Behauptung 8.3. *Sei $\rho = 1/2$. Gegeben seien die beiden Versuchspläne mit $d = 0$ und $d = 1/2$. Dann gilt, dass die relative D-Effizienz des Plans mit $d = 0$ bezüglich des Versuchsplans mit $d = 1/2$ in λ monoton wachsend ist.*

Beweis. Die quadrierte relative D-Effizienz des Versuchsplans mit $d = 0$ bezüglich des Versuchsplans mit $d = 1/2$ ist für $\rho = 1/2$ gegeben durch

$$\frac{8}{(\lambda + 2)(\lambda - 4\lambda^{1/2} + 6)}.$$

Die Substitution $\lambda = x^2$ liefert die Funktion

$$k(x) = \frac{8(2 - x^2)}{(x^2 - 4x + 6)(3 - x^4)}.$$

Die erste Ableitung der quadrierten relativen D-Effizienz lautet

$$-\frac{2\lambda}{(\lambda - 2)^2(\lambda + 2)^2}.$$

Da dieser Ausdruck negativ ist, ist die relative D-Effizienz monoton fallend in λ . \square

Wegen $\lim_{\lambda \rightarrow 0} z_3(\lambda) = (2/3)^{1/2}$, $\lim_{\lambda \rightarrow 1} z_3(\lambda) = (8/9)^{1/2}$ und der Monotonie folgt nun, dass der Versuchsplan mit $d = 1/2$ für alle λ aus dem Intervall $(0, 1)$ ein kleineres D-Kriterium als der Versuchsplan mit $d = 0$ aufweist.

Abschließend wird der Punkt diskutiert, dass bei der Modellierung eines Nugget-Effekts der Versuchsplan mit $d = 0$ ein größeres D-Kriterium aufweist als der Versuchsplan mit $d \rightarrow 0$. Der Nugget-Effekt wird in Cressie (1993) unter anderem dadurch motiviert, dass Messfehler, die bei Wiederholungsmessungen für zwei an der nahezu selben Stelle platzierten Versuchspunkte entstehen, in der Modellierung Berücksichtigung finden. Es könnte

aber die Frage gestellt werden, warum diese Ungenauigkeiten nicht auch in dem Fall auftreten, wenn an der gleichen Stelle eine weitere Messung vorgenommen wird. In der zweiten Modellverallgemeinerung wird in den beiden angesprochenen Fällen nicht unterschieden. Vom Standpunkt der zweiten Modellverallgemeinerung liegt damit bei der ersten Modellverallgemeinerung ein zusätzlicher Informationsverlust vor sobald ein Versuchspunkt an der Stelle $d = 0$ statt an der Stelle $d \rightarrow 0$ positioniert wird.

9. Zusammenfassung

In dieser Arbeit wird der Frage nachgegangen, wie die Versuchspunkte eines Versuchsplans im einfachen linearen Regressionsmodell im Sinne des D-Kriteriums optimalerweise zu wählen sind. Dabei wird der Fall betrachtet, dass korrelierte Fehler vorliegen, wobei die Stärke der Korrelation zweier Fehler von einem Parameter λ und von der räumlichen Distanz der zugrunde liegenden Versuchspunkte abhängt. Insbesondere werden Versuchspläne untersucht, die sich aus unabhängigen Versuchseinheiten zusammensetzen, wobei für jede Versuchseinheit drei Versuchspunkte aus dem Intervall $[0, 1]$ geeignet zu wählen sind. In der Arbeit von Dette *et al.* (2008) konnte gezeigt werden, dass bezüglich des D-Kriteriums, das die Schätzung der Regressionsparameter widerspiegelt, der linke und rechte Rand des Versuchsraums mit jeweils einem Versuchspunkt zu belegen sind. Darüber hinaus konnte gezeigt werden, dass der D-optimale Versuchsplan im Grenzfall hochkorrelierter Fehler gegen den äquidistanten Versuchsplan konvergiert. Die vorliegende Arbeit greift die genannten Resultate auf. Auf Grund des Resultats, dass zwei Versuchspunkte des D-optimalen Versuchsplans am Rand des Versuchsraums liegen, ist lediglich der dritte Versuchspunkt d zu bestimmen. Ergänzend zu dem zweiten Resultat aus Dette *et al.* (2008) wird hier bezüglich des D-Kriteriums in Kapitel 5 gezeigt, dass bei hochkorrelierten Fehlern der Versuchspunkt d beliebig gewählt werden kann, da jeder Versuchsplan eine D-Effizienz von 1 aufweist. Ein eigener Abschnitt in Kapitel 5 befasst sich mit der Angabe von unteren Schranken für die D-Effizienzen von Versuchsplänen mit drei Versuchspunkten, wobei erneut nur der Versuchspunkt $d \in [0, 1]$ zu bestimmen ist. Es wird gezeigt, dass für Korrelationsparameter $\lambda \in [0, 1/5]$ jeder Versuchsplan mindestens eine D-Effizienz von $(1/2)^{1/2} \approx 0.7071$ aufweist. Es gilt sogar, dass die gefundene Mindesteffizienz in $\lambda \in (0, 1/5]$ monoton wächst. Für Korrelationen $\lambda \in [1/5, 1)$ wird die Schranke 0.9 hergeleitet. Da die D-Effizienz des Versuchsplans mit $d = 0$ für $\lambda \rightarrow 0$ mit der Mindesteffizienz übereinstimmt, liefert dieser Plan im ungünstigsten Fall die kleinste D-Effizienz, die überhaupt möglich ist. Dagegen wird für den Versuchsplan mit $d = 1/2$ gezeigt, dass dessen D-Effizienzen mindestens 0.8 betragen. Numerische Untersuchungen lassen die Vermutung zu, dass seine kleinste D-Effizienz sogar einen Wert von $(3/4)^{1/2} \approx 0.8660$ aufweist. In dieser Arbeit wird zumindest belegt, dass der Versuchsplan mit $d = 1/2$ dem Versuchsplan mit $d = 0$ bezüglich der Mindesteffizienz überlegen ist. Weitere numerische Untersuchungen zeigen, dass ein Versuchsplan mit $d \approx 0.138$ der Plan ist, der die maximale Mindesteffizienz aufweist. Es wird zumindest gezeigt, dass bezüglich der Versuchspunkte $d \in \{i/8 : i = 0, 1, 2, \dots, 8\} \cup \{1/16\}$ der Plan mit $d = 1/8$ der Plan mit der

größten Mindesteffizienz ist. Ein Nachteil der bisherigen Ergebnisse ist die Annahme, dass der Korrelationsparameter λ bekannt ist. Denn falls der Korrelationsparameter unbekannt sein sollte, so kann auf Grund der Tatsache, dass der D-optimale Versuchsplan vom Korrelationsparameter abhängt, kein Versuchsplan genannt werden, der für jedes $\lambda \in [0, 1]$ den D-optimale Versuchsplan darstellt. Einen Ansatz liefert in Kapitel 6 die Schätzung sämtlicher Parameter anhand eines zweistufigen Verfahrens. Aus anderen Arbeiten ist bereits bekannt, dass das zweistufige Verfahren ein asymptotisch effizientes Verfahren zur Schätzung der Regressionsparameter darstellt. Auf die hier vorliegende Situation übertragen bedeutet dies, dass der Versuchsplan einer Versuchseinheit für unendlich viele weitere Versuchseinheiten zu wiederholen ist. Im Rahmen dieser Arbeit wird mittels Simulationsstudien nachgewiesen, dass das D-Kriterium bei bekanntem Korrelationsparameter zu dem D-Kriterium bei unbekanntem Korrelationsparameter bereits für zwei Versuchseinheiten sehr ähnlich ist. Dies hat die Konsequenz, dass sich die obigen Ergebnisse bei bekanntem Korrelationsparameter auf den Fall eines unbekanntem Korrelationsparameters übertragen lassen. Im siebten Kapitel wird die Situation betrachtet, dass neben den beiden Regressionsparametern auch der Korrelationsparameter in das D-Kriterium einfließt. Es wird gezeigt, dass es im Sinne des D-Kriteriums besser ist, den dritten Versuchspunkt in die Nähe des Rands des Versuchsraums zu legen als in die Mitte. Ein letztes Kapitel beinhaltet im Vergleich zu den anderen Kapiteln zwei Modellverallgemeinerungen. Es wird ein Spezialfall untersucht, bei dem perfekte positive Korrelationen der Fehler im Grenzfalle $d \rightarrow 0$ ausgeschlossen sind. Ein Vergleich zwischen den drei Versuchsplänen mit den Versuchspunkten $d = 0$, $d \rightarrow 0$ und $d = 1/2$ zeigt dabei für die erste Modellverallgemeinerung, dass der Versuchspunkt $d = 0$ den anderen beiden Versuchspunkten unterlegen ist. Welcher von den beiden anderen Versuchspunkten zu bevorzugen ist, hängt von der vorliegenden Korrelation ab. Die zweite Modellverallgemeinerung unterscheidet sich bezüglich des D-Kriteriums von der ersten nur durch andere Werte für das D-Kriterium für den Versuchspunkt $d = 0$. Dies hat die Konsequenz, dass manche Ergebnisse zu der ersten Modellverallgemeinerung auf die zweite übertragen werden können.

A. Lemmata

Lemma A.1. *Seien $\lambda \in [0, 1/5]$ und $d \in [0, 1]$. Dann gilt:*

$$s_1(\lambda, d) \geq s_5(\lambda, d),$$

wobei

$$s_1(\lambda, d) = d^2(2\lambda^2 - \lambda - 1) + d(-2\lambda^2 + \lambda + 1),$$

$$s_5(\lambda, d) = 23(1 - d)d/25 + \lambda d(1 - d).$$

Beweis. Zunächst sei angemerkt, dass in den folgenden Überlegungen die bivariate Funktion $s_1(\lambda, d)$ als Funktionenschar aufgefasst wird, wobei λ die Rolle der unbestimmten Variable einnimmt und d den Parameter der Funktionenschar darstellt. Daher wird nun die Notation $s_{d,1}(\lambda)$ statt $s_1(\lambda, d)$ herangezogen.

Die Grundlage für den Beweis liefert (5.37) aus Behauptung 5.10. Jede Funktion der Funktionenschar $s_{d,1}(\lambda)$ ist konkav in λ , da

$$\frac{\partial^2}{\partial \lambda^2} s_{d,1}(\lambda) = 4(d - 1)d \leq 0.$$

Auf Grund der Konkavität ist in dem kompakten Intervall $[0, 1/5]$ für jeden Wert des Parameters d die größte positive Steigung der Funktionen an der Stelle $\lambda = 0$ zu finden. Die Steigungen lauten

$$\frac{\partial}{\partial \lambda} s_{d,1}(\lambda)|_{\lambda=0} = (1 - d)d.$$

Wird nun eine Geradenschar konstruiert, die die Steigungen $(1 - d)d$ aufweist und durch die Punkte $s_1(1/5, d) = 28(1 - d)d/25$ verläuft, so liefert in Abhängigkeit des Werts des Parameters d die jeweilige Gerade der Geradenschar eine untere Schranke für die jeweilige Funktion der Funktionenschar $s_{d,1}(\lambda)$. Damit die Geradenschar als Funktion formuliert werden kann, sind lediglich die Achsenabschnitte zu bestimmen, da die Steigungen der Geradenschar bereits bekannt sind. Die Achsenabschnitte können anhand der Funktionswerte an der Stelle $\lambda = 1/5$ und der bekannten Steigung ermittelt werden. Sie lauten $s_{d,1}(1/5) - 1/5 \frac{\partial}{\partial \lambda} s_{d,1}(\lambda)|_{\lambda=0} = 23/25(1 - d)d$. Insgesamt gilt nun für jede Funktion der Funktionenschar $s_{d,1}(\lambda)$, dass

$$s_{d,1}(\lambda) \geq \frac{23}{25}(1 - d)d + \lambda(1 - d)d =: s_{d,5}(\lambda).$$

Diese Geradenschar kann nun wiederum als bivariate Funktion $s_{d,5}(\lambda) =: s_5(\lambda, d)$ aufgefasst werden. □

Lemma A.2. Seien $\lambda \in [0, 1/5]$ und $d \in [1/2, 1]$. Dann gilt:

$$s_3(\lambda, d) \geq 0,$$

wobei $s_3(\lambda, d) = -\lambda^d d - \lambda^{2d} + \lambda^d + \lambda^2$.

Beweis. Es gilt:

$$\begin{aligned} & -\lambda^d d - \lambda^{2d} + \lambda^d + \lambda^2 \geq 0 \\ \Leftrightarrow & \lambda^d(1-d) \geq \lambda^{2d} - \lambda^2 \\ \Leftrightarrow & \lambda^d(1-d) \geq \lambda^d(\lambda^d - \lambda^{2-d}) \\ \Leftrightarrow & 1-d \geq \lambda^d - \lambda^{2-d}. \end{aligned}$$

Nun wird die letzte Zeile der Äquivalenzumformungen bewiesen. Definiere $h_\lambda(d) := \lambda^d - \lambda^{2-d}$ und $p(d) := 1-d$.

Wegen

$$\frac{\partial^2}{\partial d^2} h_\lambda(d) = (\ln(\lambda))^2 (\lambda^d - \lambda)(\lambda^d + \lambda) / (\lambda^d) \geq 0$$

ist $h_\lambda(d)$ konvex in d . Außerdem gilt

$$h_\lambda(1/2) = \lambda^{1/2} - \lambda^{3/2} \leq 1/2 = p(1/2)$$

sowie

$$h_\lambda(1) = p(1) = 0.$$

Nun kann wegen (5.30) aus Behauptung 5.9 gefolgert werden, dass die Ungleichung $h_\lambda(d) \leq p(d)$ gilt. \square

Lemma A.3. Seien $\lambda \in [0, 1/5]$ und $d \in [1/2, 1]$. Dann gilt:

$$s_4(\lambda, d) \geq 0,$$

wobei $s_4(\lambda, d) = \lambda^{1-d} d - \lambda^{2-2d}$.

Beweis. Fall 1 ($d \in [1/2, 1)$):

Der Beweis nutzt mehrere Abschätzungen. Ein paar derer werden nun vorab zusammengetragen. Unter anderem wird ausgenutzt, dass

$$5^{d-1} \geq \frac{2 - \ln(5)}{2\sqrt{5}} + d \frac{\ln(5)}{\sqrt{5}}. \quad (\text{A.1})$$

Dies gilt auf Grund von (5.40) aus Behauptung 5.10. Andererseits gilt, dass

$$5^{d-1} \leq \frac{2}{\sqrt{5}} - 1 + d\left(2 - \frac{2}{\sqrt{5}}\right). \quad (\text{A.2})$$

Diese Abschätzung gilt wegen (5.30) aus Behauptung 5.9 mit $a = 0$. Außerdem wird herangezogen, dass

$$\frac{3}{2} < \ln(5) < \frac{13}{7}, \quad (\text{A.3})$$

siehe Lemma 5.16.

Nun zum eigentlichen Beweis. Die bivariate Funktion $s_4(\lambda, d)$ wird als Funktionenschar mit Parameter d aufgefasst. Daher wird im Weiteren die Notation $s_{4,d}(\lambda)$ anstelle von $s_4(\lambda, d)$ herangezogen. Es wird nun gezeigt, dass jede Funktion der Funktionenschar nur positive Werte annimmt. Es gilt

$$\frac{\partial}{\partial \lambda} s_{4,d}(\lambda) = (1-d)d\lambda^{-d} - (2-2d)\lambda^{1-2d}. \quad (\text{A.4})$$

Die Ableitung aus (A.4) gleich Null setzen und umformen dieser Gleichung nach λ liefern

$$\lambda = (d/2)^{1/(1-d)}.$$

Folglich gibt es für jede Funktion der Schar genau einen Kandidaten für eine lokale Extremstelle. Damit gibt es aber auch nur höchstens eine lokale Extremstelle. Weiter gilt

$$\begin{aligned} \frac{\partial}{\partial \lambda} s_{4,d}(\lambda)|_{\lambda=1/5} &= (1-d)d5^d - (2-2d)5^{2d-1} \\ &= (1-d)5^d(d - 2 \cdot 5^{d-1}) \\ &\leq (1-d)5^d\left(d - 2\left(\frac{2 - \ln(5)}{2\sqrt{5}} + d\frac{\ln(5)}{\sqrt{5}}\right)\right) \quad , \text{ wegen (A.1),} \\ &= (1-d)5^d\left(\frac{\sqrt{5}(\ln(5) - 2)}{5} + \frac{5 - 2\sqrt{5}\ln(5)}{5}d\right) \\ &\leq (1-d)5^d\left(\frac{\sqrt{5}(13/7 - 2)}{5} + \frac{5 - 2\sqrt{5} \cdot 3/2}{5}d\right) \quad , \text{ wegen (A.3),} \\ &= (1-d)5^d\left(-\frac{\sqrt{5}}{35} + \frac{5 - 3\sqrt{5}}{5}d\right) \\ &\leq (1-d)5^d\left(-\frac{\sqrt{5}}{35} - \frac{1}{5}d\right) \quad , \text{ wegen } \sqrt{5} > 2, \\ &< 0. \end{aligned} \quad (\text{A.5})$$

Aus (A.5) ist ersichtlich, dass $s_{4,d}(\lambda)$ für $d \in [1/2, 1)$ in $\lambda = 1/5$ streng monoton fallend

ist. Weiter gilt, dass

$$\begin{aligned}
s_{4,d}(\lambda = 1/5) &= 5^{d-1}(d - 5^{d-1}) \\
&\geq 5^{d-1}\left(d - \left(\frac{2}{\sqrt{5}} - 1 + d\left(2 - \frac{2}{\sqrt{5}}\right)\right)\right) \quad , \text{ wegen (A.2),} \\
&= 5^{d-1}\left(1 - \frac{2}{\sqrt{5}} + d\left(\frac{2}{\sqrt{5}} - 1\right)\right) \\
&\geq 0.
\end{aligned} \tag{A.6}$$

Schließlich gilt, dass

$$s_{4,d}(0) = 0. \tag{A.7}$$

Nun kann abschließend gefolgert werden, dass $s_4(\lambda, d) \geq 0$. Dies geschieht wie folgt. Zunächst hat jede Funktion $s_{4,d}(\lambda)$ höchstens eine lokale Extremstelle in $d \in [1/2, 1)$. Da aber $s_{4,d}(\lambda)$ auf dem rechten Rand streng monoton fallend ist, kann es sich bei einem eventuellen lokalen Extremum nicht um ein Minimum handeln. Falls kein Extremum vorliegt, so muss $s_{4,d}(\lambda)$ auf $d \in [1/2, 1)$ monoton fallend sein. Damit befindet sich das absolute Minimum an der Stelle $\lambda = 1/5$, also am Rand. Falls dagegen doch ein lokales Maximum vorliegt, so liegt das absolute Minimum ebenfalls am Rand. Die Ränder sind jedoch wegen (A.6) und (A.7) nicht negativ.

Fall 2 ($d = 1$):

Es gilt, dass $s_{4,d=1}(\lambda) = 0$. □

Lemma A.4. *Seien $\lambda \in [0, 1/5]$ und $d \in [0, 1]$. Dann gilt:*

$$s_3(\lambda, d) + s_4(\lambda, d) \geq 0,$$

wobei $s_3(\lambda, d)$ und $s_4(\lambda, d)$ aus Lemma A.2 und Lemma A.3 entnommen sind.

Beweis. In Lemma A.2 und Lemma A.3 wurde jeweils lediglich gezeigt, dass die Ungleichung für $d \in [1/2, 1]$ gilt. Es gilt jedoch, dass

$$s_3(\lambda, d) + s_4(\lambda, d) = -\lambda^d d + \lambda^{1-d} d - \lambda^{2d} + \lambda^d - \lambda^{2-2d} + \lambda^2 = s_3(\lambda, 1-d) + s_4(\lambda, 1-d).$$

Somit ist die Funktion $s_3(\lambda, d) + s_4(\lambda, d)$ symmetrisch um $d = 1/2$. Dies bedeutet, dass $s_3(\lambda, d) + s_4(\lambda, d)$ sogar für $d \in [0, 1]$ nur positive Funktionswerte aufweist. □

Lemma A.5. *Seien $\lambda \in [0, 1/5]$ und $c \in [-1/2, 1/2]$. Dann gilt:*

$$t_2(\lambda, c) := -\lambda c^2 - \frac{23}{25}c^2 - 2\lambda^{c+3/2}c + 2\lambda^{3/2-c}c + \lambda^{c+3/2} + \lambda^{3/2-c} - \frac{7}{4}\lambda + \frac{23}{100} \text{ ist konvex in } \lambda.$$

Beweis. Es ist zu zeigen, dass

$$\frac{\partial^2}{\partial \lambda^2} t_2(\lambda, c) \geq 0.$$

Die zweite Ableitung von $t_2(\lambda, c)$ nach λ lautet

$$\frac{\partial^2}{\partial \lambda^2} t_2(\lambda, c) = \lambda^{c-1/2} w_1(c) + \lambda^{-c-1/2} w_2(c), \quad (\text{A.8})$$

wobei $w_1(c) = -2c^3 - 3c^2 + 1/2c + 3/4$ und $w_2(c) = 2c^3 - 3c^2 - 1/2c + 3/4$.

Die Funktionen $w_1(c)$ und $w_2(c)$ sind wegen

$$\frac{\partial^2}{\partial c^2} w_1(c) = -12c - 6 \leq 0$$

und

$$\frac{\partial^2}{\partial c^2} w_2(c) = 12c - 6 \leq 0$$

konkav. Da außerdem

$$w_1\left(-\frac{1}{2}\right) = w_1\left(\frac{1}{2}\right) = w_2\left(-\frac{1}{2}\right) = w_2\left(\frac{1}{2}\right) = 0,$$

gilt mit Behauptung 5.8

$$w_1(c) \geq 0$$

sowie

$$w_2(c) \geq 0.$$

Schließlich sind die Terme $\lambda^{c-1/2}$ und $\lambda^{-c-1/2}$ ebenfalls nicht negativ. Somit gilt insgesamt, dass

$$\frac{\partial^2}{\partial \lambda^2} t_2(\lambda, c) \geq 0.$$

□

Lemma A.6. Seien $m_1(c) := 5^{-c-1/2}$ und $m_2(c) := 5^{c-1/2}$. Sei $c \in [0, 1/2]$. Dann gilt:

$$m_1(c) \leq \frac{1}{2} - \frac{3}{5}c =: n_1(c) \quad (\text{A.9})$$

und

$$m_2(c) \leq \frac{1}{2} + c =: n_2(c). \quad (\text{A.10})$$

Beweis. Wegen

$$\frac{\partial^2}{\partial c^2} m_1(c) = (\ln(5))^2 5^{-c-1/2} \geq 0$$

und

$$\frac{\partial^2}{\partial c^2} m_2(c) = (\ln(5))^2 5^{c-1/2} \geq 0$$

sind $m_1(c)$ und $m_2(c)$ konvex. Zudem gilt, dass

$$m_1(0) = 5^{-1/2} \leq 4^{-1/2} = \frac{1}{2} = n_1(0),$$

$$m_1\left(\frac{1}{2}\right) = 1/5 = n_1\left(\frac{1}{2}\right),$$

$$m_2(0) = 5^{-1/2} \leq 4^{-1/2} = \frac{1}{2} = n_2(0)$$

und

$$m_2\left(\frac{1}{2}\right) = 1 = n_2\left(\frac{1}{2}\right).$$

Die Aussagen (A.9) und (A.10) folgen daraufhin aus (5.30). \square

Lemma A.7. *Seien $m_1(c)$ und $m_2(c)$ wie in Lemma A.6. Sei $c \in [-1/2, 0]$. Dann gilt:*

$$m_1(c) \leq \frac{1}{2} - c =: n_3(c) \tag{A.11}$$

und

$$m_2(c) \leq \frac{1}{2} + \frac{3}{5}c =: n_4(c). \tag{A.12}$$

Beweis. Folgt analog zu Lemma A.6 aus (5.31) mit

$$m_1\left(-\frac{1}{2}\right) = 1 = n_3\left(-\frac{1}{2}\right),$$

$$m_1(0) = 5^{-1/2} \leq 4^{-1/2} = \frac{1}{2} = n_3(0),$$

$$m_2\left(-\frac{1}{2}\right) = \frac{1}{5} = n_4\left(-\frac{1}{2}\right)$$

und

$$m_2(0) = 5^{-1/2} \leq 4^{-1/2} = \frac{1}{2} = n_4(0).$$

\square

Lemma A.8. *Sei $c \in [-1/2, 1/2]$. Dann gilt für $t_2(\lambda, c)$ aus Lemma A.5:*

$$t_2(\lambda, c) \text{ ist in } \lambda = 1/5 \text{ monoton nicht steigend in } \lambda. \tag{A.13}$$

Beweis. Die erste Ableitung von $t_2(\lambda, c)$ nach λ lautet

$$\frac{\partial}{\partial \lambda} t_2(\lambda, c) = \lambda^{c+1/2} w_3(c) + \lambda^{-c+1/2} w_4(c) - c^2 - 7/4 =: v_1(\lambda, c),$$

wobei $w_3(c) = (1 - 2c)(3 + 2c)/2$ und $w_4(c) = (1 + 2c)(3 - 2c)/2$.

Die Funktion $v_1(\lambda, c)$ an der Stelle $\lambda = 1/5$ ausgewertet liefert

$$v_1\left(\frac{1}{5}, c\right) = m_1(c)w_3(c) + m_2(c)w_4(c) - c^2 - \frac{7}{4}, \quad (\text{A.14})$$

wobei $m_1(c) = 5^{-c-1/2}$ und $m_2(c) = 5^{c-1/2}$. Für den Nachweis von (A.13) ist somit zu zeigen, dass

$$v_1\left(\frac{1}{5}, c\right) \leq 0. \quad (\text{A.15})$$

Es ist leicht zu sehen, dass

$$w_3(c) \geq 0, \quad w_4(c) \geq 0. \quad (\text{A.16})$$

Das weitere Vorgehen erfordert eine Fallunterscheidung.

Fall 1 ($c \in [0, 1/2]$):

Bezüglich (A.14) gilt:

$$\begin{aligned} v_1\left(\frac{1}{5}, c\right) &\leq n_1(c)w_3(c) + n_2(c)w_4(c) - c^2 - \frac{7}{4} \quad , \text{ wegen (A.9), (A.10) und (A.16),} \\ &= \frac{1 - 2c}{20}(8c^2 + 2c - 5) \\ &\leq \frac{1 - 2c}{20}(2c - 3) \quad , \text{ da } 1 - 2c \geq 0 \text{ und } 8c^2 \leq 2, \\ &\leq 0. \end{aligned} \quad (\text{A.17})$$

Fall 2 ($c \in [-1/2, 0)$):

Bezüglich (A.14) gilt:

$$\begin{aligned} v_1\left(\frac{1}{5}, c\right) &\leq n_3(c)w_3(c) + n_4(c)w_4(c) - c^2 - \frac{7}{4} \quad , \text{ wegen (A.11), (A.12) und (A.16),} \\ &= \frac{1 + 2c}{20}(8c^2 - 2c - 5) \\ &\leq \frac{1 + 2c}{20}(-2c - 3) \quad , \text{ da } 1 + 2c \geq 0 \text{ und } 8c^2 \leq 2, \\ &\leq 0. \end{aligned} \quad (\text{A.18})$$

Wegen (A.17) und (A.18) gilt (A.15) und damit auch (A.13). \square

Lemma A.9. Sei $c \in [-1/2, 1/2]$. Dann gilt für $t_2(\lambda, c)$ aus Lemma A.5:

$$t_2(1/5, c) \geq 0. \quad (\text{A.19})$$

Beweis. Es gilt:

$$t_2\left(\frac{1}{5}, c\right) = (2c + 1)5^{c-3/2} + (1 - 2c)5^{-c-3/2} - \frac{28}{25}c^2 - \frac{3}{25}. \quad (\text{A.20})$$

Fall 1 ($c \in [-1/2, -1/4]$):

Wegen

$$\frac{\partial^2}{\partial c^2} 5^{c-3/2} = (\ln(5))^2 5^{c-3/2} \geq 0 \text{ und } \frac{\partial^2}{\partial c^2} 5^{-c-3/2} = (\ln(5))^2 5^{-c-3/2} \geq 0,$$

sind $5^{c-3/2}$ und $5^{-c-3/2}$ konvex.

Anwendung von (5.40) liefert

$$5^{c-3/2} \geq \frac{\ln(5) + 2}{50} + \frac{\ln(5)}{25}c \quad (\text{A.21})$$

und

$$5^{-c-3/2} \geq \frac{2 - \ln(5)}{10} - \frac{\ln(5)}{5}c. \quad (\text{A.22})$$

Daraufhin gilt, dass

$$\begin{aligned} t_2\left(\frac{1}{5}, c\right) &\geq (2c + 1)\left(\frac{\ln(5) + 2}{50} + \frac{\ln(5)}{25}c\right) + (1 - 2c)\left(\frac{2 - \ln(5)}{10} - \frac{\ln(5)}{5}c\right) - \frac{28}{25}c^2 - \frac{3}{25} \\ &= \frac{(12\ln(5) - 28)c^2 + (2\ln(5) - 8)c - 2\ln(5) + 3}{25} \\ &=: t_3(c). \end{aligned} \quad (\text{A.23})$$

Die folgenden Ausführungen zeigen, dass $t_3(c) \geq 0$.

Zum einen kann aus Lemma 5.16 leicht gefolgert werden, dass $\ln(5) < 7/3$. Daraus folgt

$$\frac{\partial^2}{\partial c^2} t_3(c) = \frac{24\ln(5) - 56}{25} < 0.$$

Da außerdem $t_3(-1/2) = 0$ und wegen Lemma 5.16 $t_3(-1/4) = (13 - 7\ln(5))/100 \geq 0$, gilt wegen (5.28) schließlich, dass $t_3(c) \geq 0$.

Fall 2 ($c \in [-1/4, 0]$):

Aus Fall 1 folgt, dass $5^{c-3/2}$ und $5^{-c-3/2}$ konvex sind. Anwendung von (5.42) liefert

$$5^{c-3/2} \geq 5^{-3/2} + \frac{\ln(5)}{5\sqrt{5}}c$$

und

$$5^{-c-3/2} \geq 5^{-3/2} - \frac{\ln(5)}{5\sqrt{5}}c.$$

Daraufhin gilt, dass

$$\begin{aligned} t_2\left(\frac{1}{5}, c\right) &\geq (2c+1)\left(5^{-3/2} + \frac{\ln(5)}{5\sqrt{5}}c\right) + (1-2c)\left(5^{-3/2} - \frac{\ln(5)}{5\sqrt{5}}c\right) - \frac{28}{25}c^2 - \frac{3}{25} \\ &= \frac{(4\sqrt{5}\ln(5) - 28)c^2 + 2\sqrt{5} - 3}{25} \\ &=: t_4(c). \end{aligned}$$

Die folgenden Ausführungen zeigen, dass $t_4(c) \geq 0$.

Wegen Lemma 5.16 und $\sqrt{5} < 3$ gilt

$$\frac{\partial^2}{\partial c^2} t_4(c) = \frac{8(\sqrt{5}\ln(5) - 7)}{25} < \frac{8(3(13/7) - 7)}{25} = -\frac{16}{35} < 0.$$

Da außerdem

$$t_4(0) = \frac{10 - 3\sqrt{5}}{25\sqrt{5}} \geq \frac{10 - 9}{25\sqrt{5}} \geq 0$$

und wegen Lemma 5.16

$$t_4\left(-\frac{1}{4}\right) = \frac{\sqrt{5}(\ln(5) + 8) - 19}{100} \geq \frac{2(3/2 + 8) - 19}{100} = 0,$$

gilt wegen (5.28) schließlich, dass $t_4(c) \geq 0$.

Fall 3 ($c \in [0, 1/2]$):

Die Funktion $t_2(1/5, c)$ ist symmetrisch um 0. Daher folgt für $c \in [0, 1/2]$ aus Symmetriegründen auf Grund der ersten beiden Fälle, dass $t_2(1/5, c) \geq 0$. \square

Lemma A.10. *Seien $\lambda \in [1/5, 1]$ und $c \in [0, 1/2]$. Dann gilt: Jede Funktion der Funktionschar*

$$s_{2,\lambda}^{(2)}(c) = 81\lambda^{c+5/2}\ln(\lambda)(2\ln(\lambda)c - \ln(\lambda) + 4) + \lambda^{5/2-c}\ln(\lambda)(-2\ln(\lambda)c - \ln(\lambda) + 4)$$

hat ein globales Maximum an der Stelle $c = 0$.

Beweis. Es gilt:

$$\begin{aligned}
\frac{\partial}{\partial c} s_{2,\lambda}^{(2)}(c) &= 81\lambda^{c+5/2} \ln(\lambda)^2 (\ln(\lambda)(2c-1) + 6) \\
&\quad + 81\lambda^{5/2-c} \ln(\lambda)^2 (\ln(\lambda)(2c+1) - 6) \\
&\leq 81\lambda^{-c+5/2} \ln(\lambda)^2 (\ln(\lambda)(2c-1) + 6) \\
&\quad + 81\lambda^{5/2-c} \ln(\lambda)^2 (\ln(\lambda)(2c+1) - 6) \\
&= 324\lambda^{5/2-c} \ln(\lambda)^3 c \\
&\leq 0.
\end{aligned}$$

Somit ist $s_{2,\lambda}^{(2)}(c)$ monoton fallend in c . □

Lemma A.11. Seien $\lambda \in [1/5, 1]$ und $c \in [0, 1/2]$. Dann gilt:

$$s_2(\lambda, c) \geq m_2(\lambda, c),$$

wobei

$$\begin{aligned}
s_2(\lambda, c) &= 81\lambda^{c+5/2}(-1+2c) + 81\lambda^{5/2-c}(-1-2c), \\
m_2(\lambda, c) &= -162\lambda^{5/2} + \frac{1}{2}(81\lambda^{5/2} \ln(\lambda)^2 - 324\lambda^{5/2} \ln(\lambda) + 648\lambda^{5/2} - 648\lambda^2)c \\
&\quad + (81\lambda^{5/2}(4 - \ln(\lambda)) \ln(\lambda))c^2.
\end{aligned}$$

Beweis. Die Funktionen $s_2(\lambda, c)$ und $m_2(\lambda, c)$ werden als Funktionenscharen $s_{2,\lambda}(c)$ und $m_{2,\lambda}(c)$ aufgefasst. Die zweiten Ableitungen lauten

$$\begin{aligned}
\frac{\partial^2}{\partial c^2} s_{2,\lambda}(c) &= 81\lambda^{c+5/2} \ln(\lambda)(2\ln(\lambda)c - \ln(\lambda) + 4) + \lambda^{5/2-c} \ln(\lambda)(-2\ln(\lambda)c - \ln(\lambda) + 4) \\
&=: s_{2,\lambda}^{(2)}(c)
\end{aligned}$$

und

$$\frac{\partial^2}{\partial c^2} m_{2,\lambda}(c) = 648\lambda^{5/2} \ln(\lambda) - 162\lambda^{5/2} \ln(\lambda)^2.$$

Wegen Lemma A.10 gilt, dass

$$s_{2,\lambda}^{(2)}(c) \leq s_{2,\lambda}^{(2)}(0) = 648\lambda^{5/2} \ln(\lambda) - 162\lambda^{5/2} \ln(\lambda)^2 = \frac{\partial^2}{\partial c^2} m_{2,\lambda}(c).$$

Weiter gilt, dass $s_{2,\lambda}(0) = m_{2,\lambda}(0)$ und $s_{2,\lambda}(1/2) = m_{2,\lambda}(1/2)$. Daraufhin folgt mit Behauptung 5.11 schließlich, dass $s_{2,\lambda}(c) \geq m_{2,\lambda}(c)$. □

Lemma A.12. Seien $\lambda \in [1/5, 1]$ und $c \in [0, 1/2]$. Dann gilt: Jede Funktion der Funktionenschar

$$s_{3,\lambda}^{(2)}(c) = 81\lambda^{c+1/2}\ln(\lambda)(\ln(\lambda)(1-2c)-4) + 81\lambda^{1/2-c}\ln(\lambda)(\ln(\lambda)(1+2c)-4)$$

hat ein globales Maximum an der Stelle $c = 1/2$.

Beweis. Es gilt:

$$\begin{aligned} \frac{\partial}{\partial c} s_{3,\lambda}^{(2)}(c) &= 81\lambda^{c+1/2}\ln(\lambda)^2(\ln(\lambda)(1-2c)-6) \\ &\quad + 81\lambda^{1/2-c}\ln(\lambda)^2(\ln(\lambda)(-1-2c)+6) \\ &\geq 81\lambda^{c+1/2}\ln(\lambda)^2(\ln(\lambda)(1-2c)-6) \\ &\quad + 81\lambda^{c+1/2}\ln(\lambda)^2(\ln(\lambda)(-1-2c)+6) \\ &= -324\lambda^{c+1/2}\ln(\lambda)^3c \\ &\geq 0. \end{aligned}$$

Somit ist $s_{3,\lambda}^{(2)}(c)$ monoton wachsend in c . □

Lemma A.13. Seien $\lambda \in [1/5, 1]$ und $c \in [0, 1/2]$. Dann gilt:

$$s_3(\lambda, c) \geq m_3(\lambda, c),$$

wobei

$$\begin{aligned} s_3(\lambda, c) &= 81(-2\lambda^{c+1/2}c + 2\lambda^{1/2-c}c + \lambda^{c+1/2} + \lambda^{1/2-c}), \\ m_3(\lambda, c) &= 162\lambda^{1/2} + \frac{81}{2}(-\ln(\lambda)^2 - (-2\lambda - 2)\ln(\lambda) - 8\lambda^{1/2} + 8)c \\ &\quad + 81(\ln(\lambda)^2 + (-2\lambda - 2)\ln(\lambda))c^2. \end{aligned}$$

Beweis. Die Funktionen $s_3(\lambda, c)$ und $m_3(\lambda, c)$ werden als Funktionenschar $s_{3,\lambda}(c)$ und $m_{3,\lambda}(c)$ aufgefasst. Die zweiten Ableitungen lauten

$$\begin{aligned} \frac{\partial^2}{\partial c^2} s_{3,\lambda}(c) &= 81\lambda^{c+1/2}\ln(\lambda)(\ln(\lambda)(1-2c)-4) + 81\lambda^{1/2-c}\ln(\lambda)(\ln(\lambda)(1+2c)-4) \\ &=: s_{3,\lambda}^{(2)}(c) \end{aligned}$$

und

$$\frac{\partial^2}{\partial c^2} m_{3,\lambda}(c) = 162\ln(\lambda)(\ln(\lambda) - 2\lambda - 2).$$

Wegen Lemma A.12 gilt, dass

$$s_{3,\lambda}^{(2)}(c) \leq s_{3,\lambda}^{(2)}(1/2) = 162\ln(\lambda)(\ln(\lambda) - 2\lambda - 2) = \frac{\partial^2}{\partial c^2} m_{3,\lambda}(c).$$

Weiter gilt, dass $s_{3,\lambda}(0) = m_{3,\lambda}(0)$ und $s_{3,\lambda}(1/2) = m_{3,\lambda}(1/2)$. Daraufhin folgt mit Behauptung 5.11 schließlich, dass $s_{3,\lambda}(c) \geq m_{3,\lambda}(c)$. \square

Lemma A.14. *Seien $\lambda \in [1/5, 1]$ und $c \in [0, 1/2]$. Dann gilt: Jede Funktion der Funktionenschar*

$$s_{4,\lambda}^{(2)}(c) = \ln(\lambda)^2(-400\lambda^{2c+1} - 400\lambda^{1-2c})$$

hat ein globales Maximum an der Stelle $c = 0$.

Beweis. Es gilt:

$$\begin{aligned} \frac{\partial}{\partial c} s_{4,\lambda}^{(2)}(c) &= 800\ln(\lambda)^3(\lambda^{1-2c} - \lambda^{2c+1}) \\ &\leq 0. \end{aligned}$$

Somit ist $s_{4,\lambda}^{(2)}(c)$ monoton fallend in c . \square

Lemma A.15. *Seien $\lambda \in [1/5, 1]$ und $c \in [0, 1/2]$. Dann gilt:*

$$s_4(\lambda, c) \geq m_4(\lambda, c),$$

wobei

$$\begin{aligned} s_4(\lambda, c) &= -100\lambda^{2c+1} - 100\lambda^{1-2c}, \\ m_4(\lambda, c) &= -200\lambda + 200(\lambda\ln(\lambda)^2 - \lambda^2 + 2\lambda - 1)c - 400\lambda\ln(\lambda)^2 c^2. \end{aligned}$$

Beweis. Die Funktionen $s_4(\lambda, c)$ und $m_4(\lambda, c)$ werden als Funktionenscharen $s_{4,\lambda}(c)$ und $m_{4,\lambda}(c)$ aufgefasst. Die zweiten Ableitungen lauten

$$\begin{aligned} \frac{\partial^2}{\partial c^2} s_{4,\lambda}(c) &= \ln(\lambda)^2(-400\lambda^{2c+1} - 400\lambda^{1-2c}) \\ &=: s_{4,\lambda}^{(2)}(c) \end{aligned}$$

und

$$\frac{\partial^2}{\partial c^2} m_{4,\lambda}(c) = -800\lambda\ln(\lambda)^2.$$

Wegen Lemma A.14 gilt, dass

$$s_{4,\lambda}^{(2)}(c) \leq s_{4,\lambda}^{(2)}(0) = -800\lambda\ln(\lambda)^2 = \frac{\partial^2}{\partial c^2} m_{4,\lambda}(c).$$

Weiter gilt, dass $s_{4,\lambda}(0) = m_{4,\lambda}(0)$ und $s_{4,\lambda}(1/2) = m_{4,\lambda}(1/2)$. Daraufhin folgt mit Behauptung 5.11 schließlich, dass $s_{4,\lambda}(c) \geq m_{4,\lambda}(c)$. \square

Lemma A.16. Sei $\lambda \in [1/2, 1]$. Dann gilt:

$$\ln(\lambda) \geq t_5(\lambda), \quad (\text{A.24})$$

wobei $t_5(\lambda) = 1/3\lambda^3 - 3/2\lambda^2 + 46/15\lambda - 19/10$.

Beweis. Es gilt:

1. $\ln(1/2) \geq t_5(1/2)$,
2. $\ln(1) = t_5(1)$,
3. $\frac{\partial^2}{\partial \lambda^2} \ln(\lambda) = -1/\lambda^2 \leq 2\lambda - 3 = \frac{\partial^2}{\partial \lambda^2} t_5(\lambda)$.

Unter Zuhilfenahme der Abschätzung $\ln(2) \leq 7/10$ aus Lemma 5.17 gilt die erste Ungleichung wegen

$$\ln(1/2) = -\ln(2) \geq -7/10 = t_5(1/2).$$

Die dritte Ungleichung gilt wegen (5.38).

Somit sind die Voraussetzungen von Behauptung 5.11 erfüllt. Folglich gilt (A.24). \square

Lemma A.17. Sei $\lambda \in [1/2, 1]$. Dann gilt:

$$\ln(\lambda)^2 \geq t_6(\lambda)^2,$$

wobei $t_6(\lambda) = \lambda^3 - 7/2\lambda^2 + 39/8\lambda - 19/8$.

Beweis. Es gilt:

1. $\ln(1/2) \leq t_6(1/2)$,
2. $\ln(1) = t_6(1)$,
3. $\frac{\partial^2}{\partial \lambda^2} \ln(\lambda) = -1/\lambda^2 \geq 6\lambda - 7 = \frac{\partial^2}{\partial \lambda^2} t_6(\lambda)$.

Unter Zuhilfenahme der Abschätzung $\ln(2) \geq 11/16$ aus Lemma 5.17 gilt die erste Ungleichung wegen

$$\ln(1/2) = -\ln(2) < -11/16 = t_6(1/2).$$

Die dritte Ungleichung gilt auf Grund von (5.32). Somit sind die Voraussetzungen von Behauptung 5.11 erfüllt. Folglich gilt, dass $\ln(\lambda) \leq t_6(\lambda)$. Wegen $t_6(\lambda) \leq t_6(\lambda) + 7/8(1 - \lambda) = 1/2(\lambda - 1)^2(2\lambda - 3) \leq 0$ gilt, dass $\ln(\lambda) \leq t_6(\lambda) \leq 0$. Daraufhin kann gefolgert werden, dass $\ln(\lambda)^2 \geq t_6(\lambda)^2$. \square

Lemma A.18. Sei $\lambda \in [1/5, 1]$. Dann gilt:

$$\ln(\lambda)^2 \leq t_7(\lambda)^2, \quad (\text{A.25})$$

wobei $t_7(\lambda) = -1/2\lambda^2 + 681/260\lambda - 551/260$.

Beweis. Es gilt:

1. $\ln(1/5) \geq t_7(1/5)$,
2. $\ln(1) = t_7(1)$,
3. $\frac{\partial^2}{\partial \lambda^2} \ln(\lambda) = -1/\lambda^2 \leq -1 = \frac{\partial^2}{\partial \lambda^2} t_7(\lambda)$.

Unter Zuhilfenahme der Abschätzung $\ln(5) \leq 21/13$ aus Lemma 5.14 gilt die erste Ungleichung wegen

$$\ln(1/5) = -\ln(5) \geq -21/13 = t_7(1/5).$$

Somit sind die Voraussetzungen von Behauptung 5.11 erfüllt. Folglich gilt $\ln(\lambda) \geq t_7(\lambda)$. Es gilt sogar $t_7(\lambda) \leq \ln(\lambda) \leq 0$. Daraufhin gilt schließlich, dass $\ln(\lambda)^2 \leq t_7(\lambda)^2$. \square

Lemma A.19. Sei $\lambda \in [1/5, 1]$. Dann gilt:

$$\ln(\lambda) \leq t_8(\lambda),$$

wobei $t_8(\lambda) = 1/3\lambda^3 - 3/2\lambda^2 + 3\lambda - 11/6$.

Beweis. Es gilt:

1. $\ln(1) = t_8(1)$,
2. $\frac{\partial}{\partial \lambda} \ln(\lambda) = 1/\lambda \geq \lambda^2 - 3\lambda + 3 = \frac{\partial}{\partial \lambda} t_8(\lambda)$.

Die Ungleichung gilt wegen

$$\lambda \left(\frac{\partial}{\partial \lambda} \ln(\lambda) - \frac{\partial}{\partial \lambda} t_8(\lambda) \right) = (1 - \lambda)^3 \geq 0.$$

Somit sind die Voraussetzungen von Behauptung 5.12 erfüllt. Folglich gilt $\ln(\lambda) \leq t_8(\lambda)$. \square

Lemma A.20. Sei $\lambda \in [1/2, 1]$. Dann gilt:

$$n_1(\lambda) + n_2(\lambda) := \ln(\lambda)^2 p_1(\lambda) + \ln(\lambda) p_2(\lambda) + p_3(\lambda) \leq 0$$

mit

$$\begin{aligned} p_1(\lambda) &= -\frac{81}{2}\lambda^{5/2} - 200\lambda + \frac{81}{2}, \\ p_2(\lambda) &= 162\lambda^{5/2} - 81\lambda - 81, \\ p_3(\lambda) &= -162\lambda^3 + 324\lambda^{5/2} - 362\lambda^2 + 562\lambda - 324\lambda^{1/2} - 38. \end{aligned}$$

Beweis. Zunächst gilt:

1. $p_1(\lambda) < 0$, denn $-\frac{81}{2}\lambda^{5/2} - 200\lambda + \frac{81}{2} \leq -\frac{81}{2}\lambda^{5/2} - 200(1/2) + \frac{81}{2} = -\frac{81}{2}\lambda^{5/2} - \frac{119}{2} < 0$,
2. $p_2(\lambda) \leq 0$, denn $162\lambda^{5/2} - 81\lambda - 81 \leq 162\lambda^2 - 81\lambda - 81 = 81(\lambda - 1)(2\lambda + 1) \leq 0$.

Daraufhin gilt mit Lemma A.16 und Lemma A.17:

$$\begin{aligned} \ln(\lambda)^2 p_1(\lambda) + \ln(\lambda) p_2(\lambda) + p_3(\lambda) &\leq t_6(\lambda)^2 p_1(\lambda) + t_5(\lambda) p_2(\lambda) + p_3(\lambda) \\ &= \frac{1}{640}(1 - \lambda)q_1(\lambda) \end{aligned} \quad (\text{A.26})$$

mit

$$\begin{aligned} q_1(\lambda) &= 25920\lambda^{15/2} - 155520\lambda^{13/2} + 128000\lambda^6 + 414720\lambda^{11/2} \\ &\quad - 793920\lambda^5 - 627480\lambda^{9/2} + 2203520\lambda^4 + 574965\lambda^{7/2} \\ &\quad - 3325440\lambda^3 - 343197\lambda^{5/2} + 2895400\lambda^2 - 207360\lambda^{3/2} \\ &\quad - 802629\lambda - 207360\lambda^{1/2} + 220381. \end{aligned}$$

Es bleibt zu zeigen, dass $q_1(\lambda) \leq 0$ für $\lambda \in (1/2, 1]$. Eine Substitution mittels $\lambda = x^2$ liefert

$$q_1(x^2) = (x - 1)q_2(x)$$

mit

$$\begin{aligned} q_2(x) &= 25920x^{14} + 25920x^{13} - 129600x^{12} - 1600x^{11} + 413120x^{10} \\ &\quad - 380800x^9 - 1008280x^8 + 1195240x^7 + 1770205x^6 - 1555235x^5 \\ &\quad - 1898432x^4 + 996968x^3 + 789608x^2 - 13021x - 220381 \\ &= 9632x^{14} + (1 - x)((1 - x)(493x^8 + 988x^7 + (633 - 633x)^2x^6 + 108x^6 \\ &\quad + (1070 - 1070x)^2x^5 + 1443x^5 + (1081 - 1081x)^2x^4 + 2064x^4 \\ &\quad + (775 - 775x)^2x^3 + (1158 - 1158x)^2x^2 + v_2(x)) + v_1(x)). \end{aligned} \quad (\text{A.27})$$

Dabei ist

$$\begin{aligned} v_1(x) &= -16288x^{13} - 42208x^{12} + 87392x^{11} + 88992x^{10} + 77054x^9 \\ v_2(x) &= 3736900x^3 - 1238541x^2 - 453783x - 220381. \end{aligned}$$

Nun ist zu zeigen, dass $q_2(x)$ für $x \in (1/\sqrt{2}, 1]$ positiv ist. Die Funktion $v_1(x)$ ist wegen Behauptung 5.13 positiv. Auf Grund der folgenden Überlegungen ist die Funktion $v_2(x)$ ebenfalls positiv. Es handelt sich um ein Polynom dritten Grades. Damit ist die erste Ableitung eine quadratische Funktion. Deren Graph ist eine nach oben geöffnete Parabel. Dies bedeutet, dass die erste Ableitung entweder stets positiv ist oder in genau einem kompakten Intervall A negativ ist und ansonsten positiv ist. Wegen

$$\frac{\partial}{\partial x} v_2(x)|_{x=0} = -453783 < 0$$

trifft letzteres zu. Es gilt also, dass $0 \in A$. Da außerdem die Ableitung von $v_2(x)$ an der Stelle $1/\sqrt{2}$ wegen

$$\begin{aligned} \frac{\partial}{\partial x} v_2(x)|_{x=1/\sqrt{2}} &= (2829334 - 1679303\sqrt{2})/(2\sqrt{2}) \\ &\geq (2829334 - 1679303(3/2))/(2\sqrt{2}) \\ &= 620759/(4\sqrt{2}) \\ &> 0 \end{aligned}$$

positiv ist, folgt, dass die rechte Intervallgrenze von A kleiner als $1/\sqrt{2}$ ist. Dabei wird ausgenutzt, dass $\sqrt{2} < 3/2$. Dies gilt wegen $2 = (\sqrt{2})^2 < (3/2)^2 = 2\frac{1}{4}$. Damit gilt, dass $v_2(x)$ in dem Intervall $[1/\sqrt{2}, 1]$ monoton wachsend ist. Nun kann wegen $v_1(x) \geq 0$, $v_2(x) \geq 0$ und anhand der Darstellung (A.27) gefolgert werden, dass $q_2(x)$ in dem relevanten Intervall positiv ist. \square

Lemma A.21. *Seien $\lambda \in [1/5, 1]$. Dann gilt:*

$$n_2(\lambda) := \ln(\lambda)^2 p_4(\lambda) + \ln(\lambda) p_5(\lambda) + p_6(\lambda) \leq 0,$$

wobei

$$\begin{aligned} p_4(\lambda) &= \frac{81}{2}\lambda^{5/2} + 200\lambda - \frac{81}{2}, \\ p_5(\lambda) &= -162\lambda^{5/2} + 81\lambda + 81, \\ p_6(\lambda) &= 324\lambda^{5/2} - 524\lambda^2 + 400\lambda - 324\lambda^{1/2} + 124. \end{aligned}$$

Beweis. Zunächst wird die Konkavität von $n_2(\lambda)$ gezeigt. Es gilt:

$$\frac{\partial^2}{\partial \lambda^2} n_2(\lambda) = \ln(\lambda)^2 p_7(\lambda) + \ln(\lambda) p_8(\lambda) + p_9(\lambda)$$

mit

$$\begin{aligned} p_7(\lambda) &= \frac{1215}{8} \lambda^{1/2}, \\ p_8(\lambda) &= -\frac{567}{2} \lambda^{1/2} + \frac{400}{\lambda} + \frac{81}{\lambda^2}, \\ p_9(\lambda) &= 648 \lambda^{1/2} + \frac{481}{\lambda} + \frac{81}{\lambda^{3/2}} - \frac{162}{\lambda^2} - 1048. \end{aligned}$$

Dabei gilt:

1. $p_7(\lambda) \geq 0$,
2. $p_8(\lambda) \geq 0$, denn $-\frac{567}{2} \lambda^{1/2} + \frac{400}{\lambda} + \frac{81}{\lambda^2} = \frac{-567\lambda^{5/2} + 800\lambda + 162}{2\lambda^2} \geq \frac{-567\lambda + 800\lambda + 162}{2\lambda^2} \geq 0$.

Daraufhin gilt mit Lemma A.18 und Lemma A.19:

$$\begin{aligned} \ln(\lambda)^2 p_7(\lambda) + \ln(\lambda) p_8(\lambda) + p_9(\lambda) &\leq t_7(\lambda)^2 p_7(\lambda) + t_8(\lambda) p_8(\lambda) + p_9(\lambda) \\ &= q_3(\lambda) / q_4(\lambda) \end{aligned}$$

mit

$$\begin{aligned} q_3(\lambda) &= 12320100\lambda^7 - 159740100\lambda^6 + 580503429\lambda^5 + 43264000\lambda^{9/2} - 823057038\lambda^4 \\ &\quad - 185927040\lambda^{7/2} + 600236649\lambda^3 + 9896640\lambda^{5/2} - 3028480\lambda^{3/2} + 26282880\lambda \\ &\quad - 100751040\lambda^{1/2}, \\ q_4(\lambda) &= 324480\lambda^{5/2}. \end{aligned}$$

Offensichtlich ist $q_4(\lambda) > 0$. Es bleibt folglich zu zeigen, dass $q_3(\lambda) \leq 0$. Die Substitution $\lambda = x^2$ liefert $q_3(x^2) = x(x-1)q_5(x)$ mit

$$\begin{aligned} q_5(x) &= 12320100x^{12} + 12320100x^{11} - 147420000x^{10} - 147420000x^9 + 433083429x^8 \\ &\quad + 476347429x^7 - 346709609x^6 - 532636649x^5 + 67600000x^4 \\ &\quad + 77496640x^3 + 77496640x^2 + 74468160x + 100751040. \end{aligned}$$

Nun ist abschließend zu zeigen, dass $q_5(x) \geq 0$ für $x \in [1/\sqrt{2}, 1]$. Dies folgt aus

$$\begin{aligned}
q_5(x) &= 12320100x^{12} + 12320100x^{11} - 147420000x^{10} - 147420000x^9 + 427897080x^8 \\
&\quad - 5186349(1-x)x^7 - 481533778(1-x)x^6 - 134824169(1-x)x^5 \\
&\quad + 397812480(1-x)x^4 + 330212480(1-x)x^3 + 252715840(1-x)x^2 \\
&\quad + 175219200(1-x)x + 100751040(1-x) \\
&\geq 12320100x^{12} + 12320100x^{11} - 147420000x^8 - 147420000x^8 + 427897080x^8 \\
&\quad - 5186349(1-x)x^7 - 481533778(1-x)x^6 - 134824169(1-x)x^5 \\
&\quad + 397812480(1-x)x^6 + 330212480(1-x)x^6 + 252715840(1-x)x^5 \\
&\quad + 175219200(1-x)x + 100751040(1-x)x^7 \\
&= 12320100x^{12} + 12320100x^{11} + 133057080x^8 + 95564691(1-x)x^7 \\
&\quad 246491182(1-x)x^6 + 117891671(1-x)x^5 + 175219200(1-x)x \\
&\geq 0.
\end{aligned}$$

Somit ist die Konkavität von $n_2(\lambda)$ gezeigt. Im nächsten Schritt wird die erste Ableitung an der Stelle $\lambda = 1$ bestimmt. Sie lautet

$$\begin{aligned}
\frac{\partial}{\partial \lambda} n_2(\lambda)|_{\lambda=1} &= (2\ln(\lambda) \frac{1}{\lambda} p_4(\lambda) + \ln(\lambda)^2 (\frac{\partial}{\partial \lambda} p_4(\lambda)) \\
&\quad + \frac{1}{\lambda} p_5(\lambda) + \ln(\lambda) (\frac{\partial}{\partial \lambda} p_5(\lambda)) + \frac{\partial}{\partial \lambda} p_6(\lambda))|_{\lambda=1} \\
&= 0.
\end{aligned}$$

Die bisherigen Resultate ermöglichen nun den Nachweis, dass $n_2(\lambda)$ keine positiven Werte annimmt. Auf Grund der Konkavität von $n_2(\lambda)$ ist die erste Ableitung von $n_2(\lambda)$ monoton fallend. Da sie zudem am rechten Rand den Wert Null annimmt, ist die erste Ableitung nicht negativ. Somit ist die Funktion $n_2(\lambda)$ stets monoton wachsend. Darüber hinaus gilt, dass $n_2(1) = 0$. □

Lemma A.22. *Sei $\lambda \in [1/5, 1/2]$. Dann gilt:*

$$\ln(\lambda)^2 \geq t_9(\lambda)^2,$$

wobei $t_9(\lambda) = 35/3\lambda^3 - 39/2\lambda^2 + 1457/120\lambda - 401/120$.

Beweis. Es gilt:

$$1. \ln(1/5) \leq t_9(1/5),$$

2. $\ln(1/2) \leq t_9(1/2)$,
3. $\frac{\partial^2}{\partial \lambda^2} \ln(\lambda) = -1/\lambda^2 \geq -39 + 70\lambda = \frac{\partial^2}{\partial \lambda^2} t_9(\lambda)$.

Unter Zuhilfenahme der Abschätzungen $\ln(5) \geq 8/5$ und $\ln(2) \geq 11/16$ aus Lemma 5.14 und Lemma 5.17 gelten die ersten beiden Ungleichungen wegen

$$\ln(1/5) = -\ln(5) < -8/5 = t_9(1/5)$$

und

$$\ln(1/2) = -\ln(2) < -11/16 = t_9(1/2).$$

Die dritte Ungleichung gilt auf Grund von (5.32). Somit sind die Voraussetzungen von Behauptung 5.11 erfüllt. Folglich gilt, dass

$$\ln(\lambda) \leq t_9(\lambda). \tag{A.28}$$

Wegen

$$\begin{aligned} \frac{\partial}{\partial \lambda} t_9(\lambda) &= (4200\lambda^2 - 4680\lambda + 1457)/120 \\ &\geq (4144\lambda^2 - 4736\lambda + 1322)/120 \\ &= 148(1 - 2\lambda)(9 - 14\lambda)/120 \\ &\geq 0 \end{aligned}$$

ist $t_9(\lambda)$ monoton wachsend in λ . Da außerdem $t_9(1/2)$ negativ ist, folgt, dass $t_9(\lambda)$ für $\lambda \in [1/5, 1/2]$ negativ ist. Daraufhin kann auf Grund von (A.28) gefolgert werden, dass $\ln(\lambda)^2 \geq t_9(\lambda)^2$. \square

Lemma A.23. *Sei $\lambda \in [1/5, 1/2]$. Dann gilt:*

$$\ln(\lambda) \geq t_{10}(\lambda), \tag{A.29}$$

wobei $t_{10}(\lambda) = 3\lambda - 9/4$.

Beweis. Es gilt:

1. $\ln(1/5) \geq t_{10}(1/5)$,
2. $\ln(1/2) \geq t_{10}(1/2)$,
3. $\frac{\partial^2}{\partial \lambda^2} \ln(\lambda) = -1/\lambda^2 \leq 0 = \frac{\partial^2}{\partial \lambda^2} t_{10}(\lambda)$.

Unter Zuhilfenahme der Abschätzungen $\ln(5) \leq 21/13$ und $\ln(2) \leq 7/10$ aus Lemma 5.14 und Lemma 5.17 gelten die ersten beiden Ungleichungen wegen

$$\ln(1/5) = -\ln(5) > -21/13 > -33/20 = t_{10}(1/5)$$

und

$$\ln(1/2) = -\ln(2) > -7/10 > -3/4 = t_{10}(1/2).$$

Somit sind die Voraussetzungen von Behauptung 5.11 erfüllt. Folglich gilt (A.29). \square

Lemma A.24. *Sei $\lambda \in [1/5, 1/2]$. Dann gilt:*

$$\ln(\lambda)^2 \leq t_{11}(\lambda)^2, \tag{A.30}$$

wobei $t_{11}(\lambda) = \frac{119}{39}\lambda - \frac{434}{195}$.

Beweis. Es gilt:

1. $\ln(1/5) \geq t_{11}(1/5)$,
2. $\ln(1/2) \geq t_{11}(1/2)$,
3. $\frac{\partial^2}{\partial \lambda^2} \ln(\lambda) = -1/\lambda^2 \leq 0 = \frac{\partial^2}{\partial \lambda^2} t_{11}(\lambda)$.

Unter Zuhilfenahme der Abschätzung $\ln(5) \leq 21/13$ aus Lemma 5.14 gilt die erste Ungleichung wegen

$$\ln(1/5) = -\ln(5) \geq -21/13 = t_{11}(1/5).$$

Mit Lemma 5.17 gilt die zweite Ungleichung, da

$$\ln(1/2) = -\ln(2) \geq -7/10 = t_{11}(1/2).$$

Somit sind die Voraussetzungen von Behauptung 5.11 erfüllt. Folglich gilt $\ln(\lambda) \geq t_{11}(\lambda)$. Es gilt sogar $t_{11}(\lambda) \leq \ln(\lambda) \leq 0$. Daraufhin gilt schließlich, dass $\ln(\lambda)^2 \leq t_{11}(\lambda)^2$. \square

Lemma A.25. *Sei $\lambda \in [1/5, 1/2]$. Dann gilt:*

$$\ln(\lambda) \leq t_{12}(\lambda),$$

wobei $t_{12}(\lambda) = -\lambda^2 + 3\lambda - \frac{31}{16}$.

Beweis. Es gilt:

1. $\ln(1/2) \leq t_{12}(1/2)$,

$$2. \quad \frac{\partial}{\partial \lambda} \ln(\lambda) = 1/\lambda \geq 3 - 2\lambda = \frac{\partial}{\partial \lambda} t_{12}(\lambda).$$

Die erste Ungleichung gilt wegen Lemma 5.17, da $t_{12}(1/2) = -11/16$. Die zweite Ungleichung gilt wegen

$$\lambda \left(\frac{\partial}{\partial \lambda} \ln(\lambda) - \frac{\partial}{\partial \lambda} t_{12}(\lambda) \right) = (1 - \lambda)(1 - 2\lambda) \geq 0.$$

Somit sind die Voraussetzungen von Behauptung 5.12 erfüllt. Folglich gilt $\ln(\lambda) \leq t_{12}(\lambda)$. \square

Lemma A.26. *Seien $\lambda \in [1/5, 1/2]$ und $c \in [0, 1/2]$. Dann gilt:*

$$s_{6a,\lambda}^{(3)}(c) \leq k_{6a,\lambda}^{(3)}(c),$$

wobei

$$\begin{aligned} s_{6a,\lambda}^{(3)}(c) &= 81\lambda^{c+1/2} \ln(\lambda)^2 (\ln(\lambda)(1 - 2c) - 6) + 81\lambda^{1/2-c} \ln(\lambda)^2 (\ln(\lambda)(-1 - 2c) + 6), \\ k_{6a,\lambda}^{(3)}(c) &= -324 \ln(\lambda)^2 (\ln(\lambda) + 3\lambda - 3)c. \end{aligned}$$

Beweis. Es gilt:

$$\begin{aligned} \frac{\partial^2}{\partial c^2} s_{6a,\lambda}^{(3)}(c) &= 81\lambda^{c+1/2} \ln(\lambda)^4 (\ln(\lambda)(1 - 2c) - 10) \\ &\quad + 81\lambda^{1/2-c} \ln(\lambda)^4 (\ln(\lambda)(-1 - 2c) + 10) \\ &\geq 81\lambda^{c+1/2} \ln(\lambda)^4 (\ln(\lambda)(1 - 2c) - 10) \\ &\quad + 81\lambda^{c+1/2} \ln(\lambda)^4 (\ln(\lambda)(-1 - 2c) + 10) \\ &= -324\lambda^{c+1/2} \ln(\lambda)^5 c \\ &\geq 0 \end{aligned}$$

und

$$\frac{\partial^2}{\partial c^2} k_{6a,\lambda}^{(3)}(c) = 0.$$

Weiter gilt, dass $s_{6a,\lambda}^{(3)}(0) = k_{6a,\lambda}^{(3)}(0)$ und $s_{6a,\lambda}^{(3)}(1/2) = k_{6a,\lambda}^{(3)}(1/2)$. Daraufhin folgt mit Behauptung 5.11 schließlich, dass $s_{6a,\lambda}^{(3)}(c) \leq k_{6a,\lambda}^{(3)}(c)$. \square

Lemma A.27. *Seien $\lambda \in [1/5, 1/2]$ und $c \in [0, 1/2]$. Dann gilt:*

$$s_{6b,\lambda}^{(3)}(c) \leq k_{6b,\lambda}^{(3)}(c),$$

wobei

$$\begin{aligned} s_{6b,\lambda}^{(3)}(c) &= 800 \ln(\lambda)^3 (\lambda^{1-2c} - \lambda^{2c+1}), \\ k_{6b,\lambda}^{(3)}(c) &= -3200 \lambda \ln(\lambda)^4 c. \end{aligned}$$

Beweis. Es gilt, dass $s_{6b,\lambda}^{(3)}(0) = k_{6b,\lambda}^{(3)}(0)$. Somit weisen die beiden Funktionen an der Stelle $c = 0$ den selben Funktionswert auf. Die Behauptung gilt nun, falls die Differenz $k_{6b,\lambda}^{(3)}(c) - s_{6b,\lambda}^{(3)}(c)$ in c monoton wächst. Dies gilt wegen

$$\frac{\partial}{\partial c}(k_{6b,\lambda}^{(3)}(c) - s_{6b,\lambda}^{(3)}(c)) = 1600\lambda^{1-2c}\ln(\lambda)^4(\lambda^c - 1)^2(\lambda^c + 1)^2 \geq 0.$$

□

Lemma A.28. *Seien $\lambda \in [1/5, 1/2]$ und $c \in [0, 1/2]$. Dann gilt: Jede Funktion der Funktionenschar*

$$\begin{aligned} s_{6,\lambda}^{(2)}(c) = & -162\lambda^{c+1/2}\ln(\lambda)^2c + 162\lambda^{1/2-c}\ln(\lambda)^2c - 400\lambda^{2c+1}\ln(\lambda)^2 + 81\lambda^{c+1/2}\ln(\lambda)^2 \\ & + 81\lambda^{1/2-c}\ln(\lambda)^2 - 400\lambda^{1-2c}\ln(\lambda)^2 - 324\lambda^{c+1/2}\ln(\lambda) - 324\lambda^{1/2-c}\ln(\lambda) \end{aligned}$$

hat ein globales Maximum an der Stelle $c = 0$.

Beweis. Die Behauptung wird gezeigt, indem der Nachweis erbracht wird, dass die Funktion $s_{6,\lambda}^{(2)}(c)$ in c monoton fallend ist. Es ist somit zu zeigen, dass

$$\frac{\partial}{\partial c}s_{6,\lambda}^{(2)}(c) \leq 0.$$

Es gilt:

$$\frac{\partial}{\partial c}s_{6,\lambda}^{(2)}(c) = s_{6a,\lambda}^{(3)}(c) + s_{6b,\lambda}^{(3)}(c) \tag{A.31}$$

mit

$$\begin{aligned} s_{6a,\lambda}^{(3)}(c) &= 81\lambda^{c+1/2}\ln(\lambda)^2(\ln(\lambda)(1-2c) - 6) + 81\lambda^{1/2-c}\ln(\lambda)^2(\ln(\lambda)(-1-2c) + 6), \\ s_{6b,\lambda}^{(3)}(c) &= 800\ln(\lambda)^3(\lambda^{1-2c} - \lambda^{2c+1}). \end{aligned}$$

Weiter gilt:

$$\begin{aligned} s_{6a,\lambda}^{(3)}(c) &\leq -324\ln(\lambda)^2(\ln(\lambda) + 3\lambda - 3)c =: k_{6a,\lambda}^{(3)}(c), \\ s_{6b,\lambda}^{(3)}(c) &\leq -3200\lambda\ln(\lambda)^4c =: k_{6b,\lambda}^{(3)}(c), \end{aligned}$$

siehe Lemma A.26 und Lemma A.27.

Damit gilt, dass

$$\frac{\partial}{\partial c}s_{6,\lambda}^{(2)}(c) \leq k_{6a,\lambda}^{(3)}(c) + k_{6b,\lambda}^{(3)}(c).$$

Nun gilt, dass

$$k_{6a,\lambda}^{(3)}(c) + k_{6b,\lambda}^{(3)}(c) = -4\ln(\lambda)^2q_6(\lambda)c$$

mit

$$q_6(\lambda) = 800\lambda \ln(\lambda)^2 + 81 \ln(\lambda) + 243\lambda - 243.$$

Wegen Lemma A.22 und Lemma A.23 gilt

$$q_6(\lambda) \geq 800\lambda t_9(\lambda)^2 + 81t_{10}(\lambda) + 243\lambda - 243 =: \frac{1}{36}q_7(\lambda).$$

Es bleibt zu zeigen, dass $q_7(\lambda)$ für $\lambda \in [1/5, 1/2]$ positiv ist. Die Substitution

$$\lambda = (3x + 2)/10$$

liefert

$$q_7((3x + 2)/10) = \frac{1}{500}q_8(x)$$

mit

$$\begin{aligned} q_8(x) = & 428652x^7 - 2776032x^6 + 8114256x^5 - 14326956x^4 \\ & + 14304627x^3 - 4898862x^2 - 2190960x + 1467900. \end{aligned}$$

Nach der Substitution ist nun zu zeigen, dass $q_8(x)$ für $x \in [0, 1]$ positiv ist. Es gilt:

$$\begin{aligned} q_8(x) = & 122625x^7 + 9(1-x)(34003(1-x)^5 + 342x^5 + 1938x^4 \\ & + (490 - 490x)^2x^3 + (53 - 53x)^2x + q_9(x)) \end{aligned}$$

mit

$$q_9(x) = 278014x^3 - 178027x^2 - 229217x + 163100.$$

Weiter gilt:

$$278014x^3 - 178027x^2 - 229217x + 163100 \geq 1000q_{10}(x)$$

mit

$$q_{10}(x) = 278x^3 - 179x^2 - 230x + 163.$$

Nun bleibt zu zeigen, dass $q_{10}(x)$ positiv ist. Es gilt:

$$q_{10}(x) = 163(1-x)^3 + 441x^2 - 668x + 259.$$

Jetzt kann alternativ nachgewiesen werden, dass die Funktion $441x^2 - 668x + 259$ positiv ist. Es handelt sich um eine quadratische Funktion mit der Minimalstelle $334/441$. Die Minimalstelle hat den Funktionswert $2663/441$. Da dieser positiv ist, ist die Funktion stets positiv.

Somit ist $s_{6,\lambda}^{(2)}(c)$ monoton fallend in c . □

Lemma A.29. Seien $\lambda \in [1/5, 1/2]$ und $c \in [0, 1/2]$. Dann gilt:

$$s_6(\lambda, c) \geq m_6(\lambda, c).$$

Dabei ist

$$\begin{aligned} s_6(\lambda, c) &= 81\lambda^{c+1/2}(1-2c) + 81\lambda^{1/2-c}(1+2c) - 100\lambda^{1-2c} - 100\lambda^{2c+1}, \\ m_6(\lambda, c) &= c^2(-400\lambda\ln(\lambda)^2 + 81\lambda^{1/2}\ln(\lambda)^2 - 324\lambda^{1/2}\ln(\lambda)) \\ &\quad c(200\lambda\ln(\lambda)^2 - (81\lambda^{1/2}\ln(\lambda)^2)/2 + 162\lambda^{1/2}\ln(\lambda) - 200\lambda^2 + 400\lambda \\ &\quad - 324\lambda^{1/2} + 124) - 200\lambda + 162\lambda^{1/2}. \end{aligned}$$

Beweis. Definiere vorab $\frac{\partial^2}{\partial \lambda^2} s_6(\lambda, c) =: s_{6,\lambda}^{(2)}(c)$ und $\frac{\partial^2}{\partial \lambda^2} m_6(\lambda, c) =: m_{6,\lambda}^{(2)}(c)$.

Es gilt:

1. $s_6(\lambda, 0) = m_6(\lambda, 0)$,
2. $s_6(\lambda, 1/2) = m_6(\lambda, 1/2)$,
3. $s_{6,\lambda}^{(2)}(c) \leq s_{6,\lambda}^{(2)}(0) = -800\lambda\ln(\lambda)^2 + 162\lambda^{1/2}\ln(\lambda)^2 - 648\lambda^{1/2}\ln(\lambda) = m_{6,\lambda}^{(2)}(c)$.

Die Ungleichung gilt auf Grund von Lemma A.28. Daraufhin folgt mit Behauptung 5.11 schließlich, dass $s_6(\lambda, c) \geq m_6(\lambda, c)$. \square

Lemma A.30. Sei $\lambda \in [1/5, 1/2]$. Dann gilt:

$$n_4(\lambda) + n_5(\lambda) := \ln(\lambda)^2 p_{10}(\lambda) + \ln(\lambda) p_{11}(\lambda) + p_{12}(\lambda) \leq 0$$

mit

$$\begin{aligned} p_{10}(\lambda) &= \lambda^{1/2} \left(-\frac{81}{2}\lambda^2 - 200\lambda^{1/2} + \frac{81}{2} \right), \\ p_{11}(\lambda) &= 162\lambda^{5/2} - 162\lambda^{1/2}, \\ p_{12}(\lambda) &= -162\lambda^3 + 324\lambda^{5/2} - 362\lambda^2 + 562\lambda - 324\lambda^{1/2} - 38. \end{aligned}$$

Beweis. Zunächst gilt:

1. $p_{10}(\lambda) < 0$, denn

$$\begin{aligned} \lambda^{1/2}\left(-\frac{81}{2}\lambda^2 - 200\lambda^{1/2} + \frac{81}{2}\right) &\leq \lambda^{1/2}\left(-200\lambda^{1/2} + \frac{81}{2}\right) \\ &\leq \lambda^{1/2}\left(-200\left(\frac{1}{16}\right)^{1/2} + \frac{81}{2}\right) \\ &= -\frac{19}{2}\lambda^{1/2} \\ &< 0, \end{aligned}$$

2. $p_{11}(\lambda) < 0$, denn $162\lambda^{5/2} - 81\lambda - 81 = -81\lambda(1 - \lambda^{3/2}) - 81(1 - \lambda^{5/2}) < 0$.

Daraufhin gilt mit Lemma A.22 und Lemma A.23:

$$\begin{aligned} \ln(\lambda)^2 p_{10}(\lambda) + \ln(\lambda) p_{11}(\lambda) + p_{12}(\lambda) &\leq t_9(\lambda)^2 p_{10}(\lambda) + t_{10}(\lambda) p_{11}(\lambda) + p_{12}(\lambda) \\ &= \frac{1}{28800} q_{11}(\lambda). \end{aligned} \tag{A.32}$$

mit

$$\begin{aligned} q_{11}(\lambda) &= -158760000\lambda^{17/2} + 530712000\lambda^{15/2} - 784000000\lambda^7 - 615211200\lambda^{13/2} \\ &\quad + 2620800000\lambda^6 + 112554360\lambda^{11/2} - 3822080000\lambda^5 + 450009351\lambda^{9/2} \\ &\quad + 3176624000\lambda^4 - 534619926\lambda^{7/2} - 1604477200\lambda^3 + 309770568\lambda^{5/2} \\ &\quad + 456980000\lambda^2 - 108646434\lambda^{3/2} - 48134800\lambda + 14191281\lambda^{1/2} - 1094400. \end{aligned}$$

Es bleibt zu zeigen, dass $q_{11}(\lambda) \leq 0$ für $\lambda \in [1/5, 1/2]$. Eine Substitution mittels $\lambda = x^2$ liefert

$$q_{11}(x^2) = -q_{12}(x)$$

mit

$$\begin{aligned} q_{12}(x) &= 158760000x^{17} - 530712000x^{15} + 784000000x^{14} + 615211200x^{13} \\ &\quad - 2620800000x^{12} - 112554360x^{11} + 3822080000x^{10} - 450009351x^9 \\ &\quad - 3176624000x^8 + 534619926x^7 + 1604477200x^6 - 309770568x^5 \\ &\quad - 456980000x^4 + 108646434x^3 + 48134800x^2 - 14191281x + 1094400. \end{aligned} \tag{A.33}$$

Die Substitution führt dazu, dass ein Polynom mit ganzzahligen Exponenten vorliegt. Dies erleichtert die kommenden Abschätzungen. Es ist nun zu zeigen, dass $q_{12}(x)$ für $x \in [1/\sqrt{5}, 1/\sqrt{2}]$ positiv ist. Eine weitere Substitution mittels $x = 3/8 + 3/8z$ liefert

$$q_{12}(3/8 + 3/8z) = q_{13}(z)$$

mit

$$\begin{aligned}
q_{13}(z) = & 35594257426875z^{17} + 605102376256875z^{16} + 3994692662079000z^{15} \\
& + 14845397782635000z^{14} + 49510737847052100z^{13} + 150029618177427300z^{12} \\
& + 82985515786957320z^{11} - 1024934708294518680z^{10} - 1681082017986676878z^9 \\
& + 3486840582133363698z^8 + 5552587982262539880z^7 \\
& - 11614799979097109496z^6 - 9566770818645973404z^5 \\
& + 22837660993658712132z^4 + 531657633323475960z^3 \\
& - 13958812572336452584z^2 - 868427187448922453z + 3276396043659615355.
\end{aligned} \tag{A.34}$$

Ist das Polynom aus (A.34) für $z \in [0, 1]$ positiv, so ist auch das Polynom aus (A.33) positiv, da im Falle einer Resubstitution wegen $3/8 = 1/\sqrt{64/9} < 1/\sqrt{5}$ und $3/4 = 1/\sqrt{16/9} > 1/\sqrt{2}$ sogar ein Intervall vorliegt, das das Intervall $[1/\sqrt{5}, 1/\sqrt{2}]$ miteinschließt.

Nun gilt:

$$q_{13}(z) \geq 10^{17}q_{14}(z).$$

mit

$$\begin{aligned}
q_{14}(z) = & z^{12} - 11z^{10} - 17z^9 + 34z^8 + 55z^7 - 117z^6 \\
& - 96z^5 + 228z^4 + 53z^3 - 140z^2 - 9z + 32.
\end{aligned}$$

Weiter gilt mit Behauptung 5.13, dass

$$q_{14}(z) \geq 49z^7 - 117z^6 - 96z^5 + 228z^4 + 53z^3 - 140z^2 - 9z + 32 =: q_{15}(z).$$

Es gilt:

$$q_{15}(z) = (1 - z)q_{16}(z)$$

mit

$$q_{16}(z) = -49z^6 + 68z^5 + 164z^4 - 64z^3 - 117z^2 + 23z + 32.$$

Nun gilt:

$$q_{16}(z) = 49z^5(1 - z) + 19z^5 + z^2 + z + 2((2 - 2z)^2 + h(z)z)$$

mit

$$h(z) = 65z^3 + 32z^2 - 155z + 75.$$

Es bleibt zu zeigen, dass $h(z)$ für $z \in [0, 1]$ positiv ist. Definiere $h(z) + 155z =: f(z)$ mit $f(z) = 65z^3 + 32z^2 + 75$. Die Funktion $f(z)$ ist konvex. Ihre Tangente an der Stelle $z = 3/4$ stellt eine untere Schranke dar. Sie lautet

$$f(3/4) - \frac{3}{4} \frac{\partial}{\partial z} f(z)|_{z=3/4} + z \frac{\partial}{\partial z} f(z)|_{z=3/4} = 69/32 + 2523/16z =: g(z).$$

Somit gilt, dass $f(z) \geq g(z)$. Dann folgt für $z \in [0, 1]$, dass

$$h(z) = f(z) - 155z \geq g(z) - 155z = 69/32 + 43/16z \geq 0.$$

□

Lemma A.31. *Seien $\lambda \in [1/5, 1/2]$. Dann gilt:*

$$n_5(\lambda) := \ln(\lambda)^2 p_{13}(\lambda) + \ln(\lambda) p_{14}(\lambda) + p_{15}(\lambda) \leq 0,$$

wobei

$$\begin{aligned} p_{13}(\lambda) &= \frac{81}{2} \lambda^{5/2} + 200\lambda - \frac{81}{2} \lambda^{1/2}, \\ p_{14}(\lambda) &= -162\lambda^{5/2} + 162\lambda^{1/2}, \\ p_{15}(\lambda) &= 324\lambda^{5/2} - 524\lambda^2 + 400\lambda - 324\lambda^{1/2} + 124. \end{aligned}$$

Beweis. Zunächst wird die Konkavität von $n_5(\lambda)$ gezeigt. Es gilt:

$$\frac{\partial^2}{\partial \lambda^2} n_5(\lambda) = \ln(\lambda)^2 p_{16}(\lambda) + \ln(\lambda) p_{17}(\lambda) + p_{18}(\lambda)$$

mit

$$\begin{aligned} p_{16}(\lambda) &= \frac{1215}{8} \lambda^{1/2} + \frac{81}{8} \lambda^{-3/2}, \\ p_{17}(\lambda) &= \frac{-567\lambda^2 + 800\lambda^{1/2} - 81}{2\lambda^{3/2}}, \\ p_{18}(\lambda) &= 648\lambda^{1/2} + 400/\lambda - 1048. \end{aligned}$$

Dabei gilt:

1. $p_{16}(\lambda) \geq 0$,
2. $p_{17}(\lambda) \geq 0$.

Die zweite Ungleichung gilt auf Grund der folgenden Überlegungen. Die Funktion $f(\lambda) = -567\lambda^2 + 800\lambda - 81$ ist eine untere Schranke des Zählers von $p_{17}(\lambda)$. Am linken Rand

des zu betrachtenden Intervalls ist die Schranke wegen $f(1/5) = 1408/25$ positiv. Da sie außerdem wegen

$$\frac{\partial}{\partial \lambda} f(\lambda) = 800 - 1134\lambda \geq 800 - 1134(1/2) = 233 > 0$$

für $\lambda \in [1/5, 1/2]$ monoton wächst, ist sie in dem Intervall stets positiv. Damit ist auch die Funktion $p_{17}(\lambda)$ positiv.

Daraufhin gilt mit Lemma A.24 und Lemma A.25:

$$\begin{aligned} \ln(\lambda)^2 p_{16}(\lambda) + \ln(\lambda) p_{17}(\lambda) + p_{18}(\lambda) &\leq t_{11}(\lambda)^2 p_{16}(\lambda) + t_{12}(\lambda) p_{17}(\lambda) + p_{18}(\lambda) \\ &= q_{17}(\lambda)/q_{18}(\lambda) \end{aligned}$$

mit

$$\begin{aligned} q_{17}(\lambda) &= 229502700\lambda^4 - 393876000\lambda^3 - 54080000\lambda^{5/2} + 281805165\lambda^2 \\ &\quad + 20550400\lambda^{3/2} - 35019360\lambda - 50700000\lambda^{1/2} + 17389791, \\ q_{18}(\lambda) &= 135200\lambda^{3/2}. \end{aligned}$$

Offensichtlich ist $q_{18}(\lambda) > 0$. Es bleibt also zu zeigen, dass $q_{17}(\lambda) \leq 0$. Die Substitution $\lambda = x^2$ liefert $q_{17}(x^2) = q_{19}(x)$ mit

$$\begin{aligned} q_{19}(x) &= 229502700x^8 - 393876000x^6 - 54080000x^5 + 281805165x^4 \\ &\quad + 20550400x^3 - 35019360x^2 - 50700000x + 17389791. \end{aligned}$$

Wie im Beweis von Lemma A.30 wird aus den gleichen Gründen hier ebenfalls die Substitution $x = 3/8 + 3/8z$ durchgeführt. Diese liefert

$$q_{19}(3/8 + 3/8z) = -\frac{q_{20}(z)}{4194304}$$

mit

$$\begin{aligned} q_{20}(z) &= -376441803675z^8 - 3011534429400z^7 - 5946200838900z^6 \\ &\quad + 8166381298200z^5 + 27598092696990z^4 - 10417914622440z^3 \\ &\quad - 58032068005620z^2 + 46886278496040z + 5441758743141. \end{aligned}$$

Es genügt nun zu zeigen, dass die Funktion $q_{20}(z)$ für $z \in [0, 1]$ positiv ist. Es gilt:

$$q_{20} \geq 10^{11} q_{21}(z)$$

mit

$$\begin{aligned}
q_{21}(z) &= -4z^8 - 31z^7 - 60z^6 + 81z^5 + 275z^4 - 105z^3 - 581z^2 + 468z + 54 \\
&= (1-z)z((1-z)((1-z)(4z^4 + 43z^3 + 177z^2 + 325z + 212) + 57) + 156) \\
&\quad + 43z + 54. \tag{A.35}
\end{aligned}$$

Es ist zu erkennen, dass die Funktion $q_{21}(z)$ für $z \in [0, 1]$ positiv ist. Somit ist die Konkavität von $n_5(\lambda)$ in dem Intervall $[1/5, 1/2]$ gezeigt.

Auf Grund der Konkavität von $n_5(\lambda)$ ist die erste Ableitung von $n_5(\lambda)$ monoton fallend. In der nächsten Überlegung gehen die Abschätzungen aus Lemma 5.17 und die Abschätzung $\sqrt{2} < \frac{10}{7}$ ein. Letztere gilt wegen $(\sqrt{2})^2 = 2 < \frac{100}{49} = (\frac{10}{7})^2$. Wegen

$$\begin{aligned}
\frac{\partial}{\partial \lambda} n_5(\lambda)|_{\lambda=1/2} &= \frac{1}{8\sqrt{2}}((1600\sqrt{2} + 81)\ln(2)^2 + (1296 - 3200\sqrt{2})\ln(2) - 992\sqrt{2} + 2592) \\
&\geq \frac{1}{8\sqrt{2}}((1600\sqrt{2} + 81)(11/16)^2 + (1296 - 3200\sqrt{2})(7/10) \\
&\quad - 992\sqrt{2} + 2592) \\
&= (4527981 - 3168960\sqrt{2})/(10240\sqrt{2}) \\
&\geq (4527981 - 3168960(10/7))/(10240\sqrt{2}) \\
&= 6267/(71680\sqrt{2}) \\
&> 0
\end{aligned}$$

und der Monotonie der ersten Ableitung gilt, dass die erste Ableitung nicht negativ ist. Da die Funktion $n_5(\lambda)$ zudem am rechten Rand wegen

$$\begin{aligned}
n_5(1/2) &= \frac{1}{8\sqrt{2}}((800\sqrt{2} - 243)\ln(2)^2 - 972\ln(2) + 1544\sqrt{2} - 1944) \\
&\leq \frac{1}{8\sqrt{2}}(800\sqrt{2} - 243)(25/36)^2 - 972(9/13) + 1544\sqrt{2} - 1944) \\
&= (32513312\sqrt{2} - 18805284)/16848 \\
&\leq (32513312(17/12) - 18805284)/16848 \\
&= -11309/50544 \\
&< 0
\end{aligned}$$

negativ ist, gilt nun insgesamt, dass sie für $\lambda \in [1/5, 1/2]$ negativ ist. Hier sind die Abschätzungen aus Lemma 5.15 und die Abschätzung $\sqrt{2} < \frac{17}{12}$ herangezogen worden. Letztere gilt wegen $(\sqrt{2})^2 = 2 < \frac{289}{144} = (\frac{17}{12})^2$. \square

Lemma A.32. Seien $\lambda \in [1/2, 1]$ und $c \in [0, 1/2]$. Dann gilt:

$$z_2(\lambda, c) = n_1(\lambda)c^2 + n_2(\lambda)c + n_3(\lambda)$$

ist monoton fallend in c .

Dabei ist

$$\begin{aligned} n_1(\lambda) &= \ln(\lambda)^2(-81\lambda^{5/2} - 400\lambda + 81) + \ln(\lambda)(324\lambda^{5/2} - 162\lambda - 162) \\ &\quad - 162(\lambda - 1)^2(\lambda + 1), \\ n_2(\lambda) &= \ln(\lambda)^2(81\lambda^{5/2})/2 + 200\lambda - 81/2 + \ln(\lambda)(-162\lambda^{5/2} + 81\lambda + 81) \\ &\quad + 324\lambda^{5/2} - 524\lambda^2 + 400\lambda - 324\lambda^{1/2} + 124, \\ n_3(\lambda) &= \frac{81}{2}\lambda^3 - 162\lambda^{5/2} + \frac{443}{2}\lambda^2 - \frac{481}{2}\lambda + 162\lambda^{1/2} - 43/2. \end{aligned}$$

Beweis. Es ist zu zeigen, dass

$$\frac{\partial}{\partial c} z_2(\lambda, c) = 2n_1(\lambda)c + n_2(\lambda) \leq 0.$$

Die Funktion $2n_1(\lambda)c + n_2(\lambda)$ ist monoton in c . Auf dem Rand $c = 0$ und $c = 1/2$ ist sie wegen Lemma A.20 und Lemma A.21 jeweils nicht positiv, so dass sie auf Grund der Monotonie sogar für $c \in [0, 1/2]$ nicht positiv ist. \square

Lemma A.33. Seien $\lambda \in [1/5, 1/2]$ und $c \in [0, 1/2]$. Dann gilt:

$$z_3(\lambda, c) = n_4(\lambda)c^2 + n_5(\lambda)c + n_6(\lambda)$$

ist monoton fallend in c .

Dabei ist

$$\begin{aligned} n_4(\lambda) &= -81\lambda^{5/2}\ln(\lambda)^2 - 400\lambda\ln(\lambda)^2 + 81\lambda^{1/2}\ln(\lambda)^2 + 324\lambda^{5/2}\ln(\lambda) \\ &\quad - 324\lambda^{1/2}\ln(\lambda) - 162\lambda^3 + 162\lambda^2 + 162\lambda - 162, \\ n_5(\lambda) &= \frac{81}{2}\lambda^{5/2}\ln(\lambda)^2 + 200\lambda\ln(\lambda)^2 - \frac{81}{2}\lambda^{1/2}\ln(\lambda)^2 - 162\lambda^{5/2}\ln(\lambda) \\ &\quad + 162\lambda^{1/2}\ln(\lambda) + 324\lambda^{5/2} - 524\lambda^2 + 400\lambda - 324\lambda^{1/2} + 124, \\ n_6(\lambda) &= \frac{81}{2}\lambda^3 - 162\lambda^{5/2} + \frac{443}{2}\lambda^2 - \frac{481}{2}\lambda + 162\lambda^{1/2} - \frac{43}{2}. \end{aligned}$$

Beweis. Es ist zu zeigen, dass

$$\frac{\partial}{\partial c} z_3(\lambda, c) = 2n_4(\lambda)c + n_5(\lambda) \leq 0.$$

Die Funktion $2n_4(\lambda)c + n_5(\lambda)$ ist monoton in c . Auf dem Rand $c = 0$ und $c = 1/2$ ist sie wegen Lemma A.30 und Lemma A.31 jeweils nicht positiv, so dass sie auf Grund der Monotonie sogar für $c \in [0, 1/2]$ nicht positiv ist. \square

Lemma A.34. *Sei $x \in (0, 1)$. Dann gilt:*

$$t_{11}(x) \geq 0$$

mit $t_{11}(x) = 750446x^4 - 3819569x^3 + 7176147x^2 - 5001470x + 1242690$.

Beweis. Wegen

$$\begin{aligned} \frac{\partial^2}{\partial x^2} t_{11}(x) &= 9005352x^2 - 22917414x + 14352294 \\ &\geq 9005352x^2 - 22917414x + 13912062 \\ &= 6(1-x)(2318677 - 1500892x) \\ &\geq 0 \end{aligned}$$

ist $t_{11}(x)$ konvex. Die Tangente an der Stelle $x = 11/20$ lautet

$$s_1(x) = \frac{10650268133}{80000} - \frac{149090709}{2000}x.$$

Diese weist unter anderem die Funktionswerte

$$s_1(0) = \frac{10650268133}{80000}$$

und

$$s_1(1) = \frac{4686639773}{80000}$$

auf. Dies bedeutet, dass die Tangente in dem Intervall $[0, 1]$ positiv ist. Da die Tangente eine untere Schranke der konvexen Funktion $t_{11}(x)$ darstellt, ist die Funktion ebenfalls in dem betrachteten Intervall positiv. \square

Lemma A.35. *Sei $x \in (0, 1)$. Dann gilt:*

$$\ln(x)^2 + \frac{f_4(x)}{f_3(x)}$$

ist monoton wachsend in x . Dabei ist

$$\begin{aligned} f_3(x) &= 2x^2(x^9 - 3x^8 + 8x^7 - 8x^6 + 6x^5 - 18x^4 + 8x^3 - 8x^2 + x - 3), \\ f_4(x) &= x^9 + x^8 - 2x^5 - 2x^4 + x + 1. \end{aligned}$$

Beweis. Definiere $\ln(x)^2 + \frac{f_4(x)}{f_3(x)} =: h_5(x)$. Es wird gezeigt, dass

$$\frac{\partial}{\partial x} h_5(x) = \frac{2\ln(x)}{x} - \frac{t_3(x)}{t_4(x)} \quad (\text{A.36})$$

mit

$$\begin{aligned} t_3(x) &= (x-1)(x+1)^2(x^2+1)(x^{13} - x^{12} - 2x^{11} + 10x^{10} - 21x^9 + 53x^8 \\ &\quad - 68x^7 + 68x^6 - 41x^5 + 73x^4 - 26x^3 + 18x^2 - 3x + 3), \\ t_4(x) &= x^3(x^9 - 3x^8 + 8x^7 - 8x^6 + 6x^5 - 18x^4 + 8x^3 - 8x^2 + x - 3)^2 \end{aligned}$$

positiv ist. In einem ersten Schritt wird die Ableitung von $h_5(x)$ mit x multipliziert. Daraufhin ist alternativ zu zeigen, dass

$$t_5(x) := 2\ln(x) - \frac{t_3(x)}{t_4(x)}x \geq 0.$$

Es gilt $t_5(1) = 0$. Damit die Behauptung gilt, genügt es nun zu zeigen, dass $t_5(x)$ monoton fallend ist. Es wird gezeigt, dass

$$\frac{\partial}{\partial x} t_5(x) = \frac{2}{x} - \frac{t_6(x)}{t_7(x)} \leq 0.$$

Diese Ableitung kann wieder mit x multipliziert werden. Äquivalenzumformungen liefern dann die zu zeigende Ungleichung

$$\frac{t_6(x)x}{t_7(x)} \geq 2$$

mit

$$\begin{aligned} t_6(x) &= -2(x+1)(x^{26} - x^{25} - 13x^{24} + 58x^{23} - 148x^{22} + 412x^{21} - 768x^{20} + 1038x^{19} \\ &\quad - 563x^{18} - 805x^{17} + 3643x^{16} - 6604x^{15} + 10096x^{14} - 12448x^{13} \\ &\quad + 13320x^{12} - 14132x^{11} + 12635x^{10} - 9419x^9 + 8329x^8 - 4750x^7 + 2980x^6 \\ &\quad - 1276x^5 + 568x^4 - 186x^3 + 87x^2 - 15x + 9), \\ t_7(x) &= x^3(x^9 - 3x^8 + 8x^7 - 8x^6 + 6x^5 - 18x^4 + 8x^3 - 8x^2 + x - 3)^3. \end{aligned}$$

An der Stelle $x = 1$ hat die die Funktion $\frac{t_6(x)x}{t_7(x)}$ den Funktionswert 2. Damit genügt es zu zeigen, dass die linke Seite monoton fallend ist. Die Ableitung lautet

$$\frac{\partial}{\partial x} \left(\frac{t_6(x)x}{t_7(x)} \right) = \frac{-2(1-x)t_8(x)}{x^3(t_9(x))^4}$$

mit

$$\begin{aligned}
t_8(x) &= 2x^{35} + 5x^{34} - 83x^{33} + 240x^{32} - 406x^{31} + 1566x^{30} - 2466x^{29} + 746x^{28} \\
&\quad + 12050x^{27} - 38350x^{26} + 104978x^{25} - 154230x^{24} + 227698x^{23} \\
&\quad - 172410x^{22} + 163014x^{21} + 35730x^{20} - 76882x^{19} + 316604x^{18} \\
&\quad - 331348x^{17} + 613874x^{16} - 308482x^{15} + 771642x^{14} - 319238x^{13} \\
&\quad + 466366x^{12} - 328314x^{11} + 179646x^{10} - 161890x^9 + 57566x^8 \\
&\quad - 24922x^7 + 13858x^6 - 1822x^5 + 3654x^4 - 456x^3 + 591x^2 - 9x + 54, \\
t_9(x) &= x^9 - 3x^8 + 8x^7 - 8x^6 + 6x^5 - 18x^4 + 8x^3 - 8x^2 + x - 3.
\end{aligned}$$

Es bleibt schließlich nur noch zu zeigen, dass die Funktion $t_8(x)$ positiv ist. Für $t_8(x)$ existiert die Darstellung

$$\begin{aligned}
t_8(x) &= t_{10}(x) + x(x(x(x(x(x(x(x(x(xt_{11}(x) + 2540281(1-x)^5) + 2848911(1-x)^6) \\
&\quad + 2313864(1-x)^7) + 1476080(1-x)^8) + 771105(1-x)^9) + 330455(1-x)^{10}) \\
&\quad + 114714(1-x)^{11}) + 30978(1-x)^{12}) + 6135(1-x)^{13}) + 801(1-x)^{14}) \\
&\quad + 54(1-x)^{15}
\end{aligned}$$

mit

$$\begin{aligned}
t_{10}(x) &= 2x^{35} + 5x^{34} - 83x^{33} + 240x^{32} - 406x^{31} + 1566x^{30} - 2466x^{29} \\
&\quad + 746x^{28} + 12050x^{27} - 38350x^{26} + 104978x^{25} - 154230x^{24} \\
&\quad + 227698x^{23} - 172410x^{22} + 163014x^{21} + 35730x^{20} - 76882x^{19} \\
&\quad + 316604x^{18} - 331348x^{17} + 613874x^{16}, \\
t_{11}(x) &= 750446x^4 - 3819569x^3 + 7176147x^2 - 5001470x + 1242690.
\end{aligned}$$

Dabei ist die Funktion $t_{10}(x)$ wegen Behauptung 5.13 positiv. Die Funktion $t_{11}(x)$ ist wegen Lemma A.34 ebenfalls positiv. Nun ist zu erkennen, dass $t_8(x)$ positiv ist. Damit gilt die Behauptung. \square

B. Tabellen

Tabelle B.1: Maximale absolute Abweichungen zwischen den Einträgen der empirischen Kovarianzmatrix und den Einträgen der Inversen der Fisher-Informationsmatrix. Es wird der Schätzvektor für die Regressionsparameter β_0 und β_1 betrachtet. Die empirischen Kovarianzmatrizen sind mit dem Faktor n multipliziert worden, damit in den Tabellen B.1 bis B.4 ein Bezug zur selben Informationsmatrix hergestellt werden kann. Es liegt der Spezialfall $n = 2$ mit 10000 Wiederholungen vor.

	d					
λ	0.1	0.2	0.3	0.4	0.5	0.6
0	0.005	0.029	0.020	0.025	0.007	0.041
0.0001	0.009	0.010	0.015	0.047	0.005	0.038
0.001	0.005	0.027	0.022	0.025	0.007	0.011
0.01	0.019	0.023	0.009	0.050	0.004	0.022
0.1	0.031	0.017	0.019	0.052	0.056	0.043
0.25	0.024	0.033	0.021	0.034	0.023	0.023
0.5	0.015	0.011	0.018	0.030	0.013	0.032
0.9	0.019	0.009	0.010	0.010	0.023	0.023

Tabelle B.2: Wie Tabelle B.1. Hier liegt jedoch der Spezialfall $n = 5$ vor.

	d					
λ	0.1	0.2	0.3	0.4	0.5	0.6
0	0.012	0.016	0.011	0.019	0.015	0.018
0.0001	0.016	0.019	0.040	0.013	0.015	0.016
0.001	0.017	0.021	0.012	0.019	0.048	0.003
0.01	0.018	0.025	0.036	0.008	0.022	0.010
0.1	0.013	0.046	0.014	0.018	0.015	0.013
0.25	0.009	0.007	0.008	0.033	0.020	0.027
0.5	0.025	0.030	0.022	0.008	0.011	0.006
0.9	0.006	0.008	0.013	0.008	0.017	0.003

Tabelle B.3: Wie Tabelle B.1. Hier liegt jedoch der Spezialfall $n = 10$ vor.

	d					
λ	0.1	0.2	0.3	0.4	0.5	0.6
0	0.011	0.007	0.016	0.015	0.026	0.002
0.0001	0.009	0.009	0.016	0.015	0.026	0.002
0.001	0.023	0.020	0.023	0.028	0.017	0.020
0.01	0.009	0.011	0.019	0.017	0.027	0.002
0.1	0.030	0.020	0.025	0.024	0.015	0.023
0.25	0.010	0.005	0.015	0.020	0.023	0.016
0.5	0.017	0.011	0.026	0.013	0.017	0.019
0.9	0.006	0.030	0.002	0.004	0.018	0.011

Tabelle B.4: *Wie Tabelle B.1. Hier liegt jedoch der Spezialfall $n = 50$ vor.*

λ	d					
	0.1	0.2	0.3	0.4	0.5	0.6
0	0.013	0.017	0.013	0.034	0.018	0.015
0.0001	0.021	0.006	0.015	0.019	0.012	0.046
0.001	0.019	0.022	0.015	0.019	0.012	0.046
0.01	0.020	0.023	0.016	0.015	0.017	0.017
0.1	0.020	0.020	0.018	0.031	0.016	0.017
0.25	0.017	0.016	0.018	0.026	0.014	0.018
0.5	0.009	0.017	0.015	0.013	0.015	0.022
0.9	0.012	0.012	0.022	0.009	0.012	0.005

Tabelle B.5: *Sämtliche Einträge der Inversen der Fisher-Informationsmatrizen, die in den Tabellen B.1 bis B.4 herangezogen werden.*

Konstellation	$\mathcal{I}_{1,1}^{-1}$	$\mathcal{I}_{2,2}^{-1}$	$\mathcal{I}_{1,2}^{-1}$	Konstellation	$\mathcal{I}_{1,1}^{-1}$	$\mathcal{I}_{2,2}^{-1}$	$\mathcal{I}_{1,2}^{-1}$
$\lambda = 0, d = 0.1$	0.555	1.648	-0.604	$\lambda = 0.1, d = 0.1$	0.965	1.793	-0.884
$\lambda = 0, d = 0.2$	0.619	1.786	-0.714	$\lambda = 0.1, d = 0.2$	0.944	1.793	-0.880
$\lambda = 0, d = 0.3$	0.690	1.899	-0.823	$\lambda = 0.1, d = 0.3$	0.935	1.796	-0.883
$\lambda = 0, d = 0.4$	0.763	1.974	-0.921	$\lambda = 0.1, d = 0.4$	0.934	1.799	-0.891
$\lambda = 0, d = 0.5$	0.833	2.000	-1.000	$\lambda = 0.1, d = 0.5$	0.940	1.800	-0.900
$\lambda = 0, d = 0.6$	0.895	1.974	-1.053	$\lambda = 0.1, d = 0.6$	0.951	1.799	-0.908
$\lambda = 0.0001, d = 0.1$	0.772	1.853	-0.817	$\lambda = 0.25, d = 0.1$	0.988	1.499	-0.746
$\lambda = 0.0001, d = 0.2$	0.711	1.863	-0.801	$\lambda = 0.25, d = 0.2$	0.980	1.499	-0.745
$\lambda = 0.0001, d = 0.3$	0.728	1.923	-0.855	$\lambda = 0.25, d = 0.3$	0.975	1.499	-0.746
$\lambda = 0.0001, d = 0.4$	0.778	1.978	-0.931	$\lambda = 0.25, d = 0.4$	0.974	1.500	-0.748
$\lambda = 0.0001, d = 0.5$	0.838	2.000	-1.000	$\lambda = 0.25, d = 0.5$	0.975	1.500	-0.750
$\lambda = 0.0001, d = 0.6$	0.894	1.978	-1.048	$\lambda = 0.25, d = 0.6$	0.978	1.500	-0.752
$\lambda = 0.001, d = 0.1$	0.827	1.900	-0.869	$\lambda = 0.5, d = 0.1$	0.998	1.000	-0.500
$\lambda = 0.001, d = 0.2$	0.764	1.901	-0.848	$\lambda = 0.5, d = 0.2$	0.996	1.000	-0.499
$\lambda = 0.001, d = 0.3$	0.764	1.941	-0.883	$\lambda = 0.5, d = 0.3$	0.995	1.000	-0.500
$\lambda = 0.001, d = 0.4$	0.798	1.982	-0.942	$\lambda = 0.5, d = 0.4$	0.994	1.000	-0.500
$\lambda = 0.001, d = 0.5$	0.847	1.998	-0.999	$\lambda = 0.5, d = 0.5$	0.995	1.000	-0.500
$\lambda = 0.001, d = 0.6$	0.897	1.982	-1.040	$\lambda = 0.5, d = 0.6$	0.995	1.000	-0.500
$\lambda = 0.01, d = 0.1$	0.893	1.934	-0.920	$\lambda = 0.9, d = 0.1$	1.000	0.200	-0.100
$\lambda = 0.01, d = 0.2$	0.843	1.933	-0.904	$\lambda = 0.9, d = 0.2$	1.000	0.200	-0.100
$\lambda = 0.01, d = 0.3$	0.831	1.952	-0.921	$\lambda = 0.9, d = 0.3$	1.000	0.200	-0.100
$\lambda = 0.01, d = 0.4$	0.845	1.972	-0.954	$\lambda = 0.9, d = 0.4$	1.000	0.200	-0.100
$\lambda = 0.01, d = 0.5$	0.874	1.980	-0.990	$\lambda = 0.9, d = 0.5$	1.000	0.200	-0.100
$\lambda = 0.01, d = 0.6$	0.908	1.972	-1.017	$\lambda = 0.9, d = 0.6$	1.000	0.200	-0.100

C. Programmcode

```
# R-Code für Tabelle 6.2:

# Größen, die für jede Tabelle anzupassen sind:
  # Anzahl Versuchseinheiten:
    n <- 2
  # Anzahl Wiederholungen:
    nn <- 10000
# Optionale Größen:
  # Anzahl der Wiederholungen des 4. Schritts:
    k.runs <- 10
  # m=9 Startwerte für die Funktion 'nlminb':
    start.werte <- seq(0.1,0.9,0.1)

##### Fall 1: d aus (0,1)
# Optionale Größen:
  # Menge der wahren Korrelationen:
    a.alle <- c(0.0001,0.001,0.01,0.1,0.25,0.5,0.9)
  # Menge der Versuchspunkte:
    d.alle <- c(0.1,0.2,0.3,0.4,0.5,0.6)
library(mvtnorm)          # zur Erzeugung von Zufallszahlen
# Kovarianzmatrix des Fehlervektors:
  Sigma <- function(a,d){matrix(c(1,a^d,a,a^d,1,a^(1-d),a,a^(1-d),1),ncol=3)}
# Inverse der obigen Kovarianzmatrix:
  Sinv <- function(a,d){
  matrix(c(-1/(a^(2*d)-1),a^d/(a^(2*d)-1),0,a^d/(a^(2*d)-1),
  (a^(2*d)*(a^2-1))/(a^(4*d)+a^2+a^(2*d)*(-a^2-1)),
  -a^(d+1)/(a^(2*d)-a^2),0,-a^(d+1)/(a^(2*d)-a^2),
  a^(2*d)/(a^(2*d)-a^2)),ncol=3)}
# Log-Likelihood-Funktion aus (6.16):
  l2.s2 <- function(a){
  -3*n/2*log(t(e.dach)%*%kronecker(diag(1,n), Sinv(a,d.))%*%e.dach)
  -n/2*log(det(Sigma(a,d.)))}
  l2.vec <- Vectorize(l2.s2) # in Log-Likelihood-Funktion
                           # können Vektoren eingesetzt werden
# Log-Likelihood-Funktion aus (6.15):
```

```

12.s2.0 <- function(a){
-3*n/2*log(t(e.dach)%*%e.dach) - n/2*log(det(Sigma(a,d.)))}

# Vorzeichenänderung der Log-Likelihood-Funktion für 'nlminb':
12.b <- function(a){-12.s2(a)}
# Leere Matrix für die Einträge von Tabelle 6.2:
D.Krit <- matrix(0,length(a.alle),length(d.alle))

for(u in 1:length(a.alle)){
a. <- a.alle[u]
for(o in 1:length(d.alle)){
d. <- d.alle[o]
# Versuchsplan der 'n' Versuchseinheiten
des <- rep(c(0,d.,1),n)
# zugehörige Designmatrix
X2 <- cbind(1,des)
# leere Matrix für die finalen ML-Schätzungen von "beta"
GLS <- matrix(0,nn,2)
# Vektor für die 'nn' erstmaligen Schätzungen von "lambda" im 2. Schritt
a.ML.start <- numeric(nn)
# Vektor für die Schätzungen von "lambda" im 2. Schritt
a.ML.verlauf <- numeric(k.runs)
# Schätzwerte für "lambda" auf Grund der verschiedenen Startwerte
a.ML.hilf <- numeric(length(start.werte))
# zugehörige Werte der Likelihood-Funktion
log.ML.hilf <- numeric(length(start.werte))
for(i in 1:nn){
# k=1 (erste Schätzung mittels gewöhnlichem KQ-Schätzer):
y <- rmvnorm(n,mean=c(0,0,0),sigma=Sigma(a.,d.)) # Zufallszahlen
e.dach <- as.vector(t(y))-X2%*%solve(t(X2)%*%X2)%*%t(X2)%*%as.vector(t(y))
for(j in 1:9){
# Maximalstellen:
a.ML.hilf[j] <- nlminb(start.werte[j],12.b,lower=0.1^6,upper=1-0.1^6)$par
# zugehörige Likelihood-Funktion:
log.ML.hilf[j] <-
12.s2(nlminb(start.werte[j],12.b,lower=0.1^6,upper=1-0.1^6)$par)}
# Beste Maximalstelle:

```

```

    welcher <- min(which.max(l2.vec(a.ML.hilf)))
# Likelihood-Funktion der besten Maximalstelle:
    a.1 <- a.ML.hilf[welcher]
# Spezialfall: lam=0
l2.0 <- -3*n/2*log(t(e.dach)%%kronecker(diag(1,n),diag(1,3))%%e.dach)
-n/2*log(det(diag(1,3)))
# Abfrage: Liefert lambda=0 eine bessere Maximalstelle?
    if(l2.0 > log.ML.hilf[welcher]){a.1 <- 0}
a.ML.start[i] <- a.1
a.ML.verlauf[1] <- a.ML.start[i]

# k>1 (weitere Schätzungen mittels verallgemeinertem KQ-Schätzer):
for(k in 2:k.runs){
# Falls im ersten Schritt lambda=0 ist -> abbrechen!
    if(a.1==0){ }
# Sonst:
    if(a.1>0){
Sigma.inv <- kronecker(diag(1,n), Sinv(a.ML.verlauf[k-1],d.))
e.dach <- as.vector(t(y)) - X2%%solve(t(X2)%%Sigma.inv%%X2)
%%t(X2)%%Sigma.inv%%as.vector(t(y))
for(j in 1:9){
a.ML.hilf[j] <- nlminb(start.werte[j],l2.b,lower=0.1^6,upper=1-0.1^6)$par
log.ML.hilf[j]
<- l2.s2(nlminb(start.werte[j],l2.b,lower=0.1^6,upper=1-0.1^6)$par)}
welcher <- min(which.max(l2.vec(a.ML.hilf)))
a.1 <- a.ML.hilf[welcher]
# Spezialfall: lam=0
l2.0 <- -3*n/2*log(t(e.dach)%%kronecker(diag(1,n), diag(1,3))%%e.dach)
-n/2*log(det(diag(1,3)))
if(l2.0 > log.ML.hilf[welcher]){a.1 <- 0}
}
a.ML.verlauf[k] <- a.1
}

# Finale Schätzwerte des zweistufigen Verfahrens:
if(a.ML.verlauf[k]==0){GLS[i,] <- solve(t(X2)%%X2)%%t(X2)%%as.vector(t(y))}
if(a.ML.verlauf[k]>0){Sigma.inv

```

```

        <- kronecker(diag(1,n),Sinv(a.ML.verlauf[k],d.))
        GLS[i,] <- solve(t(X2)%*%Sigma.inv%*%X2)%*%t(X2)
        %*%Sigma.inv%*%as.vector(t(y))}
    }
Kov <- cov(GLS)          # Empirische Kovarianzmatrix der finalen Schätzwerte
D.Krit[u,o] <- det(Kov) # Determinante der Kovarianzmatrix
}
}
round(D.Krit*n^2,3)     # Einträge für Tabelle 6.2

##### Fall 2: d = 0 oder 1
# Optionale Größen:
  # Menge der wahren Korrelationen:
    a.alle <- c(0.0001,0.001,0.01,0.1,0.25,0.5,0.9)
  # Menge der Versuchspunkte:
    d.alle <- c(0,1)

# Kovarianzmatrix des Fehlervektors:
  Sigma <- function(a,d){matrix(c(1,a,a,1),ncol=2)}
# Inverse der obigen Kovarianzmatrix:
  Sinv <- function(a,d){
    1/(1-a^2)*matrix(c(1,-a,-a,1),ncol=2)}
# Log-Likelihood-Funktion aus (6.16):
  l2.s2 <- function(a){
    -n*log(2*pi)-n*log(t(e.dach)%*%kronecker(diag(1,n), Sinv(a,d.))%*%e.dach)
    +n*log(2*n) - n/2*log(det(Sigma(a,d.)))-n}
  l2.vec <- Vectorize(l2.s2) # in Log-Likelihood-Funktion
                          # können Vektoren eingesetzt werden
# Vorzeichenänderung der Log-Likelihood-Funktion für 'nlminb':
  l2.b <- function(a){-l2.s2(a)}
# Leere Matrix für die Einträge von Tabelle 6.2:
  D.Krit <- matrix(0,length(a.alle),length(d.alle))

for(u in 1:length(a.alle)){
  a. <- a.alle[u]
  for(o in 1:length(d.alle)){

```

```

d. <- d.alle[o]
# Versuchsplan der 'n' Versuchseinheiten
  des <- rep(c(0,1),n)
# zugehörige Designmatrix
  X2 <- cbind(1,des)
# leere Matrix für die finalen ML-Schätzungen von "beta"
  GLS <- matrix(0,nn,2)
# Vektor für die 'nn' erstmaligen Schätzungen von "lambda" im 2. Schritt
  a.ML.start <- numeric(nn)
# Vektor für die Schätzungen von "lambda" im 2. Schritt
  a.ML.verlauf <- numeric(k.runs)
# Schätzwerte für "lambda" auf Grund der verschiedenen Startwerte
  a.ML.hilf <- numeric(length(start.werte))
# zugehörige Werte der Likelihood-Funktion
  log.ML.hilf <- numeric(length(start.werte))
for(i in 1:nn){
# k=1 (erste Schätzung mittels gewöhnlichem KQ-Schätzer):
  y <- rmvnorm(n,mean=c(0,0),sigma=Sigma(a.,d.)) # Zufallszahlen
  e.dach <- as.vector(t(y))-X2%*%solve(t(X2)%*%X2)%*%t(X2)%*%as.vector(t(y))
  for(j in 1:9){
# Maximalstellen:
a.ML.hilf[j] <- nlminb(start.werte[j],l2.b,lower=0.1^6,upper=0.999999)$par
# zugehörige Likelihood-Funktion:
log.ML.hilf[j] <-
l2.s2(nlminb(start.werte[j],l2.b,lower=0.000001,upper=0.999999)$par)}
# Beste Maximalstelle:
  welcher <- min(which.max(l2.vec(a.ML.hilf)))
# Likelihood-Funktion der besten Maximalstelle:
  a.1 <- a.ML.hilf[welcher]
# Spezialfall: lam=0
l2.0 <- -3*n/2*log(2*pi)-3*n/2*log((t(e.dach)%*%e.dach)/(3*n))-3*n/2
# Abfrage: Liefert lambda=0 eine bessere Maximalstelle?
  if(l2.0 > log.ML.hilf[welcher]){a.1 <- 0}
a.ML.start[i] <- a.1
a.ML.verlauf[1] <- a.ML.start[i]

# k>1 (weitere Schätzungen mittels verallgemeinertem KQ-Schätzer):

```

```

for(k in 2:k.runs){
# Falls im ersten Schritt lambda=0 ist -> abbrechen!
  if(a.1==0){ }
# Sonst:
  if(a.1>0){
Sigma.inv <- kronecker(diag(1,n), Sinv(a.ML.verlauf[k-1],d.))
e.dach <- as.vector(t(y))
-X2**solve(t(X2)**Sigma.inv**X2)**t(X2)**Sigma.inv**as.vector(t(y))
for(j in 1:9){
a.ML.hilf[j] <- nlminb(start.werte[j],l2.b,lower=0.1^6,upper=0.999999)$par
log.ML.hilf[j] <-
l2.s2(nlminb(start.werte[j],l2.b,lower=0.000001,upper=0.999999)$par)}
welcher <- min(which.max(l2.vec(a.ML.hilf)))
a.1 <- a.ML.hilf[welcher]
# Spezialfall: lam=0
l2.0 <- -3*n/2*log(2*pi)-3*n/2*log((t(e.dach)**e.dach)/(3*n)) - 3*n/2
if(l2.0 > log.ML.hilf[welcher]){a.1 <- 0}
}
a.ML.verlauf[k] <- a.1
}
# Finale Schätzwerte des zweistufigen Verfahrens:
if(a.ML.verlauf[k]==0) {GLS[i,] <- solve(t(X2)**X2)**t(X2)**as.vector(t(y))}
if(a.ML.verlauf[k]>0){Sigma.inv <- kronecker(diag(1,n),
Sinv(a.ML.verlauf[k],d.))
GLS[i,] <- solve(t(X2)**Sigma.inv**X2)**t(X2)
**Sigma.inv**as.vector(t(y))}
}
Kov <- cov(GLS)
D.Krit[u,o] <- det(Kov)
}}
round(D.Krit*n^2,3)

```

Symbolverzeichnis

Griechische Buchstaben mit Seitenzahl der ersten Erwähnung:

β_0	Parameter für Achsenabschnitt	S. 3
β_1	Parameter für Steigung	S. 3
β	Vektor der Regressionsparameter	S. 4
$\hat{\beta}_{GLS}$	verallgemeinerter Kleinste-Quadrate-Schätzer	S. 12
γ	Semi-Variogramm	S. 104
2γ	Variogramm	S. 103
$\Gamma(\lambda)$	Korrelationsmatrix	S. 5
$\Gamma(\sigma^2, \lambda)$	Kovarianzmatrix	S. 5
λ	Parameter für Korrelation	S. 4
$\Phi(\xi)$	D-Kriterium	S. 98
$\tilde{\Phi}(\xi)$	dritte Potenz des D-Kriteriums	S. 98
ρ	Parameter für Korrelation	S. 105
σ^2	Fehlervarianz	S. 4
$\Sigma(\lambda)$	Korrelationsmatrix	S. 4
$\Sigma(\sigma^2, \lambda)$	Kovarianzmatrix	S. 4
θ	Parametervektor	S. 7
$\hat{\theta}$	Schätzer für Parametervektor θ	S. 7
Θ	Parameterraum	S. 8
$\hat{\theta}_{UMVU}$	gleichmäßig bester erwartungstreuer Schätzer	S. 7
ξ	Versuchsplan	S. 3
Ξ	Menge mit Versuchsplänen	S. 6

Weitere Notationen mit Seitenzahl der ersten Erwähnung:

d	Versuchspunkt	S. 23
d^*	Versuchspunkt des D-optimalen Versuchsplans	S. 24
$D(\xi)$	D-Kriterium	S. 21
$\tilde{D}(\xi)$	quadriertes D-Kriterium	S. 24
e_j	Fehlerterm	S. 4
e_{ij}	j-ter Fehlerterm der i-ten Versuchseinheit	S. 4
e	Fehlervektor	S. 4
$\text{eff}_{D;\lambda_0}(\xi)$	D-Effizienz	S. 25

\mathcal{I}	Fisher-Informationsmatrix	S. 8
k	Länge des Vektors θ	S. 8
L	Likelihood-Funktion	S. 76
n	Anzahl Versuchseinheiten	S. 3
p	Länge des Vektors β	S. 12
q	Anzahl Versuchspunkte je Versuchseinheit	S. 3
$q_{\xi_1, \xi_2}(\lambda)$	relative D-Effizienz	S. 31
X	Designmatrix	S. 4
x_j	Versuchspunkt	S. 3
\mathcal{X}	Versuchsraum	S. 6
y_j	j-te Messung einer Zielgröße	S. 3
y_{ij}	j-te Messung einer Zielgröße der i-ten Versuchseinheit	S. 4
y	Zielgrößenvektor	S. 4
1_n	Einsenvektor der Länge n	
\otimes	Kronecker-Produkt	
Cov	Kovarianzmatrix	
det	Determinante	
E	Erwartungswert	
I_n	Einheitsmatrix der Dimension n	
Korr	Korrelation	
Kov	Kovarianz	
rg	Rang	
tr	Spur	
Var	Varianz	

Literatur

- Cressie, N. A. C. (1993). *Statistics for Spatial Data*, 2. Aufl., Wiley, New York.
- Dette, H., Kunert, J. und Pepelyshev, A. (2008). Exact optimal designs for weighted least squares analysis with correlated errors. *Statistica Sinica*, **18**, 135–154.
- Fahrmeir, L., Kneib, T. und Lang, S. (2007). *Regression*, Springer-Verlag, Berlin, Heidelberg.
- Fedorov, V. V. (1996). Design of spatial experiments: Model fitting and prediction. In *Handbook of Statistics, Volume 13*. C.R. Rao, S. Ghosh (eds), Elsevier, Amsterdam.
- Greene, W. H. (2012). *Econometric Analysis*, 7. Aufl., Pearson, Boston.
- Härdle, K. und Simar, L. (2012). *Applied Multivariate Statistical Analysis*, 3. Aufl., Springer-Verlag, Berlin, Heidelberg.
- Hildebrandt, S. (2006). *Analysis 1*, Springer, Berlin.
- Kiselák, J. und Stehlík, M. (2008). Equidistant and D-optimal designs for parameters of Ornstein-Uhlenbeck-process. *Statistics and Probability Letters*, **78**, 1388–1396.
- Krafft, O. (1978). *Lineare statistische Modelle und optimale Versuchspläne*, Vandenhoeck & Ruprecht, Göttingen.
- Lehmann, E. L. (1983). *Theory of Point Estimation*, Wiley, New York.
- Magnus, J. R. und Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 2. Aufl., Wiley, Chichester.
- Maxima.sourceforge.net (2007). *Maxima, a Computer Algebra System*. Version 5.13.0. URL <http://maxima.sourceforge.net/>
- Müller, W. und Stehlík, M. (2009). Issues in the optimal design of computer simulation experiments. *Applied Stochastic Models in Business and Industry*, **25**, 163–177.
- Oberhofer, W. und Kmenta, J. (1974). A general procedure for obtaining maximum likelihood estimates in generalized regression models. *Econometrika*, **42**, 579–590.
- Potthoff, R. F. und Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika*, **51**, 313–326.

-
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Rao, C. R., Toutenburg, H., Shalabh und Heumann, C. (2008). *Linear Models and Generalizations*, 3. Aufl., Springer, Berlin.
- Schach, S. und Schäfer, T. (1978). *Regressions- und Varianzanalyse*, Springer-Verlag, Berlin, Heidelberg.
- Schmidt, K. und Trenkler, G. (2006). *Einführung in die Moderne Matrix-Algebra*, 2. Aufl., Springer, Berlin.
- Stehlík, M. (2004). *Some Properties of D-optimal Designs for Random Fields with Different Variograms*, Report #4 of the Research Report Series of the Department of Statistics and Mathematic, Wirtschaftsuniversität Wien.
- Stöcker, H. (2009). *Taschenbuch mathematischer Formeln und moderner Verfahren*, 4. Aufl., Harri Deutsch Verlag, Frankfurt am Main.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbständig verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Dissertation ist bisher keiner anderen Fakultät vorgelegt worden. Ich erkläre, dass ich bisher kein Promotionsverfahren erfolglos beendet habe und dass keine Aberkennung eines bereits erworbenen Doktorgrades vorliegt.

Adrian Wilk