

Versuchsplanung für nichtlineare multiple Regressionsmodelle mit Anwendung auf zensierte Daten

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

von Dipl.-Math. Dennis Schmidt

geb. am 01.07.1987 in Wolfsburg

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

Gutachter: Prof. Dr. Rainer Schwabe
Prof. Dr. Stefanie Biedermann

eingereicht am: 20.02.2017

Verteidigung am: 14.06.2017

Zusammenfassung

Diese Arbeit beschäftigt sich mit der optimalen Versuchsplanung für verschiedene Regressionsmodelle, insbesondere dem Proportional Hazards Modell mit zensierten Daten. Für multiple Regressionsmodelle mit beliebig vielen Kovariablen werden allgemeine Resultate zur Struktur optimaler Designs hergeleitet. Es wird die Existenz eines optimalen Designs mit Stützpunkten nur auf den Kanten des Designraums bewiesen, wenn der Designraum ein Polytop ist. Ferner wird gezeigt, dass für viele Optimalitätskriterien die Stützpunkte eines optimalen Designs auf den Kanten des Designraums liegen müssen. Basierend auf diesen Strukturresultaten wird gezeigt, dass unter gewissen Bedingungen D -optimale Designs für multiple Regressionsmodelle aus den D -optimalen Designs in den marginalen Modellen mit einer Kovariablen konstruiert werden können. Diese neuen Methoden ermöglichen eine einfache Bestimmung optimaler Designs für viele nichtlineare Modelle, für die bisher keine optimalen Designs bei mehreren Kovariablen bekannt waren. Es werden D -optimale Designs und standardisiert Maximin D -optimale Designs für multiple Regression für eine große Klasse von Modellen bestimmt. Diese umfasst das Proportional Hazards Modell, bei dem verschiedene Zensierungsarten wie Typ I und zufällige Zensierung berücksichtigt werden können.

Für diese Modellklasse mit einer Kovariablen sind c -optimale Designs bisher nur für einzelne Vektoren \mathbf{c} bekannt. In der vorliegenden Arbeit gelingt eine vollständige Charakterisierung c -optimaler Designs für alle Vektoren \mathbf{c} . Außerdem werden für multiple Regression mit beliebig vielen Kovariablen c -optimale Designs für bestimmte Vektoren \mathbf{c} analytisch hergeleitet.

Für die Kriterien L -Optimalität und ϕ_k -Optimalität, wobei letzteres eine Verallgemeinerung vieler häufig verwendeter Optimalitätskriterien ist, werden für Modelle mit einer Kovariablen und für multiple Regression unter Verwendung der neuen Resultate optimale Designs bestimmt.

Abstract

This thesis deals with the optimal design of experiments for different regression models, in particular for the proportional hazards model with censored data. For multiple regression models with an arbitrary number of covariates general results on the structure of optimal designs are derived. The existence of an optimal design with support points only on the edges of the design region is proved, if the design region is a polytope. Moreover, it is shown that for many optimality criteria the support points of an optimal design must be located on the edges of the design region. Based on these structural results it is shown that under certain conditions D -optimal designs for multiple regression models can be constructed from the D -optimal designs in the marginal models with a single covariate. These new methods allow a simple determination of optimal designs for many nonlinear models, for which no optimal designs have been known in the case of multiple covariates. D -optimal designs and standardized maximin D -optimal designs for multiple regression are determined for a large class of models. This class includes the proportional hazards model, for which different types of censoring such as type I and random censoring can be considered.

For this class of models with one covariate c -optimal designs are only known for particular vectors \mathbf{c} . In this thesis a complete characterization of c -optimal designs for all vectors \mathbf{c} is established. In addition, for multiple regression with an arbitrary number of covariates c -optimal designs for certain vectors \mathbf{c} are derived analytically.

For the criteria L -optimality and ϕ_k -optimality, the latter being a generalization of many frequently used optimality criteria, optimal designs are determined for models with one covariate and for multiple regression using the new results.

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlagen	5
2.1	Zensierte Daten	5
2.2	Grundbegriffe der Ereigniszeitanalyse	6
2.3	Das Proportional Hazards Modell	7
3	Optimale Versuchsplanung	15
3.1	Grundlagen der Versuchsplanung	15
3.2	Approximative Designtheorie	16
3.3	Optimalitätskriterien	18
3.4	Äquivalenzsätze	22
3.5	Resultate zu Optimalitätskriterien	25
3.6	Optimales Design für multiple Regression	28
4	<i>D</i>-optimale Designs	33
4.1	Allgemeine Ergebnisse	33
4.2	Modell mit einer Kovariablen	37
4.3	Multiple Regression	39
4.4	Standardisiert Maximin <i>D</i> -optimale Designs	44
4.5	Bayessche <i>D</i> -optimale Designs	48
4.6	Additives Modell mit einem Effektparameter je erklärender Variable	53
4.7	Modell mit Interaktionsterm	54
5	<i>c</i>-optimale Designs	61
5.1	Modell mit einer Kovariablen	61
5.2	Multiple Regression	66
5.3	Bayessche <i>c</i> -optimale Designs	72
6	<i>L</i>-optimale Designs	77
6.1	Modell mit einer Kovariablen	77

6.2	Multiple Regression	80
7	ϕ_k-optimale Designs	89
7.1	ϕ_k -optimale Designs für $k \in (-\infty, 1)$	89
7.2	E -optimale Designs	91
8	Diskussion	95
A	Anhang	99
	Abbildungsverzeichnis	121
	Symbolverzeichnis	122
	Literaturverzeichnis	124

1 Einleitung

Ereignisdaten entstehen, wenn für verschiedene Individuen oder Objekte die Zeit bis zum Eintritt eines interessierenden Ereignisses beobachtet wird. Die Analyse solcher Ereignisdaten spielt in vielen Bereichen wie in der Medizin, Biologie, Soziologie und in den Wirtschafts- und Ingenieurwissenschaften eine wichtige Rolle. In der Medizin sind Heilung oder Tod von Patienten einer klinischen Studie oder eines wissenschaftlichen Experiments interessierende Ereignisse. Bei Zuverlässigkeitsanalysen sind Objekte wie z. B. Maschinen Gegenstand solcher Untersuchungen. Dann ist häufig die Zeit bis zum Ausfall gewisser Komponenten oder des gesamten Systems von Interesse.

Ein besonderes Merkmal von Ereignisdaten ist Zensierung. Zensierte Daten sind Daten, bei denen einige Werte nicht vollständig bekannt sind. In der Praxis tritt Zensierung typischerweise in Form der Rechtszensierung auf, wenn beim Beenden einer Studie bzw. eines Experiments das interessierende Ereignis noch nicht bei allen Individuen beobachtet wurde. Dann liegen für einige Individuen unvollständige Beobachtungen vor. Ihr exakter Ereigniszeitpunkt ist unbekannt, muss jedoch größer sein als die Beobachtungsdauer. Solche zensierten Beobachtungen beinhalten somit Informationen und sind bei der statistischen Analyse zu berücksichtigen.

Zensierung kann auf verschiedene Arten auftreten. Einen ganz gewöhnlichen Zensierungsmechanismus stellt das Beenden eines Experiments beispielsweise aus Zeit- oder Kostengründen dar. Dann liegt ein fester Zensierungszeitpunkt vor, der gleich dem Endzeitpunkt des Experiments ist. Scheiden Teilnehmer aus anderen Gründen als dem zu untersuchenden Ereignis aus dem Experiment aus, so erfolgt eine zufällige Zensierung. Beispiele hierfür sind der Abbruch der Teilnahme an dem Experiment aufgrund eines Umzugs oder wegen Desinteresses. Auch der Tod durch eine andere Krankheit führt zum Verlust des Individuums für die weitere Beobachtung.

Die Ereigniszeiten der Individuen hängen häufig von Einflussfaktoren ab. Dies können quantitative Variablen wie Medikamentendosis oder qualitative Variablen wie Behandlungsindikatoren sein. Von Interesse ist die Bestimmung des Zusammenhangs zwischen

den Ereigniszeiten und diesen Kovariablen. Das Proportional Hazards Modell von Cox ist ein weit verbreitetes Modell in der Ereigniszeitanalyse, das den Einfluss von Kovariablen auf die Ereigniszeiten modelliert. Verschiedene Zensierungsarten lassen sich in diesem Modell berücksichtigen.

Zur Quantifizierung des Zusammenhangs werden die unbekanntes Modellparameter auf Basis der Beobachtungen aus dem Experiment geschätzt. Ziel ist die interessierenden Parameter so präzise wie möglich zu schätzen, um möglichst genaue Schlussfolgerungen ziehen zu können. Die Parameterschätzungen hängen von den gewählten Versuchseinstellungen ab. Daher sind vor Durchführung des Experiments die Werte der Versuchseinstellungen sowie die Anzahl der Wiederholungen so zu bestimmen, dass die Qualität der Schätzungen in einem gewünschten Sinne optimiert wird. Diesbezüglich existieren verschiedene Optimalitätskriterien, die je nach Fragestellung oder Zielsetzung ausgewählt werden. Man spricht bei den gefundenen optimalen Versuchsplänen auch von optimalen Designs.

Die Verwendung optimaler Designs gewährleistet eine minimale Anzahl von Versuchsdurchführungen. Dies bewirkt eine Verringerung der für das Experiment benötigten Ressourcen und führt somit auch zu Kostensenkungen. Optimale Designs sollten auch aus ethischen Gründen z. B. bei Menschen- und Tierversuchen verwendet werden, damit die Anzahl der für das Experiment benötigten Individuen möglichst gering gehalten werden kann.

Die Theorie der optimalen Versuchsplanung stellt viele Hilfsmittel zur Lösung dieser Optimierungsprobleme und zur Bestimmung optimaler Designs zur Verfügung. Für nichtlineare Modelle mit mehreren Kovariablen sind die Identifikation optimaler Designs und der Nachweis der Optimalität dennoch mit erheblichen Schwierigkeiten verbunden. Im Fokus dieser Arbeit steht die Entwicklung neuer Methoden und Resultate zur Bestimmung optimaler Designs für multiple Regressionsmodelle. Besondere Aufmerksamkeit wird der Untersuchung der Struktur optimaler Designs für diese Modelle gewidmet, um mit Hilfe der gewonnenen Aussagen konkrete Designs anzugeben.

In der Regel können die Kovariablen Werte aus einem stetigen Intervall annehmen. In der Praxis kommt es aber auch vor, dass die Werte der Kovariablen aus einer vorgegebenen diskreten Menge gewählt werden. So gibt es bei der Verabreichung von Arzneimitteln häufig eine vorgegebene Menge von diskreten Dosisstufen, die typischerweise äquidistant im logarithmischen Maßstab sind. Der Vergleich zweier Gruppen mit unterschiedlichen Behandlungen kann mit einer binären Kovariablen codiert werden. Dies motiviert die

Betrachtung sowohl stetiger als auch diskreter Designräume.

Die Arbeit ist folgendermaßen gegliedert. Kapitel 2 präsentiert die mathematischen Grundlagen zu zensierten Daten, wichtigen Größen der Ereigniszeitanalyse und dem Proportional Hazards Modell. In Kapitel 3 erfolgt eine Einführung in die Theorie der optimalen Versuchsplanung. Nach der Darstellung des Konzepts der approximativen Designtheorie werden verschiedene Optimalitätskriterien definiert und zugehörige Äquivalenzsätze vorgestellt. Abschließend werden einige Strukturresultate für multiple Regressionsmodelle hergeleitet.

Die folgenden Kapitel widmen sich der Berechnung optimaler Designs für die eingeführten Optimalitätskriterien. Kapitel 4 befasst sich mit D -Optimalität. Zunächst werden allgemeine Ergebnisse zur Bestimmung D -optimaler Designs für multiple Regressionsmodelle hergeleitet, die dann auf eine große Klasse von Modellen angewendet werden. Diese Modellklasse umfasst das Proportional Hazards Modell mit verschiedenen Zensierungsarten sowie das Poisson- und Negativ-Binomial-Modell. Ferner werden standardisiert Maximin D -optimale Designs und Pseudo-Bayessche D -optimale Designs für multiple Regression sowie D -optimale Designs für Modelle mit zwei Kovariablen und einem Interaktionsterm bestimmt, die in einer gewissen Klasse von Designs optimal sind.

In Kapitel 5 steht die Herleitung c -optimaler Designs für Modelle mit einer oder mehreren Kovariablen im Vordergrund. Auch Pseudo-Bayessche c -optimale Designs werden für spezielle Proportional Hazards Modelle mit zensierten Daten berechnet. Kapitel 6 und 7 behandeln L - und ϕ_k -Optimalität für Modelle mit einer Kovariablen und multiple Regression. Als Spezialfall der ϕ_k -Kriterien werden E -optimale Designs für Modelle mit einer Kovariablen bestimmt.

2 Grundlagen

In diesem Kapitel werden die Grundlagen zu zensierten Daten vorgestellt, Grundbegriffe der Ereigniszeitanalyse und ihre Zusammenhänge erläutert und das Proportional Hazards Modell eingeführt. Es erfolgt die Herleitung der Fisher-Informationsmatrix für verschiedene Zensierungsarten.

2.1 Zensierte Daten

In der Praxis ist die Rechtszensierung die am häufigsten auftretende Art der Zensierung. Das interessierende Ereignis wird bei rechtszensierten Daten bis zum Ende des Experiments nicht beobachtet, sodass der tatsächliche Ereigniszeitpunkt unbekannt ist.

Es seien Y_1, \dots, Y_n unabhängige, nichtnegative Zufallsvariablen, die die Zeit bis zum Eintritt des interessierenden Ereignisses der Individuen beschreiben. Diese Ereigniszeiten werden im Folgenden als Lebensdauern bezeichnet. Die Zensierungszeiten der Individuen werden mit C_1, \dots, C_n bezeichnet. Sie geben die maximale Beobachtungsdauer an und können zufällig oder fest sein. Ist das interessierende Ereignis für das i -te Individuum bis zum Zensierungszeitpunkt C_i noch nicht eingetreten, so wird die Lebensdauer bei C_i zensiert. Aufgrund der Zensierung können nur die Paare (T_i, δ_i) , $i = 1, \dots, n$, beobachtet werden (vgl. Klein und Moeschberger, 2003, S. 65):

$$T_i = \min(Y_i, C_i), \quad (2.1)$$

$$\delta_i = \begin{cases} 1 & \text{falls } Y_i \leq C_i \\ 0 & \text{sonst} \end{cases}. \quad (2.2)$$

Dabei gibt die Zufallsvariable T_i den Ereignis- bzw. Zensierungszeitpunkt an und δ_i ist ein Zensierungsindikator mit Wert 0 bei Zensierung und Wert 1 bei einer exakten Beobachtung der Lebensdauer.

Es werden zwei verschiedene Zensierungsarten betrachtet: Typ I Zensierung und zufällige Zensierung.

Bei der Typ I Zensierung treten alle Individuen dem Experiment zur gleichen Zeit bei. Das Experiment wird zu einem vorgegebenen festen Zeitpunkt c beendet, sodass die Zensierungszeiten $C_i = c$ fest und für alle Individuen gleich sind.

Auch bei der zufälligen Zensierung treten die Individuen dem Experiment gleichzeitig bei. Die Zensierungszeiten C_i der Individuen sind jedoch Zufallsvariablen und keine festen Werte. Sie werden als unabhängig und identisch verteilt angenommen. Außerdem sollen sie unabhängig von den Lebensdauern Y_i und den interessierenden Parametern der Verteilung der Lebensdauern sein.

2.2 Grundbegriffe der Ereigniszeitanalyse

Eine der wichtigsten Größen in der Ereigniszeitanalyse ist die Überlebensfunktion $S_Y(t)$ von Y . Sie ist definiert als die Wahrscheinlichkeit, dass das interessierende Ereignis bis zum Zeitpunkt t noch nicht eingetreten ist:

$$S_Y(t) := P(Y > t) = 1 - F_Y(t). \quad (2.3)$$

Dabei bezeichnet $F_Y(t)$ die Verteilungsfunktion von Y .

Eine weitere zentrale Größe in der Ereigniszeitanalyse ist die Hazardfunktion, die auch als Hazardrate bezeichnet wird. Für absolut stetiges Y ist die Hazardfunktion $\lambda_Y(t)$ definiert als

$$\lambda_Y(t) := \lim_{\Delta t \downarrow 0} \frac{P(t \leq Y < t + \Delta t | Y \geq t)}{\Delta t} \quad \forall t \text{ mit } P(Y \geq t) > 0. \quad (2.4)$$

Das Produkt $\lambda_Y(t)\Delta t$ gibt approximativ die Wahrscheinlichkeit dafür an, dass das interessierende Ereignis im Zeitintervall Δt eintritt, unter der Bedingung, dass es bis zum Zeitpunkt t nicht eingetreten ist (vgl. Klein und Moeschberger, 2003, S. 27).

Zwischen der Hazardfunktion $\lambda_Y(t)$, der Überlebensfunktion $S_Y(t)$ und der Dichtefunktion $f_Y(t)$ von Y gilt der Zusammenhang

$$\lambda_Y(t) = \frac{f_Y(t)}{S_Y(t)}, \quad (2.5)$$

der auch als alternative äquivalente Definition der Hazardfunktion verwendet wird. Eine weitere Größe ist die kumulierte Hazardfunktion

$$H_Y(t) := \int_0^t \lambda_Y(u) du. \quad (2.6)$$

Die Überlebensfunktion ist mit der Hazardfunktion und kumulierten Hazardfunktion verknüpft über die Beziehung (vgl. Klein und Moeschberger, 2003, S. 27)

$$S_Y(t) = \exp\left(-\int_0^t \lambda_Y(u) \, du\right) = \exp(-H_Y(t)). \quad (2.7)$$

2.3 Das Proportional Hazards Modell

Eines der am häufigsten verwendeten Modelle für Ereignisdaten ist das Proportional Hazards Modell von Cox (1972). Es handelt sich um ein semiparametrisches Regressionsmodell, das den Einfluss von Kovariablen auf die Lebensdauern der Individuen beschreibt. Für $i = 1, \dots, n$ bezeichne $\mathbf{x}_i = (x_{i,1}, \dots, x_{i,p-1})^T$ den Vektor der Kovariablen des i -ten Individuums. Dies sind die Versuchseinstellungen des Experiments und somit feste Werte, die vom Versuchsleiter auszuwählen sind.

Das Proportional Hazards Modell stellt den folgenden Zusammenhang zwischen der Hazardfunktion und den Kovariablen her (Cox, 1972):

$$\lambda_{Y_i}(t; \mathbf{x}_i) = \lambda_0(t) \exp(\tilde{\mathbf{f}}(\mathbf{x}_i)^T \tilde{\boldsymbol{\beta}}). \quad (2.8)$$

Dabei ist $\lambda_0(t)$ eine Baseline-Hazardfunktion, $\tilde{\mathbf{f}} = (f_1, \dots, f_{p-1})^T$ ein Vektor von bekannten Regressionsfunktionen und $\tilde{\boldsymbol{\beta}} = (\beta_1, \dots, \beta_{p-1})^T \in \mathbb{R}^{p-1}$ der Vektor der unbekannt Parameter. Der Name des Modells rührt daher, dass die Hazardfunktionen zweier Individuen proportional zueinander sind, da ihr Quotient zeitunabhängig ist (vgl. Klein und Moeschberger, 2003, S. 48).

Gängige Parametrisierungen der Baseline-Hazardfunktion beruhen auf der Exponentialverteilung mit $\lambda_0(t) = \lambda$, der Weibullverteilung mit $\lambda_0(t) = \lambda \rho t^{\rho-1}$ und der Gompertzverteilung mit $\lambda_0(t) = \lambda \exp(\gamma t)$, die zu entsprechenden Verteilungen der Lebensdauern führen (vgl. Duchateau und Janssen, 2008, Abschnitt 1.4.2). In dieser Arbeit wird die folgende Parametrisierung der Baseline-Hazardfunktion gewählt (vgl. Cox und Oakes (1984, S. 40), Konstantinou et al. (2014)):

$$\lambda_0(t) = \lambda h_0(t) = \exp(\beta_0) h_0(t), \quad (2.9)$$

wobei der Parameter $\lambda > 0$ zu $\exp(\beta_0)$ umparametrisiert wurde. Die Funktion $h_0(t) > 0$ ist eine beliebige bekannte positive Funktion, die nicht von den interessierenden Parametern β_i abhängt. In $h_0(t)$ auftretende Parameter werden als bekannt vorausgesetzt.

Mit $\mathbf{f} = (1, f_1, \dots, f_{p-1})^T$ und $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^T \in \mathbb{R}^p$ lässt sich das parametrische Proportional Hazards Modell wie folgt darstellen:

$$\lambda_{Y_i}(t; \mathbf{x}_i) = h_0(t) \exp(\mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta}). \quad (2.10)$$

Mit Hilfe der Zusammenhänge des vorigen Abschnitts lassen sich die kumulierte Hazardfunktion, Überlebensfunktion und Dichtefunktion der Lebensdauern berechnen. Die kumulierte Hazardfunktion berechnet sich nach (2.6) zu

$$H_{Y_i}(t) = \int_0^t \lambda_{Y_i}(u; \mathbf{x}_i) du = \exp(\mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta}) \int_0^t h_0(u) du \quad (2.11)$$

und die Überlebensfunktion ergibt sich über die Beziehung (2.7) zu

$$S_{Y_i}(t) = \exp(-H_{Y_i}(t)) = \exp\left(-\exp(\mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta}) \int_0^t h_0(u) du\right). \quad (2.12)$$

Mit $f_{Y_i}(t) = \lambda_{Y_i}(t; \mathbf{x}_i) \cdot S_{Y_i}(t)$ nach Gleichung (2.5) berechnet sich die Dichtefunktion zu

$$f_{Y_i}(t) = h_0(t) \exp(\mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta}) \exp\left(-\exp(\mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta}) \int_0^t h_0(u) du\right). \quad (2.13)$$

In den folgenden Abschnitten wird die Fisher-Informationsmatrix für das Proportional Hazards Modell für die beschriebenen Zensierungsarten hergeleitet.

2.3.1 Fisher-Information für Typ I Zensierung

In diesem Abschnitt wird die Typ I Zensierung betrachtet, bei der die Zensierungszeiten $C_i = c$ fest sind. Es gilt $T_i = \min(Y_i, c)$ für $i = 1, \dots, n$. Die gemeinsame Dichte von (T_i, δ_i) bezüglich des Produktmaßes aus Lebesgue-Maß und Zählmaß ist gegeben durch (vgl. Klein und Moeschberger, 2003, S. 75)

$$f_{(T_i, \delta_i)}(t, d) = f_{Y_i}(t)^d \cdot S_{Y_i}(c)^{1-d}. \quad (2.14)$$

Aus der Unabhängigkeit der Y_i folgt, dass auch die Zufallsvariablen $(T_1, \delta_1), \dots, (T_n, \delta_n)$ unabhängig sind. Es seien $t_i = \min(y_i, c)$ die beobachteten Ereignis- oder Zensierungszeiten und d_i die beobachteten Zensierungsindikatoren. Die Likelihood-Funktion ist gegeben durch

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n f_{Y_i}(t_i)^{d_i} \cdot S_{Y_i}(t_i)^{1-d_i}. \quad (2.15)$$

Es folgt die Log-Likelihood-Funktion, die unter Verwendung von (2.5) und (2.7) umgeschrieben werden kann (vgl. Cox und Oakes, 1984, S. 33):

$$\begin{aligned}\log L(\boldsymbol{\beta}) &= \sum_{i=1}^n d_i \log f_{Y_i}(t_i) + (1 - d_i) \log S_{Y_i}(t_i) = \sum_{i=1}^n d_i \log \lambda_{Y_i}(t_i; \mathbf{x}_i) + \log S_{Y_i}(t_i) \\ &= \sum_{i=1}^n d_i \log \lambda_{Y_i}(t_i; \mathbf{x}_i) - H_{Y_i}(t_i).\end{aligned}\quad (2.16)$$

Für das betrachtete Proportional Hazards Modell gilt $\log \lambda_{Y_i}(t; \mathbf{x}_i) = \log h_0(t) + \mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta}$. Dieser Ausdruck ist linear in $\boldsymbol{\beta}$ und verschwindet bei zweimaliger Ableitung nach $\boldsymbol{\beta}$. Für die kumulierte Hazardfunktion (2.11) gilt $\partial H_{Y_i}(t)/\partial \beta_j = H_{Y_i}(t) f_j(\mathbf{x}_i)$. Somit ergeben sich die zweiten partiellen Ableitungen der Log-Likelihood-Funktion nach den Parametern für $j, k = 0, 1, \dots, p - 1$ zu

$$\frac{\partial^2 \log L(\boldsymbol{\beta})}{\partial \beta_k \partial \beta_j} = - \sum_{i=1}^n H_{Y_i}(t_i) f_j(\mathbf{x}_i) f_k(\mathbf{x}_i).\quad (2.17)$$

Mit partieller Integration, dem Zusammenhang $f_Y(t) = -S'_Y(t)$ und Anwendung von Gleichung (2.5) lässt sich die folgende Identität zeigen (vgl. Cox und Oakes, 1984, S. 46):

$$\begin{aligned}E(H_Y(T)) &= \int_0^c H_Y(t) f_Y(t) dt + H_Y(c) P(Y > c) \\ &= [-H_Y(t) S_Y(t)]_0^c + \int_0^c \lambda_Y(t) S_Y(t) dt + H_Y(c) S_Y(c) \\ &= -H_Y(c) S_Y(c) + \int_0^c f_Y(t) dt + H_Y(c) S_Y(c) = F_Y(c).\end{aligned}\quad (2.18)$$

Das Element (j, k) der Fisher-Informationsmatrix berechnet sich zu

$$\begin{aligned}\mathbf{I}_{(j,k)}(\boldsymbol{\beta}) &= -E \left(- \sum_{i=1}^n H_{Y_i}(T_i) f_j(\mathbf{x}_i) f_k(\mathbf{x}_i) \right) = \sum_{i=1}^n F_{Y_i}(c) f_j(\mathbf{x}_i) f_k(\mathbf{x}_i) \\ &= \sum_{i=1}^n \left[1 - \exp \left(- \exp(\mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta}) \int_0^c h_0(u) du \right) \right] f_j(\mathbf{x}_i) f_k(\mathbf{x}_i).\end{aligned}\quad (2.19)$$

Somit ist die Fisher-Informationsmatrix gegeben durch

$$\mathbf{I}(\boldsymbol{\beta}) = \sum_{i=1}^n \left[1 - \exp \left(- \exp(\mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta}) \int_0^c h_0(u) du \right) \right] \mathbf{f}(\mathbf{x}_i) \mathbf{f}(\mathbf{x}_i)^T.\quad (2.20)$$

Sie ist die Summe der Fisher-Informationsmatrizen $\mathbf{I}(\mathbf{x}_i, \boldsymbol{\beta})$ für eine Beobachtung. Die Fisher-Informationsmatrix lässt sich in der Form

$$\mathbf{I}(\boldsymbol{\beta}) = \sum_{i=1}^n \mathbf{I}(\mathbf{x}_i, \boldsymbol{\beta}) = \sum_{i=1}^n Q(\mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta}) \mathbf{f}(\mathbf{x}_i) \mathbf{f}(\mathbf{x}_i)^T \quad (2.21)$$

darstellen mit Intensitätsfunktion (vgl. Fedorov, 1972, S. 39)

$$Q(\theta) = 1 - \exp\left(-\exp(\theta) \int_0^c h_0(u) du\right). \quad (2.22)$$

Diese Funktion Q für Typ I Zensierung erfüllt die folgenden vier Bedingungen (A1)–(A4) (vgl. Konstantinou et al., 2014):

- (A1) $Q(\theta)$ ist positiv für alle $\theta \in \mathbb{R}$ und zweimal stetig differenzierbar.
- (A2) $Q'(\theta)$ ist positiv für alle $\theta \in \mathbb{R}$.
- (A3) Die zweite Ableitung $g''(\theta)$ der Funktion $g(\theta) = 1/Q(\theta)$ ist injektiv.
- (A4) Die Funktion $Q(\theta)/Q'(\theta)$ ist monoton wachsend.

Die Bedingung (A4) ist äquivalent dazu, dass die Funktion Q log-konkav ist. Dies bedeutet, dass $\log Q(\theta)$ eine konkave Funktion ist. Die vier Bedingungen (A1)–(A4) sind von wesentlicher Bedeutung in dieser Arbeit. Bei der späteren Berechnung optimaler Designs werden Modelle mit diesen Bedingungen für die Intensitätsfunktion betrachtet.

2.3.2 Fisher-Information für zufällige Zensierung

In diesem Abschnitt wird die zufällige Zensierung betrachtet, bei der die Zensierungszeiten C_i Zufallsvariablen sind. Die gemeinsame Dichte von (T_i, δ_i) bezüglich des Produktmaßes aus Lebesgue-Maß und Zählmaß ist gegeben durch (vgl. Duchateau und Janssen, 2008, S. 19–20)

$$f_{(T_i, \delta_i)}(t, d) = [S_C(t) f_{Y_i}(t)]^d \cdot [S_{Y_i}(t) f_C(t)]^{1-d}, \quad (2.23)$$

wobei S_C die Überlebensfunktion und f_C die Dichtefunktion der Zensierungszeiten C_i bezeichnen. Die Zufallsvariablen $(T_1, \delta_1), \dots, (T_n, \delta_n)$ sind aufgrund der Unabhängigkeit der Zufallsvariablen Y_i und C_i ebenfalls unabhängig. Es folgt die Likelihood-Funktion

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n [S_C(t_i) f_{Y_i}(t_i)]^{d_i} \cdot [S_{Y_i}(t_i) f_C(t_i)]^{1-d_i}. \quad (2.24)$$

Die Faktoren $S_C(t_i)^{d_i}$ und $f_C(t_i)^{1-d_i}$ werden bei der Log-Likelihood-Funktion zu additiven Summanden, die nicht von dem Parametervektor $\boldsymbol{\beta}$ abhängen. Sie fallen daher beim Ableiten weg, sodass die ersten und zweiten partiellen Ableitungen der Log-Likelihood-Funktion mit denen bei der Typ I Zensierung übereinstimmen. Der zur Berechnung der Fisher-Informationsmatrix zu bestimmende Erwartungswert $E(H_{Y_i}(T_i))$ unterscheidet sich zur Typ I Zensierung dadurch, dass die Zensierungszeiten diesmal Zufallsvariablen sind. Da die Zufallsvariablen Y_i und C_i unabhängig sind, folgt unter Verwendung von (2.18)

$$E(H_{Y_i}(T_i) | C_i = c) = E(H_{Y_i}(\min(Y_i, C_i)) | C_i = c) = E(H_{Y_i}(\min(Y_i, c))) = F_{Y_i}(c)$$

und der Erwartungswert von $H_{Y_i}(T_i)$ berechnet sich zu (vgl. Konstantinou et al., 2014)

$$\begin{aligned} E(H_{Y_i}(T_i)) &= E(E(H_{Y_i}(T_i) | C_i)) = \int_0^\infty E(H_{Y_i}(T_i) | C_i = c) \cdot f_C(c) dc \\ &= \int_0^\infty \left[1 - \exp\left(-\exp(\mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta}) \int_0^c h_0(u) du\right) \right] \cdot f_C(c) dc. \end{aligned} \quad (2.25)$$

Es folgt, dass die Fisher-Informationsmatrix von der Form (2.21) ist mit Intensitätsfunktion

$$Q(\theta) = \int_0^\infty \left[1 - \exp\left(-\exp(\theta) \int_0^c h_0(u) du\right) \right] \cdot f_C(c) dc. \quad (2.26)$$

Konstantinou (2013) zeigte, dass diese Intensitätsfunktion für stetig gleichverteilte Zensierungszeiten die Bedingungen (A1), (A2) und (A4) erfüllt, wobei die Baseline-Hazardfunktion $\lambda_0(t) = \lambda h_0(t)$ gewisse Voraussetzungen erfüllen muss. Schmidt und Schwabe (2017) haben unter Annahme geeigneter Voraussetzungen gezeigt, dass für viele Verteilungen der Zensierungszeiten und eine konstante Baseline-Hazardfunktion die Intensitätsfunktion die Bedingungen (A1), (A2) und (A4) erfüllt. Die folgenden Lemmata führen diese Resultate zusammen.

Lemma 2.1

Die Dichte f_C der Zensierungszeiten sei stetig. Dann erfüllt die Funktion Q aus (2.26) für die zufällige Zensierung die Bedingungen (A1) und (A2).

Beweis:

Der Integrand in (2.26) ist zweimal differenzierbar bezüglich θ . Seine Ableitungen werden durch die integrierbare Funktion $M \cdot f_C(c)$ majorisiert, wobei M eine genügend große Konstante ist. Nach dem Satz von der majorisierten Konvergenz können Differentiation

(bezüglich θ) und Integration (bezüglich c) vertauscht werden. Daher ist Q zweimal differenzierbar und

$$Q'(\theta) = \int_0^\infty \exp(\theta) \int_0^c h_0(u) du \cdot \exp\left(-\exp(\theta) \int_0^c h_0(u) du\right) \cdot f_C(c) dc > 0. \quad \square$$

Lemma 2.2

Die Dichte f_C der Zensierungszeiten sei stetig und log-konkav. Außerdem sei die Funktion $h(x) = \int_0^x h_0(u) du$ log-konkav mit stetig differenzierbarer Funktion h_0 . Dann erfüllt die Funktion Q aus (2.26) für die zufällige Zensierung die Bedingung (A4).

Beweis:

Die Funktion $m(\theta, c) = 1 - \exp(-\exp(\theta) \int_0^c h_0(u) du)$ ist eine log-konkave Funktion in θ und c (Konstantinou, 2013). Das Produkt $m(\theta, c) \cdot f_C(c)$ ist ebenfalls log-konkav, da $\log(m(\theta, c) \cdot f_C(c)) = \log m(\theta, c) + \log f_C(c)$ gilt. Nach Theorem 6 von Prékopa (1973) folgt, dass $Q(\theta)$ eine log-konkave Funktion ist. \square

Für die Log-Konkavität der Funktion $h(x) = \int_0^x h_0(u) du$ ist es hinreichend, wenn die Funktion h_0 log-konkav ist (vgl. Bagnoli und Bergstrom, 2005). Diese Bedingung für h_0 ist allerdings nicht notwendig. So ist die Funktion $h_0(t) = \rho t^{\rho-1}$ für $\rho \in (0, 1)$ nicht log-konkav, aber die Funktion $h(x) = \int_0^x h_0(u) du = x^\rho$ ist log-konkav.

Viele Wahrscheinlichkeitsverteilungen besitzen log-konkave Dichten. Bagnoli und Bergstrom (2005) haben eine Liste von Wahrscheinlichkeitsverteilungen mit log-konkaven Dichten zusammengestellt, zu denen unter anderem die Normalverteilung, logistische Verteilung, Gleichverteilung, Exponentialverteilung und Chi-Quadrat-Verteilung gehören sowie für bestimmte Parameterwerte die Weibullverteilung, Gammaverteilung und Betaverteilung. Für die Zensierung können auch abgeschnittene Verteilungen von Interesse sein. Die abgeschnittene Verteilung einer Wahrscheinlichkeitsverteilung mit log-konkaver Dichte hat ebenfalls eine log-konkave Dichte (vgl. Bagnoli und Bergstrom, 2005). Spätere Resultate sind somit für eine ganze Klasse von Modellen mit zufälliger Zensierung gültig.

Im Folgenden wird der Einfachheit halber $h_0(t) = 1$ gewählt. Dann gilt $\int_0^c h_0(u) du = c$. Werden die Zensierungszeiten C_i als exponentialverteilt mit Parameter $\lambda > 0$ angenommen, so berechnet sich die Intensitätsfunktion Q zu

$$Q(\theta) = 1 - \int_0^\infty e^{-ce^\theta} \cdot \lambda e^{-\lambda c} dc = 1 - \left[-\frac{\lambda e^{-c(e^\theta + \lambda)}}{e^\theta + \lambda} \right]_0^\infty = 1 - \frac{\lambda}{e^\theta + \lambda} = \frac{e^\theta}{e^\theta + \lambda}. \quad (2.27)$$

Da die Dichte der Exponentialverteilung stetig und log-konkav ist, gelten die Bedingun-

gen (A1), (A2) und (A4) nach Lemmata 2.1 und 2.2. Für $g(\theta) := 1/Q(\theta) = 1 + \lambda e^{-\theta}$ ist die zweite Ableitung gegeben durch $g''(\theta) = \lambda e^{-\theta}$. Sie ist injektiv und somit ist auch die Bedingung (A3) erfüllt. Die Intensitätsfunktion ist proportional zur Intensitätsfunktion eines Negativ-Binomial-Modells (vgl. Rodríguez-Torreblanca und Rodríguez-Díaz, 2007) bzw. Poisson-Gamma-Modells (vgl. Graßhoff et al., 2016).

Konstantinou et al. (2014) haben Zensierungszeiten betrachtet, die auf dem Intervall $[0, c]$ stetig gleichverteilt sind. Die resultierende Intensitätsfunktion ist gegeben durch

$$Q(\theta) = 1 - \frac{1 - \exp(-c \exp(\theta))}{c \exp(\theta)}. \quad (2.28)$$

Es kann gezeigt werden, dass sie die Bedingung (A3) erfüllt (vgl. Schmidt, 2013). Außerdem gelten (A1), (A2) und (A4) nach Lemmata 2.1 und 2.2.

Das folgende Lemma liefert eine Eigenschaft der betrachteten Intensitätsfunktionen, die hilfreich bei der Anwendung der später eingeführten Äquivalenzsätze ist.

Lemma 2.3

Es sei Q eine Funktion, die die Bedingungen (A1), (A2) und (A4) erfüllt. Dann fällt $Q(\theta)$ exponentiell schnell auf null für $\theta \rightarrow -\infty$, d. h. es gilt

$$\lim_{\theta \rightarrow -\infty} p(\theta)Q(\theta) = 0 \quad (2.29)$$

für ein beliebiges Polynom $p(\theta)$.

Beweis:

Es sei n der Grad des Polynoms $p(\theta)$. Zunächst sei festgestellt, dass die Funktion $h(\theta) := \theta + (n + 1)Q(\theta)/Q'(\theta)$ wegen (A4) eine streng monoton wachsende und somit bijektive Funktion ist. Sie hat daher genau eine Nullstelle, die sich bei $\tilde{\theta} < 0$ befindet, da $Q(\theta)$ und $Q'(\theta)$ positiv sind.

Es wird nun die Funktion $r(\theta) := \theta^{n+1}Q(\theta)$ mit $n \geq 0$ betrachtet. Ihre Ableitung berechnet sich zu $r'(\theta) = (n + 1)\theta^n Q(\theta) + \theta^{n+1}Q'(\theta) = \theta^n Q'(\theta)h(\theta)$. Diese Ableitung hat eine Nullstelle bei 0 und eine weitere Nullstelle bei $\tilde{\theta} < 0$. Die Funktion r kann keinen Sattelpunkt bei $\tilde{\theta}$ haben, weil ihre Ableitung bei $\tilde{\theta}$ das Vorzeichen wechselt. Für ungerades n handelt es sich bei $\tilde{\theta}$ um ein Maximum von r , da $r(0) = 0$ und $r(\theta) > 0$ für alle $\theta \neq 0$ gilt. Für gerades n handelt es sich bei $\tilde{\theta}$ um ein Minimum von r , da $r(0) = 0$ und $r(\theta) < 0$ für alle $\theta < 0$ gilt. Daher ist die Funktion $r(\theta)$ monoton für $\theta < \tilde{\theta}$ und durch null beschränkt für $\theta \rightarrow -\infty$. Es folgt die Konvergenz von $r(\theta)$ für $\theta \rightarrow -\infty$ und damit $\lim_{\theta \rightarrow -\infty} p(\theta)Q(\theta) = 0$. □

3 Optimale Versuchsplanung

Die Qualität der Parameterschätzungen hängt von der Wahl der Versuchseinstellungen ab. Um möglichst präzise Schätzungen der Parameter zu erhalten, werden optimale Werte der Kovariablen benötigt. Dies führt auf die Berechnung von optimalen Designs. In diesem Kapitel wird die Theorie der optimalen Versuchsplanung vorgestellt.

3.1 Grundlagen der Versuchsplanung

Die Versuchseinstellungen $\mathbf{x}_1, \dots, \mathbf{x}_n$ zum Stichprobenumfang n können aus dem Versuchsbereich $\mathcal{X} \subset \mathbb{R}^{p-1}$ gewählt werden. Diese Menge wird auch Designraum genannt. Die einzelnen Versuchseinstellungen können auch mehrfach verwendet werden. Es seien $\mathbf{x}_1, \dots, \mathbf{x}_m$ mit $m \leq n$ unterschiedliche Werte der Versuchseinstellungen und r_1, \dots, r_m die zugehörigen Häufigkeiten mit $\sum_{i=1}^m r_i = n$. Ein exaktes Design ξ_n hat die folgende Gestalt (vgl. Atkinson et al., 2007, S. 120):

$$\xi_n = \left\{ \begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_m \\ r_1/n & r_2/n & \dots & r_m/n \end{array} \right\}. \quad (3.1)$$

Das Design ξ_n lässt sich als diskretes Wahrscheinlichkeitsmaß auf \mathcal{X} auffassen, das jeder Versuchseinstellung \mathbf{x}_i das Gewicht $\omega_i = r_i/n$ zuordnet. Dieser Wert gibt den Anteil der Beobachtungen bei \mathbf{x}_i an.

Die Fisher-Informationsmatrix hängt von der Wahl der Versuchseinstellungen, d. h. vom gewählten Design ab:

$$\mathbf{I}(\xi_n, \boldsymbol{\beta}) = \sum_{i=1}^m r_i \mathbf{I}(\mathbf{x}_i, \boldsymbol{\beta}). \quad (3.2)$$

Eine untere Schranke für die Kovarianzmatrix eines erwartungstreuen Schätzers $\hat{\boldsymbol{\beta}}$ für $\boldsymbol{\beta}$ wird unter Regularitätsbedingungen durch die Informationsungleichung geliefert. Diese

Schranke ist durch die Inverse der Fisher-Informationsmatrix gegeben:

$$\text{Cov}_{\beta}(\hat{\beta}) \geq \mathbf{I}(\xi_n, \beta)^{-1}. \quad (3.3)$$

Dabei ist „ \geq “ für Matrizen im positiv semidefiniten Sinne gemeint, d. h. die Differenz $\text{Cov}_{\beta}(\hat{\beta}) - \mathbf{I}(\xi_n, \beta)^{-1}$ ist positiv semidefinit.

Unter geeigneten Regularitätsbedingungen ist der Maximum-Likelihood-Schätzer für β asymptotisch effizient und somit ist die Inverse der Fisher-Informationsmatrix gleich der asymptotischen Kovarianzmatrix. Daher soll die Inverse der Informationsmatrix in einem gewissen Sinne minimiert werden bzw. die Informationsmatrix maximiert werden, um die optimalen Werte der Versuchseinstellungen zu finden und somit die Parameter möglichst präzise schätzen zu können.

3.2 Approximative Designtheorie

Da es sich bei den Gewichten um Vielfache von $1/n$ handelt, liegt ein Optimierungsproblem mit ganzzahligen Variablen vor. Die Berechnung optimaler Designs kann sehr kompliziert sein. Daher wird das von Elfving (1952) und Kiefer (1959) eingeführte Konzept approximativer Designs verwendet.

Definition 3.1 (Approximatives Design)

Ein approximatives Design ξ ist ein Wahrscheinlichkeitsmaß auf \mathcal{X} mit endlichem Träger.

Es entfällt die Forderung, dass die Produkte $n\omega_i$ ganzzahlig sein müssen. Somit gilt für die Gewichte $0 \leq \omega_i \leq 1$ für $i = 1, \dots, m$ und $\sum_{i=1}^m \omega_i = 1$. Ein approximatives Design ξ hat die folgende Gestalt:

$$\xi = \left\{ \begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_m \\ \omega_1 & \omega_2 & \dots & \omega_m \end{array} \right\}. \quad (3.4)$$

Insbesondere ist das approximative Designproblem unabhängig von n . Ein optimales approximatives Design ξ stellt in der Regel eine gute Approximation für das ganzzahlige Optimierungsproblem dar. Sind die Produkte $n\omega_i$ für $i = 1, \dots, m$ ganzzahlig, so liegt ein exaktes Design, wie im vorangegangenen Abschnitt beschrieben, vor. Sind sie nicht ganzzahlig, so muss mit Hilfe geeigneter Rundungsverfahren ein exaktes Design konstruiert werden, um es in der Praxis umsetzen zu können. Solche Rundungsverfahren werden beispielsweise von Pukelsheim und Rieder (1992) vorgestellt.

Es lassen sich auch stetige Designs betrachten, die durch stetige Wahrscheinlichkeitsmaße gegeben sind. Die Informationsmatrix $\mathbf{M}(\xi, \boldsymbol{\beta})$ eines beliebigen Designs ξ wird definiert als (vgl. Silvey, 1980, S. 53)

$$\mathbf{M}(\xi, \boldsymbol{\beta}) := \int_{\mathcal{X}} \mathbf{I}(\mathbf{x}, \boldsymbol{\beta}) \xi(d\mathbf{x}). \quad (3.5)$$

Für ein approximatives Design ergibt sich die Informationsmatrix zu

$$\mathbf{M}(\xi, \boldsymbol{\beta}) = \sum_{i=1}^m \omega_i \mathbf{I}(\mathbf{x}_i, \boldsymbol{\beta}) \quad (3.6)$$

und damit zu einer gewichteten Version der Fisher-Informationsmatrix. Für ein exaktes Design gilt $\mathbf{M}(\xi_n, \boldsymbol{\beta}) = \frac{1}{n} \mathbf{I}(\xi_n, \boldsymbol{\beta})$.

Es werden Modelle mit Fisher-Informationsmatrix $\mathbf{I}(\mathbf{x}, \boldsymbol{\beta}) = Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \mathbf{f}(\mathbf{x}) \mathbf{f}(\mathbf{x})^T$ betrachtet, wie sie für das Proportional Hazards Modell mit Typ I und zufälliger Zensierung und häufig bei verallgemeinerten linearen Modellen vorkommen. Die Informationsmatrizen haben somit die Form

$$\mathbf{M}(\xi, \boldsymbol{\beta}) = \sum_{i=1}^m \omega_i Q(\mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta}) \mathbf{f}(\mathbf{x}_i) \mathbf{f}(\mathbf{x}_i)^T. \quad (3.7)$$

Es sei Ξ die Menge aller Wahrscheinlichkeitsmaße auf \mathcal{X} und $\mathcal{M}_\beta = \{\mathbf{M}(\xi, \boldsymbol{\beta}) : \xi \in \Xi\}$ die Menge aller Informationsmatrizen für festes $\boldsymbol{\beta}$. Es gilt $\mathcal{M}_\beta = \text{conv}\{\mathbf{I}(\mathbf{x}, \boldsymbol{\beta}) : \mathbf{x} \in \mathcal{X}\}$ (vgl. Silvey, 1980, S. 54), wobei $\text{conv}(S)$ die konvexe Hülle der Menge S bezeichnet. Die Informationsmatrizen sind symmetrische positiv semidefinite $p \times p$ -Matrizen, die als Vektoren im $\mathbb{R}^{p(p+1)/2}$ dargestellt werden können. Dies erlaubt die Anwendung des Satzes von Carathéodory, was zu folgendem Resultat führt (vgl. Silvey (1980, S. 16, 72), Fedorov (1972, S. 66)).

Satz 3.2

Jede Informationsmatrix $\mathbf{M} \in \mathcal{M}_\beta$ lässt sich darstellen als Konvexkombination

$$\mathbf{M} = \sum_{i=1}^m \omega_i \mathbf{I}(\mathbf{x}_i, \boldsymbol{\beta}) \quad (3.8)$$

mit $\mathbf{x}_i \in \mathcal{X}$ für $i = 1, \dots, m$ und $m \leq \frac{1}{2}p(p+1) + 1$. Für Randpunkte $\mathbf{M} \in \mathcal{M}_\beta$ gilt $m \leq \frac{1}{2}p(p+1)$.

Nach Satz 3.2 existiert zu jeder Informationsmatrix $\mathbf{M} \in \mathcal{M}_\beta$ ein Design ξ , das einen

endlichen Träger hat, mit $\mathbf{M} = \mathbf{M}(\xi, \boldsymbol{\beta})$. Insbesondere existiert also immer ein optimales approximatives Design, sodass man sich bei der Suche nach optimalen Designs auf approximative Designs beschränken kann. Satz 3.2 liefert ferner eine obere Schranke für die Anzahl verschiedener Stützpunkte eines optimalen Designs.

3.3 Optimalitätskriterien

Ein Design ξ^* , für das $\mathbf{M}(\xi^*, \boldsymbol{\beta}) \geq \mathbf{M}(\xi, \boldsymbol{\beta})$ für alle Designs $\xi \in \Xi$ gilt, existiert im Allgemeinen nicht (vgl. Silvey, 1980, S. 10). Es werden daher reellwertige Funktionen Φ von der Informationsmatrix des Designs betrachtet, die bezüglich des Designs maximiert oder minimiert werden. Jedes Minimierungsproblem kann in ein Maximierungsproblem überführt werden und umgekehrt. Um die Existenz optimaler Designs zu sichern, wird der Designraum \mathcal{X} als kompakt und die Intensitätsfunktion Q sowie die Regressionsfunktionen $f_i, i = 1, \dots, p$, als stetig vorausgesetzt.

Eine Besonderheit bei den hier betrachteten nichtlinearen Modellen besteht darin, dass die optimalen Designs von den Werten der unbekannt Parameter abhängen. Solche Designs, die optimal für einen vorgegebenen Parameterwert sind, werden lokal optimal genannt (Chernoff, 1953). Im Folgenden wird auf die Bezeichnung „lokal“ verzichtet.

Definition 3.3 (*D-Optimalität*)

Ein Design $\xi^* = \xi_{\boldsymbol{\beta}}^*$ mit regulärer Informationsmatrix $\mathbf{M}(\xi^*, \boldsymbol{\beta})$ heißt *D-optimal*, wenn $\det(\mathbf{M}(\xi^*, \boldsymbol{\beta})) \geq \det(\mathbf{M}(\xi, \boldsymbol{\beta}))$ für alle $\xi \in \Xi$ gilt (vgl. Silvey, 1980, S. 54).

Bei *D-Optimalität* handelt es sich um eines der am häufigsten verwendeten Optimalitätskriterien. Ein *D-optimales* Design minimiert das Volumen des Konfidenzellipsoids für die Parameter. Außerdem ist *D-Optimalität* invariant unter linearen Umparametrisierungen (vgl. Schwabe, 1996, S. 9).

Bei *c-Optimalität* für einen Vektor $\mathbf{c} \in \mathbb{R}^p, \mathbf{c} \neq \mathbf{0}_p$, liegt Interesse darin, den linearen Aspekt $\mathbf{c}^T \boldsymbol{\beta}$ mit minimaler Varianz zu schätzen. Der lineare Aspekt $\mathbf{c}^T \boldsymbol{\beta}$ muss identifizierbar sein, d. h. es muss $\mathbf{c} = \mathbf{M}(\xi, \boldsymbol{\beta}) \mathbf{y}$ für einen Vektor $\mathbf{y} \in \mathbb{R}^p$ gelten (vgl. Silvey, 1980, S. 25). Designs mit singulärer Informationsmatrix sind als *c-optimale* Designs möglich.

Definition 3.4 (*c-Optimalität*)

Ein Design ξ^* heißt *c-optimal*, wenn der lineare Aspekt $\mathbf{c}^T \boldsymbol{\beta}$ identifizierbar ist und wenn $\mathbf{c}^T \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-} \mathbf{c} \leq \mathbf{c}^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-} \mathbf{c}$ für alle $\xi \in \Xi$ gilt, für die $\mathbf{c}^T \boldsymbol{\beta}$ identifizierbar ist. Dabei ist $\mathbf{M}(\xi, \boldsymbol{\beta})^{-}$ eine verallgemeinerte Inverse von $\mathbf{M}(\xi, \boldsymbol{\beta})$.

Definition 3.5 (*L*-Optimalität)

Ein Design ξ^* mit regulärer Informationsmatrix $\mathbf{M}(\xi^*, \boldsymbol{\beta})$ heißt *L*-optimal, wenn für alle $\xi \in \Xi$ mit regulärer Informationsmatrix $\text{Spur}(\mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{B}) \leq \text{Spur}(\mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1} \mathbf{B})$ gilt. Dabei ist \mathbf{B} eine symmetrische positiv definite Matrix.

L-Optimalität kann als Verallgemeinerung von *c*-Optimalität aufgefasst werden, wenn positiv semidefinite Matrizen \mathbf{B} zugelassen werden. Hier wird die Matrix \mathbf{B} als positiv definit und damit regulär vorausgesetzt. Somit muss ein *L*-optimales Design eine reguläre Informationsmatrix haben (vgl. Silvey, 1980, S. 50). Wird die Matrix \mathbf{B} als Einheitsmatrix \mathbf{I}_p gewählt, so ergibt sich *A*-Optimalität. Ein weiterer Spezialfall von *L*-Optimalität ist *V*-Optimalität, bei dem der Integrated Mean Squared Error minimiert wird. Gesucht ist ein Design, das den Ausdruck

$$\int_{\mathcal{X}} \mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}) \mu(d\mathbf{x}) \quad (3.9)$$

minimiert, wobei μ eine Wahrscheinlichkeitsverteilung auf \mathcal{X} ist (vgl. Atkinson et al., 2007, S. 143). Dieser Ausdruck kann umgeschrieben werden zu $\text{Spur}(\mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{B})$ mit

$$\mathbf{B} = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x}) \mathbf{f}(\mathbf{x})^T \mu(d\mathbf{x}). \quad (3.10)$$

V-Optimalität ist ebenso wie *D*-Optimalität invariant unter Umparametrisierungen.

Eine allgemeine Klasse von Optimalitätskriterien bilden die ϕ_k -Kriterien von Kiefer (1974) mit $k \in [-\infty, 1]$. Zunächst wird eine beliebige Potenz einer Matrix definiert. Jede symmetrische Matrix \mathbf{M} ist orthogonal diagonalisierbar, d. h. es existiert eine Diagonalmatrix \mathbf{D} und eine orthogonale Matrix \mathbf{U} mit $\mathbf{M} = \mathbf{U} \mathbf{D} \mathbf{U}^T$. Die Matrixpotenz \mathbf{M}^k mit $k \in \mathbb{R}$ ist definiert als $\mathbf{M}^k = \mathbf{U} \mathbf{D}^k \mathbf{U}^T$. Für eine positiv definite $p \times p$ -Matrix \mathbf{M} sei

$$\phi_k(\mathbf{M}) = \begin{cases} \left(\frac{1}{p} \text{Spur}(\mathbf{M}^k) \right)^{1/k} & \text{für } k \neq -\infty, 0, \\ \det(\mathbf{M})^{1/p} & \text{für } k = 0, \\ \lambda_{\min}(\mathbf{M}) & \text{für } k = -\infty, \end{cases} \quad (3.11)$$

wobei $\lambda_{\min}(\mathbf{M})$ den kleinsten Eigenwert der Matrix \mathbf{M} bezeichnet (vgl. Pukelsheim, 1993, S. 141).

Definition 3.6 (ϕ_k -Optimalität)

Ein Design ξ_k^* mit regulärer Informationsmatrix $\mathbf{M}(\xi_k^*, \boldsymbol{\beta})$ heißt ϕ_k -optimal, wenn für alle $\xi \in \Xi$ mit regulärer Informationsmatrix $\phi_k(\mathbf{M}(\xi_k^*, \boldsymbol{\beta})) \geq \phi_k(\mathbf{M}(\xi, \boldsymbol{\beta}))$ gilt.

Die Fälle $k = 0$, $k = -1$ und $k = -\infty$ entsprechen D -, A - und E -Optimalität.

Zum Vergleich verschiedener Designs wird die Effizienz berechnet, die den Wert der Kriteriumsfunktion eines Designs ins Verhältnis zum Wert der Kriteriumsfunktion für das optimale Design setzt.

Definition 3.7 (Effizienz)

Die (a) D -Effizienz, (b) c -Effizienz, (c) L -Effizienz, (d) ϕ_k -Effizienz eines Designs ξ ist definiert als

$$(a) \quad \text{eff}_D(\xi, \boldsymbol{\beta}) = \left(\frac{\det(\mathbf{M}(\xi, \boldsymbol{\beta}))}{\det(\mathbf{M}(\xi_{\boldsymbol{\beta}}^*, \boldsymbol{\beta}))} \right)^{\frac{1}{p}}, \quad (3.12)$$

$$(b) \quad \text{eff}_c(\xi, \boldsymbol{\beta}) = \frac{\mathbf{c}^T \mathbf{M}(\xi_{\boldsymbol{\beta}}^*, \boldsymbol{\beta})^{-1} \mathbf{c}}{\mathbf{c}^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{c}}, \quad (3.13)$$

$$(c) \quad \text{eff}_L(\xi, \boldsymbol{\beta}) = \frac{\text{Spur}(\mathbf{M}(\xi_{\boldsymbol{\beta}}^*, \boldsymbol{\beta})^{-1} \mathbf{B})}{\text{Spur}(\mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{B})}, \quad (3.14)$$

$$(d) \quad \text{eff}_{\phi_k}(\xi, \boldsymbol{\beta}) = \frac{\phi_k(\mathbf{M}(\xi, \boldsymbol{\beta}))}{\phi_k(\mathbf{M}(\xi_{\boldsymbol{\beta}}^*, \boldsymbol{\beta}))}. \quad (3.15)$$

Dabei ist $\xi_{\boldsymbol{\beta}}^*$ das lokal optimale Design für das jeweilige Optimalitätskriterium.

Die Effizienz nimmt Werte zwischen 0 und 1 an, wobei der Wert 1 für optimale Designs angenommen wird.

Infolge der Parameterabhängigkeit der optimalen Designs können sich bei Fehlspezifikation der unbekannt Parameter ineffiziente Designs ergeben. Eine Herangehensweise zur Lösung dieser Problematik ist die Bestimmung von standardisiert Maximin optimalen Designs (vgl. Müller (1995), Dette (1997)). Dafür werden zusätzliche Informationen in Form eines Wertebereichs Θ für die Parameter benötigt. Die Menge Θ wird als kompakt vorausgesetzt. Die Idee besteht darin, die minimale Effizienz zu maximieren und damit robuste Designs zu erhalten.

Definition 3.8 (Standardisierte Maximin D -Optimalität)

Ein Design ξ^* mit regulärer Informationsmatrix $\mathbf{M}(\xi^*, \boldsymbol{\beta})$ für alle $\boldsymbol{\beta} \in \Theta$ heißt standardisiert Maximin D -optimal, wenn es die bezüglich $\boldsymbol{\beta} \in \Theta$ minimale D -Effizienz maximiert:

$$\min_{\boldsymbol{\beta} \in \Theta} \left(\frac{\det(\mathbf{M}(\xi^*, \boldsymbol{\beta}))}{\det(\mathbf{M}(\xi_{\boldsymbol{\beta}}^*, \boldsymbol{\beta}))} \right)^{\frac{1}{p}} = \max_{\xi \in \Xi} \min_{\boldsymbol{\beta} \in \Theta} \left(\frac{\det(\mathbf{M}(\xi, \boldsymbol{\beta}))}{\det(\mathbf{M}(\xi_{\boldsymbol{\beta}}^*, \boldsymbol{\beta}))} \right)^{\frac{1}{p}}. \quad (3.16)$$

Dabei ist $\xi_{\boldsymbol{\beta}}^*$ das lokal D -optimale Design.

Eine weitere Möglichkeit zur Vermeidung der Problematik, dass optimale Designs von den unbekanntem Parametern abhängen, sind Bayessche Vorbewertungen für die Parameter. Bei diesen Pseudo-Bayesschen Optimalitätskriterien wird eine Gewichtung in Form einer Wahrscheinlichkeitsverteilung auf einer Parametermenge Θ für die Parameter angenommen. Es wird der Erwartungswert bezüglich dieser vorgegebenen Gewichtung von einer Funktion der Informationsmatrix maximiert oder minimiert. Im Folgenden wird bei diesen Kriterien auf den Begriff „Pseudo“ verzichtet.

Es existieren verschiedene Bayessche Versionen für D -Optimalität (vgl. Firth und Hinde (1997), Atkinson et al. (2007, S. 293)). Hier werden die folgenden zwei Versionen betrachtet.

Definition 3.9 (Bayessche D -Optimalität)

Ein Design ξ^* mit regulärer Informationsmatrix $\mathbf{M}(\xi^*, \boldsymbol{\beta})$ für alle $\boldsymbol{\beta} \in \Theta$ heißt Bayessches D -optimales Design nach Version (I) bzw. (II) bezüglich einer Gewichtung π für die Parameter, wenn

$$(I) \quad E_{\pi} [\log \det(\mathbf{M}(\xi^*, \boldsymbol{\beta}))] \geq E_{\pi} [\log \det(\mathbf{M}(\xi, \boldsymbol{\beta}))] \quad (3.17)$$

$$(II) \quad \log E_{\pi} [\det(\mathbf{M}(\xi^*, \boldsymbol{\beta}))^{-1}] \leq \log E_{\pi} [\det(\mathbf{M}(\xi, \boldsymbol{\beta}))^{-1}] \quad (3.18)$$

für alle $\xi \in \Xi$ gilt.

Definition 3.10 (Bayessche c -Optimalität)

Ein Design ξ^* heißt Bayessches c -optimales Design bezüglich einer Gewichtung π für die Parameter, wenn der lineare Aspekt $\mathbf{c}^T \boldsymbol{\beta}$ für alle $\boldsymbol{\beta} \in \Theta$ identifizierbar ist, d. h. es muss $\mathbf{c} = \mathbf{M}(\xi^*, \boldsymbol{\beta}) \mathbf{y}(\boldsymbol{\beta})$ mit einem Vektor $\mathbf{y}(\boldsymbol{\beta}) \in \mathbb{R}^p$ für alle $\boldsymbol{\beta} \in \Theta$ gelten, und wenn

$$E_{\pi} [\mathbf{c}^T \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1} \mathbf{c}] \leq E_{\pi} [\mathbf{c}^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{c}] \quad (3.19)$$

für alle $\xi \in \Xi$ gilt, für die $\mathbf{c}^T \boldsymbol{\beta}$ für alle $\boldsymbol{\beta} \in \Theta$ identifizierbar ist.

Für die Bayesschen Kriterien kann der Satz von Carathéodory bei nichtlinearen Modellen nicht mehr angewendet werden (vgl. Chaloner und Verdinelli, 1995). Somit ist die Anzahl der Stützpunkte eines Bayesschen optimalen Designs nicht durch $\frac{1}{2}p(p+1)+1$ begrenzt.

3.4 Äquivalenzsätze

Eines der wichtigsten Werkzeuge zur Konstruktion optimaler Designs und zum Nachweis der Optimalität ist der Äquivalenzsatz. Im Folgenden wird angenommen, dass die Kriteriumsfunktion Φ maximiert werden soll. Die Kriteriumsfunktion Φ sei konkav auf \mathcal{M}_β , d. h. für alle $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{M}_\beta$ und $\alpha \in [0, 1]$ gilt

$$\Phi(\alpha\mathbf{M}_1 + (1 - \alpha)\mathbf{M}_2) \geq \alpha\Phi(\mathbf{M}_1) + (1 - \alpha)\Phi(\mathbf{M}_2). \quad (3.20)$$

Gilt in (3.20) „>“ für alle $\alpha \in (0, 1)$ und Matrizen $\mathbf{M}_1 \neq \mathbf{M}_2$, so heißt Φ streng konkav. Da die Kriteriumsfunktion Φ eine Funktion von Matrizen ist, wird eine Ableitung für Matrizen betrachtet. Die Fréchet-Ableitung von Φ bei \mathbf{M}_1 in Richtung von \mathbf{M}_2 ist definiert als (vgl. Silvey, 1980, S. 18)

$$F_\Phi(\mathbf{M}_1, \mathbf{M}_2) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} [\Phi((1 - \varepsilon)\mathbf{M}_1 + \varepsilon\mathbf{M}_2) - \Phi(\mathbf{M}_1)]. \quad (3.21)$$

Die Funktion Φ ist differenzierbar bei \mathbf{M}_1 genau dann, wenn $F_\Phi(\mathbf{M}_1, \cdot)$ linear in der zweiten Komponente ist. Es gilt der folgende allgemeine Äquivalenzsatz (vgl. Silvey, 1980, S. 54):

Satz 3.11 (Äquivalenzsatz)

Es sei β fest. Die Kriteriumsfunktion Φ sei konkav auf \mathcal{M}_β und differenzierbar bei $\mathbf{M}(\xi^*, \beta)$. Das Design ξ^* maximiert $\Phi(\mathbf{M}(\xi, \beta))$ genau dann, wenn die Ungleichung $F_\Phi(\mathbf{M}(\xi^*, \beta), \mathbf{I}(\mathbf{x}, \beta)) \leq 0$ für alle $\mathbf{x} \in \mathcal{X}$ gilt. An den Stützpunkten von ξ^* gilt Gleichheit.

Viele Kriterien lassen sich mit Hilfe geeigneter monotoner Transformationen der Kriteriumsfunktion in konkave Versionen überführen. So ist $\Phi(\mathbf{M}) = \log \det(\mathbf{M})$ sogar eine streng konkave Version für D -Optimalität auf der Menge $\text{PD}(p)$ der symmetrischen positiv definiten $p \times p$ -Matrizen (vgl. Fedorov, 1972, S. 71), woraus die Eindeutigkeit der optimalen Informationsmatrix für D -Optimalität folgt (vgl. Silvey, 1980, S. 17). Die ϕ_k -Kriterien sind streng konkav auf $\text{PD}(p)$ für $k \in (-\infty, 1)$ in dem Sinne, dass in (3.20) „>“ für alle $\mathbf{M}_1, \mathbf{M}_2 \in \text{PD}(p)$ mit $\mathbf{M}_2 \neq \alpha\mathbf{M}_1$, $\alpha \geq 0$, gilt (vgl. Pukelsheim, 1993, S. 116, 151). Auch für sie ist die optimale Informationsmatrix eindeutig. Die Kriterien

für L - und E -Optimalität sind konkav auf $\text{PD}(p)$ (vgl. Pronzato und Pázman, 2013, S. 116–117).

Für die verschiedenen Optimalitätskriterien lassen sich die folgenden Äquivalenzsätze angeben (vgl. Kiefer (1974), Pukelsheim (1993, S. 180)):

Satz 3.12 (Äquivalenzsätze)

Ein Design ξ^* (bzw. ξ_k^*) mit regulärer Informationsmatrix $\mathbf{M}(\xi^*, \boldsymbol{\beta})$ (bzw. $\mathbf{M}(\xi_k^*, \boldsymbol{\beta})$) ist

(a) D -optimal, (b) L -optimal, (c) ϕ_k -optimal für $k \in (-\infty, 1]$ genau dann, wenn für alle $\mathbf{x} \in \mathcal{X}$ gilt:

$$(a) \quad Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}) \leq p, \quad (3.22)$$

$$(b) \quad Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1} \mathbf{B} \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}) \leq \text{Spur}(\mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1} \mathbf{B}), \quad (3.23)$$

$$(c) \quad Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{k-1} \mathbf{f}(\mathbf{x}) \leq \text{Spur}(\mathbf{M}(\xi_k^*, \boldsymbol{\beta})^k). \quad (3.24)$$

An den Stützpunkten von ξ^* (bzw. ξ_k^*) gilt Gleichheit.

Die jeweilige Funktion auf der linken Seite dieser Ungleichungen wird als Sensitivitätsfunktion bezeichnet. Da c -Optimalität mit $\mathbf{B} = \mathbf{c}\mathbf{c}^T$ als Spezialfall von L -Optimalität aufgefasst werden kann, gilt der Äquivalenzsatz auch für c -Optimalität bei regulärer Informationsmatrix. In der Bedingung (3.23) muss $\mathbf{B} = \mathbf{c}\mathbf{c}^T$ gesetzt werden.

Der Fall $k = -\infty$ für E -Optimalität ist bei den ϕ_k -Kriterien ausgenommen. Hier existiert der folgende Äquivalenzsatz (vgl. Pukelsheim, 1993, S. 182):

Satz 3.13 (Äquivalenzsatz für E -Optimalität)

Ein Design ξ^* mit regulärer Informationsmatrix $\mathbf{M}(\xi^*, \boldsymbol{\beta})$ ist genau dann E -optimal, wenn eine positiv semidefinite Matrix \mathbf{E} mit $\text{Spur}(\mathbf{E}) = 1$ existiert, sodass für alle $\mathbf{x} \in \mathcal{X}$ gilt:

$$Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{E} \mathbf{f}(\mathbf{x}) \leq \lambda_{\min}(\mathbf{M}(\xi^*, \boldsymbol{\beta})). \quad (3.25)$$

An den Stützpunkten von ξ^* gilt Gleichheit.

Liegt Interesse darin, nicht den gesamten Parametervektor, sondern nur einzelne Linearkombinationen $\mathbf{A}^T \boldsymbol{\beta}$ der Parameter bestmöglich zu schätzen, wobei \mathbf{A} eine $p \times s$ -Matrix mit $\text{Rang}(\mathbf{A}) = s < p$ ist, so sind optimale Designs mit singulärer Informationsmatrix möglich. Es bezeichnen $\text{PSD}(p)$ die Menge der symmetrischen positiv semidefiniten $p \times p$ -Matrizen und $\mathcal{N}_{\mathbf{A}} = \{\mathbf{M} \in \text{PSD}(p) : \exists \mathbf{Y} \in \mathbb{R}^{p \times s} \text{ mit } \mathbf{A} = \mathbf{M}\mathbf{Y}\}$ die Menge der

symmetrischen positiv semidefiniten Matrizen, die die Identifizierbarkeitsbedingung erfüllen. Typische Kriteriumsfunktionen sind von der Form

$$\Phi(\mathbf{M}) = \begin{cases} \psi(\mathbf{A}^T \mathbf{M}^{-1} \mathbf{A}) & \text{für } \mathbf{M} \in \mathcal{N}_{\mathbf{A}} \\ -\infty & \text{sonst} \end{cases}, \quad (3.26)$$

wobei ψ eine endliche Funktion auf $\text{PD}(s)$ ist und die Definition von Φ hier von \mathcal{M}_{β} auf die Menge $\text{PSD}(p)$ erweitert wurde. Die Funktion Φ wird als konkav angenommen. Für viele Kriterien ist sie differenzierbar bei regulären Matrizen und nicht differenzierbar bei singulären Matrizen. So erfordern die Äquivalenzsätze 3.12 eine reguläre Informationsmatrix. Sie lassen sich nicht anwenden, wenn die Informationsmatrix eines optimalen Designs singulär ist, wie es beispielsweise bei c -Optimalität geschehen kann. Silvey (1978) und Pukelsheim und Titterington (1983) haben folgenden Äquivalenzsatz für den Fall singulärer Informationsmatrizen hergeleitet.

Satz 3.14 (Äquivalenzsatz für singuläre Matrizen)

Es sei β fest. Die Kriteriumsfunktion Φ sei konkav auf $\mathcal{N}_{\mathbf{A}}$ und differenzierbar bei regulären Matrizen. Ein Design ξ^* , dessen Informationsmatrix $\mathbf{M}(\xi^*, \beta) \in \mathcal{M}_{\beta} \cap \mathcal{N}_{\mathbf{A}}$ singulär mit Rang $r < p$ ist, maximiert $\Phi(\mathbf{M}(\xi, \beta))$ auf $\mathcal{M}_{\beta} \cap \mathcal{N}_{\mathbf{A}}$ genau dann, wenn eine $p \times (p - r)$ -Matrix \mathbf{H} mit $\text{Rang}(\mathbf{H}) = p - r$ existiert, so dass $\mathbf{M}(\xi^*, \beta) + \mathbf{H}\mathbf{H}^T$ regulär ist und $F_{\Phi}(\mathbf{M}(\xi^*, \beta) + \mathbf{H}\mathbf{H}^T, \mathbf{I}(\mathbf{x}, \beta)) \leq 0$ für alle $\mathbf{x} \in \mathcal{X}$ gilt. An den Stützpunkten von ξ^* gilt Gleichheit.

Dieser Satz gilt insbesondere für c -Optimalität, deren Kriteriumsfunktion konkav auf $\mathcal{N}_{\mathbf{A}}$ ist (vgl. Pronzato und Pázman, 2013, S. 116).

Für standardisierte Maximin D -Optimalität gilt der folgende Äquivalenzsatz (vgl. Dette et al., 2007).

Satz 3.15 (Äquivalenzsatz für standardisierte Maximin D -Optimalität)

Es sei ξ^* ein Design mit regulärer Informationsmatrix $\mathbf{M}(\xi^*, \beta)$ für alle $\beta \in \Theta$ und es sei

$$\mathcal{N}(\xi^*) := \left\{ \beta \in \Theta \left| \left(\frac{\det(\mathbf{M}(\xi^*, \beta))}{\det(\mathbf{M}(\xi_{\beta}^*, \beta))} \right)^{\frac{1}{p}} = \min_{\beta \in \Theta} \left(\frac{\det(\mathbf{M}(\xi^*, \beta))}{\det(\mathbf{M}(\xi_{\beta}^*, \beta))} \right)^{\frac{1}{p}} \right. \right\} \quad (3.27)$$

die Menge aller Parameterwerte β , bei denen die D -Effizienz des Designs ξ^* minimal ist. Das Design ξ^* ist standardisiert Maximin D -optimal genau dann, wenn eine a-priori-

Verteilung π auf der Menge $\mathcal{N}(\xi^*)$ existiert, sodass die Ungleichung

$$\int_{\mathcal{N}(\xi^*)} Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}) \pi(d\boldsymbol{\beta}) \leq p \quad (3.28)$$

für alle $\mathbf{x} \in \mathcal{X}$ gilt. An den Stützpunkten von ξ^* gilt Gleichheit.

Auch für die Bayesschen Optimalitätskriterien lassen sich Äquivalenzsätze herleiten (vgl. Chaloner und Larntz (1989), Firth und Hinde (1997)).

Satz 3.16 (Äquivalenzsätze für Bayes-Kriterien)

Ein Design ξ^* mit regulärer Informationsmatrix $\mathbf{M}(\xi^*, \boldsymbol{\beta})$ für alle $\boldsymbol{\beta} \in \Theta$ ist (a) ein Bayessches D -optimales Design nach Version (I), (b) ein Bayessches D -optimales Design nach Version (II), (c) ein Bayessches c -optimales Design bezüglich einer Gewichtung π genau dann, wenn

$$(a) \quad E_\pi [Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x})] - p \leq 0 \quad (3.29)$$

$$(b) \quad \frac{E_\pi \left[\det(\mathbf{M}(\xi^*, \boldsymbol{\beta}))^{-1} \cdot Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}) \right]}{E_\pi \left[\det(\mathbf{M}(\xi^*, \boldsymbol{\beta}))^{-1} \right]} - p \leq 0 \quad (3.30)$$

$$(c) \quad E_\pi \left[Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot (\mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1} \mathbf{c})^2 \right] - E_\pi [\mathbf{c}^T \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1} \mathbf{c}] \leq 0 \quad (3.31)$$

für alle $\mathbf{x} \in \mathcal{X}$ gilt. An den Stützpunkten von ξ^* gilt Gleichheit.

3.5 Resultate zu Optimalitätskriterien

In diesem Abschnitt werden einige Charakterisierungen optimaler Designs, Eigenschaften von und Zusammenhänge zwischen Optimalitätskriterien dargestellt.

Für D -Optimalität lässt sich zeigen, dass die Determinante bei einem Randpunkt von \mathcal{M}_β maximiert wird (vgl. Fedorov, 1972, S. 79). Daher existiert ein D -optimales Design mit höchstens $\frac{1}{2} p(p+1)$ Stützpunkten. Bei c -Optimalität existiert ein optimales Design mit höchstens p Stützpunkten (vgl. Pukelsheim, 1993, S. 190).

Eine geometrische Charakterisierung c -optimaler Designs stammt von Elfving (1952). Dabei wird die konvexe Hülle

$$\mathcal{R} := \text{conv} \left(\{ \tilde{\mathbf{f}}(x) \mid x \in \mathcal{X} \} \cup \{ -\tilde{\mathbf{f}}(x) \mid x \in \mathcal{X} \} \right) \quad (3.32)$$

mit $\tilde{\mathbf{f}}(\mathbf{x}) := \sqrt{Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta})} \cdot \mathbf{f}(\mathbf{x})$ betrachtet, die als Elfving-Menge bezeichnet wird. Es gilt der folgende Satz (vgl. Schwabe, 1996, S. 14).

Satz 3.17 (Satz von Elfving)

Ein Design ξ^* mit Stützpunkten $\mathbf{x}_1, \dots, \mathbf{x}_m$ und Gewichten $\omega_1, \dots, \omega_m$ ist c -optimal für einen Aspekt $\mathbf{c}^T \boldsymbol{\beta}$ genau dann, wenn Zahlen $\gamma > 0$ und $\varepsilon_1, \dots, \varepsilon_m \in \{-1, 1\}$ existieren, so dass $\gamma \mathbf{c}$ ein Randpunkt der Elfving-Menge \mathcal{R} ist und

$$\gamma \mathbf{c} = \sum_{i=1}^m \omega_i \varepsilon_i \tilde{\mathbf{f}}(\mathbf{x}_i). \quad (3.33)$$

Die Bedingung (3.33) besagt, dass der Randpunkt $\gamma \mathbf{c}$ die Konvexkombination von zu den Designpunkten zugehörigen Punkten der Elfving-Menge \mathcal{R} ist und die Gewichte dieser Konvexkombination die c -optimalen Gewichte sind.

Der kleinste Eigenwert λ_{\min} einer symmetrischen Matrix \mathbf{A} lässt sich über die Formel

$$\lambda_{\min} = \min_{\mathbf{y}: \|\mathbf{y}\|=1} \mathbf{y}^T \mathbf{A} \mathbf{y} \quad (3.34)$$

berechnen. Dabei bezeichnet $\|\cdot\|$ die euklidische Norm. Bei E -Optimalität handelt es sich somit um ein maximin-Kriterium. Es wird die Varianz der am schlechtesten geschätzten Linearkombination $\mathbf{y}^T \hat{\boldsymbol{\beta}}$ minimiert (vgl. Atkinson et al., 2007, S. 135–136). Der folgende Satz zeigt, dass das E -optimale Design mit einem c -optimalen Design übereinstimmt, wenn der kleinste Eigenwert Vielfachheit 1 besitzt (vgl. Pukelsheim und Studden, 1993).

Satz 3.18

Es sei ξ^* ein Design mit regulärer Informationsmatrix $\mathbf{M}(\xi^*, \boldsymbol{\beta})$ und es sei \mathbf{z} ein Eigenvektor zum kleinsten Eigenwert λ_{\min} von $\mathbf{M}(\xi^*, \boldsymbol{\beta})$. Hat λ_{\min} Vielfachheit 1, so ist das Design ξ^* genau dann E -optimal, wenn es c -optimal für $\mathbf{c} = \mathbf{z}$ ist.

Auch für E -Optimalität existiert eine geometrische Interpretation, falls der kleinste Eigenwert Vielfachheit 1 hat. So zeigt der Vektor $\sqrt{\lambda_{\min}} \cdot \mathbf{z} / \|\mathbf{z}\|$ auf den Berührungspunkt der größten Kugel innerhalb der Elfving-Menge \mathcal{R} mit dem Rand der Elfving-Menge, d. h. die Menge $\{\mathbf{y} : \|\mathbf{y}\| \leq \sqrt{\lambda_{\min}}\}$ ist die größte Kugel innerhalb der Elfving-Menge und der Vektor $\sqrt{\lambda_{\min}} \cdot \mathbf{z} / \|\mathbf{z}\|$ liegt auf dem Rand der Elfving-Menge (vgl. Dette und Studden, 1993). Solch ein Vektor wird als Inkugel-Vektor bezeichnet.

Lemma 3.19

Die Informationsmatrix lässt sich zerlegen in $\mathbf{M}(\xi, \boldsymbol{\beta}) = \tilde{\mathbf{X}}^T \mathbf{D}_\omega \tilde{\mathbf{X}}$ mit Diagonalmatrix

$D_\omega = \text{diag}(\omega_1, \dots, \omega_m)$, $\tilde{\mathbf{X}} = (\tilde{\mathbf{f}}(\mathbf{x}_1), \dots, \tilde{\mathbf{f}}(\mathbf{x}_m))^T$ und $\tilde{\mathbf{f}}(\mathbf{x}) := \sqrt{Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta})} \cdot \mathbf{f}(\mathbf{x})$. Die Matrix $\tilde{\mathbf{X}}$ lässt sich schreiben als $\tilde{\mathbf{X}} = \mathbf{Q}^{\frac{1}{2}} \mathbf{X}$ mit $\mathbf{X} = (\mathbf{f}(\mathbf{x}_1), \dots, \mathbf{f}(\mathbf{x}_m))^T$ und $\mathbf{Q}^{\frac{1}{2}} = \text{diag}(\sqrt{Q(\mathbf{f}(\mathbf{x}_1)^T \boldsymbol{\beta})}, \dots, \sqrt{Q(\mathbf{f}(\mathbf{x}_m)^T \boldsymbol{\beta})})$. Folglich gilt:

$$\mathbf{M}(\xi, \boldsymbol{\beta}) = \tilde{\mathbf{X}}^T \mathbf{D}_\omega \tilde{\mathbf{X}} = \mathbf{X}^T \mathbf{Q}^{\frac{1}{2}} \mathbf{D}_\omega \mathbf{Q}^{\frac{1}{2}} \mathbf{X}. \quad (3.35)$$

Basierend auf dieser Zerlegung der Informationsmatrix liefert der folgende Satz optimale Gewichte für verschiedene Optimalitätskriterien (vgl. Silvey (1980, S. 42), Pukelsheim und Torsney (1991)). Dabei wird vorausgesetzt, dass die Vektoren der Regressionsfunktionen linear unabhängig sind, d. h. die Zeilen der Matrix \mathbf{X} sind linear unabhängig.

Satz 3.20

Die (a) D -optimalen Gewichte eines Designs mit minimalem Träger (d. h. p Stützpunkte), (b) c -optimalen Gewichte für ein Design mit $m \leq p$ Stützpunkten, (c) L -optimalen Gewichte für ein Design mit minimalem Träger, (d) ϕ_k -optimalen Gewichte, $k \in (-\infty, 1]$, für ein Design mit minimalem Träger sind gegeben durch

$$(a) \quad \omega_i^* = \frac{1}{p} \quad \text{für } i = 1, \dots, p, \quad (3.36)$$

$$(b) \quad \omega_i^* = \frac{|v_i|}{\sum_{j=1}^m |v_j|} \quad \text{für } i = 1, \dots, m \quad \text{mit } \mathbf{v} := (\tilde{\mathbf{X}} \tilde{\mathbf{X}}^T)^{-1} \tilde{\mathbf{X}} \mathbf{c}, \quad (3.37)$$

$$(c) \quad \omega_i^* = \frac{\sqrt{s_{ii}}}{\sum_{j=1}^p \sqrt{s_{jj}}} \quad \text{für } i = 1, \dots, p \quad \text{mit } \mathbf{S} := (\tilde{\mathbf{X}}^T)^{-1} \mathbf{B} \tilde{\mathbf{X}}^{-1}, \quad (3.38)$$

$$(d) \quad \omega_i^* = \frac{\sqrt{s_{ii}}}{\sum_{j=1}^p \sqrt{s_{jj}}} \quad \text{für } i = 1, \dots, p \quad \text{mit } \mathbf{S} := (\tilde{\mathbf{X}}^T)^{-1} \mathbf{M}(\xi, \boldsymbol{\beta})^{k+1} \tilde{\mathbf{X}}^{-1}. \quad (3.39)$$

Dabei sind v_i die Einträge des Vektors \mathbf{v} und s_{ii} die Diagonaleinträge der Matrix \mathbf{S} .

Da bei ϕ_k -Optimalität für $k \neq -1$ die Informationsmatrix $\mathbf{M}(\xi, \boldsymbol{\beta})$ und somit die Gewichte in der Matrix \mathbf{S} auftreten, muss (3.39) als Gleichungssystem aufgefasst und nach den Gewichten ω_i^* aufgelöst werden. Es handelt sich hierbei um eine Fixpunktgleichung. Für D -Optimalität mit $k = 0$ erhält man $\mathbf{S} = \text{diag}(\omega_1^*, \dots, \omega_p^*)$ und damit die Lösungen $\omega_i^* = 1/p$ für $i = 1, \dots, p$ in Übereinstimmung mit (3.36) (vgl. Pukelsheim und Torsney, 1991). Ferner hat Atwood (1973) gezeigt, dass für die Gewichte eines D -optimalen Designs immer $\omega_i^* \leq 1/p$ gilt.

Wie bei D -Optimalität sind die optimalen Gewichte eines standardisiert Maximin D -optimalen oder Bayesschen D -optimalen Designs mit minimalem Träger gegeben durch $1/p$.

3.6 Optimales Design für multiple Regression

Für das in diesem Abschnitt betrachtete multiple Regressionsmodell mit $p - 1$ Kovariablen, $p \geq 3$, sind die Informationsmatrizen von der Form (3.7). Zunächst gelte $\mathbf{f}(\mathbf{x}) = (1, \mathbf{x}^T)^T$ mit $\mathbf{x} = (x_1, \dots, x_{p-1})^T \in \mathcal{X} \subset \mathbb{R}^{p-1}$. Als Designraum wird ein mehrdimensionales Polytop betrachtet, welches als konvexe Hülle endlich vieler Punkte definiert ist und somit insbesondere beschränkt ist. Für solche Designräume lässt sich das allgemeine Resultat herleiten, dass für viele Optimalitätskriterien die Stützpunkte eines optimalen Designs auf den Kanten des Designraums liegen (Schmidt und Schwabe, 2017).

Lemma 3.21

Die Menge \mathcal{X} sei ein mehrdimensionales Polytop und die Funktion $Q \not\equiv 0$ sei nichtnegativ. Das Maximum der Funktion

$$d(\mathbf{x}) = Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{A} \mathbf{f}(\mathbf{x}) \tag{3.40}$$

mit positiv definiten Matrix \mathbf{A} wird nur auf den Kanten von \mathcal{X} angenommen.

Beweis:

Es sei $\mathbf{A} = (a_{ij})_{i,j=1,\dots,p}$ und \mathbf{A}_{11} die Teilmatrix von \mathbf{A} , die durch Streichen der ersten Zeile und der ersten Spalte entsteht. Es gilt

$$p(\mathbf{x}) := \mathbf{f}(\mathbf{x})^T \mathbf{A} \mathbf{f}(\mathbf{x}) = \mathbf{x}^T \mathbf{A}_{11} \mathbf{x} + 2\mathbf{a}^T \mathbf{x} + a_{11}$$

mit $\mathbf{a} = (a_{12}, \dots, a_{1p})^T$. Da \mathbf{A} positiv definit ist, ist auch \mathbf{A}_{11} positiv definit und somit ist die Funktion $p(\mathbf{x})$ streng konvex. Für beliebiges $\eta \in \mathbb{R}$ wird die Funktion d auf der Hyperebene $H_\eta = \{\mathbf{x} \in \mathbb{R}^{p-1} : \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} = \eta\}$ betrachtet. Auf H_η ist $Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) = Q(\eta)$ konstant und $d|_{H_\eta}(\mathbf{x}) = Q(\eta)p(\mathbf{x})$ wird aufgrund der strengen Konvexität der Funktion $p(\mathbf{x})$ an den Ecken von $\mathcal{X} \cap H_\eta$ maximal. Es folgt, dass d an den Kanten von \mathcal{X} maximal wird. □

Aus dem Beweis folgt für einen beliebigen Designraum \mathcal{X} , dass die Funktion d auf dem Rand des Designraums maximal wird. Solche Funktionen treten in den Äquivalenzsätzen vieler Optimalitätskriterien auf, wie z. B. bei D -, L - und ϕ_k -Optimalität in Satz 3.12. Es kann nun direkt der nächste Satz gefolgert werden, der ein allgemeines Strukturresultat für ein multiples Regressionsmodell angibt.

Satz 3.22

Die Voraussetzungen von Lemma 3.21 seien erfüllt. Ist die Bedingung im Äquivalenzsatz von der Form

$$Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{A} \mathbf{f}(\mathbf{x}) \leq c \quad (3.41)$$

mit positiv definiten Matrix \mathbf{A} und Konstante $c > 0$, so müssen die Stützpunkte eines optimalen Designs auf den Kanten des Designraums \mathcal{X} liegen.

In Satz 3.22 dürfen die Matrix $\mathbf{A} = \mathbf{A}(\xi^*)$ und die Konstante $c = c(\xi^*)$ von dem optimalen Design ξ^* abhängen. So gilt beispielsweise für A -Optimalität $\mathbf{A} = \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-2}$ und $c = \text{Spur}(\mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1})$. Satz 3.22 ist somit gültig für viele Optimalitätskriterien wie die allgemeine Klasse der ϕ_k -Kriterien für $k \in (-\infty, 1]$, die auch D -Optimalität einschließen.

D -optimale Designs für multiple Regressionsmodelle mit Stützpunkten auf den Kanten des Designraums wurden von Russell et al. (2009) für das Poisson-Modell mit beliebig vielen Kovariablen und von Kabera et al. (2015) für das logistische Regressionsmodell mit zwei Kovariablen berechnet.

Satz 3.22 kann auf unbeschränkte Designräume erweitert werden, d. h. auf Polyeder, die als Durchschnitt endlich vieler Halbräume definiert sind. Ein typisches Beispiel für solch ein Polyeder ist ein Quadrant. Dann liegen die Stützpunkte eines optimalen Designs auf den Achsen.

Korollar 3.23

Der Designraum \mathcal{X} sei ein mehrdimensionales Polyeder und die Bedingung im Äquivalenzsatz sei von der Form (3.41). Sind die Hyperebenen $H_\eta = \{\mathbf{x} \in \mathbb{R}^{p-1} : \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} = \eta\}$ auf \mathcal{X} beschränkt, so ist ein Design ξ^* genau dann optimal, wenn die Bedingung (3.41) auf den Kanten des Designraums erfüllt ist. Die Stützpunkte von ξ^* liegen dann auf den Kanten des Designraums.

Die Forderung der Beschränktheit von $\mathcal{X} \cap H_\eta$ sichert im Fall eines unbeschränkten Polyeders, dass die Sensitivitätsfunktion auf den Kanten des Designraums maximal wird. Sonst existiert kein optimales Design, da die Sensitivitätsfunktion nach oben unbeschränkt ist.

Die bisherigen Resultate setzen die positive Definitheit der Matrix \mathbf{A} im Äquivalenzsatz voraus. Ist diese Matrix nur positiv semidefinit, so können optimale Designs mit Stützpunkten im Inneren des Designraums existieren. Dies ist bei c -Optimalität der Fall

und wird in Abschnitt 5.2 gezeigt. Das folgende Resultat zeigt, dass immer ein optimales Design mit Stützpunkten nur auf den Kanten existiert.

Satz 3.24

Der Designraum \mathcal{X} sei ein mehrdimensionales Polytop. Zu jedem Design ξ existiert ein Design ξ' , dessen Stützpunkte nur auf den Kanten liegen, mit $\mathbf{M}(\xi', \boldsymbol{\beta}) \geq \mathbf{M}(\xi, \boldsymbol{\beta})$.

Beweis:

Es sei $\tilde{\mathbf{x}}$ ein Stützpunkt, der nicht auf einer Kante liegt und es sei $\eta = \mathbf{f}(\tilde{\mathbf{x}})^T \boldsymbol{\beta}$. Es seien $\mathbf{x}_1, \dots, \mathbf{x}_m$ die Schnittpunkte der Hyperebene $H_\eta = \{\mathbf{x} \in \mathbb{R}^{p-1} : \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} = \eta\}$ mit den Kanten des Designraums. Dann ist $\tilde{\mathbf{x}} = \sum_{i=1}^m \alpha_i \mathbf{x}_i$ eine Konvexkombination der \mathbf{x}_i mit $\alpha_i \in (0, 1)$ für mindestens zwei $i \in \{1, \dots, m\}$ und es gilt $\mathbf{f}(\tilde{\mathbf{x}}) = \sum_{i=1}^m \alpha_i \mathbf{f}(\mathbf{x}_i)$. Nach Lemma 8.4 in Pukelsheim (1993, S. 190–191) folgt

$$\mathbf{I}(\tilde{\mathbf{x}}, \boldsymbol{\beta}) = Q(\eta) \mathbf{f}(\tilde{\mathbf{x}}) \mathbf{f}(\tilde{\mathbf{x}})^T \leq Q(\eta) \sum_{i=1}^m \alpha_i \mathbf{f}(\mathbf{x}_i) \mathbf{f}(\mathbf{x}_i)^T = \sum_{i=1}^m \alpha_i \mathbf{I}(\mathbf{x}_i, \boldsymbol{\beta})$$

und insbesondere $\mathbf{I}(\tilde{\mathbf{x}}, \boldsymbol{\beta}) \neq \sum_{i=1}^m \alpha_i \mathbf{I}(\mathbf{x}_i, \boldsymbol{\beta})$. Ersetzt man die Stützpunkte eines Designs ξ , die sich nicht auf einer Kante befinden, durch solche Konvexkombinationen von Punkten auf den Kanten wie oben, so erhält man ein Design ξ' mit $\mathbf{M}(\xi', \boldsymbol{\beta}) \geq \mathbf{M}(\xi, \boldsymbol{\beta})$. \square

Alle Kriteriumsfunktionen Φ sind isoton, d. h. es gilt $\Phi(\mathbf{M}_1) \geq \Phi(\mathbf{M}_2)$ für $\mathbf{M}_1 \geq \mathbf{M}_2$ (vgl. Pronzato und Pázman, 2013, S. 118), wenn die Kriteriumsfunktion maximiert werden soll. Gilt sogar $\Phi(\mathbf{M}_1) > \Phi(\mathbf{M}_2)$ für $\mathbf{M}_1 \geq \mathbf{M}_2$ mit $\mathbf{M}_1 \neq \mathbf{M}_2$, so heißt die Kriteriumsfunktion Φ streng isoton. Beispielsweise sind D -Optimalität und die ϕ_k -Kriterien, $k \in (-\infty, 1]$, streng isoton auf $\text{PD}(p)$ (vgl. Pukelsheim, 1993, S. 151). Für streng isotone Kriterien folgt aus dem Beweis von Satz 3.24 das folgende Korollar.

Korollar 3.25

Der Designraum \mathcal{X} sei ein mehrdimensionales Polytop. Ist die Kriteriumsfunktion streng isoton, so kann ein optimales Design Stützpunkte nur auf den Kanten des Designraums haben.

Unter Verwendung geeigneter Bedingungen für die Intensitätsfunktion Q kann das folgende Resultat zur Anzahl der Stützpunkte je Kante gezeigt werden, das auch für den Fall $p = 2$ mit einer Kovariablen gültig ist.

Lemma 3.26

Der Designraum \mathcal{X} sei ein Polyeder und die Bedingungen (A1) und (A3) seien erfüllt. Ist die Bedingung im Äquivalenzsatz von der Form (3.41) mit positiv semidefiniter Matrix \mathbf{A} und Konstante $c > 0$, so hat ein optimales Design höchstens zwei Stützpunkte je Kante.

Beweis:

Es seien $p(\mathbf{x}) := \mathbf{f}(\mathbf{x})^T \mathbf{A} \mathbf{f}(\mathbf{x})$ und $h(\mathbf{x}) := p(\mathbf{x}) - c/Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta})$. Dann ist Bedingung 3.41 äquivalent zu $h(\mathbf{x}) \leq 0$. Es wird die Funktion h auf einer beliebigen Kante $K := \{\mathbf{x} \in \mathbb{R}^{p-1} : \mathbf{x} = \mathbf{s} + y\mathbf{r}, y \geq 0\}$ betrachtet. Dabei ist \mathbf{s} eine zur Kante K inzidente Ecke und \mathbf{r} gibt die Richtung der Kante an. Bei $\tilde{p}(y) := p(\mathbf{s} + y\mathbf{r})$ handelt es sich um ein quadratisches Polynom, dessen zweite Ableitung konstant ist. Da die zweite Ableitung von $1/Q$ nach (A3) injektiv ist, ist auch die zweite Ableitung der Funktion $\tilde{h}(y) := h(\mathbf{s} + y\mathbf{r})$ injektiv und kann höchstens eine Nullstelle haben. Nach dem Satz von Rolle kann \tilde{h}' höchstens zwei Nullstellen haben und folglich weist die Funktion \tilde{h} höchstens zwei Extrema auf. Ein optimales Design kann daher höchstens zwei Stützpunkte je Kante haben. \square

In diesem Abschnitt wurden multiple Regressionsmodelle mit $\mathbf{f}(\mathbf{x}) = (1, \mathbf{x}^T)^T$ betrachtet. Nun gelte allgemeiner $\mathbf{f}(\mathbf{x}) = (1, f_1(x_1), \dots, f_{p-1}(x_{p-1}))^T$ mit stetigen Funktionen $f_i, i = 1, \dots, p-1$. Der Designraum sei das Polytop $\mathcal{X} = [u_1, v_1] \times \dots \times [u_{p-1}, v_{p-1}]$. Dieser Fall kann mit der Transformation $z_i = f_i(x_i), i = 1, \dots, p-1$, auf den vorigen zurückgeführt werden. Im transformierten Modell gilt $\mathbf{f}(\mathbf{z}) = (1, \mathbf{z}^T)^T$ mit $\mathbf{z} = (z_1, \dots, z_{p-1})^T$. Der transformierte Designraum $Z = [\tilde{u}_1, \tilde{v}_1] \times \dots \times [\tilde{u}_{p-1}, \tilde{v}_{p-1}]$ behält seine Struktur. Dabei gilt $\tilde{u}_i = \min \{f_i(x_i) : x_i \in [u_i, v_i]\}$ und $\tilde{v}_i = \max \{f_i(x_i) : x_i \in [u_i, v_i]\}$ für $i = 1, \dots, p-1$. Somit lässt sich aus dem optimalen Design für das transformierte Modell das optimale Design für das ursprüngliche Modell herleiten. Sind die Funktionen $f_i: [u_i, v_i] \rightarrow [\tilde{u}_i, \tilde{v}_i]$ bijektiv, so können die Stützpunkte eindeutig zurücktransformiert werden. Sind die Funktionen f_i jedoch nicht bijektiv, so kann es mehrere optimale Designs im ursprünglichen Modell geben. Dies wird anhand eines Beispiels in Abschnitt 4.6 gezeigt. Für diese additiven Modelle müssen die Stützpunkte eines optimalen Designs nicht auf den Kanten des Designraums liegen. Sie können sich auch im Inneren des Designraums befinden, wenn die Extrema der Funktionen f_i im Inneren des Intervalls $[u_i, v_i]$ liegen.

4 D -optimale Designs

Dieses Kapitel beschäftigt sich mit der Bestimmung von D -optimalen Designs. In Abschnitt 4.1 wird zunächst allgemein gezeigt, wie für ein multiples Regressionsmodell D -optimale Designs konstruiert werden können. In den darauf folgenden Abschnitten werden für die in dieser Arbeit betrachtete Modellklasse D -optimale Designs berechnet. Abschnitt 4.2 behandelt D -optimale Designs für das Modell mit einer Kovariablen. Anschließend werden für ein multiples Regressionsmodell in Abschnitt 4.3 D -optimale Designs und in Abschnitt 4.4 standardisiert Maximin D -optimale Designs hergeleitet. Bayessche D -optimale Designs werden in Abschnitt 4.5 behandelt. In Abschnitt 4.6 werden die am Ende von Abschnitt 3.6 diskutierten additiven Modelle mit nur jeweils einem Effektparameter je erklärender Variable betrachtet und in Abschnitt 4.7 wird ein Interaktionsterm berücksichtigt.

4.1 Allgemeine Ergebnisse

In diesem Abschnitt wird ein multiples Regressionsmodell mit $p - 1$ Kovariablen, $p \geq 3$, und rechteckigem Designraum betrachtet. Die Informationsmatrizen seien von der Form (3.7) mit $\mathbf{f}(\mathbf{x}) = (1, \mathbf{x}^T)^T$ und $\mathbf{x} = (x_1, \dots, x_{p-1})^T \in \mathcal{X}$. Der Parametervektor sei gegeben durch $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^T$.

Zur Berechnung D -optimaler Designs sind die D -optimalen Designs in den marginalen Modellen mit einer einzigen Kovariablen hilfreich. Für die marginalen Modelle gilt $\mathbf{f}_{M,i}(x_i) = (1, x_i)^T$ für $i = 1, \dots, p - 1$. Der Parametervektor im i -ten marginalen Modell ist zweidimensional und wird mit $\tilde{\boldsymbol{\beta}}_i$ bezeichnet. Das D -optimale Design im Gesamtmodell mit $p - 1$ Kovariablen kann aus den D -optimalen Designs in den marginalen Modellen mit einer einzigen Kovariablen konstruiert werden (Schmidt und Schwabe, 2017). Der folgende Satz gibt das Resultat zunächst für den Fall identischer Designräume in den marginalen Modellen an. Dabei bezeichnet $\mathbf{0}_{p-1} = (0, \dots, 0)^T$ den $(p - 1)$ -dimensionalen Nullvektor.

Satz 4.1

Sei

$$\xi_i^* = \begin{Bmatrix} x_i^* & 0 \\ 1/2 & 1/2 \end{Bmatrix} \quad (4.1)$$

ein D -optimales Design im marginalen Modell mit einer Kovariablen, Parametervektor $\tilde{\boldsymbol{\beta}}_i = (\beta_0, \beta_i)^T$ und Designraum $\mathcal{X}_i = (-\infty, 0]$, $i = 1, \dots, p-1$. Es sei \mathbf{x}_i^* die Einbettung von x_i^* in den $(p-1)$ -dimensionalen Designraum $\mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_{p-1}$ mit Komponenten $x_{ij}^* = 0$ für $j \neq i$ und $x_{ii}^* = x_i^*$. Unter den Bedingungen von Korollar 3.23 ist das Design

$$\xi^* = \begin{Bmatrix} \mathbf{x}_1^* & \dots & \mathbf{x}_{p-1}^* & \mathbf{0}_{p-1} \\ 1/p & \dots & 1/p & 1/p \end{Bmatrix} \quad (4.2)$$

D -optimal im Gesamtmodell mit $p-1$ Kovariablen und Designraum \mathcal{X} .

Beweis:

Es seien $\mathbf{x}^* = (x_1^*, \dots, x_{p-1}^*)^T$ und $\mathbf{1}_{p-1} = (1, \dots, 1)^T$ $(p-1)$ -dimensionale Vektoren. Die Informationsmatrix des Designs ξ^* ist gegeben durch $\mathbf{M}(\xi^*, \boldsymbol{\beta}) = \mathbf{X}^T \mathbf{Q} \mathbf{X} / p$ mit

$$\mathbf{X} = \begin{pmatrix} 1 & \mathbf{0}_{p-1}^T \\ \mathbf{1}_{p-1} & \text{diag}(\mathbf{x}^*) \end{pmatrix}$$

und $\mathbf{Q} = \text{diag}(Q(\beta_0), Q(\beta_0 + \beta_1 x_1^*), \dots, Q(\beta_0 + \beta_{p-1} x_{p-1}^*))$, wobei $\text{diag}(\mathbf{x})$ die Diagonalmatrix ist, deren Diagonaleinträge gleich den Komponenten von \mathbf{x} sind. Es bezeichne $d(\xi^*, \mathbf{x})$ die Sensitivitätsfunktion für das Design ξ^* . Nach Korollar 3.23 und Satz 3.12 ist das Design ξ^* genau dann D -optimal, wenn für alle $x \leq 0$

$$\begin{aligned} p \geq d(\xi^*, \mathbf{x}e_i) &= Q(\mathbf{f}(\mathbf{x}e_i)^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x}e_i)^T \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}e_i) \\ &= p \cdot Q(\beta_0 + \beta_i x) \cdot (1, \mathbf{x}e_i^T) \mathbf{X}^{-1} \mathbf{Q}^{-1} (\mathbf{X}^T)^{-1} (1, \mathbf{x}e_i^T)^T \end{aligned}$$

für $i = 1, \dots, p-1$ gilt. Dabei bezeichnet \mathbf{e}_i den i -ten Einheitsvektor. Die Inverse der Matrix \mathbf{X} berechnet sich zu

$$\mathbf{X}^{-1} = \begin{pmatrix} 1 & \mathbf{0}_{p-1}^T \\ -\mathbf{y}^* & \text{diag}(\mathbf{y}^*) \end{pmatrix}$$

mit $\mathbf{y}^* = (1/x_1^*, \dots, 1/x_{p-1}^*)^T$. Es gilt $(1, \mathbf{x}e_i^T) \mathbf{X}^{-1} = (1 - x/x_i^*, (x/x_i^*)\mathbf{e}_i^T)$ und somit

folgt

$$d(\xi^*, x\mathbf{e}_i) = p \cdot Q(\beta_0 + \beta_i x) \cdot \left(\frac{\left(1 - \frac{x}{x_i^*}\right)^2}{Q(\beta_0)} + \frac{\left(\frac{x}{x_i^*}\right)^2}{Q(\beta_0 + \beta_i x_i^*)} \right) = p \cdot \frac{1}{2} \cdot d_{M,i}(\xi_i^*, x),$$

wobei $d_{M,i}(\xi_i^*, x)$ die Sensitivitätsfunktion für das Design ξ_i^* im marginalen Modell bezeichnet. Da ξ_i^* im marginalen Modell D -optimal ist, gilt $d_{M,i}(\xi_i^*, x) \leq 2$ für alle $x \leq 0$ und daher $d(\xi^*, x\mathbf{e}_i) \leq p$. Es folgt, dass das Design ξ^* D -optimal ist. \square

Im Folgenden seien $S_1, S_2 \subseteq \{1, \dots, p-1\}$ Indexmengen (nicht notwendigerweise nicht-leer) mit $S_1 \cap S_2 = \emptyset$ und $S_1 \cup S_2 = \{1, \dots, p-1\}$. Der nächste Satz liefert die D -optimalen Designs für den allgemeinen Fall, bei dem die Designräume der marginalen Modelle linksseitig unbeschränkt oder rechtsseitig unbeschränkt sein können.

Satz 4.2

Für $i \in S_1$ und $j \in S_2$ seien $\mathcal{X}_i = (-\infty, a_i]$ und $\mathcal{X}_j = [a_j, \infty)$ die Designräume in den marginalen Modellen mit einer Kovariablen. Die Parameter für die marginalen Modelle seien definiert durch $\tilde{\boldsymbol{\beta}}_k = (\beta_0 + \sum_{l \neq k} \beta_l a_l, \beta_k)^T$ für $k = 1, \dots, p-1$. Es seien

$$\xi_i^* = \begin{Bmatrix} x_i^* & a_i \\ 1/2 & 1/2 \end{Bmatrix} \quad \text{und} \quad \xi_j^* = \begin{Bmatrix} a_j & x_j^* \\ 1/2 & 1/2 \end{Bmatrix} \quad (4.3)$$

D -optimale Designs in den entsprechenden marginalen Modellen, $i \in S_1, j \in S_2$. Für $k = 1, \dots, p-1$ sei \mathbf{x}_k^* definiert durch $x_{kl}^* = a_l$ für $l \neq k$ und $x_{kk}^* = x_k^*$. Ferner sei $\mathbf{a} = (a_1, \dots, a_{p-1})^T$. Unter den Bedingungen von Korollar 3.23 ist das folgende Design D -optimal im Gesamtmodell mit $p-1$ Kovariablen und Designraum $\mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_{p-1}$:

$$\xi^* = \begin{Bmatrix} \mathbf{x}_1^* & \dots & \mathbf{x}_{p-1}^* & \mathbf{a} \\ 1/p & \dots & 1/p & 1/p \end{Bmatrix}. \quad (4.4)$$

Beweis:

Mit der Transformation $z_i = x_i - a_i$ für $i \in S_1$ und $z_j = a_j - x_j$ für $j \in S_2$ lässt sich das Design-Problem auf eine kanonische Version zurückführen, wobei $z_k \in (-\infty, 0]$ für $k = 1, \dots, p-1$ (vgl. Ford et al., 1992). Der Parametervektor $\boldsymbol{\beta}$ wird transformiert zu $\boldsymbol{\beta}_z = (\beta_0 + \sum_{l=1}^{p-1} \beta_l a_l, \beta_{z,1}, \dots, \beta_{z,p-1})^T$, wobei $\beta_{z,i} = \beta_i$ für $i \in S_1$ und $\beta_{z,j} = -\beta_j$ für $j \in S_2$ gilt. Die marginalen Modelle können in der gleichen Weise transformiert werden. Dabei werden die Parametervektoren der marginalen Modelle transformiert zu $\tilde{\boldsymbol{\beta}}_{z,i} = (\beta_0 + \sum_{l=1}^{p-1} \beta_l a_l, \beta_i)^T$ für $i \in S_1$ und $\tilde{\boldsymbol{\beta}}_{z,j} = (\beta_0 + \sum_{l=1}^{p-1} \beta_l a_l, -\beta_j)^T$ für $j \in S_2$. Die

D -optimalen Designs in den transformierten marginalen Modellen sind gegeben durch

$$\xi_{z,i}^* = \begin{Bmatrix} x_i^* - a_i & 0 \\ 1/2 & 1/2 \end{Bmatrix} \quad \text{und} \quad \xi_{z,j}^* = \begin{Bmatrix} a_j - x_j^* & 0 \\ 1/2 & 1/2 \end{Bmatrix}$$

für $i \in S_1$ und $j \in S_2$. Satz 4.1 liefert das D -optimale Design für das kanonische Modell. Die Rücktransformation liefert das D -optimale Design (4.4). \square

Bemerkung 4.3

Gilt $\beta_i > 0$ für $i \in S_1$ und $\beta_j < 0$ für $j \in S_2$ bzw. $\beta_i < 0$ für $i \in S_1$ und $\beta_j > 0$ für $j \in S_2$, so sind die Hyperebenen $H_\eta = \{\mathbf{x} \in \mathbb{R}^{p-1} : \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} = \eta\}$ auf \mathcal{X} beschränkt und somit die Bedingungen von Korollar 3.23 erfüllt.

Bemerkung 4.4

Satz 4.2 kann auch für beschränkte Designräume formuliert werden. Die Designräume der marginalen Modelle sind gegeben durch $\mathcal{X}_i = (-\infty, v_i]$ für $i \in S_1$ und $\mathcal{X}_j = [u_j, \infty)$ für $j \in S_2$. Der Designraum des Gesamtmodells $\mathcal{X} = [u_1, v_1] \times \dots \times [u_{p-1}, v_{p-1}]$ ist beschränkt. Es sei $a_i = v_i$ für $i \in S_1$ und $a_j = u_j$ für $j \in S_2$. Mit dieser Definition können die Vektoren \mathbf{a} und \mathbf{x}_k^* , $k = 1, \dots, p-1$, wie in Satz 4.2 gewählt werden. Dann ist das Design (4.4) D -optimal, wenn zusätzlich $u_i \leq x_i^*$ für $i \in S_1$ und $x_j^* \leq v_j$ für $j \in S_2$ gilt.

Ist der Designraum \mathcal{X} so klein, dass die Stützpunkte für das Gesamtmodell, die aus den optimalen Stützpunkten für die marginalen Modelle konstruiert wurden, außerhalb des Designraums liegen, dann muss es kein D -optimales Design mit minimalem Träger geben. Designs mit zusätzlichen Stützpunkten auf den anderen Kanten können D -optimal werden.

Das folgende Beispiel demonstriert für ein häufig verwendetes Modell die Anwendung von Satz 4.2. Es zeigt, dass die Einschränkung des Designraums bzw. die Wahl eines geeigneten Designraums zur Erfüllung der Voraussetzungen von Satz 4.2 führen kann.

Beispiel 4.5

Es wird das logistische Regressionsmodell betrachtet, bei dem die Funktion Q gegeben ist durch $Q(\theta) = e^\theta / (1 + e^\theta)^2$. Das D -optimale Design für den unbeschränkten Designraum $\mathcal{X}_1 = \mathbb{R}$ im marginalen Modell mit einer Kovariablen und Parametervektor $\tilde{\boldsymbol{\beta}}_i = (0, 1)^T$ hat zwei Stützpunkte $x_1^* = -1,543$ und $x_2^* = 1,543$ mit gleichen Gewichten. Für den beschränkten Designraum $\tilde{\mathcal{X}}_1 = [0, \infty)$ hat das D -optimale Design die zwei Stützpunkte $x_1^* = 0$ und $x_2^* = 2,399$ mit gleichen Gewichten (vgl. Ford et al. (1992), Sebastiani und Settini (1997)). Im letzteren Fall ist Satz 4.2 anwendbar. Für das Modell mit zwei

Kovariablen, Parametervektor $\boldsymbol{\beta} = (0, 1, 1)^T$ und Designraum $\mathcal{X} = [0, \infty) \times [0, \infty)$ ist das folgende Design D -optimal:

$$\xi^* = \left\{ \begin{array}{ccc} (2,399; 0) & (0; 2,399) & (0; 0) \\ 1/3 & 1/3 & 1/3 \end{array} \right\}.$$

Dies stimmt mit den Resultaten von Kabera et al. (2015) überein, die das logistische Regressionsmodell mit zwei Kovariablen betrachtet haben. Die Erweiterung auf eine beliebige Anzahl von Kovariablen ist leicht möglich.

4.2 Modell mit einer Kovariablen

Eine Möglichkeit zur Erfüllung der Voraussetzungen von Satz 4.2 ist die Annahme geeigneter Bedingungen an die Intensitätsfunktion Q . In diesem und den folgenden Abschnitten wird angenommen, dass die Funktion Q die vier Bedingungen (A1)–(A4) erfüllt, wie es bei dem Proportional Hazards Modell mit verschiedenen Zensierungsarten der Fall ist. Diese Bedingungen garantieren, dass die D -optimalen Designs in den marginalen Modellen immer 2-Punkt-Designs sind, die einen Randpunkt des Designraums als Stützpunkt haben.

In diesem Abschnitt wird das Modell mit einer Kovariablen betrachtet, d. h. es gilt $p = 2$, $\mathbf{f}(x) = (1, x)^T$ und $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$. Konstantinou et al. (2014) haben D -optimale Designs für ein Modell mit ähnlichen Bedingungen an die Intensitätsfunktion Q bestimmt. Das Resultat ist im folgenden Satz für die hier verwendeten Bedingungen und den Designraum $\mathcal{X} = [u, v]$ angegeben.

Satz 4.6

Es sei $\mathcal{X} = [u, v]$ und die Bedingungen (A1)–(A4) seien erfüllt.

a) Für $\beta_1 > 0$ ist das eindeutige D -optimale Design gegeben durch

$$\xi^* = \left\{ \begin{array}{cc} x_1^* & v \\ 1/2 & 1/2 \end{array} \right\}. \quad (4.5)$$

Dabei ist $x_1^* = u$, falls $\beta_1(v - u) < 2Q(\beta_0 + \beta_1 u)/Q'(\beta_0 + \beta_1 u)$ gilt. Ansonsten ist x_1^* die eindeutige Lösung der Gleichung $\beta_1(v - x_1) - 2Q(\beta_0 + \beta_1 x_1)/Q'(\beta_0 + \beta_1 x_1) = 0$.

b) Für $\beta_1 < 0$ ist das eindeutige D -optimale Design gegeben durch

$$\xi^* = \left\{ \begin{array}{cc} u & x_2^* \\ 1/2 & 1/2 \end{array} \right\}. \quad (4.6)$$

Dabei ist $x_2^* = v$, falls $\beta_1(v - u) > -2Q(\beta_0 + \beta_1 v)/Q'(\beta_0 + \beta_1 v)$ gilt. Ansonsten ist x_2^* die eindeutige Lösung der Gleichung $\beta_1(x_2 - u) + 2Q(\beta_0 + \beta_1 x_2)/Q'(\beta_0 + \beta_1 x_2) = 0$.

Es wird nun die Funktion $\phi_{\mathbf{a}}(\cdot): \mathbb{R} \rightarrow \mathbb{R}$ mit $\mathbf{a} = (a_1, \dots, a_{p-1})^T$ allgemein eingeführt. Sie sei definiert als (vgl. Schmidt und Schwabe, 2017)

$$\phi_{\mathbf{a}}(x) := x - 2 \cdot \frac{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x)}{Q'(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x)}. \quad (4.7)$$

Das folgende Lemma liefert einige wichtige Eigenschaften der Funktion $\phi_{\mathbf{a}}$ (vgl. Schmidt und Schwabe, 2017).

Lemma 4.7

Die Bedingungen (A1), (A2) und (A4) seien erfüllt. Die Funktion $\phi_{\mathbf{a}}$ ist streng monoton wachsend, stetig und bijektiv. Daher existiert die Umkehrfunktion $\phi_{\mathbf{a}}^{-1}$ mit denselben Eigenschaften.

Da in diesem Abschnitt $p = 2$ gilt, ist der Vektor $\mathbf{a} = a$ eindimensional. Mit Hilfe der Funktion ϕ_a lassen sich die Stützpunkte des D -optimalen Designs aus Satz 4.6 kompakt formulieren. Für $\beta_1 > 0$ sind sie gegeben durch $x_1^* = \max(v - \phi_v^{-1}(0)/\beta_1, u)$ und $x_2^* = v$. Für $\beta_1 < 0$ sind sie gegeben durch $x_1^* = u$ und $x_2^* = \min(u - \phi_u^{-1}(0)/\beta_1, v)$. Im Falle $\beta_1 = 0$ ist die Informationsmatrix proportional zu der für das lineare Modell. Das D -optimale Design hat die Stützpunkte u und v mit gleichen Gewichten.

Nun wird ein diskreter Designraum $\mathcal{X} = \{x_1, \dots, x_k\}$ mit $x_1 < x_2 < \dots < x_k$ betrachtet. Für $a \in (x_1, x_k)$ sei $x^-(a) := \max\{x_i \in \mathcal{X} : x_i < a\}$ das nächstkleinere Element des Designraums und $x^+(a) := \min\{x_i \in \mathcal{X} : x_i \geq a\}$ das nächstgrößere Element des Designraums, wobei hier Gleichheit mit a zugelassen wird.

Schmidt und Schwabe (2015) haben gezeigt, dass es bei einem diskreten Designraum drei verschiedene Typen von D -optimalen Designs geben kann. Die Ergebnisse lassen sich in der hier verwendeten Notation mit der Funktion ϕ_a folgendermaßen zusammenfassen.

Satz 4.8

Es sei $\mathcal{X} = \{x_1, \dots, x_k\}$ mit $x_1 < \dots < x_k$. Die Bedingungen (A1)–(A4) seien erfüllt.
a) Es sei $\beta_1 > 0$ und $x^* := v - \phi_v^{-1}(0)/\beta_1$. Gilt $x^* \leq x_1$, so hat das D -optimale Design die Stützpunkte x_1 und x_k . Für $x^* > x_1$ sind als D -optimales Design nur drei Designs mit folgenden Stützpunkten möglich:

- i) $\{x^+(x^*), x_k\}$
- ii) $\{x^-(x^*), x_k\}$
- iii) $\{x^-(x^*), x^+(x^*), x_k\}$.

b) Es sei $\beta_1 < 0$ und $x^* := u - \phi_u^{-1}(0)/\beta_1$. Gilt $x^* \geq x_k$, so hat das D -optimale Design die Stützpunkte x_1 und x_k . Für $x^* < x_k$ sind als D -optimales Design nur drei Designs mit folgenden Stützpunkten möglich:

$$\text{i) } \{x_1, x^-(x^*)\} \quad \text{ii) } \{x_1, x^+(x^*)\} \quad \text{iii) } \{x_1, x^-(x^*), x^+(x^*)\}.$$

c) Für $\beta_1 = 0$ hat das D -optimale Design die Stützpunkte x_1 und x_k .

Schmidt und Schwabe (2015) haben ferner weitere Bedingungen für die D -Optimalität der einzelnen Designs sowie die optimalen Gewichte des 3-Punkt-Designs angegeben.

4.3 Multiple Regression

Mit Hilfe von Satz 4.2 lassen sich aus den D -optimalen Designs aus Satz 4.6 D -optimale Designs für ein multiples Regressionsmodell mit $p - 1$ Kovariablen und Parametervektor $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^T$ herleiten, wobei $p \geq 3$ gelte (Schmidt und Schwabe, 2017).

Satz 4.9

Es sei $\mathcal{X} = [u_1, v_1] \times \dots \times [u_{p-1}, v_{p-1}]$ und die Bedingungen (A1)–(A4) seien erfüllt. Für $i = 1, \dots, p-1$ sei $a_i = v_i$, falls $\beta_i > 0$ und $a_i = u_i$, falls $\beta_i < 0$. Es sei $\mathbf{a} = (a_1, \dots, a_{p-1})^T$. Gilt $\phi_{\mathbf{a}}^{-1}(0) \leq |\beta_i| (v_i - u_i)$ für $i = 1, \dots, p - 1$, so ist das Design

$$\boldsymbol{\xi}^* = \left\{ \begin{array}{cccc} \mathbf{x}_1^* & \mathbf{x}_2^* & \dots & \mathbf{x}_p^* \\ 1/p & 1/p & \dots & 1/p \end{array} \right\} \quad (4.8)$$

mit Stützpunkten $\mathbf{x}_i^* = \mathbf{a} - (\phi_{\mathbf{a}}^{-1}(0)/\beta_i) \mathbf{e}_i$, $i = 1, \dots, p - 1$, und $\mathbf{x}_p^* = \mathbf{a}$ D -optimal.

Beweis:

Für $\beta_i > 0$ hat das D -optimale Design im marginalen Modell mit einer Kovariablen, Parametervektor $\tilde{\boldsymbol{\beta}}_i = (\beta_0 + \sum_{k \neq i} \beta_k a_k, \beta_i)^T$ und Designraum $\mathcal{X}_i = (-\infty, v_i]$ zwei Stützpunkte x_1^* und $x_2^* = v_i$ mit gleichen Gewichten. Der Stützpunkt x_1^* ist nach Satz 4.6 die eindeutige Lösung der Gleichung

$$\beta_i \cdot (v_i - x_1) - 2 \cdot \frac{Q(\beta_0 + \sum_{k \neq i} \beta_k a_k + \beta_i x_1)}{Q'(\beta_0 + \sum_{k \neq i} \beta_k a_k + \beta_i x_1)} = 0.$$

Mit $z = v_i - x_1$ ergibt sich die Gleichung zu $\phi_{\mathbf{a}}(\beta_i z) = 0$ mit Lösung $z = \phi_{\mathbf{a}}^{-1}(0)/\beta_i$. Es folgt $x_1^* = v_i - \phi_{\mathbf{a}}^{-1}(0)/\beta_i$.

Für $\beta_i < 0$ hat das D -optimale Design im marginalen Modell mit einer Kovariablen,

Parametervektor $\tilde{\boldsymbol{\beta}}_i = (\beta_0 + \sum_{k \neq i} \beta_k a_k, \beta_i)^T$ und Designraum $\mathcal{X}_i = [u_i, \infty)$ zwei Stützpunkte $x_1^* = u_i$ und $x_2^* = u_i - \phi_{\mathbf{a}}^{-1}(0)/\beta_i$ mit gleichen Gewichten.

Die Ungleichungen $\phi_{\mathbf{a}}^{-1}(0) \leq |\beta_i|(v_i - u_i)$ für $i = 1, \dots, p-1$ sichern in beiden Fällen, dass die Stützpunkte innerhalb des Designraums liegen. Satz 4.9 folgt aus Satz 4.2 und Bemerkungen 4.3 und 4.4. \square

Bemerkung 4.10

Es genügt, dass die Intensitätsfunktion Q die Bedingungen (A1)–(A4) nur auf dem Intervall $[\mathbf{f}(\bar{\mathbf{a}})^T \boldsymbol{\beta}, \mathbf{f}(\mathbf{a})^T \boldsymbol{\beta}]$ erfüllt, wobei $\bar{\mathbf{a}} = (\bar{a}_1, \dots, \bar{a}_{p-1})^T$ mit $\bar{a}_i \in \{u_i, v_i\} \setminus \{a_i\}$, d. h. \bar{a}_i bezeichnet den anderen Randpunkt des Intervalls $[u_i, v_i]$. Somit lassen sich die Resultate auch für die Funktionen $Q(\theta) = e^\theta/(1 + e^\theta)^2$, $\theta \in (-\infty, 0)$, für logistische Regression und $Q(\theta) = \theta^n$, $\theta \in (0, \infty)$ angeben.

Bemerkung 4.11

Die Nullstelle $\phi_{\mathbf{a}}^{-1}(0)$ der Funktion $\phi_{\mathbf{a}}$ liegt im Intervall $(0, 2Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})/Q'(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})]$. Dies folgt wegen $\phi_{\mathbf{a}}(0) < 0$ und $\phi_{\mathbf{a}}(2Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})/Q'(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})) \geq 0$ (nach (A4)) aus dem Zwischenwertsatz.

Die Aussage von Satz 4.9 ist nur gültig, falls alle $\beta_1, \dots, \beta_{p-1}$ von null verschieden sind. Ist eines dieser β_i gleich null, dann muss es kein optimales Design mit minimalem Träger geben. Optimale Designs lassen sich dann als Produktdesign von einem Design mit gleichen Gewichten auf den Ecken für diejenigen Komponenten, für die $\beta_k = 0$ ist, und einem Design von der Struktur der optimalen Designs aus Satz 4.9 für diejenigen Komponenten mit $\beta_j \neq 0$ konstruieren. Für letztere Komponenten können die Stützpunkte von den optimalen marginalen Designpunkten aus Satz 4.9 abweichen. Dies wird in Beispiel 4.13 gezeigt.

Satz 4.9 besagt, dass ein Stützpunkt des D -optimalen Designs die Ecke \mathbf{a} ist, die in Abhängigkeit von den Vorzeichen der Parameter $\beta_1, \dots, \beta_{p-1}$ bestimmt wird. Sind alle diese Parameter positiv, so gilt $\mathbf{a} = \mathbf{v} = (v_1, \dots, v_{p-1})^T$. Die anderen $p-1$ Stützpunkte befinden sich auf den $p-1$ Kanten, die inzident zur Ecke \mathbf{a} sind, mit Abstand $\phi_{\mathbf{a}}^{-1}(0)/|\beta_i|$ zu \mathbf{a} . Die Ecke \mathbf{a} ist der Punkt im Designraum mit dem höchsten Wert für die Intensitätsfunktion Q . Der Designraum darf auch unbeschränkt sein, sofern die Ecke \mathbf{a} endlich bleibt.

Zur Berechnung des D -optimalen Designs muss also lediglich $\phi_{\mathbf{a}}^{-1}(0)$ bestimmt und somit nur eine einzige Gleichung gelöst werden. Für $Q(\theta) = e^\theta$ erhält man $\phi_{\mathbf{a}}^{-1}(0) = 2$ und damit das Resultat von Russell et al. (2009) für das Poisson-Modell. Für $Q(\theta) = e^\theta/(e^\theta + \lambda)$

mit einer Konstante λ , wobei es sich um die Intensitätsfunktion eines Proportional Hazards Modells mit exponentialverteilten Zensierungszeiten oder eines Negativ-Binomial-Modells bzw. Poisson-Gamma-Modells handelt, hat die Gleichung $\phi_{\mathbf{a}}(x) = 0$ ebenfalls eine analytische Lösung. Es gilt $Q'(\theta) = \lambda e^\theta / (e^\theta + \lambda)^2$ und somit $Q(\theta)/Q'(\theta) = (e^\theta + \lambda)/\lambda$. Es folgt:

$$\begin{aligned} \phi_{\mathbf{a}}(x) &= x - 2 \cdot \frac{e^{\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x} + \lambda}{\lambda} = 0 \\ \iff & (x - 2) \cdot e^{x-2} = \frac{2}{\lambda} \cdot e^{\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - 2} \\ \iff & x - 2 = W\left(\frac{2}{\lambda} \cdot e^{\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - 2}\right) \\ \iff & x = \phi_{\mathbf{a}}^{-1}(0) = 2 + W\left(\frac{2}{\lambda} \cdot e^{\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - 2}\right). \end{aligned}$$

Hierbei bezeichnet W den oberen Zweig der Lambertschen W -Funktion, die als Umkehrfunktion von $g(w) = we^w$ für $w \geq -1$ definiert ist (vgl. Corless et al., 1996).

Beispiel 4.12

a) Für das Poisson-Gamma-Modell haben Graßhoff et al. (2016) D -optimale Designs für den Fall zweier binärer Kovariablen bestimmt. Sie zeigen, dass unter einer bestimmten Bedingung für die Parameter das Design ξ^* mit Designpunkten $(1, 0)$, $(0, 1)$ und $(0, 0)$ und gleichen Gewichten D -optimal ist. Für den Parametervektor $\boldsymbol{\beta} = (4, -4, -4)^T$ und $\lambda = 1$ ist dieses Design nach Satz 4.9 auch D -optimal auf dem stetigen Designraum $\mathcal{X} = [0, \infty) \times [0, \infty)$. Die D -Effizienz des Produktdesigns mit gleichen Gewichten auf allen vier binären Designpunkten $(0, 0)$, $(1, 0)$, $(0, 1)$ und $(1, 1)$, bei dem es sich um das D -optimale Design im linearen Modell mit $\tilde{\mathcal{X}} = [0, 1] \times [0, 1]$ handelt, ist gegeben durch 0,772. Somit ist das D -optimale 3-Punkt-Design ξ^* dem Produktdesign überlegen, dessen Effizienzverlust durchaus beträchtlich ist.

b) Für das lineare Modell mit drei Kovariablen und $\tilde{\mathcal{X}} = [0, 1]^3$ sind die beiden Designs ξ_1 mit Stützpunkten $(0, 0, 0)$, $(0, 1, 1)$, $(1, 0, 1)$ und $(1, 1, 0)$ und ξ_2 mit Stützpunkten $(1, 1, 1)$, $(1, 0, 0)$, $(0, 1, 0)$ und $(0, 0, 1)$ und jeweils gleichen Gewichten D -optimale Designs mit minimalem Träger. Das Design $\xi_3 = \frac{1}{2}(\xi_1 + \xi_2)$ mit gleichen Gewichten auf allen acht Eckpunkten des Designraums ist ebenfalls D -optimal.

Für das Poisson-Gamma-Modell mit Parametervektor $\boldsymbol{\beta} = (4, -4, -4, -4)^T$ und Designraum $\mathcal{X} = [0, \infty)^3$ ist das Design ξ^* mit Stützpunkten $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ und $(0, 0, 0)$ und gleichen Gewichten nach Satz 4.9 D -optimal. Die optimalen Designs im linearen Modell weisen einen erheblichen Effizienzverlust im Vergleich zum D -optimalen

Design ξ^* auf. Die D -Effizienzen der Designs ξ_1 und ξ_2 sind gegeben durch 0,117 und 0,192. Das 8-Punkt-Design ξ_3 erreicht noch eine D -Effizienz von 0,533.

Beispiel 4.13

Zur Illustration des Falls, wenn einige $\beta_i, i > 0$, gleich null sind, wird das Poisson-Modell mit drei Kovariablen und $\boldsymbol{\beta} = (0, -1, -1, 0)^T$ betrachtet. Für das Poisson-Modell ist das D -optimale Design unabhängig von β_0 (vgl. Russell et al., 2009). Das numerisch berechnete D -optimale Design auf dem Designraum $\mathcal{X} = [0, 10]^3$ ist gegeben durch das Produktdesign

$$\begin{aligned} \xi^* &= \xi_{12}^* \otimes \xi_3^* \\ &= \left\{ \begin{array}{cccccc} (0; 0; 0) & (1,86; 0; 0) & (0; 1,86; 0) & (0; 0; 10) & (1,86; 0; 10) & (0; 1,86; 10) \\ 0,23 & 0,13 & 0,13 & 0,23 & 0,13 & 0,13 \end{array} \right\}. \end{aligned}$$

Dabei weist für die ersten zwei Komponenten das Design ξ_{12}^* der Ecke $(0; 0)$ das Gewicht 0,46 und beiden Achsenpunkten $(0; 1,86)$ und $(1,86; 0)$ das Gewicht 0,27 zu, und für die dritte Komponente weist das Design ξ_3^* den Endpunkten 0 und 10 des Designraums $[0, 10]$ gleiche Gewichte 0,5 zu. Die Stützpunkte und Gewichte unterscheiden sich im Vergleich zu dem Produktdesign

$$\xi_{12} \otimes \xi_3^* = \left\{ \begin{array}{cccccc} (0, 0, 0) & (2, 0, 0) & (0, 2, 0) & (0, 0, 10) & (2, 0, 10) & (0, 2, 10) \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{array} \right\},$$

das aus dem optimalen marginalen Design ξ_{12} für $\beta_1 = \beta_2 = -1$ für die ersten zwei Komponenten mit gleichen Gewichten $1/3$ auf der Ecke $(0, 0)$ und den Achsenpunkten $(0, 2)$ und $(2, 0)$ und dem optimalen Design ξ_3^* für $\beta_3 = 0$ für die dritte Komponente konstruiert wurde. Für das optimale marginale Design ξ_{12}^* verschieben sich die Designpunkte auf den Achsen zur Ecke hin, die ein größeres Gewicht erhält. Der Effizienzverlust ist jedoch nicht erheblich. Das aus den optimalen marginalen Designs konstruierte Produktdesign $\xi_{12} \otimes \xi_3^*$ hat eine D -Effizienz von 0,965. Weitere Berechnungen lassen vermuten, dass das Design ξ_{12}^* unabhängig von dem Designraum der dritten Komponente ist.

Für das Produktdesign $\xi_{12} \otimes \xi_3^*$ und die analog konstruierten Produktdesigns für Modelle mit weiteren Kovariablen können Effizienzschranken angegeben werden. Die D -Effizienz eines Designs ξ kann ohne Kenntnis des D -optimalen Designs abgeschätzt werden über (vgl. Atwood (1969), Dette (1996))

$$\text{eff}_D(\xi, \boldsymbol{\beta}) \geq \frac{p}{\max_{\mathbf{x} \in \mathcal{X}} Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x})}. \quad (4.9)$$

Für das Poisson-Modell mit $p - 1$ Kovariablen, Parametervektor $\boldsymbol{\beta} = (0, -\mathbf{1}_{p-2}^T, 0)^T$ und Designraum $\mathcal{X} = [0, 10]^{p-1}$ werden die Produktdesigns

$$\xi_p = \left\{ \begin{array}{cccc} \mathbf{0}_{p-2} & 2\mathbf{e}_{1,p-2} & \cdots & 2\mathbf{e}_{p-2,p-2} \\ \frac{1}{p-1} & \frac{1}{p-1} & \cdots & \frac{1}{p-1} \end{array} \right\} \otimes \left\{ \begin{array}{cc} 0 & 10 \\ 1/2 & 1/2 \end{array} \right\} \quad (4.10)$$

betrachtet, wobei $\mathbf{e}_{i,k}$ den i -ten Einheitsvektor im \mathbb{R}^k bezeichnet und $p \geq 3$ gilt. Numerisch stellt sich heraus, dass das Maximum im Nenner von (4.9) bei $\mathbf{x} = \mathbf{0}_{p-1}$ und $\mathbf{x} = 10\mathbf{e}_{p-1,p-1}$ angenommen wird. Für die Designs ξ_p folgt nach einigen Rechenschritten die Effizienzschranke

$$\text{eff}_D(\xi_p, \boldsymbol{\beta}) \geq \frac{p(e^2 + p - 2)}{(p - 1)(2e^2 + p - 2)}. \quad (4.11)$$

Einzelne Werte sind in Tabelle 4.1 festgehalten. Die tatsächliche D -Effizienz des Designs $\xi_4 = \xi_{12} \otimes \xi_3^*$ ist mit 0,965 deutlich höher als die berechnete Effizienzschranke.

Anzahl Kovariablen	2	3	4	5	6	10	20	50
Effizienzschranke	0,798	0,746	0,730	0,728	0,731	0,758	0,820	0,902

Tabelle 4.1: Effizienzschranken

Beispiel 4.14

Dieses Beispiel behandelt den Fall, wenn die Ungleichungen $\phi_{\mathbf{a}}^{-1}(0) \leq |\beta_i|(v_i - u_i)$ in Satz 4.9 nicht erfüllt sind. Dazu wird das Proportional Hazards Modell mit Typ I Zensurierung, Experimentdauer $c = 1$ und zwei Kovariablen betrachtet. Es sei $\boldsymbol{\beta} = (-7, 3, 3)^T$ und $\mathcal{X} = [1, 2] \times [1, 2]$. Dann gilt $\phi_{\mathbf{v}}^{-1}(0) = 4,604 > 3 \cdot (2 - 1)$ und Satz 4.9 kann nicht angewendet werden.

Das D -optimale Design in der Klasse der 3-Punkt-Designs hat die Stützpunkte $(2, 2)$, $(1, 2)$ und $(2, 1)$ mit gleichen Gewichten. In der Klasse der 4-Punkt-Designs sind die beiden folgenden Designs D -optimal:

$$\xi_1 = \left\{ \begin{array}{cccc} (2; 2) & (1; 2) & (2; 1) & (1; 1,513) \\ 0,318 & 0,255 & 0,318 & 0,108 \end{array} \right\}, \quad (4.12)$$

$$\xi_2 = \left\{ \begin{array}{cccc} (2; 2) & (1; 2) & (2; 1) & (1,513; 1) \\ 0,318 & 0,318 & 0,255 & 0,108 \end{array} \right\}. \quad (4.13)$$

Das global D -optimale Design ist das folgende 5-Punkt-Design:

$$\xi^* = \left\{ \begin{array}{ccccc} (2; 2) & (1; 2) & (2; 1) & (1, 531; 1) & (1; 1, 531) \\ 0,317 & 0,281 & 0,281 & 0,061 & 0,061 \end{array} \right\}. \quad (4.14)$$

In diesem Beispiel sind die D -Effizienzen der optimalen Designs in der Klasse der Designs mit drei bzw. vier Stützpunkten sehr hoch. Das 3-Punkt-Design hat eine D -Effizienz von 0,988. Die beiden 4-Punkt-Designs weisen eine D -Effizienz von 0,9991 auf.

4.4 Standardisiert Maximin D -optimale Designs

In diesem Abschnitt wird weiterhin ein multiples Regressionsmodell mit $p-1$ Kovariablen betrachtet, $p \geq 2$. Für die Parameter β_i , $i = 1, \dots, p-1$, sei ein Wertebereich $[b_{u,i}, b_{o,i}]$ bekannt, der nicht den Wert 0 enthält. Der Parameter β_0 sei fest. Die Designräume $[u_i, v_i]$ für die einzelnen Kovariablen seien derart, dass stets ein Randpunkt des Intervalls bei null liegt. Dabei handele es sich um den Randpunkt, bei dem die Intensitätsfunktion Q maximal wird.

Es werden Designs von der Gestalt

$$\xi = \left\{ \begin{array}{cccc} x_1 \mathbf{e}_1 & \dots & x_{p-1} \mathbf{e}_{p-1} & \mathbf{0}_{p-1} \\ 1/p & \dots & 1/p & 1/p \end{array} \right\}, \quad (4.15)$$

betrachtet, von der auch das lokal D -optimale Design nach Satz 4.9 ist. Es müssen also die optimalen Werte für x_1, \dots, x_{p-1} ermittelt werden.

Für das Modell mit einer Kovariablen haben Konstantinou et al. (2014) das standardisiert Maximin D -optimale Design in der Klasse der 2-Punkt-Designs bestimmt, das von der Gestalt (4.15) ist. Sie haben eine Gleichung zur Berechnung des optimalen Werts für x_1 angegeben. Im nächsten Satz wird das standardisiert Maximin D -optimale Design in der Klasse der Designs von der Gestalt (4.15) bestimmt, wobei gezeigt wird, dass sich die von Konstantinou et al. (2014) angegebene Gleichung vereinfachen lässt.

Satz 4.15

Es sei $\mathcal{X} = [u_1, v_1] \times \dots \times [u_{p-1}, v_{p-1}]$ und die Bedingungen (A1)–(A4) seien erfüllt. Es gelte $\beta_i \in [b_{u,i}, b_{o,i}]$ für $i = 1, \dots, p-1$ und β_0 sei fest. Es gelte entweder $b_{u,i} > 0$ und $v_i = 0$ oder $b_{o,i} < 0$ und $u_i = 0$. Ferner gelte $\phi_{\mathbf{0}_{p-1}}^{-1}(0) \leq |\beta_i| (v_i - u_i)$ für alle $\beta_i \in [b_{u,i}, b_{o,i}]$ und $i = 1, \dots, p-1$.

Für das standardisiert Maximin D -optimale Design in der Klasse der Designs von der Gestalt (4.15) sind die optimalen Werte x_i^* für $i = 1, \dots, p-1$ die Lösungen der folgenden

Gleichungen:

$$\frac{Q(\beta_0 + b_{u,i}x_i)}{Q(\beta_0 + b_{o,i}x_i)} = \frac{b_{o,i}^2}{b_{u,i}^2}. \quad (4.16)$$

Beweis:

Es bezeichne $x_i(\beta_i) := -\phi_{\mathbf{0}_{p-1}}^{-1}(0)/\beta_i$. Unter den gegebenen Voraussetzungen ist das Design mit Stützpunkten $x_i(\beta_i)\mathbf{e}_i$, $i = 1, \dots, p-1$, und $\mathbf{0}_{p-1}$ und gleichen Gewichten nach Satz 4.9 für alle $\beta_i \in [b_{u,i}, b_{o,i}]$ lokal D -optimal. Die Informationsmatrix eines Designs ξ von der Gestalt (4.15) ergibt sich zu $\mathbf{M}(\xi, \boldsymbol{\beta}) = \mathbf{X}^T \mathbf{Q} \mathbf{X} / p$ mit

$$\mathbf{X} = \begin{pmatrix} 1 & \mathbf{0}_{p-1}^T \\ \mathbf{1}_{p-1} & \text{diag}(\mathbf{x}) \end{pmatrix}$$

und $\mathbf{Q} = \text{diag}(Q(\beta_0), Q(\beta_0 + \beta_1 x_1), \dots, Q(\beta_0 + \beta_{p-1} x_{p-1}))$, wobei $\mathbf{x} = (x_1, \dots, x_{p-1})^T$. Die Determinante der Informationsmatrix von ξ berechnet sich zu

$$\det(\mathbf{M}(\xi, \boldsymbol{\beta})) = \frac{1}{p^p} \det(\mathbf{Q}) \det(\mathbf{X})^2 = \frac{1}{p^p} Q(\beta_0) \prod_{i=1}^{p-1} Q(\beta_0 + \beta_i x_i) x_i^2.$$

Damit ergibt sich die D -Effizienz von ξ zu

$$\text{eff}_D(\xi, \boldsymbol{\beta}) = \left(\prod_{i=1}^{p-1} \frac{Q(\beta_0 + \beta_i x_i) x_i^2}{Q(\beta_0 + \beta_i x_i(\beta_i)) x_i(\beta_i)^2} \right)^{\frac{1}{p}}. \quad (4.17)$$

Es sei $k_i(x_i) := Q(\beta_0 + \beta_i x_i) x_i^2 / [Q(\beta_0 + \beta_i x_i(\beta_i)) x_i(\beta_i)^2]$. Bei $k_i(x_i)^{1/2}$ handelt es sich um die D -Effizienz im marginalen Modell mit einer Kovariablen für ein Design mit Stützpunkten x_i und 0 und gleichen Gewichten (vgl. Konstantinou et al., 2014). Da die D -Effizienz von ξ in (4.17) bis auf Exponenten das Produkt der D -Effizienzen in den marginalen Modellen ist, lässt sich das Design-Problem auf das für eine Kovariable zurückführen. Konstantinou et al. (2014) haben für das marginale Modell mit einer Kovariablen gezeigt, dass sich der optimale Wert für x_i als Lösung der folgenden Gleichung ergibt:

$$Q(\beta_0 + b_{u,i}x_i)Q(\beta_0 + b_{o,i}x_i(b_{o,i}))x_i(b_{o,i})^2 = Q(\beta_0 + b_{o,i}x_i)Q(\beta_0 + b_{u,i}x_i(b_{u,i}))x_i(b_{u,i})^2. \quad (4.18)$$

Es gilt $Q(\beta_0 + b_{o,i}x_i(b_{o,i})) = Q(\beta_0 - \phi_{\mathbf{0}_{p-1}}^{-1}(0)) = Q(\beta_0 + b_{u,i}x_i(b_{u,i}))$. Beide Faktoren sind identisch und können aus Gleichung (4.18) eliminiert werden. Unter Berücksichtigung von $x_i(\beta_i) = -\phi_{\mathbf{0}_{p-1}}^{-1}(0)/\beta_i$ lässt sich (4.18) weiter vereinfachen zu Gleichung (4.16). \square

Zur Bestimmung des standardisiert Maximin D -optimalen Designs ist also die Berechnung der Stützpunkte lokal D -optimaler Designs für bestimmte Parameterwerte nicht nötig.

In typischen Anwendungsfällen besitzen die Parameter $\beta_1, \dots, \beta_{p-1}$ das gleiche Vorzeichen. Sind sie positiv, so gilt in Satz 4.15 der Fall $b_{u,i} > 0$ und $v_i = 0$, d. h. es ist $\mathcal{X} = [u_1, 0] \times \dots \times [u_{p-1}, 0]$.

Bemerkung 4.16

Die Ungleichung $\phi_{\mathbf{0}_{p-1}}^{-1}(0) \leq |\beta_i| (v_i - u_i)$ aus Satz 4.15 ist genau dann für alle Parameterwerte $\beta_i \in [b_{u,i}, b_{o,i}]$ erfüllt, wenn für $b_{u,i} > 0$ die Ungleichung $\phi_{\mathbf{0}_{p-1}}^{-1}(0) \leq -b_{u,i}u_i$ gilt bzw. für $b_{o,i} < 0$ die Ungleichung $\phi_{\mathbf{0}_{p-1}}^{-1}(0) \leq -b_{o,i}v_i$ gilt.

Da für das Poisson-Modell optimale Designs unabhängig von β_0 sind, kann für dieses Modell auch ein Wertebereich für β_0 zugelassen werden.

Das folgende Beispiel zeigt einen Fall, bei dem die Designs in den marginalen Modellen standardisiert Maximin D -optimal in der Klasse aller Designs sind und das aus diesen Designs konstruierte Design aus Satz 4.15 für das multiple Regressionsmodell ebenfalls standardisiert Maximin D -optimal ist.

Beispiel 4.17

Es wird das Proportional Hazards Modell mit Typ I Zensierung und Experimentdauer $c = 1$ betrachtet. Der Designraum im marginalen Modell mit einer Kovariablen sei $\mathcal{X}_1 = (-\infty, 0]$. Für den Parameter β_1 sei der Wertebereich $\Theta_1 = [b_{u,1}, b_{o,1}] = [\frac{2}{3}, \frac{3}{2}]$ gegeben und es sei $\beta_0 = 0$. Die Lösung der Gleichung (4.16) berechnet sich numerisch zu $x_1^* = -2,067$. Das Design ξ_1^* mit Stützpunkten $-2,067$ und 0 und gleichen Gewichten ist standardisiert Maximin D -optimal in der Klasse der 2-Punkt-Designs. Konstantinou et al. (2014) haben gezeigt, dass die minimale D -Effizienz bei den Randpunkten von Θ_1 angenommen wird. Es sei π_1 die a-priori-Verteilung auf $b_{u,1} = \frac{2}{3}$ und $b_{o,1} = \frac{3}{2}$ mit zugehörigen Gewichten $0,567$ und $0,433$. Graphisch lässt sich zeigen, dass das Design ξ_1^* standardisiert Maximin D -optimal in der Klasse aller Designs ist. In Abbildung 4.1 ist die Funktion $d(\xi_1^*, x) = \int Q(\mathbf{f}(x)^T \boldsymbol{\beta}) \cdot \mathbf{f}(x)^T \mathbf{M}(\xi_1^*, \boldsymbol{\beta})^{-1} \mathbf{f}(x) \pi_1(d\beta_1)$ aus dem Äquivalenzsatz graphisch dargestellt, wobei die Variable x mittels der nichtlinearen Transformation $z = x/(1-x)$ auf das endliche Intervall $[-1, 0]$ abgebildet wurde.

Nun wird das Modell mit zwei Kovariablen und Designraum $\mathcal{X} = [-5, 0] \times [-5, 0]$ betrachtet. Es sei $\beta_0 = 0$ und für die Parameter β_1 und β_2 seien die Parameterbereiche $\Theta_1 = \Theta_2 = [\frac{2}{3}, \frac{3}{2}]$ gegeben. Die Ungleichungen $\phi_{\mathbf{0}_{p-1}}^{-1}(0) \leq |\beta_i| (v_i - u_i)$ aus Satz 4.15 sind nach Bemerkung 4.16 wegen $\phi_{\mathbf{0}_{p-1}}^{-1}(0) = 2,124 \leq -\frac{2}{3} \cdot (-5) = -b_{u,i}u_i$ für alle $\beta_i \in \Theta_i$,

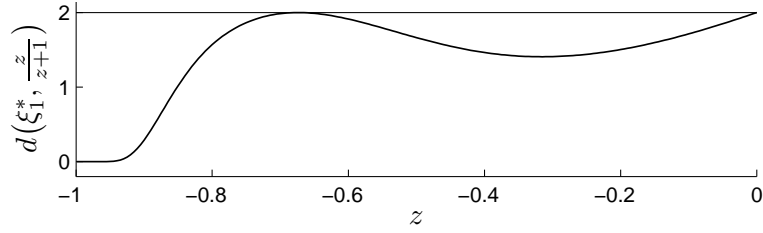


Abbildung 4.1: Graphische Darstellung der Funktion $d(\xi_1^*, \frac{z}{z+1})$

$i = 1, 2$, erfüllt. Es soll nun demonstriert werden, dass das Design

$$\xi^* = \begin{Bmatrix} (-2,067; 0) & (0; -2,067) & (0; 0) \\ 1/3 & 1/3 & 1/3 \end{Bmatrix} \quad (4.19)$$

aus Satz 4.15 standardisiert Maximin D -optimal ist. Da die D -Effizienz des Designs ξ^* nach Gleichung (4.17) bis auf Exponenten das Produkt der D -Effizienzen der standardisiert Maximin D -optimalen Designs $\xi_1^* = \xi_2^*$ in den beiden identischen marginalen Modellen ist, ergibt sich die Menge $\mathcal{N}(\xi^*)$ aus (3.27), die alle Parameterwerte enthält, bei denen die D -Effizienz des Designs ξ^* minimal ist, zu $\mathcal{N}(\xi^*) = \mathcal{N}(\xi_1^*) \times \mathcal{N}(\xi_2^*)$. Es sei π das Produktmaß $\pi = \pi_1 \otimes \pi_2$ auf $\mathcal{N}(\xi^*)$, wobei $\pi_1 = \pi_2$ ist. In Abbildung 4.2 ist die Funktion $d(\xi^*, \mathbf{x}) = \int Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}) \pi(d(\beta_1, \beta_2))$ aus dem Äquivalenzsatz graphisch dargestellt. Es folgt, dass das Design ξ^* standardisiert Maximin D -optimal ist.

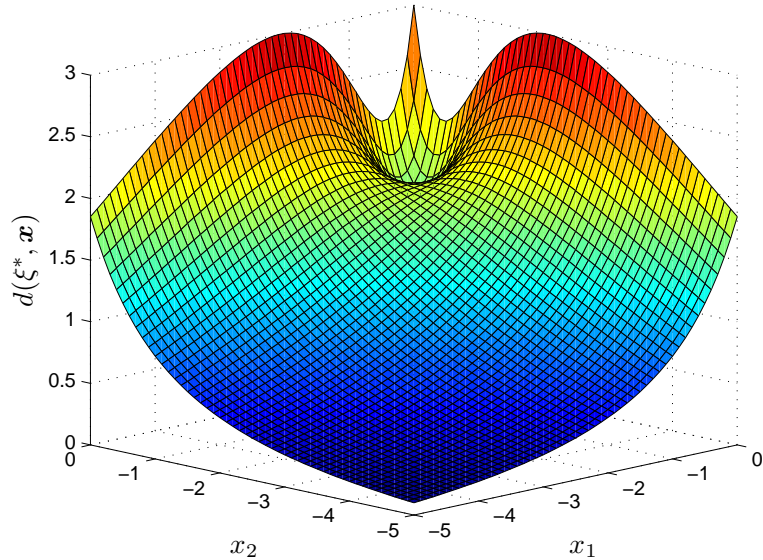


Abbildung 4.2: Graphische Darstellung der Funktion $d(\xi^*, \mathbf{x})$

Eine naheliegende Vermutung ist, dass im multiplen Regressionsmodell die Stützpunkte eines standardisiert Maximin D -optimalen Designs auf den Kanten des Designraums

liegen und sich die optimalen Designs analog zum Vorgehen in Abschnitt 4.1 für D -Optimalität aus den optimalen Designs in den marginalen Modellen konstruieren lassen. Die linke Seite der Ungleichung (3.28) des Äquivalenzsatzes muss jedoch nicht auf den Kanten des Designraums maximal sein. Die Richtung der Hyperebenen H_η , auf denen die Intensitätsfunktion Q konstant ist, hängt von den Parameterwerten ab. Daher können bei der Integration über die Parameter Maxima im Inneren des Designraums auftreten. Entsprechende Resultate wie bei D -Optimalität lassen sich somit nicht analog zeigen. Als Beispiel hierzu wird das Proportional Hazards Modell mit Typ I Zensierung, Experimentdauer $c = 1$ und zwei Kovariablen betrachtet. Es sei π die Verteilung, die den vier Parameterwerten $(0; 0,1; 0,1)$, $(0; 0,1; 5)$, $(0; 5; 0,1)$ und $(0; 5; 5)$ gleiche Gewichte $1/4$ zuweist. Die Funktion $d(\xi, \mathbf{x}) = \int Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}) \pi(d\boldsymbol{\beta})$ für das Design ξ mit Stützpunkten $(-2; 0)$, $(0; -2)$ und $(0; 0)$ und gleichen Gewichten $1/3$ ist in Abbildung 4.3 graphisch dargestellt. Sie wird nicht auf den Kanten, sondern im Inneren des Designraums maximal.

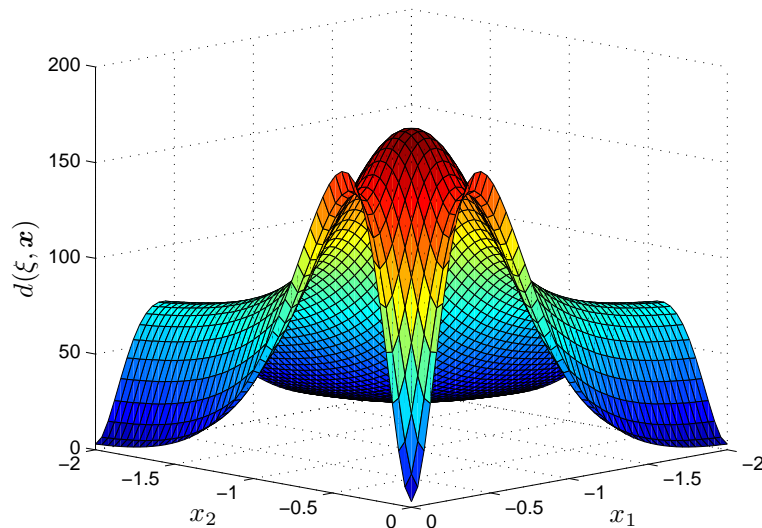


Abbildung 4.3: Graphische Darstellung der Funktion $d(\xi, \mathbf{x})$

4.5 Bayessche D -optimale Designs

In diesem Abschnitt werden Bayessche D -optimale Designs für spezielle Modelle mit zensierten Daten bestimmt. Zunächst werden Bayessche D -optimale Designs nach Version (I) gesucht, die die Kriteriumsfunktion $E_\pi[\log \det(\mathbf{M}(\xi, \boldsymbol{\beta}))]$ maximieren. Dabei wird das Proportional Hazards Modell mit Typ I Zensierung und Experimentdauer $c = 1$ betrachtet.

Für das Modell mit einer Kovariablen seien voneinander unabhängige Gewichtungen $N(0, 1)$ für β_0 und $N(\mu_1, \sigma_1^2)$ für β_1 gegeben, wobei $N(\mu, \sigma^2)$ die Normalverteilung mit Parametern μ und σ^2 bezeichnet. Eine Schwierigkeit ergibt sich dadurch, dass die Integrale der Kriteriumsfunktion nicht analytisch berechnet werden können. Die Berechnung optimaler Designs erfolgt daher numerisch.

Die Integrale werden mittels Quadraturformeln in einer Umgebung von fünf Standardabweichungen um den Mittelwert näherungsweise berechnet. Der Verlust an Wahrscheinlichkeitsmasse beträgt damit nur $0,573 \cdot 10^{-6}$ pro Integral. Es wird die abgeschlossene Newton-Cotes-Formel

$$\int_a^b f(x) dx \approx (b - a) \cdot \sum_{i=0}^n w_i f(x_i) \quad (4.20)$$

verwendet. Dabei sind x_i Stützstellen und w_i Gewichte. Die Stützstellen x_i sind äquidistant auf dem Integrationsintervall $[a, b]$. Sie sind für $i = 0, \dots, n$ gegeben durch $x_i = a + i \cdot (b - a) / n$. Die hier verwendete Newton-Cotes-Formel vom Grad $n = 4$ ist auch als Boole- oder Milne-Regel bekannt und für Polynome bis zum Grad $n + 1 = 5$ exakt. Die Gewichte sind gegeben durch $w_1 = 7/90$, $w_2 = 32/90$, $w_3 = 12/90$, $w_4 = 32/90$ und $w_5 = 7/90$. Um eine bessere Approximationsgenauigkeit zu erreichen, wird das Integrationsintervall in 20 gleich große Teilintervalle unterteilt, auf denen jeweils die Boole-Regel angewendet wird. Anschließend werden die Näherungen addiert.

In Tabelle 4.2 sind die numerisch bestimmten Bayesschen D -optimalen Designs nach Version (I) auf dem Designraum $\mathcal{X} = [0, 1]$ zusammengestellt.

	$\mu_1 = 0$	$\mu_1 = 5$	$\mu_1 = 8$
$\sigma_1^2 = 1$	$\begin{Bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{Bmatrix}$	$\begin{Bmatrix} 0,072 & 1 \\ 1/2 & 1/2 \end{Bmatrix}$	$\begin{Bmatrix} 0,116 & 1 \\ 1/2 & 1/2 \end{Bmatrix}$
$\sigma_1^2 = 3$	$\begin{Bmatrix} 0 & 0,267 & 1 \\ 0,488 & 0,023 & 0,490 \end{Bmatrix}$	$\begin{Bmatrix} 0,059 & 1 \\ 1/2 & 1/2 \end{Bmatrix}$	$\begin{Bmatrix} 0,106 & 1 \\ 1/2 & 1/2 \end{Bmatrix}$
$\sigma_1^2 = 4,5$	$\begin{Bmatrix} 0 & 0,255 & 1 \\ 0,426 & 0,147 & 0,427 \end{Bmatrix}$	$\begin{Bmatrix} 0 & 0,084 & 1 \\ 0,214 & 0,290 & 0,496 \end{Bmatrix}$	$\begin{Bmatrix} 0,094 & 1 \\ 1/2 & 1/2 \end{Bmatrix}$

Tabelle 4.2: Bayessche D -optimale Designs nach Version (I) für das Proportional Hazards Modell mit Typ I Zensierung

Es lässt sich das in der Literatur häufig genannte Phänomen beobachten, dass die Anzahl der Stützpunkte des optimalen Designs mit größer werdender Varianz der Gewichtsverteilung steigt (vgl. Chaloner und Larntz (1989), Atkinson et al. (2007, S. 298–299)). Die aus dem Satz von Carathéodory gewonnene Aussage, dass ein optimales Design mit

höchstens $\frac{1}{2} p(p+1)$ Stützpunkten existiert, gilt nicht mehr. Die Anzahl der Stützpunkte eines optimalen Designs kann sogar beliebig groß werden (vgl. Braess und Dette, 2007). In Abbildung 4.4 ist zu den optimalen Designs aus Tabelle 4.2 die Fréchet-Ableitung graphisch dargestellt.

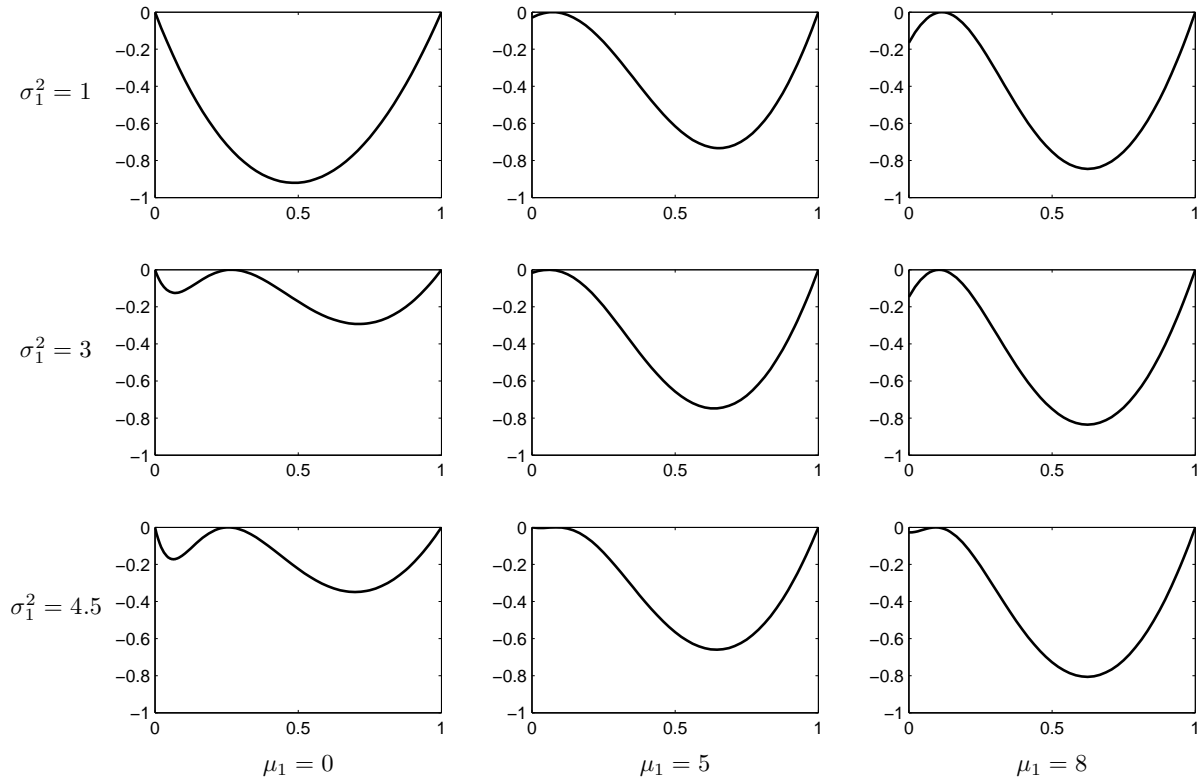


Abbildung 4.4: Fréchet-Ableitung für die optimalen Designs aus Tabelle 4.2

Es können auch andere Gewichtungen für die Parameter verwendet werden. Im folgenden Beispiel wird das Modell mit einer und zwei Kovariablen betrachtet und als Gewichtung eine Gleichverteilung angesetzt.

Beispiel 4.18

Wie in Beispiel 4.17 sei $\mathcal{X}_1 = (-\infty, 0]$ und $\beta_0 = 0$ im marginalen Modell. Für den Parameter β_1 sei als Gewichtung eine Gleichverteilung auf $\Theta_1 = [\frac{2}{3}, \frac{3}{2}]$ gegeben. Das numerisch bestimmte Bayessche D -optimale Design nach Version (I) hat die Stützpunkte $-1,961$ und 0 mit Gewichten $1/2$.

Die Berechnung Bayesscher D -optimaler Designs für Modelle mit mehreren Kovariablen ist in der Regel sehr schwierig. Für das Modell mit zwei Kovariablen und Designraum $\mathcal{X} = [-5, 0] \times [-5, 0]$ seien als Gewichtungen für β_1 und β_2 voneinander unabhängige Gleichverteilungen auf $\Theta_1 = \Theta_2 = [\frac{2}{3}, \frac{3}{2}]$ gegeben. Es sei $\beta_0 = 0$. Analog zum Vorgehen

bei D -Optimalität wird das folgende Design betrachtet:

$$\xi^* = \left\{ \begin{array}{ccc} (-1,961; 0) & (0; -1,961) & (0; 0) \\ 1/3 & 1/3 & 1/3 \end{array} \right\}. \quad (4.21)$$

In Abbildung 4.5 ist die Fréchet-Ableitung für das Design ξ^* graphisch dargestellt. Es zeigt sich, dass das Design ξ^* Bayes-optimal ist.

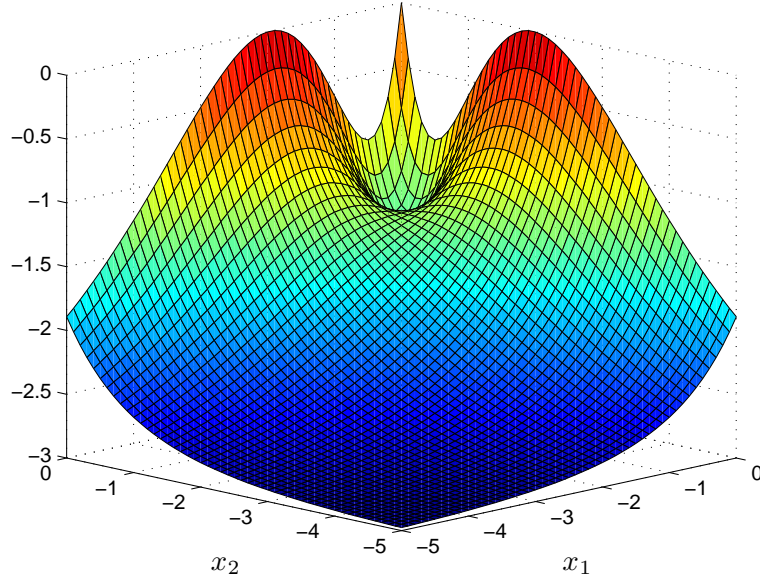


Abbildung 4.5: Fréchet-Ableitung für das Design ξ^* in (4.21)

Wie bei standardisierter Maximin D -Optimalität liegt die Vermutung nahe, dass sich im multiplen Regressionsmodell Bayessche D -optimale Designs nach Version (I) wie in Abschnitt 4.1 aus den optimalen Designs in den marginalen Modellen konstruieren lassen, falls diese 2-Punkt-Designs mit dem Randpunkt 0 des Designraums als Stützpunkt sind. Das Beispiel am Ende von Abschnitt 4.4 für das Vorliegen eines Maximums im Inneren des Designraums der Funktion $d(\xi, \mathbf{x}) = \int Q(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}) \pi(d\boldsymbol{\beta})$ aus dem Äquivalenzsatz zeigt, dass sich auch für das Bayes-Kriterium entsprechende Resultate wie bei D -Optimalität nicht analog beweisen lassen.

Für Bayessche D -Optimalität nach Version (II) ist die zu minimierende Kriteriumsfunktion gegeben durch $\log E_\pi [\det(\mathbf{M}(\xi, \boldsymbol{\beta}))^{-1}]$. Es kann das folgende analytische Resultat für das Proportional Hazards Modell mit exponentialverteilten Zensierungszeiten bzw. Negativ-Binomial-Modell mit einer Kovariablen gezeigt werden, dessen Beweis sich im Anhang befindet. Dabei bezeichnet W den oberen Zweig der Lambertschen W -Funktion.

Satz 4.19

Es sei $Q(\theta) = e^\theta / (e^\theta + \lambda)$ und $\mathcal{X} = [u, v]$. Es seien die voneinander unabhängigen Gewichtungen $N(\mu_0, \sigma_0^2)$ für β_0 und $N(0, \sigma_1^2)$ für β_1 gegeben. Es sei

$$x^* = \frac{1}{\sigma_1} \cdot \sqrt{2 + 2W\left(\frac{1}{2\lambda} (1 + \lambda^2 e^{-2\mu_0 + 2\sigma_0^2}) \cdot e^{\mu_0 - \frac{1}{2}\sigma_0^2 - 1}\right)}. \quad (4.22)$$

Das Bayessche D -optimale Design nach Version (II) in der Klasse der 2-Punkt-Designs ist gegeben durch

$$\xi^* = \left\{ \begin{array}{cc} x_1^* & x_2^* \\ 1/2 & 1/2 \end{array} \right\}. \quad (4.23)$$

Dabei ist $x_1^* = -x^*$ und $x_2^* = x^*$, falls $\pm x^* \in \mathcal{X}$.

Gilt $-x^* \notin \mathcal{X}$ oder $x^* \notin \mathcal{X}$, so seien $C_1 := \lambda e^{-\mu_0 + \frac{1}{2}\sigma_0^2}$, $C_2 := \lambda^2 e^{-2\mu_0 + 2\sigma_0^2}$ und

$$D_a(x) := C_1 e^{\frac{\sigma_1^2 x^2}{2}} (\sigma_1^2 x(x-a) - 2) + C_2 e^{\frac{\sigma_1^2 (a+x)^2}{2}} (\sigma_1^2 (a+x)(x-a) - 2) - 2 - 2C_1 e^{\frac{\sigma_1^2 a^2}{2}}.$$

Für $u+v \geq 0$ ist $x_1^* = u$ und $x_2^* = \min(\tilde{x}, v)$, wobei \tilde{x} die eindeutige Lösung der Gleichung $D_u(x) = 0$ im Intervall (u, ∞) ist.

Für $u+v \leq 0$ ist $x_2^* = v$ und $x_1^* = \max(u, \tilde{x})$, wobei \tilde{x} die eindeutige Lösung der Gleichung $D_v(x) = 0$ im Intervall $(-\infty, v)$ ist.

Bemerkung 4.20

Ist $u = 0$ oder $v = 0$, so hat die Gleichung $D_0(x) = 0$ analytische Lösungen:

$$\begin{aligned} D_0(x) &= 0 \\ \iff (C_1 + C_2) \cdot e^{\frac{1}{2}\sigma_1^2 x^2} (\sigma_1^2 x^2 - 2) &= 2 + 2C_1 \\ \iff e^{\frac{1}{2}\sigma_1^2 x^2 - 1} \left(\frac{1}{2}\sigma_1^2 x^2 - 1 \right) &= \frac{1 + C_1}{e(C_1 + C_2)} \\ \iff \frac{1}{2}\sigma_1^2 x^2 - 1 &= W\left(\frac{1 + C_1}{e(C_1 + C_2)} \right) \\ \iff x &= \pm \frac{1}{\sigma_1} \cdot \sqrt{2 + 2W\left(\frac{1 + C_1}{e(C_1 + C_2)} \right)}. \end{aligned}$$

Ist der Designraum groß genug, so liegen die Stützpunkte des Designs aus Satz 4.19 symmetrisch um den Erwartungswert $\mu_1 = 0$ für β_1 . Numerische Berechnungen zeigen, dass für $\mu_1 \neq 0$ die optimalen Stützpunkte nicht mehr symmetrisch um μ_1 liegen.

Beispiel 4.21

Es wird das Proportional Hazards Modell mit Exp(1)-verteilten Zensierungszeiten und $\mathcal{X} = [0; 1,5]$ betrachtet. Es seien voneinander unabhängige Gewichtungen $N(0, 1)$ für β_0 und $N(0, 1)$ für β_1 gegeben. Dann berechnet sich x^* aus (4.22) zu $x^* = 1,757$. Somit gilt $-x^* \notin \mathcal{X}$ und $x^* \notin \mathcal{X}$. Das Bayessche D -optimale Design nach Version (II) in der Klasse der 2-Punkt-Designs hat nach Satz 4.19 und Bemerkung 4.20 die Stützpunkte $x_1^* = 0$ und $x_2^* = 1,482$. Insbesondere liegt der zweite Stützpunkt nicht beim Randpunkt 1,5.

4.6 Additives Modell mit einem Effektparameter je erklärender Variable

In diesem Abschnitt soll mit Hilfe eines konkreten Beispiels dargestellt werden, wie man D -optimale Designs für Modelle mit Informationsmatrizen von der Form (3.7) mit $\mathbf{f}(\mathbf{x}) = (1, f_1(x_1), \dots, f_{p-1}(x_{p-1}))^T$ berechnen kann.

Es sei $\mathbf{f}(\mathbf{x}) = (1, x_1^2, x_2^2)^T$. Betrachtet wird das Poisson-Modell mit Parametervektor $\beta = (0, -1, -1)^T$ und Designraum $\mathcal{X} = [-\sqrt{v}, \sqrt{v}] \times [-\sqrt{v}, \sqrt{v}]$ mit $v > 2$. Der durch die Transformation $z_i = f_i(x_i) = x_i^2$, $i = 1, 2$, induzierte Designraum ist gegeben durch $\mathcal{Z} = [0, v] \times [0, v]$. Satz 4.9 liefert für das transformierte Modell das D -optimale Design

$$\xi^* = \left\{ \begin{array}{ccc} (2, 0) & (0, 2) & (0, 0) \\ 1/3 & 1/3 & 1/3 \end{array} \right\}. \quad (4.24)$$

Da die Funktionen f_i , $i = 1, 2$, nicht bijektiv sind, ist keine eindeutige Rücktransformation der Stützpunkte möglich. Sowohl $x_i = \sqrt{2}$ als auch $x_i = -\sqrt{2}$ liefern $x_i^2 = 2 = z_i$. Es gibt daher mehrere D -optimale Designs. Da die Kriteriumsfunktion konkav ist, sind alle Konvexkombinationen D -optimaler Designs ebenfalls D -optimal. Somit folgt, dass alle Designs der Gestalt

$$\xi_{\alpha_1, \alpha_2}^* = \left\{ \begin{array}{ccccc} (\sqrt{2}, 0) & (-\sqrt{2}, 0) & (0, \sqrt{2}) & (0, -\sqrt{2}) & (0, 0) \\ \frac{\alpha_1}{3} & \frac{1-\alpha_1}{3} & \frac{\alpha_2}{3} & \frac{1-\alpha_2}{3} & \frac{1}{3} \end{array} \right\} \quad (4.25)$$

mit $\alpha_i \in [0, 1]$, $i = 1, 2$, D -optimal sind und wegen der strengen Konkavität der Kriteriumsfunktion die gleiche Informationsmatrix haben. Alle Stützpunkte dieser D -optimalen Designs befinden sich im Inneren des Designraums. Die Sensitivitätsfunktion für die Designs $\xi_{\alpha_1, \alpha_2}^*$ ist in Abbildung 4.6 für $v = 9$ graphisch dargestellt.

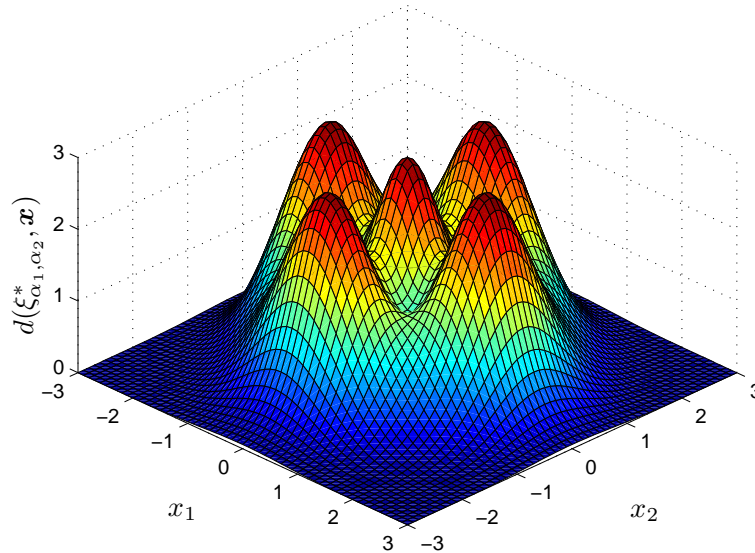


Abbildung 4.6: Sensitivitätsfunktion $d(\xi_{\alpha_1, \alpha_2}^*, \mathbf{x})$ für die Designs $\xi_{\alpha_1, \alpha_2}^*$ in (4.25)

4.7 Modell mit Interaktionsterm

In diesem Abschnitt wird ein Modell mit zwei Kovariablen und einem Interaktionsterm betrachtet. Es gilt $\mathbf{f}(\mathbf{x}) = (1, x_1, x_2, x_1x_2)^T$ und $\mathbf{x} = (x_1, x_2)^T \in \mathcal{X}$. Der Designraum ist gegeben durch $\mathcal{X} = [u_1, v_1] \times [u_2, v_2]$ und der Parametervektor lautet $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$.

Für nichtlineare Modelle mit Interaktionstermen sind kaum analytische Resultate bekannt. Für das Poisson-Modell, das in der hier betrachteten Klasse von Modellen enthalten ist, haben Wang et al. (2006) D -optimale Designs numerisch berechnet und Könnner (2011) hat die Stützpunkte eines optimalen 4-Punkt-Designs für einzelne Fälle bestimmt.

Das folgende Beispiel zeigt, dass die Resultate aus Abschnitt 3.6 nicht für Modelle mit Interaktionstermen gültig sind. Insbesondere müssen die Stützpunkte eines D -optimalen Designs nicht auf den Kanten liegen, sondern können sich auch im Inneren des Designraums befinden.

Beispiel 4.22

Es wird das Proportional Hazards Modell mit Typ I Zensierung und Experimentdauer $c = 1$ betrachtet. Es sei $\mathcal{X} = [-4, 0] \times [-4, 0]$ und der Parametervektor sei gegeben durch $\boldsymbol{\beta} = (0, 1, 1, -1)^T$. Das folgende Design ist D -optimal:

$$\xi^* = \left\{ \begin{array}{cccc} (0; 0) & (-2,124; 0) & (0; -2,124) & (-1,016; -1,016) \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array} \right\}. \quad (4.26)$$

Dies kann mit Hilfe des Äquivalenzsatzes graphisch verifiziert werden. Die Sensitivitätsfunktion ist in Abbildung 4.7 dargestellt.

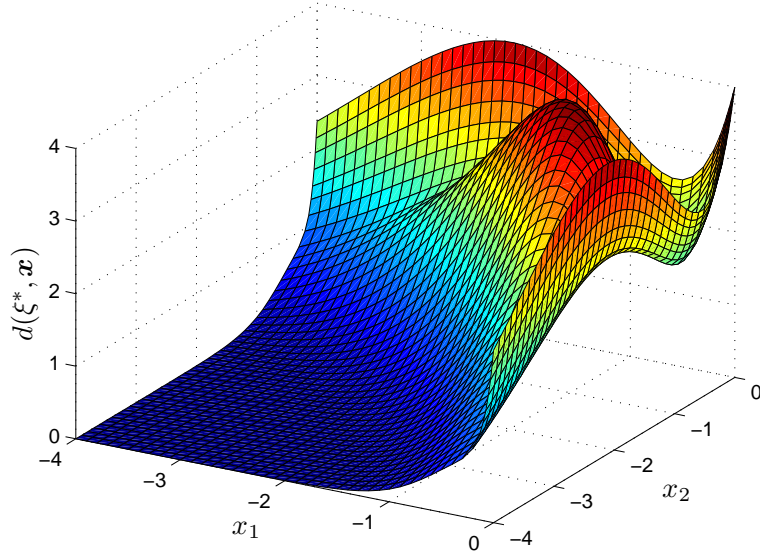


Abbildung 4.7: Sensitivitätsfunktion $d(\xi^*, \mathbf{x})$ für das Design ξ^* in (4.26)

Die Struktur des D -optimalen Designs aus Beispiel 4.22 weist Gemeinsamkeiten und Unterschiede zum D -optimalen Design für das Modell ohne Interaktionsterm aus Satz 4.9 auf. So hat es den gleichen Eckpunkt des Designraums als Stützpunkt und zwei weitere Stützpunkte auf den zu dieser Ecke inzidenten Kanten des Designraums. Zusätzlich gibt es allerdings noch einen vierten Stützpunkt, der sich im Inneren des Designraums befindet. Numerische Berechnungen mit anderen Parameterwerten und Designräumen für den in Beispiel 4.22 betrachteten Fall $\beta_1 > 0$, $\beta_2 > 0$, $\beta_{12} \leq 0$, $v_1 \leq 0$, $v_2 \leq 0$ lassen vermuten, dass das D -optimale Design immer von dieser Gestalt ist, sofern der Designraum groß genug ist. Im vorliegenden Fall mit $\beta_{12} \leq 0$ liegt ein Synergieeffekt vor. Die einzelnen Terme des linearen Prädiktors $\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}$ können gleichzeitig maximiert werden, d. h. der Term mit Interaktion $\beta_{12} x_1 x_2$ wird für die Werte der Kovariablen maximal, die auch die Terme ohne Interaktion $\beta_1 x_1$ und $\beta_2 x_2$ maximieren. Im Folgenden wird der Fall $\beta_1 > 0$, $\beta_2 > 0$, $\beta_{12} \leq 0$, $v_1 \leq 0$, $v_2 \leq 0$ betrachtet und der Ansatz eines Designs von der folgenden Gestalt gemacht:

$$\xi = \left\{ \begin{array}{cccc} (v_1, v_2) & (v_1 - x_1, v_2) & (v_1, v_2 - x_2) & (v_1 - x_3, v_2 - x_4) \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array} \right\}. \quad (4.27)$$

Es müssen also die optimalen Werte für $x_1, \dots, x_4 > 0$ bestimmt werden. Zunächst werden zwei Funktionen eingeführt. Die Funktion $\phi_{\mathbf{a}}(\cdot): \mathbb{R} \rightarrow \mathbb{R}$ mit $\mathbf{a} = (a_1, a_2)^T$ sei

definiert wie in (4.7):

$$\phi_{\mathbf{a}}(x) := x - 2 \cdot \frac{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x)}{Q'(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x)}. \quad (4.28)$$

Der Unterschied zu (4.7) besteht nur in der Definition der Funktion \mathbf{f} . Somit gelten die Aussagen aus Lemma 4.7. Es sei $C := (\beta_2 + \beta_{12}v_1) \cdot (\beta_1 + \beta_{12}v_2)$ und $\mathbf{v} = (v_1, v_2)^T$. Die Funktion $\zeta(\cdot): \mathbb{R} \rightarrow \mathbb{R}$ sei definiert als

$$\zeta(x) := x \cdot \left(1 - \frac{\beta_{12}x}{C}\right) - 2 \cdot \frac{Q(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - x \cdot (2 - \frac{\beta_{12}x}{C}))}{Q'(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - x \cdot (2 - \frac{\beta_{12}x}{C}))}. \quad (4.29)$$

Der folgende Satz, dessen Beweis sich im Anhang befindet, liefert das D -optimale Design in der Klasse der Designs von der Gestalt (4.27).

Satz 4.23

Es sei $\mathcal{X} = [u_1, v_1] \times [u_2, v_2]$ und die Bedingungen (A1), (A2) und (A4) seien erfüllt. Es gelte $\beta_1 > 0$, $\beta_2 > 0$, $\beta_{12} \leq 0$, $v_1 \leq 0$, $v_2 \leq 0$. Es sei \tilde{x} die eindeutige Lösung der Gleichung $\zeta(x) = 0$ im Intervall $(0, \infty)$ und es sei $\mathbf{v} = (v_1, v_2)^T$.

Für das D -optimale Design in der Klasse der Designs von der Gestalt (4.27) gilt

$$x_1^* = \min \left\{ v_1 - u_1, \frac{\phi_{\mathbf{v}}^{-1}(0)}{\beta_1 + \beta_{12}v_2} \right\}, \quad (4.30)$$

$$x_2^* = \min \left\{ v_2 - u_2, \frac{\phi_{\mathbf{v}}^{-1}(0)}{\beta_2 + \beta_{12}v_1} \right\}. \quad (4.31)$$

Für $\tilde{x}/(\beta_1 + \beta_{12}v_2) \leq v_1 - u_1$ und $\tilde{x}/(\beta_2 + \beta_{12}v_1) \leq v_2 - u_2$ gilt $x_3^* = \tilde{x}/(\beta_1 + \beta_{12}v_2)$ und $x_4^* = \tilde{x}/(\beta_2 + \beta_{12}v_1)$.

Für $\tilde{x}/(\beta_1 + \beta_{12}v_2) > v_1 - u_1$ gilt $x_3^* = v_1 - u_1$ und

$$x_4^* = \min \left\{ v_2 - u_2, \frac{\phi_{(u_1, v_2)^T}^{-1}(0)}{\beta_2 + \beta_{12}u_1} \right\}. \quad (4.32)$$

Für $\tilde{x}/(\beta_2 + \beta_{12}v_1) > v_2 - u_2$ gilt $x_4^* = v_2 - u_2$ und

$$x_3^* = \min \left\{ v_1 - u_1, \frac{\phi_{(v_1, u_2)^T}^{-1}(0)}{\beta_1 + \beta_{12}u_2} \right\}. \quad (4.33)$$

Wie im Beweis gezeigt wurde, lässt sich die Gleichung $\zeta(x) = 0$ auch so skalieren, dass als Lösung direkt die optimalen Werte von x_3^* oder x_4^* erzielt werden. Für x_4^* ist sie

gegeben durch

$$\tilde{\zeta}(x) := x\tilde{C}(\beta_1 + \beta_{12}v_2 - \beta_{12}x) - 2 \cdot \frac{Q(\mathbf{f}(\mathbf{v})^T\boldsymbol{\beta} - x\tilde{C}(2\beta_1 + 2\beta_{12}v_2 - \beta_{12}x))}{Q'(\mathbf{f}(\mathbf{v})^T\boldsymbol{\beta} - x\tilde{C}(2\beta_1 + 2\beta_{12}v_2 - \beta_{12}x))} \quad (4.34)$$

mit $\tilde{C} := (\beta_2 + \beta_{12}v_1)/(\beta_1 + \beta_{12}v_2)$. Die Funktion $\zeta(x)$ nimmt aber in der in (4.29) gegebenen Form eine einfachere Gestalt an und die optimalen Werte für x_1^* und x_3^* bzw. für x_2^* und x_4^* haben eine ähnliche Struktur. Es folgen einige Bemerkungen zu Satz 4.23.

Bemerkung 4.24

Es gilt $x_3^* \leq x_1^*$ und $x_4^* \leq x_2^*$.

Beweis:

Es wird nur $x_4^* \leq x_2^*$ gezeigt, da die Ungleichung $x_3^* \leq x_1^*$ analog folgt. Für $x_2^* = v_2 - u_2$ ist die Aussage klar. Sei also $x_2^* = \phi_{\mathbf{v}}^{-1}(0)/(\beta_2 + \beta_{12}v_1)$. Einsetzen von x_2^* in $\tilde{\zeta}$ liefert mit (A4) und $C = (\beta_2 + \beta_{12}v_1) \cdot (\beta_1 + \beta_{12}v_2) > 0$ unter Beachtung von $\beta_{12}\phi_{\mathbf{v}}^{-1}(0)/C < 0$ die folgende Abschätzung:

$$\begin{aligned} \tilde{\zeta}(x_2^*) &= \phi_{\mathbf{v}}^{-1}(0) \cdot \left(1 - \frac{\beta_{12}\phi_{\mathbf{v}}^{-1}(0)}{C}\right) - 2 \cdot \frac{Q(\mathbf{f}(\mathbf{v})^T\boldsymbol{\beta} - \phi_{\mathbf{v}}^{-1}(0) \cdot (2 - \frac{\beta_{12}\phi_{\mathbf{v}}^{-1}(0)}{C}))}{Q'(\mathbf{f}(\mathbf{v})^T\boldsymbol{\beta} - \phi_{\mathbf{v}}^{-1}(0) \cdot (2 - \frac{\beta_{12}\phi_{\mathbf{v}}^{-1}(0)}{C}))} \\ &\geq \phi_{\mathbf{v}}^{-1}(0) - 2 \cdot \frac{Q(\mathbf{f}(\mathbf{v})^T\boldsymbol{\beta} - \phi_{\mathbf{v}}^{-1}(0))}{Q'(\mathbf{f}(\mathbf{v})^T\boldsymbol{\beta} - \phi_{\mathbf{v}}^{-1}(0))} = \phi_{\mathbf{v}}(\phi_{\mathbf{v}}^{-1}(0)) = 0 = \tilde{\zeta}(x_4^*). \end{aligned}$$

Da die Funktion $\tilde{\zeta}(x)$ für $x \geq 0$ streng monoton wachsend ist, folgt $x_2^* \geq x_4^*$.

Es muss noch der Fall $\tilde{x}/(\beta_1 + \beta_{12}v_2) > v_1 - u_1$ aus Satz 4.23 betrachtet werden. Es gilt:

$$\begin{aligned} \phi_{(u_1, v_2)^T}(\phi_{\mathbf{v}}^{-1}(0)) &= \phi_{\mathbf{v}}^{-1}(0) - 2 \cdot \frac{Q(\beta_0 + \beta_1 u_1 + \beta_2 v_2 + \beta_{12} u_1 v_2 - \phi_{\mathbf{v}}^{-1}(0))}{Q'(\beta_0 + \beta_1 u_1 + \beta_2 v_2 + \beta_{12} u_1 v_2 - \phi_{\mathbf{v}}^{-1}(0))} \\ &\geq \phi_{\mathbf{v}}^{-1}(0) - 2 \cdot \frac{Q(\beta_0 + \beta_1 v_1 + \beta_2 v_2 + \beta_{12} v_1 v_2 - \phi_{\mathbf{v}}^{-1}(0))}{Q'(\beta_0 + \beta_1 v_1 + \beta_2 v_2 + \beta_{12} v_1 v_2 - \phi_{\mathbf{v}}^{-1}(0))} \\ &= \phi_{\mathbf{v}}(\phi_{\mathbf{v}}^{-1}(0)) = 0 = \phi_{(u_1, v_2)^T}(\phi_{(u_1, v_2)^T}^{-1}(0)). \end{aligned}$$

Da $\phi_{(u_1, v_2)^T}$ streng monoton wachsend ist, folgt $\phi_{(u_1, v_2)^T}^{-1}(0) \leq \phi_{\mathbf{v}}^{-1}(0)$ und somit

$$x_4^* = \frac{\phi_{(u_1, v_2)^T}^{-1}(0)}{\beta_2 + \beta_{12}u_1} \leq \frac{\phi_{\mathbf{v}}^{-1}(0)}{\beta_2 + \beta_{12}v_1} = x_2^*.$$

Folglich gilt in allen Fällen $x_4^* \leq x_2^*$. □

Aus dem Beweis wird ersichtlich, dass in Bemerkung 4.24 Gleichheit nur für das Poisson-Modell mit $\beta_{12} = 0$ eintreten kann oder wenn der Designraum zu klein ist.

Bemerkung 4.25

a) Gelten die beiden Ungleichungen $\tilde{x}/(\beta_1 + \beta_{12}v_2) > v_1 - u_1$ und $\tilde{x}/(\beta_2 + \beta_{12}v_1) > v_2 - u_2$, so folgt $x_3^* = v_1 - u_1$ und $x_4^* = v_2 - u_2$ und somit sind wegen Bemerkung 4.24 alle Eckpunkte des Designraums Stützpunkte des D -optimalen Designs.

b) Ist der Designraum groß genug, so gilt die folgende Beziehung:

$$\frac{x_1^*}{x_2^*} = \frac{x_3^*}{x_4^*} = \frac{\beta_2 + \beta_{12}v_1}{\beta_1 + \beta_{12}v_2}.$$

In der folgenden Bemerkung sind weitere Kombinationen von Parametern und Designräumen gegeben, für die ein Synergieeffekt vorliegt und ein Resultat wie in Satz 4.23 erzielt werden kann.

Bemerkung 4.26

Die folgenden Fälle lassen sich mit der Transformation $\mathbf{f}(\mathbf{x}) \rightarrow \mathbf{f}(\mathbf{z}) = \mathbf{T}\mathbf{f}(\mathbf{x})$ mit $\mathbf{T} = \text{diag}(1, \beta_1, \beta_2, \beta_1\beta_2)$ auf den Fall in Satz 4.23 zurückführen.

- i) $\beta_1 < 0, \beta_2 < 0, \beta_{12} \leq 0, u_1 \geq 0, u_2 \geq 0,$
- ii) $\beta_1 > 0, \beta_2 < 0, \beta_{12} \geq 0, v_1 \leq 0, u_2 \geq 0,$
- iii) $\beta_1 < 0, \beta_2 > 0, \beta_{12} \geq 0, u_1 \geq 0, v_2 \leq 0.$

Numerische Untersuchungen sowie die Analogie zum D -optimalen Design für ein multiples Regressionsmodell lassen die folgende Vermutung anstellen.

Vermutung 4.27

Gilt $\phi_v^{-1}(0)/(\beta_1 + \beta_{12}v_2) \leq v_1 - u_1$ und $\phi_v^{-1}(0)/(\beta_2 + \beta_{12}v_1) \leq v_2 - u_2$, so ist das Design aus Satz 4.23 D -optimal in der Klasse aller Designs.

Es gibt weitere Parameter / Designraum-Kombinationen mit Synergieeffekten. Das folgende Beispiel zeigt für den Fall $\beta_1 > 0, \beta_2 > 0, \beta_{12} > 0, u_1 \geq 0, u_2 \geq 0$, dass ein Resultat analog zu Satz 4.23 nicht gilt.

Beispiel 4.28

Es wird das Proportional Hazards Modell mit Typ I Zensierung, Experimentdauer $c = 1$ und $\beta = (-3, 1, 1, 1)^T$ betrachtet. Für den Designraum $\mathcal{X}_1 = [0; 1,944] \times [0; 1,944]$ ist

das folgende Design D -optimal in der Klasse der 4-Punkt-Designs:

$$\xi_1 = \left\{ \begin{array}{cccc} (1,944; 1,944) & (1,944; 0,471) & (0,471; 1,944) & (0,468; 0,468) \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array} \right\}. \quad (4.35)$$

Für den geringfügig verkleinerten Designraum $\mathcal{X}_2 = [0; 1,943] \times [0; 1,943]$ ist das folgende Design D -optimal in der Klasse der 4-Punkt-Designs:

$$\xi_2 = \left\{ \begin{array}{cccc} (1,943; 1,943) & (1,943; 0,471) & (0,471; 1,943) & (0; 0) \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array} \right\}. \quad (4.36)$$

Es ist jedoch keines der beiden Designs D -optimal in der Klasse aller Designs. Für den Designraum \mathcal{X}_1 ist das folgende 5-Punkt-Design global D -optimal:

$$\xi^* = \left\{ \begin{array}{ccccc} (1,944; 1,944) & (1,944; 0,471) & (0,471; 1,944) & (0,58; 0,58) & (0; 0) \\ 0,25 & 0,249 & 0,249 & 0,111 & 0,142 \end{array} \right\}. \quad (4.37)$$

Zum Nachweis der D -Optimalität von ξ^* ist die Sensitivitätsfunktion in Abbildung 4.8 graphisch dargestellt.

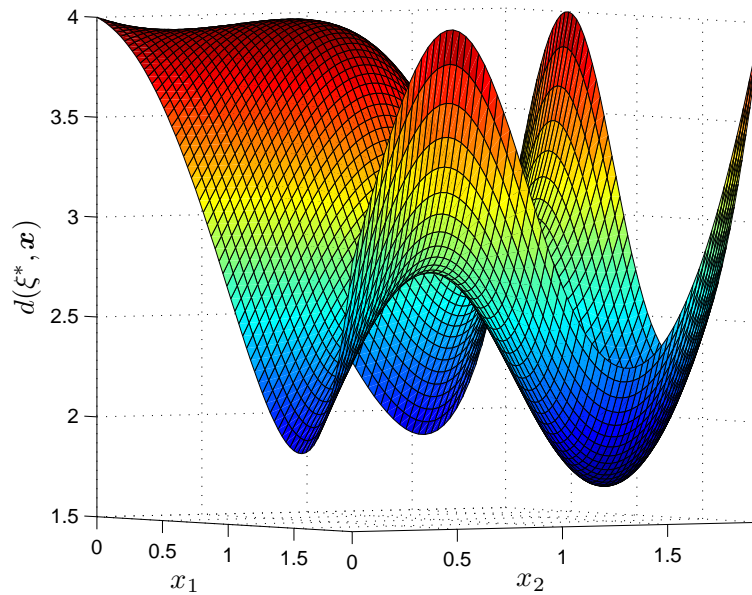


Abbildung 4.8: Sensitivitätsfunktion $d(\xi^*, \mathbf{x})$ für das Design ξ^* in (4.37)

5 c -optimale Designs

In diesem Kapitel werden c -optimale Designs für einen linearen Aspekt $\mathbf{c}^T \boldsymbol{\beta}$ berechnet. In Abschnitt 5.1 werden c -optimale Designs für das Modell mit einer Kovariablen für einen stetigen oder diskreten Designraum bestimmt. In Abschnitt 5.2 wird ein multiples Regressionsmodell betrachtet. Abschnitt 5.3 befasst sich mit Bayesschen c -optimalen Designs.

5.1 Modell mit einer Kovariablen

In diesem Abschnitt wird das Modell mit einer Kovariablen betrachtet. Es gilt also $p = 2$, $\mathbf{f}(x) = (1, x)^T$ und $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$.

Rodríguez-Torreblanca und Rodríguez-Díaz (2007) haben c -optimale Designs für das Poisson-Modell und das Negativ-Binomial-Modell für die Vektoren $\mathbf{c} = (1, 0)^T$ und $\mathbf{c} = (0, 1)^T$ bestimmt. Für die hier betrachtete Modellklasse, die das Poisson-Modell und das Negativ-Binomial-Modell beinhaltet, haben Konstantinou et al. (2014) c -optimale Designs für den Vektor $\mathbf{c} = (0, 1)^T$ berechnet. Nun werden c -optimale Designs für einen beliebigen Vektor $\mathbf{c} = (c_1, c_2)^T$ bestimmt.

Dazu werden die c -optimalen Gewichte benötigt. Für ein 2-Punkt-Design mit Stützpunkten x_1 und x_2 ergeben sich die c -optimalen Gewichte aus (3.37) zu

$$\omega_1^* = \frac{\sqrt{Q(\beta_0 + \beta_1 x_2) \cdot (c_1 x_2 - c_2)^2}}{\sqrt{Q(\beta_0 + \beta_1 x_1) \cdot (c_1 x_1 - c_2)^2} + \sqrt{Q(\beta_0 + \beta_1 x_2) \cdot (c_1 x_2 - c_2)^2}}, \quad (5.1a)$$

$$\omega_2^* = \frac{\sqrt{Q(\beta_0 + \beta_1 x_1) \cdot (c_1 x_1 - c_2)^2}}{\sqrt{Q(\beta_0 + \beta_1 x_1) \cdot (c_1 x_1 - c_2)^2} + \sqrt{Q(\beta_0 + \beta_1 x_2) \cdot (c_1 x_2 - c_2)^2}}. \quad (5.1b)$$

Zunächst wird die Funktion $\psi_{\mathbf{a}}(\cdot): \mathbb{R} \rightarrow \mathbb{R}$ mit $\mathbf{a} = (a_1, \dots, a_{p-1})^T$ allgemein eingeführt. Sie sei definiert als

$$\psi_{\mathbf{a}}(x) := x - 2 \cdot \frac{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x)}{Q'(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x)} \cdot \left(\frac{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x)}}{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}} + 1 \right). \quad (5.2)$$

Analog zu Lemma 4.7 können für die Funktion $\psi_{\mathbf{a}}$ die gleichen Eigenschaften wie für die Funktion $\phi_{\mathbf{a}}$ angegeben werden.

Lemma 5.1

Die Bedingungen (A1), (A2) und (A4) seien erfüllt. Die Funktion $\psi_{\mathbf{a}}$ ist streng monoton wachsend, stetig und bijektiv. Daher existiert die Umkehrfunktion $\psi_{\mathbf{a}}^{-1}$ mit denselben Eigenschaften.

Satz 5.2

Es sei $\mathcal{X} = [u, v]$ und die Bedingungen (A1), (A2) und (A4) seien erfüllt. Die Funktion g sei definiert als $g(x) := (c_1x - c_2)^2$.

a) Es sei $\beta_1 > 0$ und $x^* = \max \{u, v - \psi_v^{-1}(0)/\beta_1\}$.

Gilt $c_2/c_1 \notin [x^*, v]$, so ist das eindeutige c -optimale Design gegeben durch

$$\xi^* = \left\{ \begin{array}{cc} x^* & v \\ \frac{\sqrt{Q(\beta_0 + \beta_1 v)g(v)}}{\sqrt{Q(\beta_0 + \beta_1 x^*)g(x^*)} + \sqrt{Q(\beta_0 + \beta_1 v)g(v)}} & \frac{\sqrt{Q(\beta_0 + \beta_1 x^*)g(x^*)}}{\sqrt{Q(\beta_0 + \beta_1 x^*)g(x^*)} + \sqrt{Q(\beta_0 + \beta_1 v)g(v)}} \end{array} \right\}. \quad (5.3)$$

Gilt $c_2/c_1 \in [x^*, v]$, dann ist das 1-Punkt-Design mit Stützpunkt c_2/c_1 das eindeutige c -optimale Design.

b) Es sei $\beta_1 < 0$ und $x^* = \min \{v, u - \psi_u^{-1}(0)/\beta_1\}$.

Gilt $c_2/c_1 \notin [u, x^*]$, so ist das eindeutige c -optimale Design gegeben durch

$$\xi^* = \left\{ \begin{array}{cc} u & x^* \\ \frac{\sqrt{Q(\beta_0 + \beta_1 x^*)g(x^*)}}{\sqrt{Q(\beta_0 + \beta_1 u)g(u)} + \sqrt{Q(\beta_0 + \beta_1 x^*)g(x^*)}} & \frac{\sqrt{Q(\beta_0 + \beta_1 u)g(u)}}{\sqrt{Q(\beta_0 + \beta_1 u)g(u)} + \sqrt{Q(\beta_0 + \beta_1 x^*)g(x^*)}} \end{array} \right\}. \quad (5.4)$$

Gilt $c_2/c_1 \in [u, x^*]$, dann ist das 1-Punkt-Design mit Stützpunkt c_2/c_1 das eindeutige c -optimale Design.

Beweis:

Es sei $\mathbf{T} = \begin{pmatrix} 1 & 0 \\ \beta_0 & \beta_1 \end{pmatrix}$. Mit der kanonischen Transformation $\mathbf{f}(x) \rightarrow \mathbf{f}(z) = \mathbf{T}\mathbf{f}(x)$ ist die Minimierung von $\mathbf{c}^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{c}$ äquivalent zur Minimierung von $\mathbf{c}_z^T \mathbf{M}_z(\xi)^{-1} \mathbf{c}_z$ mit Informationsmatrix

$$\mathbf{M}_z(\xi) = \sum_{i=1}^m \omega_i Q(z_i) \mathbf{f}(z_i) \mathbf{f}(z_i)^T,$$

$\mathbf{c}_z = \mathbf{T}\mathbf{c}$, transformiertem Parametervektor $\boldsymbol{\beta}_z = (0, 1)^T$ und induziertem Designraum $\mathcal{Z} = \beta_0 + \beta_1 \mathcal{X} =: [\tilde{u}, \tilde{v}]$ (vgl. Ford et al., 1992). Zunächst wird gezeigt, dass die Kurve

$\tilde{\mathbf{f}}(z) = \sqrt{Q(z)} \cdot \mathbf{f}(z) = \sqrt{Q(z)} \cdot (1, z)^T$ streng konvex ist. Dazu seien $x(z) := \sqrt{Q(z)}$ und $y(z) := \sqrt{Q(z)}z$. Es gilt:

$$\frac{dy}{dx} = \frac{\frac{dy(z)}{dz}}{\frac{dx(z)}{dz}} = z + \frac{\sqrt{Q(z)}}{\left(\frac{Q'(z)}{2\sqrt{Q(z)}}\right)} = z + \frac{2Q(z)}{Q'(z)},$$

$$\frac{d^2y}{dx^2} = \frac{d}{dz} \left(z + \frac{2Q(z)}{Q'(z)} \right) \cdot \frac{1}{\frac{dx(z)}{dz}} = \left(1 + 2 \cdot \frac{d}{dz} \frac{Q(z)}{Q'(z)} \right) \cdot \frac{2\sqrt{Q(z)}}{Q'(z)}.$$

Die Funktion $Q(z)/Q'(z)$ ist nach (A4) monoton wachsend und somit ist ihre Ableitung nichtnegativ. Da $Q(z)$ und $Q'(z)$ positiv sind, folgt $\frac{d^2y}{dx^2} > 0$ und damit die strenge Konvexität. Es gibt also einen eindeutigen Punkt $\tilde{\mathbf{f}}(z^*)$, sodass der Rand der Elfving-Menge aus den beiden Verbindungsstrecken von $-\tilde{\mathbf{f}}(\tilde{v})$ zu $\tilde{\mathbf{f}}(z^*)$ und von $\tilde{\mathbf{f}}(\tilde{v})$ zu $-\tilde{\mathbf{f}}(z^*)$ und aus den Kurven $\tilde{\mathbf{f}}(z)$ und $-\tilde{\mathbf{f}}(z)$, $z \in [z^*, \tilde{v}]$, besteht. In Abbildung 5.1 ist die Elfving-Menge graphisch dargestellt.

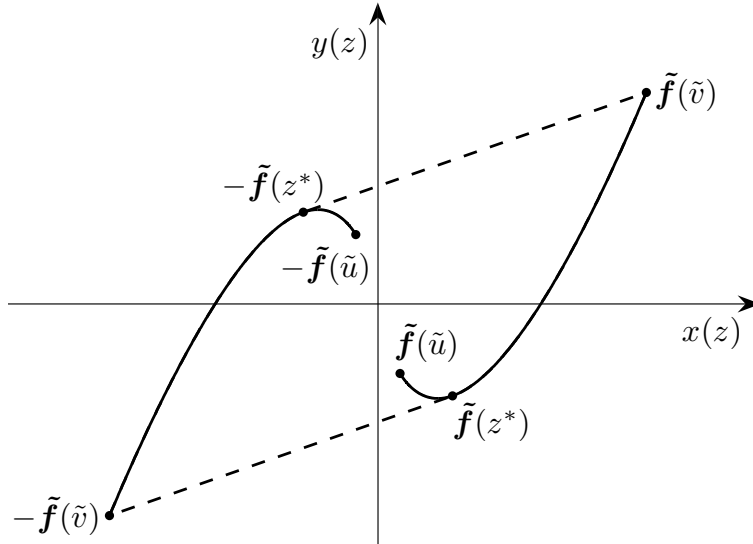


Abbildung 5.1: Elfving-Menge

Die Verbindungsstrecke von $-\tilde{\mathbf{f}}(\tilde{v})$ zu $\tilde{\mathbf{f}}(z^*)$ muss die Tangente an den Punkt $\tilde{\mathbf{f}}(z^*)$ sein, sofern der Designraum groß genug ist. Dann berechnet sich z^* durch

$$z^* + \frac{2Q(z^*)}{Q'(z^*)} = \frac{\sqrt{Q(z^*)}z^* - (-\sqrt{Q(\tilde{v})}\tilde{v})}{\sqrt{Q(z^*)} - (-\sqrt{Q(\tilde{v})})}$$

$$\Leftrightarrow z^* \sqrt{Q(\tilde{v})} + \frac{2Q(z^*)}{Q'(z^*)} \cdot (\sqrt{Q(z^*)} + \sqrt{Q(\tilde{v})}) = \sqrt{Q(\tilde{v})}\tilde{v}$$

$$\Leftrightarrow \tilde{v} - z^* - 2 \cdot \frac{Q(z^*)}{Q'(z^*)} \cdot \left(\frac{\sqrt{Q(z^*)}}{\sqrt{Q(\tilde{v})}} + 1 \right) = 0.$$

Mit $z = \tilde{v} - z^*$ ergibt sich die Gleichung zu $\psi_{\tilde{v}}(z) = 0$ mit Lösung $z = \psi_{\tilde{v}}^{-1}(0)$. Es folgt $z^* = \tilde{v} - \psi_{\tilde{v}}^{-1}(0)$.

Schneidet die durch den Vektor \mathbf{c}_z beschriebene Gerade den Rand der Elfving-Menge auf der Verbindungsstrecke von $-\tilde{\mathbf{f}}(\tilde{v})$ zu $\tilde{\mathbf{f}}(z^*)$ oder von $\tilde{\mathbf{f}}(\tilde{v})$ zu $-\tilde{\mathbf{f}}(z^*)$, so ist das Design mit Stützpunkten z^* und \tilde{v} nach dem Satz von Elfving c -optimal für den Vektor \mathbf{c}_z . Dies ist genau dann der Fall, wenn $c_{z,2}/c_{z,1} \notin [z^*, \tilde{v}]$ gilt.

Für $c_{z,2}/c_{z,1} \in [z^*, \tilde{v}]$ ist ein 1-Punkt-Design c -optimal. Wegen $\mathbf{c}_z \parallel \tilde{\mathbf{f}}(c_{z,2}/c_{z,1})$ hat es den Stützpunkt $c_{z,2}/c_{z,1}$.

Gilt $\tilde{v} - \psi_{\tilde{v}}^{-1}(0) < \tilde{u}$, so besteht der Rand der Elfving-Menge aus den beiden Verbindungsstrecken von $-\tilde{\mathbf{f}}(\tilde{v})$ zu $\tilde{\mathbf{f}}(\tilde{u})$ und von $\tilde{\mathbf{f}}(\tilde{v})$ zu $-\tilde{\mathbf{f}}(\tilde{u})$ und aus den Kurven $\tilde{\mathbf{f}}(z)$ und $-\tilde{\mathbf{f}}(z)$, $z \in [\tilde{u}, \tilde{v}]$. Mit gleicher Argumentation wie oben folgen die c -optimalen Designs. Die Rücktransformation liefert die c -optimalen Designs aus Satz 5.2. \square

Bemerkung 5.3

Für $c_1 = 0$ sind in Satz 5.2 für $\beta_1 > 0$ der Fall $c_2/c_1 \notin [x^*, v]$ und für $\beta_1 < 0$ der Fall $c_2/c_1 \notin [u, x^*]$ anzuwenden.

Satz 5.2 stimmt mit den Resultaten von Konstantinou et al. (2014) und Rodríguez-Torreblanca und Rodríguez-Díaz (2007) überein. Für das Poisson-Modell mit $Q(\theta) = e^\theta$ hat die Gleichung $\psi_a(x) = 0$ eine analytische Lösung:

$$\begin{aligned} \psi_a(x) &= x - 2 \cdot \left(e^{-\frac{1}{2}x} + 1 \right) = 0 \\ \iff \left(\frac{1}{2}x - 1 \right) e^{\frac{1}{2}x-1} &= e^{-1} \\ \iff \frac{1}{2}x - 1 &= W(e^{-1}) \\ \iff x = \psi_a^{-1}(0) &= 2 \cdot (1 + W(e^{-1})) \approx 2,557. \end{aligned}$$

Hier bezeichnet W wieder den oberen Zweig der Lambertschen W -Funktion.

Für einen diskreten Designraum $\mathcal{X} = \{x_1, \dots, x_k\}$ mit $x_1 < x_2 < \dots < x_k$ bezeichnen $x^-(a)$ und $x^+(a)$ wie in Abschnitt 4.2 das nächstkleinere bzw. das nächstgrößere Element des Designraums von $a \in (x_1, x_k)$, wobei bei $x^+(a)$ Gleichheit mit a zugelassen wird.

Satz 5.4

Es sei $\mathcal{X} = \{x_1, \dots, x_k\}$ mit $x_1 < \dots < x_k$. Die Bedingungen (A1), (A2) und (A4) seien erfüllt und die Funktionen g und h seien definiert als $g(x) := (c_1x - c_2)^2$ und

$$h(x_1, x_2) := \frac{\sqrt{Q(\beta_0 + \beta_1 x_1)}x_1 + \sqrt{Q(\beta_0 + \beta_1 x_2)}x_2}{\sqrt{Q(\beta_0 + \beta_1 x_1)} + \sqrt{Q(\beta_0 + \beta_1 x_2)}}. \quad (5.5)$$

a) Es sei $\beta_1 > 0$ und $\tilde{x} = x_k - \psi_{x_k}^{-1}(0)/\beta_1$. Ferner sei

$$x^* := \begin{cases} x_1 & \text{falls } \tilde{x} \leq x_1 \\ x^-(\tilde{x}) & \text{falls } \tilde{x} > x_1 \text{ und } h(x^+(\tilde{x}), x_k) \geq h(x^-(\tilde{x}), x_k) \\ x^+(\tilde{x}) & \text{falls } \tilde{x} > x_1 \text{ und } h(x^+(\tilde{x}), x_k) < h(x^-(\tilde{x}), x_k) \end{cases}$$

Gilt $c_2/c_1 \in [x^*, x_k]$ und $c_2/c_1 \in \mathcal{X}$, so ist das 1-Punkt-Design mit Stützpunkt c_2/c_1 c -optimal. Ansonsten ist das folgende 2-Punkt-Design c -optimal:

$$\xi^* = \left\{ \begin{array}{cc} x_1^* & x_2^* \\ \frac{\sqrt{Q(\beta_0 + \beta_1 x_2^*)g(x_2^*)}}{\sqrt{Q(\beta_0 + \beta_1 x_1^*)g(x_1^*)} + \sqrt{Q(\beta_0 + \beta_1 x_2^*)g(x_2^*)}} & \frac{\sqrt{Q(\beta_0 + \beta_1 x_1^*)g(x_1^*)}}{\sqrt{Q(\beta_0 + \beta_1 x_1^*)g(x_1^*)} + \sqrt{Q(\beta_0 + \beta_1 x_2^*)g(x_2^*)}} \end{array} \right\}. \quad (5.6)$$

Für $c_2/c_1 \in [x^*, x_k]$ und $c_2/c_1 \notin \mathcal{X}$ gilt $x_1^* = x^-(c_2/c_1)$ und $x_2^* = x^+(c_2/c_1)$.

Für $c_2/c_1 \notin [x^*, x_k]$ gilt $x_1^* = x^*$ und $x_2^* = x_k$.

b) Es sei $\beta_1 < 0$ und $\tilde{x} = x_1 - \psi_{x_1}^{-1}(0)/\beta_1$. Ferner sei

$$x^* := \begin{cases} x_k & \text{falls } \tilde{x} \geq x_k \\ x^+(\tilde{x}) & \text{falls } \tilde{x} < x_k \text{ und } h(x^+(\tilde{x}), x_1) \leq h(x^-(\tilde{x}), x_1) \\ x^-(\tilde{x}) & \text{falls } \tilde{x} < x_k \text{ und } h(x^+(\tilde{x}), x_1) > h(x^-(\tilde{x}), x_1) \end{cases}$$

Gilt $c_2/c_1 \in [x_1, x^*]$ und $c_2/c_1 \in \mathcal{X}$, so ist das 1-Punkt-Design mit Stützpunkt c_2/c_1 c -optimal. Ansonsten ist das 2-Punkt-Design (5.6) c -optimal.

Für $c_2/c_1 \in [x_1, x^*]$ und $c_2/c_1 \notin \mathcal{X}$ gilt $x_1^* = x^-(c_2/c_1)$ und $x_2^* = x^+(c_2/c_1)$.

Für $c_2/c_1 \notin [x_1, x^*]$ gilt $x_1^* = x_1$ und $x_2^* = x^*$.

Beweis:

Es wird wieder das kanonische Design-Problem betrachtet. Der induzierte Designraum ist gegeben durch $\mathcal{Z} = \{z_i = \beta_0 + \beta_1 x_i \mid x_i \in \mathcal{X}, i = 1, \dots, k\}$. Es sei $\tilde{z} = \tilde{v} - \psi_{\tilde{v}}^{-1}(0)$ mit $\tilde{v} := \max \{z_1, \dots, z_k\}$ und es sei $\tilde{u} := \min \{z_1, \dots, z_k\}$. Da für den stetigen Designraum die Punkte $\pm \tilde{f}(z)$ für $z \in [\max \{\tilde{u}, \tilde{z}\}, \tilde{v}]$ auf dem Rand der Elfving-Menge liegen, gilt

dies im Fall $\tilde{z} \leq \tilde{u}$ für alle Vektoren $\pm \tilde{\mathbf{f}}(z_i)$, $i = 1, \dots, k$, und im Fall $\tilde{z} > \tilde{u}$ für die Vektoren $\pm \tilde{\mathbf{f}}(z_i)$ mit $z_i \geq z^+(\tilde{z})$. Auch die Vektoren $\pm \tilde{\mathbf{f}}(z^-(\tilde{z}))$ können auf dem Rand der Elfving-Menge liegen. Dies ist genau dann der Fall, wenn die Steigung der Geraden durch die Punkte $-\tilde{\mathbf{f}}(\tilde{v})$ und $\tilde{\mathbf{f}}(z^-(\tilde{z}))$ kleiner oder gleich der Steigung der Geraden durch die Punkte $-\tilde{\mathbf{f}}(\tilde{v})$ und $\tilde{\mathbf{f}}(z^+(\tilde{z}))$ ist. Dasjenige Geradenstück mit kleinerer Steigung bildet den Rand der Elfving-Menge. Die c -optimalen Designs folgen nun wie beim Beweis von Satz 5.2 aus dem Satz von Elfving. Die Steigung ist nach Rücktransformation gegeben durch (5.5). \square

Bemerkung 5.5

Gilt bei a) $h(x^+(\tilde{x}), x_k) = h(x^-(\tilde{x}), x_k)$, so gibt es für $c_2/c_1 \notin [x^+(\tilde{x}), x_k]$ ein weiteres c -optimales 2-Punkt-Design mit den Stützpunkten $x_1^* = x^+(\tilde{x})$ und $x_2^* = x_k$ sowie ein c -optimales 3-Punkt-Design mit den Stützpunkten $x^-(\tilde{x})$, $x^+(\tilde{x})$ und x_k .

Gilt bei b) $h(x^+(\tilde{x}), x_1) = h(x^-(\tilde{x}), x_1)$, so gibt es für $c_2/c_1 \notin [x_1, x^-(\tilde{x})]$ ein weiteres c -optimales 2-Punkt-Design mit den Stützpunkten $x_1^* = x_1$ und $x_2^* = x^-(\tilde{x})$ sowie ein c -optimales 3-Punkt-Design mit den Stützpunkten x_1 , $x^-(\tilde{x})$ und $x^+(\tilde{x})$.

Beweis:

Bei gleichen Steigungen der Geraden von $-\tilde{\mathbf{f}}(\tilde{v})$ zu $\tilde{\mathbf{f}}(z^-(\tilde{z}))$ bzw. zu $\tilde{\mathbf{f}}(z^+(\tilde{z}))$ bildet das Geradenstück durch die drei Punkte $-\tilde{\mathbf{f}}(\tilde{v})$, $\tilde{\mathbf{f}}(z^-(\tilde{z}))$ und $\tilde{\mathbf{f}}(z^+(\tilde{z}))$ den Rand der Elfving-Menge. Daher lässt sich der Vektor \mathbf{c}_z im Fall $c_{z,2}/c_{z,1} \notin [z^+(\tilde{z}), \tilde{v}]$ zusätzlich zu den Konvexkombinationen mit den in Satz 5.4 angegebenen Stützpunkten als Konvexkombination der Punkte $-\tilde{\mathbf{f}}(\tilde{v})$ und $\tilde{\mathbf{f}}(z^+(\tilde{z}))$ bzw. aller drei Punkte schreiben. \square

5.2 Multiple Regression

In diesem Abschnitt werden c -optimale Designs für ein multiples Regressionsmodell mit $p-1$ Kovariablen und rechteckigem Designraum bestimmt. Es gilt $p \geq 3$, $\mathbf{f}(\mathbf{x}) = (1, \mathbf{x}^T)^T$ und $\mathbf{x} = (x_1, \dots, x_{p-1})^T \in \mathcal{X}$.

Satz 5.6

Es sei $\mathcal{X} = [u_1, v_1] \times \dots \times [u_{p-1}, v_{p-1}]$ und die Bedingungen (A1), (A2) und (A4) seien erfüllt. Der Vektor \mathbf{c} sei gegeben durch $\mathbf{c} = (0, \tilde{\mathbf{c}}^T)^T$ mit $\tilde{\mathbf{c}} = (c_1, \dots, c_{p-1})^T$. Für $i = 1, \dots, p-1$ gelte $\beta_i \neq 0$. Es gelte entweder $\beta_i c_i \geq 0$ für $i = 1, \dots, p-1$ oder $\beta_i c_i \leq 0$ für $i = 1, \dots, p-1$. Es sei $a_i = v_i$, falls $\beta_i > 0$ und $a_i = u_i$, falls $\beta_i < 0$. Ferner sei $x^* = \mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - \psi_{\mathbf{a}}^{-1}(0)$ mit $\mathbf{a} = (a_1, \dots, a_{p-1})^T$.

Gilt $|c_i| \psi_{\mathbf{a}}^{-1}(0) \leq |\mathbf{c}^T \boldsymbol{\beta}| (v_i - u_i)$ für $i = 1, \dots, p-1$, so ist das folgende 2-Punkt-Design c -optimal:

$$\xi^* = \left\{ \begin{array}{cc} \mathbf{a} - \frac{\psi_{\mathbf{a}}^{-1}(0)}{\mathbf{c}^T \boldsymbol{\beta}} \cdot \tilde{\mathbf{c}} & \mathbf{a} \\ \frac{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}}{\sqrt{Q(x^*) + \sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}}} & \frac{\sqrt{Q(x^*)}}{\sqrt{Q(x^*) + \sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}}} \end{array} \right\}. \quad (5.7)$$

Gilt zudem $c_i \neq 0$ für mindestens zwei $i \in \{1, \dots, p-1\}$, so sind die folgenden m -Punkt-Designs

$$\xi^* = \left\{ \begin{array}{cccc} \mathbf{x}_1^* & \dots & \mathbf{x}_{m-1}^* & \mathbf{a} \\ \alpha_1 \cdot \frac{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}}{\sqrt{Q(x^*) + \sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}}} & \dots & \alpha_{m-1} \cdot \frac{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}}{\sqrt{Q(x^*) + \sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}}} & \frac{\sqrt{Q(x^*)}}{\sqrt{Q(x^*) + \sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}}} \end{array} \right\}$$

mit $m \geq 3$ ebenfalls c -optimal, wobei

$$\mathbf{x}_i^* \in \mathcal{X} \cap \left\{ \mathbf{x} \in \mathbb{R}^{p-1} : \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} = x^* \right\} \subseteq \text{conv} \left\{ \mathbf{a} - \frac{\psi_{\mathbf{a}}^{-1}(0)}{\beta_i} \cdot \mathbf{e}_i, i = 1, \dots, p-1 \right\}$$

für $i = 1, \dots, m-1$ und die Konvexkombination der Stützpunkte \mathbf{x}_i^* den ersten Stützpunkt des c -optimalen 2-Punkt-Designs (5.7) ergibt:

$$\sum_{i=1}^{m-1} \alpha_i \mathbf{x}_i^* = \mathbf{a} - \frac{\psi_{\mathbf{a}}^{-1}(0)}{\mathbf{c}^T \boldsymbol{\beta}} \cdot \tilde{\mathbf{c}}. \quad (5.8)$$

Beweis:

Der Beweis wird für den Fall $\beta_i c_i \geq 0$, $i = 1, \dots, p-1$, geführt. Da c -Optimalität unabhängig vom Vorzeichen des Vektors \mathbf{c} ist, sind die Ergebnisse auch gültig, wenn $\beta_i c_i \leq 0$ für $i = 1, \dots, p-1$ gilt.

Die Matrix \mathbf{T} sei definiert durch

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ \beta_0 & \beta_1 & \beta_2 & \beta_3 & \dots & \beta_{p-1} \\ 0 & 0 & \beta_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ 0 & 0 & \dots & \dots & 0 & \beta_{p-1} \end{pmatrix}.$$

Mit der kanonischen Transformation $\mathbf{f}(\mathbf{x}) \rightarrow \mathbf{f}(\mathbf{z}) = \mathbf{T} \mathbf{f}(\mathbf{x})$ ist die Minimierung von

$\mathbf{c}^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{c}$ äquivalent zur Minimierung von $\mathbf{c}_z^T \mathbf{M}_z(\xi)^{-1} \mathbf{c}_z$ mit Informationsmatrix

$$\mathbf{M}_z(\xi) = \sum_{j=1}^m \omega_j Q(z_{j,1}) \mathbf{f}(z_j) \mathbf{f}(z_j)^T,$$

$\mathbf{c}_z = \mathbf{T} \mathbf{c} = (0, \sum_{i=1}^{p-1} \beta_i c_i, \beta_2 c_2, \dots, \beta_{p-1} c_{p-1})^T$ und transformiertem Parametervektor $\boldsymbol{\beta}_z = (0, \mathbf{e}_1^T)^T$ (vgl. Ford et al., 1992). Im transformierten Modell können die Komponenten z_i , $i = 2, \dots, p-1$, einer Versuchseinstellung \mathbf{z} Werte aus dem Intervall $[\tilde{u}_i, \tilde{v}_i]$ annehmen, wobei $[\tilde{u}_i, \tilde{v}_i] := [\beta_i u_i, \beta_i v_i]$ für $\beta_i > 0$ und $[\tilde{u}_i, \tilde{v}_i] := [\beta_i v_i, \beta_i u_i]$ für $\beta_i < 0$. Für die Komponente z_1 gilt $z_1 \in [\sum_{i=1}^{p-1} \tilde{u}_i, \sum_{i=1}^{p-1} \tilde{v}_i]$, wobei $[\tilde{u}_1, \tilde{v}_1] := [\beta_0 + \beta_1 u_1, \beta_0 + \beta_1 v_1]$ für $\beta_1 > 0$ und $[\tilde{u}_1, \tilde{v}_1] := [\beta_0 + \beta_1 v_1, \beta_0 + \beta_1 u_1]$ für $\beta_1 < 0$. Dabei ist zu beachten, dass die einzelnen Komponenten nicht unabhängig voneinander sind, denn es gilt $z_1 = \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} = \beta_0 + \beta_1 x_1 + \sum_{i=2}^{p-1} z_i$.

Die Fläche $\tilde{\mathbf{f}}(\mathbf{z}) = \sqrt{Q(z_1)} \cdot \mathbf{f}(\mathbf{z}) = \sqrt{Q(z_1)} \cdot (1, z_1, \dots, z_{p-1})^T$ stimmt in den ersten beiden Komponenten mit der Kurve für den zweidimensionalen Fall mit einer Kovariablen überein. Daher kann analog zum Beweis von Satz 5.2 gefolgert werden, dass die Verbindungsfläche vom Punkt $\tilde{\mathbf{f}}(\sum_{i=1}^{p-1} \tilde{v}_i, \tilde{v}_2, \dots, \tilde{v}_{p-1})$ zu den Punkten $-\tilde{\mathbf{f}}(z^*, z_2, \dots, z_{p-1})$ auf dem Rand der Elfving-Menge liegt, wobei $z^* = \sum_{i=1}^{p-1} \tilde{v}_i - \psi_{\tilde{v}_1 + \dots + \tilde{v}_{p-1}}^{-1}(0)$ ist mit der Funktion ψ für das Modell mit einer Kovariablen für $\boldsymbol{\beta} = (0, 1)^T$. Für die Komponenten z_i , $i = 2, \dots, p-1$, muss wegen der Abhängigkeit zur ersten Komponente, die den Wert z^* hat, $z^* - \sum_{i=2}^{p-1} z_i \in [\tilde{u}_1, \tilde{v}_1]$ gelten. Es sei

$$z_i^* := \tilde{v}_i - \frac{\beta_i c_i}{\sum_{j=1}^{p-1} \beta_j c_j} \cdot \psi_{\tilde{v}_1 + \dots + \tilde{v}_{p-1}}^{-1}(0)$$

für $i = 2, \dots, p-1$ und

$$\omega_1^* = \frac{\sqrt{Q(\sum_{i=1}^{p-1} \tilde{v}_i)}}{\sqrt{Q(z^*)} + \sqrt{Q(\sum_{i=1}^{p-1} \tilde{v}_i)}},$$

$$\omega_2^* = \frac{\sqrt{Q(z^*)}}{\sqrt{Q(z^*)} + \sqrt{Q(\sum_{i=1}^{p-1} \tilde{v}_i)}}.$$

Damit der Punkt $(z^*, z_2^*, \dots, z_{p-1}^*)$ im Designraum liegt, muss $z_i^* > \tilde{u}_i$ für $i = 2, \dots, p-1$ und

$$z^* - \sum_{i=2}^{p-1} z_i^* = \tilde{v}_1 - \frac{\beta_1 c_1}{\sum_{j=1}^{p-1} \beta_j c_j} \cdot \psi_{\tilde{v}_1 + \dots + \tilde{v}_{p-1}}^{-1}(0) > \tilde{u}_1$$

gelten. Dann liegt der Punkt $(z^*, z_2^*, \dots, z_{p-1}^*)$ sowie jeder Punkt auf der Verbindungsstrecke zu $(\sum_{i=1}^{p-1} \tilde{v}_i, \tilde{v}_2, \dots, \tilde{v}_{p-1})$ auf dem Rand der Elfving-Menge. Es folgt

$$\begin{aligned} & \omega_1^* \cdot (-\tilde{\mathbf{f}}(z^*, z_2^*, \dots, z_{p-1}^*)) + \omega_2^* \cdot \tilde{\mathbf{f}}\left(\sum_{i=1}^{p-1} \tilde{v}_i, \tilde{v}_2, \dots, \tilde{v}_{p-1}\right) \\ &= \frac{\sqrt{Q(z^*)Q(\sum_{i=1}^{p-1} \tilde{v}_i)}}{\sqrt{Q(z^*)} + \sqrt{Q(\sum_{i=1}^{p-1} \tilde{v}_i)}} \cdot \frac{\psi_{\tilde{v}_1 + \dots + \tilde{v}_{p-1}}^{-1}(0)}{\sum_{j=1}^{p-1} \beta_j c_j} \cdot \begin{pmatrix} 0 \\ \sum_{j=1}^{p-1} \beta_j c_j \\ \beta_2 c_2 \\ \vdots \\ \beta_{p-1} c_{p-1} \end{pmatrix} \\ &= \gamma \mathbf{c}_z, \end{aligned}$$

wobei γ gleich dem Vorfaktor ist. Nach dem Satz von Elfving ist das Design mit Stützpunkten $(z^*, z_2^*, \dots, z_{p-1}^*)$ und $(\sum_{i=1}^{p-1} \tilde{v}_i, \tilde{v}_2, \dots, \tilde{v}_{p-1})$ und zugehörigen Gewichten ω_1^* und ω_2^* c -optimal für den Vektor \mathbf{c}_z .

Es gibt beliebig viele weitere c -optimale Designs mit $m \geq 3$ Stützpunkten, sofern $c_i \neq 0$ für mindestens zwei $i \in \{1, \dots, p-1\}$ gilt. Dann lässt sich der Stützpunkt $(z^*, z_2^*, \dots, z_{p-1}^*)$ als Konvexkombination von Punkten $(z^*, \mathbf{z}_1^T), \dots, (z^*, \mathbf{z}_{m-1}^T)$ aus dem Designraum, welche auf der Hyperebene $\{\mathbf{z} \in \mathbb{R}^{p-1} : \mathbf{f}(\mathbf{z})^T \boldsymbol{\beta}_z = z_1 = z^*\}$ liegen, darstellen:

$$\sum_{i=1}^{m-1} \alpha_i (z^*, \mathbf{z}_i^T) = (z^*, z_2^*, \dots, z_{p-1}^*).$$

Folglich sind alle Designs

$$\xi^* = \left\{ \begin{array}{ccc} (z^*, \mathbf{z}_1^T) & \dots & (z^*, \mathbf{z}_{m-1}^T) & (\sum_{i=1}^{p-1} \tilde{v}_i, \tilde{v}_2, \dots, \tilde{v}_{p-1}) \\ \alpha_1 \omega_1^* & \dots & \alpha_{m-1} \omega_1^* & \omega_2^* \end{array} \right\}$$

ebenfalls c -optimal. Die Rücktransformation liefert das Resultat aus Satz 5.6. Die Bedingungen, dass der Stützpunkt $(z^*, z_2^*, \dots, z_{p-1}^*)$ im transformierten Designraum liegt, ergeben sich zu $|c_i| \psi_{\mathbf{a}}^{-1}(0) \leq |\mathbf{c}^T \boldsymbol{\beta}| (v_i - u_i)$ für $i = 1, \dots, p-1$. \square

Aus dem Beweis folgt, dass neben diskreten auch c -optimale Designs mit stetigem Anteil existieren.

Satz 5.7

Die Voraussetzungen von Satz 5.6 seien gegeben. Es bezeichnen ω_1^* und ω_2^* die Gewichte

des c -optimalen 2-Punkt-Designs (5.7). Es sei ξ_1 ein beliebiges, nicht notwendigerweise diskretes Design auf $\mathcal{X} \cap \{\mathbf{x} \in \mathbb{R}^{p-1} : \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} = x^*\}$ mit $E(\xi_1) = \mathbf{a} - \frac{\psi_{\mathbf{a}}^{-1}(0)}{\mathbf{c}^T \boldsymbol{\beta}} \cdot \tilde{\mathbf{c}}$ und es bezeichne ξ_2 das 1-Punkt-Design mit Stützpunkt \mathbf{a} . Dann ist das Design $\xi^* = \omega_1^* \xi_1 + \omega_2^* \xi_2$ c -optimal.

Bemerkung 5.8

Die Nullstelle $\psi_{\mathbf{a}}^{-1}(0)$ der Funktion $\psi_{\mathbf{a}}$ liegt im Intervall $(0, 4Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})/Q'(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta}))$.

Mit dem Vektor $\mathbf{c} = (0, \mathbf{e}_i)^T$ erhält man aus Satz 5.6 für $i = 1, \dots, p-1$ die c -optimalen Designs für β_i mit Stützpunkten $\mathbf{a} - (\psi_{\mathbf{a}}^{-1}(0)/\beta_i) \mathbf{e}_i$ und \mathbf{a} . Diese Designs sind eindeutig, denn der Vektor \mathbf{c} hat nur eine Komponente ungleich null. Der erste Stützpunkt liegt somit auf einer Kante des Designraums und lässt sich daher nicht als Konvexkombination weiterer Stützpunkte darstellen.

Korollar 5.9

Es gelte $\psi_{\mathbf{a}}^{-1}(0) \leq |\beta_i| (v_i - u_i)$ für $i = 1, \dots, p-1$ und die Voraussetzungen von Satz 5.6 seien erfüllt. Es seien ξ_i^* die eindeutigen c -optimalen Designs für β_i mit Stützpunkten $\mathbf{a} - (\psi_{\mathbf{a}}^{-1}(0)/\beta_i) \mathbf{e}_i$ und \mathbf{a} und es seien $\alpha_i = \beta_i c_i / (\mathbf{c}^T \boldsymbol{\beta})$ für $i = 1, \dots, p-1$. Dann ist das Design $\xi^* = \sum_{i=1}^{p-1} \alpha_i \xi_i^*$, dessen Stützpunkte nur auf den Kanten des Designraums liegen, c -optimal für den Vektor $\mathbf{c} = (0, c_1, \dots, c_{p-1})^T$. Dies folgt aus Satz 5.6, da die Konvexkombination der Stützpunkte $\mathbf{a} - (\psi_{\mathbf{a}}^{-1}(0)/\beta_i) \mathbf{e}_i$ der Designs ξ_i^* den ersten Stützpunkt des c -optimalen 2-Punkt-Designs (5.7) ergibt.

Abbildung 5.2 zeigt die Struktur der c -optimalen Designs für das Modell mit zwei Kovariablen ($p = 3$). Die von \mathbf{a} verschiedenen Stützpunkte der c -optimalen Designs aus Satz 5.6 liegen auf der durch $\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} = x^*$ definierten Hyperebene, die im vorliegenden Fall eine Gerade ist.

Die Informationsmatrix des c -optimalen 2-Punkt-Designs ist singulär. Um alle Parameter schätzen zu können, werden Designs mit regulärer Informationsmatrix benötigt. Daher sind die m -Punkt-Designs aus Satz 5.6 von Interesse. Die folgende Bemerkung liefert genaue Formeln für die c -optimalen 3-Punkt-Designs zu Abbildung 5.2.

Bemerkung 5.10

Die c -optimalen 3-Punkt-Designs für das Modell mit zwei Kovariablen ($p = 3$) sind von der Gestalt

$$\xi^* = \left\{ \begin{array}{ccc} \mathbf{x}_1^* & \mathbf{x}_2^* & \mathbf{a} \\ \frac{s}{r+s} \cdot \frac{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}}{\sqrt{Q(x^*) + \sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}}} & \frac{r}{r+s} \cdot \frac{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}}{\sqrt{Q(x^*) + \sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}}} & \frac{\sqrt{Q(x^*)}}{\sqrt{Q(x^*) + \sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}}} \end{array} \right\}$$

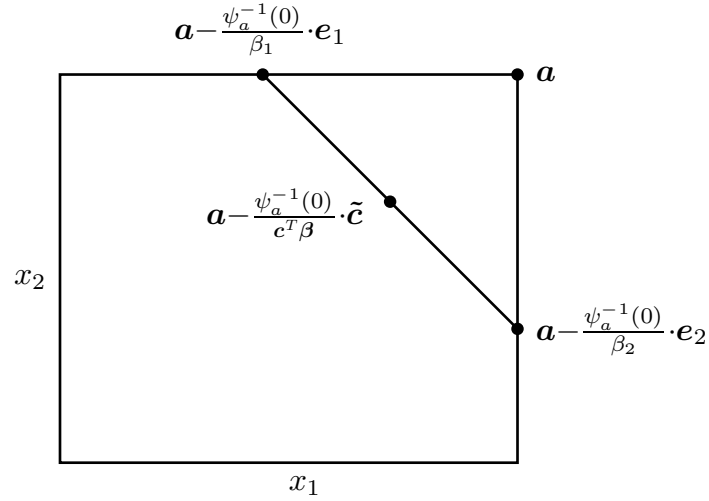


Abbildung 5.2: Struktur der c -optimalen Designs

mit

$$\mathbf{x}_1^* = \left(a_1 - \frac{\psi_{\mathbf{a}}^{-1}(0)}{\mathbf{c}^T \boldsymbol{\beta}} c_1 - \frac{r}{\beta_1}, a_2 - \frac{\psi_{\mathbf{a}}^{-1}(0)}{\mathbf{c}^T \boldsymbol{\beta}} c_2 + \frac{r}{\beta_2} \right)$$

$$\mathbf{x}_2^* = \left(a_1 - \frac{\psi_{\mathbf{a}}^{-1}(0)}{\mathbf{c}^T \boldsymbol{\beta}} c_1 + \frac{s}{\beta_1}, a_2 - \frac{\psi_{\mathbf{a}}^{-1}(0)}{\mathbf{c}^T \boldsymbol{\beta}} c_2 - \frac{s}{\beta_2} \right)$$

für $r \in \left(0, \frac{\beta_2 c_2}{\beta_1 c_1 + \beta_2 c_2} \cdot \psi_{\mathbf{a}}^{-1}(0)\right]$ und $s \in \left(0, \frac{\beta_1 c_1}{\beta_1 c_1 + \beta_2 c_2} \cdot \psi_{\mathbf{a}}^{-1}(0)\right]$.

Beweis:

Im transformierten Modell wie im Beweis von Satz 5.6 sind die c -optimalen 3-Punkt-Designs für zwei Kovariablen von der Gestalt

$$\xi^* = \left\{ \begin{array}{ccc} (z^*, z_2^* + r) & (z^*, z_2^* - s) & (\tilde{v}_1 + \tilde{v}_2, \tilde{v}_2) \\ \frac{s}{r+s} \cdot \omega_1^* & \frac{r}{r+s} \cdot \omega_1^* & \omega_2^* \end{array} \right\},$$

wobei $r \in (0, \tilde{v}_2 - z_2^*]$ und $s \in (0, z_2^* - \tilde{v}_2 + \psi_{\tilde{v}_1 + \dots + \tilde{v}_{p-1}}^{-1}(0)]$. Die Rücktransformation liefert die Formeln in Bemerkung 5.10. \square

Das c -optimale 2-Punkt-Design kann bei der vorliegenden multiplen Regression einen Stützpunkt im Inneren des Designraums haben. Es können auch c -optimale m -Punkt-Designs, $m \geq 3$, mit Stützpunkten im Inneren des Designraums existieren. Dies stellt keinen Widerspruch zu Satz 3.22 dar, weil die Matrix $\mathbf{A} = \mathbf{M} \mathbf{c} \mathbf{c}^T \mathbf{M}$ im Äquivalenzsatz für c -Optimalität, wobei $\mathbf{M} = (\mathbf{M}(\xi^*, \boldsymbol{\beta}) + \mathbf{H} \mathbf{H}^T)^{-1}$ für Designs mit singulärer Informationsmatrix und $\mathbf{M} = \mathbf{M}(\xi^*, \boldsymbol{\beta})^{-1}$ für Designs mit regulärer Informationsmatrix gilt, nicht positiv definit, sondern nur positiv semidefinit ist. Daher ist die Funktion

$p(\mathbf{x}) := \mathbf{f}(\mathbf{x})^T \mathbf{A} \mathbf{f}(\mathbf{x})$ für c -Optimalität nicht streng konvex und kann somit auch im Inneren des Designraums maximal sein. Sie wird allerdings auch immer an den Ecken maximal und nach Satz 3.24 existiert auch ein c -optimales Design mit Stützpunkten nur auf den Kanten des Designraums. Es sind unendlich viele c -optimale Designs möglich. Sitter und Torsney (1995) haben für andere Modelle mit 2 Kovariablen mit Hilfe des Satzes von Elfving optimale Designs bestimmt und ebenfalls festgestellt, dass es verschiedene c -optimale Designs geben kann.

Beispiel 5.11

Es wird das Proportional Hazards Modell mit Typ I Zensierung, Experimentdauer $c = 1$ und zwei Kovariablen betrachtet. Der Designraum sei $\mathcal{X} = [0, 10] \times [0, 10]$ und der Parametervektor sei gegeben durch $\boldsymbol{\beta} = (0, -1, -1)^T$. Gesucht ist das c -optimale Design für $\mathbf{c} = (0, 1, 1)^T$. Es gilt $\psi_{\mathbf{0}_2}^{-1}(0) = 2,723$ und nach Satz 5.6 ist das folgende 2-Punkt-Design c -optimal:

$$\xi^* = \left\{ \begin{array}{cc} (1,361; 1,361) & (0; 0) \\ 0,759 & 0,241 \end{array} \right\}.$$

Es existiert ein c -optimales Design mit Stützpunkten nur auf den Kanten des Designraums. Dazu seien ξ_i^* für $i = 1, 2$ die eindeutigen c -optimalen Designs für β_i mit Stützpunkten $2,723e_i$ und $\mathbf{0}_2$ und Gewichten 0,759 und 0,241. Dann ist nach Korollar 5.9 das folgende Design mit Stützpunkten nur auf den Kanten c -optimal für $\mathbf{c} = (0, 1, 1)^T$:

$$\xi^* = \frac{1}{2}\xi_1^* + \frac{1}{2}\xi_2^* = \left\{ \begin{array}{ccc} (2,723; 0) & (0; 2,723) & (0; 0) \\ 0,380 & 0,380 & 0,241 \end{array} \right\}.$$

Insbesondere besitzt dieses Design eine reguläre Informationsmatrix.

5.3 Bayessche c -optimale Designs

In diesem Abschnitt werden für spezielle Modelle mit zensierten Daten und mit einer Kovariablen Bayessche c -optimale Designs für den Effektparameter β_1 berechnet, d. h. es gilt $\mathbf{c} = (0, 1)^T$. Zunächst wird das Proportional Hazards Modell mit Typ I Zensierung, Experimentdauer $c = 1$ und voneinander unabhängigen Gewichtungen $N(0, 1)$ für β_0 und $N(\mu_1, \sigma_1^2)$ für β_1 betrachtet. Die Bestimmung der optimalen Designs erfolgt numerisch. Dabei werden die Integrale in der Kriteriumsfunktion wie in Abschnitt 4.5 beschrieben durch Quadraturformeln näherungsweise berechnet. Die Bayesschen c -optimalen Designs für β_1 auf dem Designraum $\mathcal{X} = [0, 1]$ sind in Tabelle 5.1 angegeben.

	$\mu_1 = 0$	$\mu_1 = 4$	$\mu_1 = 10$
$\sigma_1^2 = 1$	$\begin{Bmatrix} 0 & 1 \\ 0,451 & 0,549 \end{Bmatrix}$	$\begin{Bmatrix} 0 & 1 \\ 0,600 & 0,400 \end{Bmatrix}$	$\begin{Bmatrix} 0,088 & 1 \\ 0,541 & 0,459 \end{Bmatrix}$
$\sigma_1^2 = 3$	$\begin{Bmatrix} 0 & 0,421 & 1 \\ 0,256 & 0,440 & 0,304 \end{Bmatrix}$	$\begin{Bmatrix} 0 & 0,262 & 1 \\ 0,378 & 0,149 & 0,473 \end{Bmatrix}$	$\begin{Bmatrix} 0,082 & 1 \\ 0,546 & 0,454 \end{Bmatrix}$
$\sigma_1^2 = 4,5$	$\begin{Bmatrix} 0 & 0,320 & 1 \\ 0,249 & 0,579 & 0,172 \end{Bmatrix}$	$\begin{Bmatrix} 0 & 0,325 & 1 \\ 0,246 & 0,405 & 0,349 \end{Bmatrix}$	$\begin{Bmatrix} 0 & 0,136 & 1 \\ 0,245 & 0,297 & 0,458 \end{Bmatrix}$

Tabelle 5.1: Bayessche c -optimale Designs für das Proportional Hazards Modell mit Typ I Zensierung

Die Aussage aus Abschnitt 3.5, dass ein c -optimales Design mit höchstens zwei Stützpunkten existiert, gilt für das Bayes-Kriterium nicht mehr. In Abbildung 5.3 ist zu den optimalen Designs aus Tabelle 5.1 die Fréchet-Ableitung graphisch dargestellt.

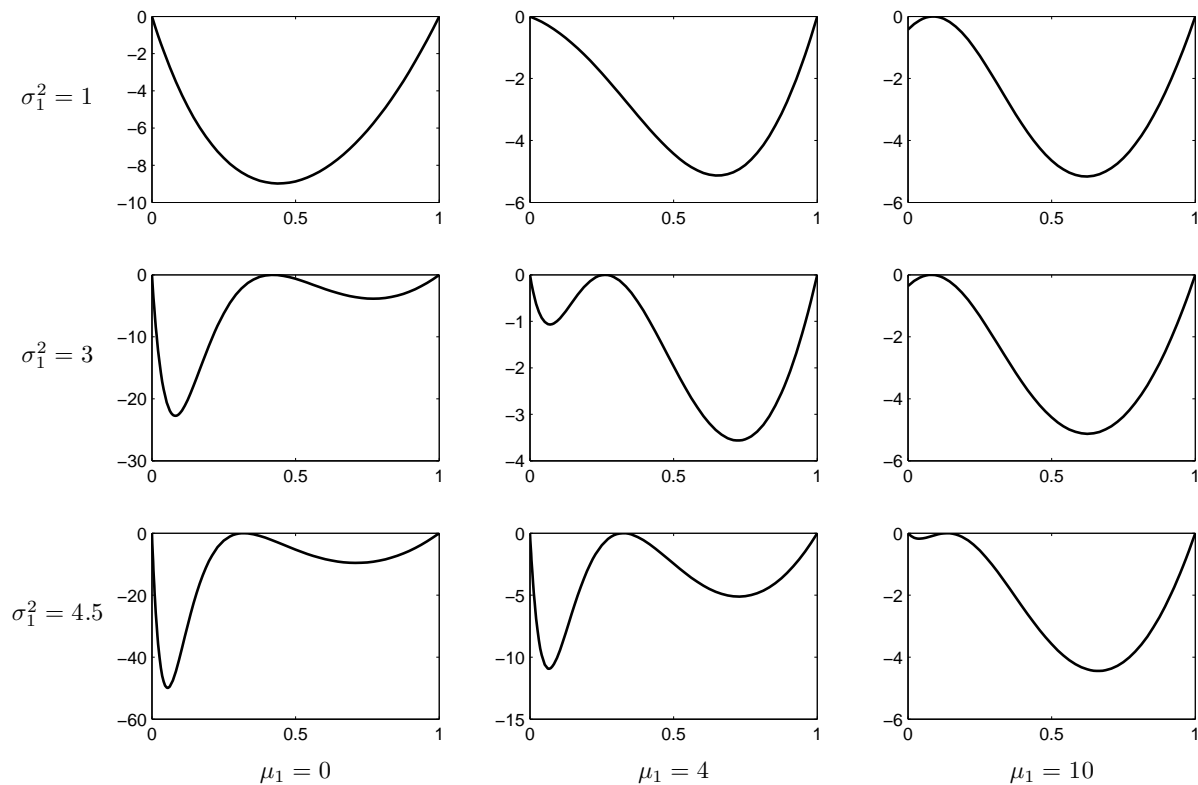


Abbildung 5.3: Fréchet-Ableitung für die optimalen Designs aus Tabelle 5.1

Im Folgenden wird das Proportional Hazards Modell mit exponentialverteilten Zensierungszeiten bzw. Negativ-Binomial-Modell betrachtet. Die Beweise der folgenden Resultate befinden sich im Anhang.

Satz 5.12

Es sei $Q(\theta) = e^\theta/(e^\theta + \lambda)$ und $\mathcal{X} = [u, v]$. Es seien die voneinander unabhängigen Gewichtungen $N(\mu_0, \sigma_0^2)$ für β_0 und $N(\mu_1, \sigma_1^2)$ für β_1 gegeben. Es sei

$$x^* = \frac{1}{\sigma_1} \cdot \sqrt{2 + 2W\left(\frac{1}{\lambda} \exp\left(\mu_0 - \frac{\sigma_0^2}{2} + \frac{\mu_1^2}{2\sigma_1^2} - 1\right)\right)}. \quad (5.9)$$

Gilt $\mu_1/\sigma_1^2 \pm x^* \in \mathcal{X}$, so ist das Bayessche c -optimale Design für β_1 in der Klasse der 2-Punkt-Designs gegeben durch

$$\xi^* = \left\{ \begin{array}{cc} \mu_1/\sigma_1^2 - x^* & \mu_1/\sigma_1^2 + x^* \\ 1/2 & 1/2 \end{array} \right\}. \quad (5.10)$$

Für $\mu_1/\sigma_1^2 - x^* < u$ bzw. $\mu_1/\sigma_1^2 + x^* > v$ ist es gegeben durch

$$\xi^* = \left\{ \begin{array}{cc} x_1^* & x_2^* \\ \frac{w(x_1^*)}{w(x_1^*)+w(x_2^*)} & \frac{w(x_2^*)}{w(x_1^*)+w(x_2^*)} \end{array} \right\}, \quad (5.11)$$

wobei die Funktion w definiert sei wie folgt:

$$w(x) = \sqrt{1 + \lambda \exp\left(-\mu_0 + \frac{\sigma_0^2}{2} - x\mu_1 + \frac{x^2\sigma_1^2}{2}\right)}. \quad (5.12)$$

Gilt $\mu_1/\sigma_1^2 - x^* < u$, so ist $x_1^* = u$ und $x_2^* = \min(\tilde{x}, v)$, wobei \tilde{x} die eindeutige Lösung der Gleichung $w'(x)(x - u) - w(u) - w(x) = 0$ im Intervall (u, ∞) ist.

Gilt $\mu_1/\sigma_1^2 + x^* > v$, so ist $x_2^* = v$ und $x_1^* = \max(u, \tilde{x})$, wobei \tilde{x} die eindeutige Lösung der Gleichung $w'(x)(x - v) - w(v) - w(x) = 0$ im Intervall $(-\infty, v)$ ist.

Für den nächsten Satz werden Gewichtungen der Parameter verwendet, die nur auf der Menge der positiven reellen Zahlen definiert sind. Hierbei handelt es sich um die Gamma- und Exponentialverteilung. Für die Dichtefunktion der Gammaverteilung wird die folgende Parametrisierung gewählt, wobei $\Gamma(\cdot)$ die Gammafunktion bezeichnet:

$$f_{\text{Gamma}(a,b)}(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}, \quad x > 0. \quad (5.13)$$

Als Spezialfall der Gammaverteilung erhält man für $a = 1$ eine Exponentialverteilung mit Parameter b .

Satz 5.13

Es sei $Q(\theta) = e^\theta/(e^\theta + \lambda)$ und $\mathcal{X} = [u, v]$. Es seien die voneinander unabhängigen Gewichtungen $\text{Gamma}(a, b)$ für β_0 und $\text{Exp}(\vartheta)$ für β_1 mit $u > -\vartheta$ gegeben. Die Funktion w sei definiert wie folgt:

$$w(x) = \sqrt{1 + \frac{\lambda b^a \vartheta}{(b+1)^a (\vartheta + x)}}. \quad (5.14)$$

Das Bayessche c -optimale Design für β_1 in der Klasse der 2-Punkt-Designs ist gegeben durch

$$\xi^* = \left\{ \begin{array}{cc} x_1^* & v \\ \frac{w(x_1^*)}{w(x_1^*)+w(v)} & \frac{w(v)}{w(x_1^*)+w(v)} \end{array} \right\}. \quad (5.15)$$

Dabei ist

$$x_1^* = \max \left\{ u, -\vartheta - \frac{\lambda b^a \vartheta}{2(b+1)^a} + \frac{1}{2} \sqrt{\left(\frac{\lambda b^a \vartheta}{(b+1)^a} \right)^2 + \frac{\lambda b^a \vartheta (\vartheta + v)}{(b+1)^a}} \right\}. \quad (5.16)$$

6 L -optimale Designs

In diesem Kapitel werden L -optimale Designs bestimmt. In Abschnitt 6.1 wird das Modell mit einer Kovariablen betrachtet. Dabei kann der Designraum stetig oder diskret sein. In Abschnitt 6.2 werden L -optimale Designs für ein multiples Regressionsmodell berechnet.

Die Matrix $\mathbf{B} = (b_{ij})_{i,j=1,\dots,p}$ wird als positiv definit und damit regulär angenommen. Daher muss ein L -optimales Design eine reguläre Informationsmatrix und somit mindestens p Stützpunkte haben.

6.1 Modell mit einer Kovariablen

Für das in diesem Abschnitt betrachtete Modell mit einer Kovariablen gilt $p = 2$, $\mathbf{f}(x) = (1, x)^T$ und $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$.

Ein L -optimales Design muss mindestens zwei Stützpunkte haben. Die L -optimalen Gewichte aus (3.38) berechnen sich für ein 2-Punkt-Design mit Stützpunkten x_1 und x_2 zu

$$\omega_1^* = \frac{\sqrt{Q(\beta_0 + \beta_1 x_2) \cdot g(x_2)}}{\sqrt{Q(\beta_0 + \beta_1 x_1) \cdot g(x_1)} + \sqrt{Q(\beta_0 + \beta_1 x_2) \cdot g(x_2)}}$$

$$\omega_2^* = \frac{\sqrt{Q(\beta_0 + \beta_1 x_1) \cdot g(x_1)}}{\sqrt{Q(\beta_0 + \beta_1 x_1) \cdot g(x_1)} + \sqrt{Q(\beta_0 + \beta_1 x_2) \cdot g(x_2)}}$$

mit $g(x) := b_{11}x^2 - 2b_{12}x + b_{22}$. Die Funktion $\gamma_a(\cdot): \mathbb{R} \rightarrow \mathbb{R}$ sei definiert als

$$\gamma_a(x) := \beta_1(a - x) - 2 \cdot \frac{Q(\beta_0 + \beta_1 x)}{Q'(\beta_0 + \beta_1 x)} \cdot \left(1 + \frac{\sqrt{Q(\beta_0 + \beta_1 x)}}{\sqrt{Q(\beta_0 + \beta_1 a)}} \cdot \frac{b_{11}xa - b_{12}(x + a) + b_{22}}{\sqrt{g(x)g(a)}} \right).$$

Satz 6.1

Es sei $\mathcal{X} = [u, v]$ und die Bedingungen (A1)–(A4) seien erfüllt. Die Matrix \mathbf{B} sei positiv definit und die Funktion g sei definiert als $g(x) = b_{11}x^2 - 2b_{12}x + b_{22}$.

a) Es sei $\beta_1 > 0$. Das eindeutige L -optimale Design ist gegeben durch

$$\xi^* = \left\{ \begin{array}{c} x_1^* \\ \frac{\sqrt{Q(\beta_0 + \beta_1 v)g(v)}}{\sqrt{Q(\beta_0 + \beta_1 x_1^*)g(x_1^*) + \sqrt{Q(\beta_0 + \beta_1 v)g(v)}}} \end{array} \quad \begin{array}{c} v \\ \frac{\sqrt{Q(\beta_0 + \beta_1 x_1^*)g(x_1^*)}}{\sqrt{Q(\beta_0 + \beta_1 x_1^*)g(x_1^*) + \sqrt{Q(\beta_0 + \beta_1 v)g(v)}}} \end{array} \right\}. \quad (6.1)$$

Dabei ist $x_1^* = \max\{u, x^*\}$, wobei x^* die eindeutige Lösung der Gleichung $\gamma_v(x) = 0$ im Intervall $(-\infty, v)$ ist.

b) Es sei $\beta_1 < 0$. Das eindeutige L -optimale Design ist gegeben durch

$$\xi^* = \left\{ \begin{array}{c} u \\ \frac{\sqrt{Q(\beta_0 + \beta_1 x_2^*)g(x_2^*)}}{\sqrt{Q(\beta_0 + \beta_1 u)g(u) + \sqrt{Q(\beta_0 + \beta_1 x_2^*)g(x_2^*)}}} \end{array} \quad \begin{array}{c} x_2^* \\ \frac{\sqrt{Q(\beta_0 + \beta_1 u)g(u)}}{\sqrt{Q(\beta_0 + \beta_1 u)g(u) + \sqrt{Q(\beta_0 + \beta_1 x_2^*)g(x_2^*)}}} \end{array} \right\}. \quad (6.2)$$

Dabei ist $x_2^* = \min\{v, x^*\}$, wobei x^* die eindeutige Lösung der Gleichung $\gamma_u(x) = 0$ im Intervall (u, ∞) ist.

Der Beweis wird später im Abschnitt über multiple Regressionsmodelle erbracht. Für $\beta_1 = 0$ ist die Informationsmatrix proportional zu der für das lineare Modell und das L -optimale Design hat die beiden Randpunkte u und v des Designraums als Stützpunkte.

Es wird nun ein diskreter Designraum $\mathcal{X} = \{x_1, \dots, x_k\}$ mit $x_1 < x_2 < \dots < x_k$ betrachtet. Der folgende Satz zeigt, dass es für $\beta_1 > 0$ bzw. $\beta_1 < 0$ vier verschiedene Typen von L -optimalen Designs gibt. Der Beweis befindet sich im Anhang.

Satz 6.2

Es sei $\mathcal{X} = \{x_1, \dots, x_k\}$ mit $x_1 < \dots < x_k$. Die Bedingungen (A1)–(A4) seien erfüllt und die Matrix \mathbf{B} sei positiv definit.

a) Es sei $\beta_1 > 0$ und x^* die eindeutige Lösung der Gleichung $\gamma_{x_k}(x) = 0$ im Intervall $(-\infty, x_k)$. Gilt $x^* \leq x_1$, so hat das L -optimale Design die Stützpunkte x_1 und x_k .

Für $x^* > x_1$ sind als L -optimales Design nur vier Designs mit folgenden Stützpunkten möglich:

$$\text{i) } \{x^+(x^*), x_k\} \quad \text{ii) } \{x^-(x^*), x_k\} \quad \text{iii) } \{x^-(x^*), x^+(x^*)\} \quad \text{iv) } \{x^-(x^*), x^+(x^*), x_k\}.$$

b) Es sei $\beta_1 < 0$ und x^* die eindeutige Lösung der Gleichung $\gamma_{x_1}(x) = 0$ im Intervall (x_1, ∞) . Gilt $x^* \geq x_k$, so hat das L -optimale Design die Stützpunkte x_1 und x_k .

Für $x^* < x_k$ sind als L -optimales Design nur vier Designs mit folgenden Stützpunkten

möglich:

$$\text{i) } \{x_1, x^-(x^*)\} \quad \text{ii) } \{x_1, x^+(x^*)\} \quad \text{iii) } \{x^-(x^*), x^+(x^*)\} \quad \text{iv) } \{x_1, x^-(x^*), x^+(x^*)\}.$$

c) Für $\beta_1 = 0$ hat das L -optimale Design die Stützpunkte x_1 und x_k .

Satz 6.2 liefert eine Auswahl von möglichen L -optimalen Designs. Diese Auswahl kann nicht weiter verringert werden, da sich zu jedem der angegebenen Designs ein Beispiel finden lässt, sodass das Design auch tatsächlich L -optimal ist. Für die beiden 2-Punkt-Designs vom Typ i) und ii) ist dies ersichtlich, wenn $x^-(x^*)$ bzw. $x^+(x^*)$ nahe bei x^* liegt. Dass auch die anderen beiden Designs L -optimal sein können, wird im nächsten Beispiel gezeigt. Die optimalen Gewichte der 2-Punkt-Designs wurden bereits angegeben. Die Gewichte des 3-Punkt-Designs müssen numerisch bestimmt werden.

Die Struktur der L -optimalen Designs auf dem diskreten Designraum ähnelt der von den D -optimalen Designs aus Satz 4.8. Im Gegensatz zu D -Optimalität kann jedoch auf dem diskreten Designraum ein Design vom Typ iii) ohne maximales oder minimales Element des Designraums L -optimal sein. Für einen stetigen Designraum ist dies nicht möglich, da das L -optimale Design nach Satz 6.1 immer einen Randpunkt des Designraums besitzt.

Beispiel 6.3

a) Es wird das Proportional Hazards Modell mit Typ I Zensierung, Experimentdauer $c = 1$ und diskretem Designraum $\mathcal{X}_1 = \{-4, -3, 2, 3\}$ betrachtet. Für $\beta = (0, 1)^T$ und $\mathbf{B} = \mathbf{I}_2$ ist $x^* = -0,239$ die eindeutige Lösung von $\gamma_3(x) = 0$ im Intervall $(-\infty, 3)$. In diesem Fall ist L -Optimalität gleich A -Optimalität und das A -optimale Design ist gegeben durch

$$\xi_1^* = \begin{Bmatrix} -3 & 2 \\ 0,762 & 0,238 \end{Bmatrix}. \quad (6.3)$$

b) Für das gleiche Modell wie in a) wird der Designraum zu $\mathcal{X}_2 = \{-4, -3, 1, 3\}$ verändert. Das A -optimale Design ist nun das folgende 3-Punkt-Design:

$$\xi_2^* = \begin{Bmatrix} -3 & 1 & 3 \\ 0,437 & 0,431 & 0,132 \end{Bmatrix}. \quad (6.4)$$

Die Optimalität der beiden Designs in a) und b) kann mit dem Äquivalenzsatz verifiziert werden. Die Fréchet-Ableitung wird in Abbildung 6.1 für eine stetige Variable x auf dem Intervall $[-4, 4]$ graphisch dargestellt.

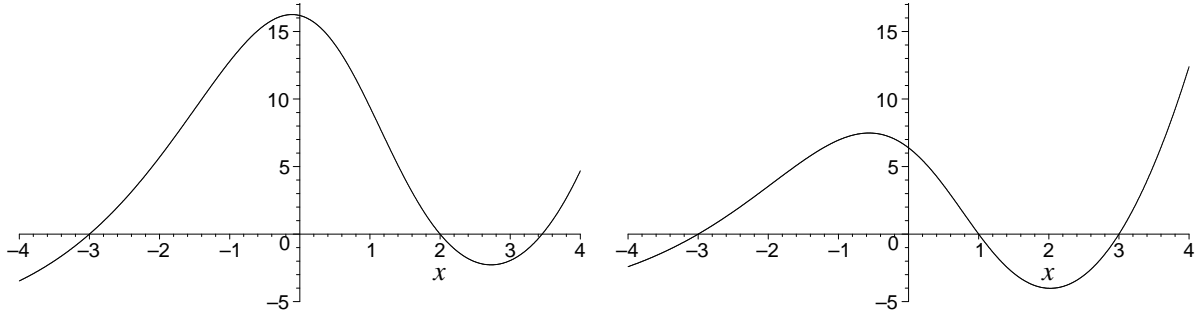


Abbildung 6.1: Fréchet-Ableitung: links für ξ_1^* , rechts für ξ_2^*

6.2 Multiple Regression

In diesem Abschnitt wird ein multiples Regressionsmodell mit $p - 1$ Kovariablen, $p \geq 2$, betrachtet. Es gilt $\mathbf{f}(\mathbf{x}) = (1, \mathbf{x}^T)^T$ und $\mathbf{x} = (x_1, \dots, x_{p-1})^T \in \mathcal{X}$.

Zunächst werden L -optimale Designs für ein kanonisches Design-Problem bestimmt.

Satz 6.4

Es sei $\mathcal{X} = [u_1, 0] \times \dots \times [u_{p-1}, 0]$ und die Bedingungen (A1)–(A4) seien erfüllt. Der Parametervektor sei gegeben durch $\boldsymbol{\beta} = (\beta_0, 1, \dots, 1)^T$ und die Matrix $\mathbf{S} = (s_{i,j})_{i,j=1,\dots,p}$ sei definiert wie in (3.38) als $\mathbf{S} = (\tilde{\mathbf{X}}^T)^{-1} \mathbf{B} \tilde{\mathbf{X}}^{-1}$ für ein Design mit Stützpunkten $x_1 \mathbf{e}_1, \dots, x_{p-1} \mathbf{e}_{p-1}$ und $\mathbf{0}_{p-1} = (0, \dots, 0)^T$.

Für $i = 1, \dots, p - 1$ seien x_i^* die eindeutigen Lösungen des folgenden Gleichungssystems im Intervall $(-\infty, 0)$:

$$x_i + 2 \cdot \frac{Q(\beta_0 + x_i)}{Q'(\beta_0 + x_i)} \cdot \left(1 - \frac{\sqrt{Q(\beta_0 + x_i)}}{\sqrt{Q(\beta_0)}} \cdot \frac{s_{1,i+1}}{\sqrt{s_{1,1} s_{i+1,i+1}}} \right) = 0. \quad (6.5)$$

Gilt $x_i^* \geq u_i$ für $i = 1, \dots, p - 1$, so ist das Design ξ^* mit Stützpunkten $x_1^* \mathbf{e}_1, \dots, x_{p-1}^* \mathbf{e}_{p-1}$ und $\mathbf{0}_{p-1}$ und zugehörigen Gewichten aus (3.38) das eindeutige L -optimale Design.

Beweis:

Es wird das Design-Problem auf dem unbeschränkten Designraum mit $x_i \in (-\infty, 0]$ für $i = 1, \dots, p - 1$ betrachtet. Es seien

$$p_i(x) := \mathbf{f}(x \mathbf{e}_i)^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{B} \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{f}(x \mathbf{e}_i).$$

Nach Satz 3.12, Korollar 3.23 und Bemerkung 4.3 ist ein Design ξ genau dann L -optimal,

wenn

$$Q(\mathbf{f}(x_i \mathbf{e}_i)^T \boldsymbol{\beta}) \cdot p_i(x_i) - \text{Spur}(\mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{B}) \leq 0$$

für $i = 1, \dots, p-1$ gilt. Diese Bedingung ist äquivalent zu

$$h_i(x_i) := p_i(x_i) - \frac{\text{Spur}(\mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{B})}{Q(\mathbf{f}(x_i \mathbf{e}_i)^T \boldsymbol{\beta})} \leq 0. \quad (6.6)$$

Aus der positiven Definitheit der Matrix $\mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{B} \mathbf{M}(\xi, \boldsymbol{\beta})^{-1}$ folgt, dass die quadratischen Polynome $p_i(x_i)$ positiv sind für alle $x_i \in \mathbb{R}$. Sie streben somit für $x_i \rightarrow \infty$ gegen unendlich. Dies gilt auch für $h_i(x_i)$, da die Funktion Q streng monoton wachsend ist. Nach Lemma 2.3 folgt

$$\lim_{x_i \rightarrow -\infty} h_i(x_i) = \lim_{x_i \rightarrow -\infty} \frac{p_i(x_i)Q(\beta_0 + x_i) - \text{Spur}(\mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{B})}{Q(\beta_0 + x_i)} = -\infty.$$

Aus dem Beweis von Lemma 3.26 folgt, dass $h_i(x_i)$ höchstens zwei Extrema hat und ein L -optimales Design kann daher höchstens zwei Stützpunkte je Kante haben. Da für ein L -optimales Design in (6.6) Gleichheit an den Stützpunkten gilt, muss infolge der Grenzwertbetrachtungen einer der beiden Stützpunkte auf einer Kante bei der Ecke $\mathbf{0}_{p-1}$ liegen. Bei dem anderen Stützpunkt auf der Kante muss ein Maximum der Funktion $h_i(x_i)$ vorliegen. Zwischen diesen beiden Stützpunkten existiert ein Minimum von $h_i(x_i)$. Weitere Extrema der Funktion $h_i(x_i)$ gibt es nicht. Ein L -optimales Design ξ^* muss also Stützpunkte $x_1^* \mathbf{e}_1, \dots, x_{p-1}^* \mathbf{e}_{p-1}$ und $\mathbf{0}_{p-1}$ haben mit L -optimalen Gewichten in (3.38), wobei $x_i^* < 0$ für $i = 1, \dots, p-1$. Die Informationsmatrix eines solchen Designs ξ^* lässt sich gemäß (3.35) in der Form

$$\mathbf{M}(\xi^*, \boldsymbol{\beta}) = \tilde{\mathbf{X}}^T \mathbf{D}_\omega \tilde{\mathbf{X}} = \mathbf{X}^T \mathbf{Q}^{\frac{1}{2}} \mathbf{D}_\omega \mathbf{Q}^{\frac{1}{2}} \mathbf{X}$$

darstellen, wobei die Matrizen $\mathbf{Q}^{\frac{1}{2}} = \text{diag}(\sqrt{Q(\beta_0)}, \sqrt{Q(\beta_0 + x_1^*)}, \dots, \sqrt{Q(\beta_0 + x_{p-1}^*)})$ und $\mathbf{D}_\omega = \text{diag}(\omega_1^*, \dots, \omega_p^*)$ Diagonalmatrizen sind und die Matrix \mathbf{X} gegeben ist durch

$$\mathbf{X} = \begin{pmatrix} 1 & \mathbf{0}_{p-1}^T \\ \mathbf{1}_{p-1} & \text{diag}(\mathbf{x}^*) \end{pmatrix}$$

mit Vektoren $\mathbf{x}^* = (x_1^*, \dots, x_{p-1}^*)^T$ und $\mathbf{1}_{p-1} = (1, \dots, 1)^T$.

Die Polynome $p_i(x_i)$ berechnen sich unter Verwendung der Matrix $\mathbf{S} = (\tilde{\mathbf{X}}^T)^{-1} \mathbf{B} \tilde{\mathbf{X}}^{-1}$

aus (3.38) zu

$$\begin{aligned}
p_i(x_i) &= \mathbf{f}(x_i \mathbf{e}_i)^T (\tilde{\mathbf{X}}^T \mathbf{D}_\omega \tilde{\mathbf{X}})^{-1} \mathbf{B} (\tilde{\mathbf{X}}^T \mathbf{D}_\omega \tilde{\mathbf{X}})^{-1} \mathbf{f}(x_i \mathbf{e}_i) \\
&= (1, x_i \mathbf{e}_i^T) \tilde{\mathbf{X}}^{-1} \mathbf{D}_\omega^{-1} (\tilde{\mathbf{X}}^T)^{-1} \mathbf{B} \tilde{\mathbf{X}}^{-1} \mathbf{D}_\omega^{-1} (\tilde{\mathbf{X}}^T)^{-1} (1, x_i \mathbf{e}_i^T)^T \\
&= (1, x_i \mathbf{e}_i^T) \mathbf{X}^{-1} \mathbf{Q}^{-\frac{1}{2}} \mathbf{D}_\omega^{-1} \mathbf{S} \mathbf{D}_\omega^{-1} \mathbf{Q}^{-\frac{1}{2}} (\mathbf{X}^T)^{-1} (1, x_i \mathbf{e}_i^T)^T.
\end{aligned}$$

Die Inverse der Matrix \mathbf{X} ist wie im Beweis von Satz 4.1 gegeben durch

$$\mathbf{X}^{-1} = \begin{pmatrix} 1 & \mathbf{0}_{p-1}^T \\ -\mathbf{y}^* & \text{diag}(\mathbf{y}^*) \end{pmatrix}$$

mit $\mathbf{y}^* = (1/x_1^*, \dots, 1/x_{p-1}^*)^T$. Mit $(1, x_i \mathbf{e}_i^T) \mathbf{X}^{-1} = (1 - x_i/x_i^*, (x_i/x_i^*) \mathbf{e}_i^T)$ folgt

$$\begin{aligned}
p_i(x_i) &= \frac{\left(1 - \frac{x_i}{x_i^*}\right)^2 \cdot s_{1,1}}{(\omega_1^*)^2 Q(\beta_0)} + \frac{\left(\frac{x_i}{x_i^*}\right)^2 \cdot s_{i+1,i+1}}{(\omega_{i+1}^*)^2 Q(\beta_0 + x_i^*)} + 2 \cdot \frac{\left(1 - \frac{x_i}{x_i^*}\right) \frac{x_i}{x_i^*} \cdot s_{1,i+1}}{\omega_1^* \omega_{i+1}^* \sqrt{Q(\beta_0)} \sqrt{Q(\beta_0 + x_i^*)}} \\
&= \left(\sum_{j=1}^p \sqrt{s_{jj}}\right)^2 \cdot \left[\frac{\left(1 - \frac{x_i}{x_i^*}\right)^2}{Q(\beta_0)} + \frac{\left(\frac{x_i}{x_i^*}\right)^2}{Q(\beta_0 + x_i^*)} + 2 \cdot \frac{\left(1 - \frac{x_i}{x_i^*}\right) \frac{x_i}{x_i^*} \cdot \frac{s_{1,i+1}}{\sqrt{s_{1,1} s_{i+1,i+1}}}}{\sqrt{Q(\beta_0)} \sqrt{Q(\beta_0 + x_i^*)}} \right].
\end{aligned}$$

Außerdem gilt:

$$\begin{aligned}
\text{Spur}(\mathbf{M}(\xi^*, \beta)^{-1} \mathbf{B}) &= \text{Spur}((\tilde{\mathbf{X}}^T \mathbf{D}_\omega \tilde{\mathbf{X}})^{-1} \mathbf{B}) = \text{Spur}(\tilde{\mathbf{X}}^{-1} \mathbf{D}_\omega^{-1} (\tilde{\mathbf{X}}^T)^{-1} \mathbf{B}) \\
&= \text{Spur}(\mathbf{D}_\omega^{-1} (\tilde{\mathbf{X}}^T)^{-1} \mathbf{B} \tilde{\mathbf{X}}^{-1}) = \text{Spur}(\mathbf{D}_\omega^{-1} \mathbf{S}) \\
&= \sum_{i=1}^p \frac{s_{ii}}{\omega_i^*} = \sum_{j=1}^p \sqrt{s_{jj}} \cdot \sum_{i=1}^p \sqrt{s_{ii}} = \left(\sum_{j=1}^p \sqrt{s_{jj}}\right)^2.
\end{aligned}$$

Es gilt $h_i(x_i^*) = h_i(0) = 0$. Es muss nun gezeigt werden, dass die Funktionen $h_i(x_i)$ bei $x_i = x_i^*$ ein Maximum haben. Dazu genügt es wegen der Grenzwertbetrachtungen zu zeigen, dass die Gleichungen $h_i'(x_i^*) = 0$ für $i = 1, \dots, p-1$ erfüllt sind. Es gilt

$$h_i'(x_i^*) = \left(\sum_{j=1}^p \sqrt{s_{jj}}\right)^2 \cdot \left[\frac{\frac{2}{x_i^*}}{Q(\beta_0 + x_i^*)} + \frac{2 \cdot \left(-\frac{1}{x_i^*}\right) \cdot \frac{s_{1,i+1}}{\sqrt{s_{1,1} s_{i+1,i+1}}}}{\sqrt{Q(\beta_0)} \sqrt{Q(\beta_0 + x_i^*)}} + \frac{Q'(\beta_0 + x_i^*)}{Q(\beta_0 + x_i^*)^2} \right].$$

Die Gleichungen $h_i'(x_i^*) = 0$ sind äquivalent zu den Gleichungen $k_i(x_i^*) = 0$ mit

$$k_i(x_i) := 2 \cdot \frac{Q(\beta_0 + x_i)}{Q'(\beta_0 + x_i)} \cdot \left(1 - \frac{\sqrt{Q(\beta_0 + x_i)}}{\sqrt{Q(\beta_0)}} \cdot \frac{s_{1,i+1}}{\sqrt{s_{1,1} s_{i+1,i+1}}}\right) + x_i \quad (6.7)$$

für $i = 1, \dots, p-1$, da $k_i(x_i^*) = C \cdot h'_i(x_i^*)$ mit $C = x_i^* Q(\beta_0 + x_i^*)^2 / Q'(\beta_0 + x_i^*) < 0$ gilt. Lösen die x_i^* dieses Gleichungssystem, so ist das Design ξ^* L -optimal.

Zunächst wird die Existenz einer Lösung gezeigt. Nach Lemma A.1 im Anhang gilt $\lim_{x_i \rightarrow 0} s_{1,i+1} / \sqrt{s_{1,1} s_{i+1,i+1}} = -1$ und daher folgt $\lim_{x_i \rightarrow 0} k_i(x_i) = 4Q(\beta_0) / Q'(\beta_0) > 0$. Für die Variable x_i existiert somit eine Umgebung um 0, in der $k_i(x_i)$ positiv ist.

Da die Matrix \mathbf{S} für $x_i \in (-\infty, 0)$ positiv definit ist, gilt $s_{1,i+1}^2 < s_{1,1} s_{i+1,i+1}$. Es folgt $s_{1,i+1} / \sqrt{s_{1,1} s_{i+1,i+1}} \in (-1, 1)$ und der gesamte Ausdruck in runden Klammern in (6.7) nimmt somit Werte im Intervall $(0, 2)$ an. Es gilt mit (A4)

$$k_i \left(\frac{-4Q(\beta_0)}{Q'(\beta_0)} \right) < 2 \cdot \frac{Q(\beta_0)}{Q'(\beta_0)} \cdot 2 - \frac{4Q(\beta_0)}{Q'(\beta_0)} = 0$$

für $i = 1, \dots, p-1$. Aus dem Satz von Poincaré-Miranda (vgl. Kulpa, 1997) folgt die Existenz einer Lösung des Gleichungssystems $k_i(x_i^*) = 0$, $i = 1, \dots, p-1$, die sich im Inneren des Würfels $[-4Q(\beta_0)/Q'(\beta_0), 0]^{p-1}$ befinden muss.

Angenommen, dieses Gleichungssystem hätte mehrere Lösungen. Dann erfüllen alle aus diesen Lösungen konstruierten Designs den Äquivalenzsatz und wären daher L -optimal. Sind ξ_1 und ξ_2 zwei solche L -optimale Designs, dann wäre auch das Design $\xi_3 = \frac{1}{2}(\xi_1 + \xi_2)$ L -optimal, da die Kriteriumsfunktion konkav ist. Es existiert eine Kante, auf der das Design ξ_3 drei Stützpunkte hat. Dies liefert einen Widerspruch zu Lemma 3.26 und es folgt die Eindeutigkeit der Lösung des Gleichungssystems.

Nun wird der Fall $p = 2$ betrachtet, d. h. das Modell mit einer Kovariablen. Im Beweis von Lemma A.1 wurde der Ausdruck $s_{1,i+1} / \sqrt{s_{1,1} s_{i+1,i+1}}$ berechnet. Somit ist die Gleichung für x_1^* gegeben durch

$$k_1(x) := 2 \cdot \frac{Q(\beta_0 + x)}{Q'(\beta_0 + x)} \cdot \left(1 - \frac{\sqrt{Q(\beta_0 + x)}}{\sqrt{Q(\beta_0)}} \cdot \frac{b_{12}x - b_{22}}{\sqrt{g(0)g(x)}} \right) + x = 0$$

mit $g(x) := b_{11}x^2 - 2b_{12}x + b_{22}$. Es wird der Fall betrachtet, dass die Lösung der Gleichung $k_1(x) = 0$ kleiner als u_1 ist. Für das Design mit Stützpunkten u_1 und 0 gilt $h'_1(u_1) < 0$ genau dann, wenn $k_1(u_1) > 0$ ist, da $k_1(u_1) = \tilde{C} \cdot h'_1(u_1)$ mit $\tilde{C} < 0$ gilt. Da die Gleichung $k_1(x) = 0$ nach Annahme keine Lösung im Intervall $[u_1, 0]$ hat, haben $k_1(u_1)$ und $k_1(0)$ das gleiche Vorzeichen. Wegen $k_1(0) = 4Q(\beta_0) / Q'(\beta_0) > 0$ ist es positiv und somit folgt $h'_1(u_1) < 0$. Das Maximum \tilde{x} von h_1 liegt also bei $\tilde{x} < u_1$ und es gilt $h_1(x_1) \leq 0$ für alle $x_1 \in [u_1, 0]$. Nach dem Äquivalenzsatz ist das Design mit Stützpunkten u_1 und 0 L -optimal. \square

Mit Hilfe von Satz 6.4 lassen sich L -optimale Designs für beliebige Designräume und Parametervektoren herleiten.

Satz 6.5

Es sei $\mathcal{X} = [u_1, v_1] \times \dots \times [u_{p-1}, v_{p-1}]$ und die Bedingungen (A1)–(A4) seien erfüllt. Für $i = 1, \dots, p-1$ sei $a_i = v_i$, falls $\beta_i > 0$ und $a_i = u_i$, falls $\beta_i < 0$. Die Matrix \mathbf{S} sei definiert wie in (3.38) als $\mathbf{S} = (\tilde{\mathbf{X}}^T)^{-1} \mathbf{B} \tilde{\mathbf{X}}^{-1}$ für ein Design mit Stützpunkten $\mathbf{a} - (x_1/\beta_1)\mathbf{e}_1, \dots, \mathbf{a} - (x_{p-1}/\beta_{p-1})\mathbf{e}_{p-1}$ und $\mathbf{a} = (a_1, \dots, a_{p-1})^T$.

Für $i = 1, \dots, p-1$ seien x_i^* die eindeutigen Lösungen des folgenden Gleichungssystems im Intervall $(0, \infty)$:

$$x_i - 2 \cdot \frac{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i)}{Q'(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i)} \cdot \left(1 - \frac{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i)}}{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}} \cdot \frac{s_{1,i+1}}{\sqrt{s_{1,1} s_{i+1,i+1}}} \right) = 0. \quad (6.8)$$

Gilt $x_i^* \leq |\beta_i|(v_i - u_i)$ für $i = 1, \dots, p-1$, so ist das Design mit den Stützpunkten $\mathbf{a} - (x_1^*/\beta_1)\mathbf{e}_1, \dots, \mathbf{a} - (x_{p-1}^*/\beta_{p-1})\mathbf{e}_{p-1}$ und \mathbf{a} und zugehörigen Gewichten aus (3.38) das eindeutige L -optimale Design.

Beweis:

Es sei $\mathbf{d} = (\beta_1 a_1, \dots, \beta_{p-1} a_{p-1})^T$. Die Matrix \mathbf{T} sei definiert durch

$$\mathbf{T} = \begin{pmatrix} 1 & \mathbf{0}_{p-1}^T \\ -\mathbf{d} & \text{diag}(\beta_1, \dots, \beta_{p-1}) \end{pmatrix}.$$

Mit der kanonischen Transformation $\mathbf{f}(\mathbf{x}) \rightarrow \mathbf{f}(\mathbf{z}) = \mathbf{T} \mathbf{f}(\mathbf{x})$ folgt der Zusammenhang $\mathbf{M}(\xi, \boldsymbol{\beta}) = \mathbf{T}^{-1} \mathbf{M}_z(\xi_z, \boldsymbol{\beta}_z) (\mathbf{T}^{-1})^T$ mit transformiertem Design ξ_z , Informationsmatrix

$$\mathbf{M}_z(\xi_z, \boldsymbol{\beta}_z) = \sum_{i=1}^m \omega_i Q(\mathbf{f}(\mathbf{z}_i)^T \boldsymbol{\beta}_z) \mathbf{f}(\mathbf{z}_i) \mathbf{f}(\mathbf{z}_i)^T,$$

und transformiertem Parametervektor $\boldsymbol{\beta}_z = (\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta}, 1, \dots, 1)^T$. Es folgt

$$\text{Spur}(\mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{B}) = \text{Spur}(\mathbf{T}^T \mathbf{M}_z(\xi_z, \boldsymbol{\beta}_z)^{-1} \mathbf{T} \mathbf{B}) = \text{Spur}(\mathbf{M}_z(\xi_z, \boldsymbol{\beta}_z)^{-1} \mathbf{B}_z)$$

mit $\mathbf{B}_z = \mathbf{T} \mathbf{B} \mathbf{T}^T$. Nach Lemma 3.19 gilt $\mathbf{M}_z(\xi_z, \boldsymbol{\beta}_z) = \tilde{\mathbf{Z}}^T \mathbf{D}_\omega \tilde{\mathbf{Z}}$, wobei $\tilde{\mathbf{Z}}^T = \mathbf{T} \tilde{\mathbf{X}}^T$ ist. Die Matrix \mathbf{S} bleibt unter linearen Transformationen invariant:

$$\begin{aligned} \mathbf{S}_z &= (\tilde{\mathbf{Z}}^T)^{-1} \mathbf{B}_z \tilde{\mathbf{Z}}^{-1} = (\tilde{\mathbf{Z}}^T)^{-1} \mathbf{T} \mathbf{B} \mathbf{T}^T \tilde{\mathbf{Z}}^{-1} = (\mathbf{T}^{-1} \tilde{\mathbf{Z}}^T)^{-1} \mathbf{B} (\mathbf{T}^{-1} \tilde{\mathbf{Z}}^T)^{-T} \\ &= (\tilde{\mathbf{X}}^T)^{-1} \mathbf{B} \tilde{\mathbf{X}}^{-1} = \mathbf{S}. \end{aligned}$$

Es gilt $z_i \in [-|\beta_i|(v_i - u_i), 0]$ für $i = 1, \dots, p-1$. Damit liegen die Voraussetzungen von Satz 6.4 vor. Das L -optimale Design folgt aus der Rücktransformation des L -optimalen Designs aus Satz 6.4. Dabei wird das Gleichungssystem (6.5) transformiert zu

$$0 = \beta_i(x_i - a_i) + 2 \cdot \frac{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} + \beta_i(x_i - a_i))}{Q'(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} + \beta_i(x_i - a_i))} \cdot \left(1 - \frac{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} + \beta_i(x_i - a_i))}}{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}} \cdot \frac{s_{1,i+1}}{\sqrt{s_{1,1}s_{i+1,i+1}}} \right).$$

Mit der Umparametrisierung $x_i = a_i - \tilde{x}_i/\beta_i$ folgt das Gleichungssystem in (6.8). \square

Bemerkung 6.6

Die optimalen Gewichte des L -optimalen Designs aus Satz 6.5 sind nach Satz 3.20 für $i = 1, \dots, p$ gegeben durch

$$\omega_i^* = \frac{\sqrt{s_{ii}}}{\sum_{j=1}^p \sqrt{s_{jj}}},$$

wobei $\mathbf{S} = (\tilde{\mathbf{X}}^T)^{-1} \mathbf{B} \tilde{\mathbf{X}}^{-1}$. Mit $\mathbf{y}^* = (\beta_1/x_1^*, \dots, \beta_{p-1}/x_{p-1}^*)^T$ berechnen sich die Diagonaleinträge von \mathbf{S} für $j = 2, \dots, p-1$ zu

$$s_{1,1} = \frac{1}{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})} \left[b_{1,1}(1 - \mathbf{a}^T \mathbf{y}^*)^2 + 2(1 - \mathbf{a}^T \mathbf{y}^*) \sum_{k=2}^p b_{1,k} \frac{\beta_{k-1}}{x_{k-1}^*} + \sum_{l=2}^p \sum_{k=2}^p b_{l,k} \frac{\beta_{l-1} \beta_{k-1}}{x_{l-1}^* x_{k-1}^*} \right],$$

$$s_{j,j} = \frac{\beta_{j-1}^2}{(x_{j-1}^*)^2} \cdot \frac{b_{1,1} a_{j-1}^2 - 2b_{1,j} a_{j-1} + b_{j,j}}{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_{j-1}^*)}.$$

Der Beweis zu Bemerkung 6.6 befindet sich im Anhang.

Aus dem Beweis von Satz 6.4 folgt, dass die eindeutigen Lösungen x_i^* des Gleichungssystems (6.8) im Intervall $(0, 4Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})/Q'(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta}))$ liegen. Dieses Gleichungssystem muss numerisch gelöst werden. Es lässt sich als Fixpunktgleichung $\mathbf{x} = \mathbf{k}(\mathbf{x})$ schreiben. Ausgehend von einem Startvektor \mathbf{x}_0 kann das Fixpunktverfahren $\mathbf{x}_{n+1} = \mathbf{k}(\mathbf{x}_n)$ verwendet werden, um iterativ eine Lösung zu bestimmen. Bei numerischen Untersuchungen konvergierte dieses Verfahren immer gegen den Fixpunkt.

Die Struktur der L -optimalen Designs aus Satz 6.5 und der D -optimalen Designs aus Satz 4.9 ist in Bezug auf die Lage der Stützpunkte ähnlich. Ein Stützpunkt ist die Ecke des Designraums, bei der die Intensitätsfunktion Q maximal wird, und die weiteren Stützpunkte befinden sich auf den zu dieser Ecke inzidenten Kanten des Designraums.

Bezüglich der Gewichte unterscheiden sich die L -optimalen Designs von den D -optimalen Designs. Während bei D -Optimalität gleiche Gewichte optimal sind, sind die optimalen Gewichte bei L -Optimalität im Allgemeinen unterschiedlich, wie das folgende Beispiel zeigt.

Beispiel 6.7

Es wird die Situation aus Beispiel 5.11 betrachtet, d. h. das Proportional Hazards Modell mit Typ I Zensierung, Experimentdauer $c = 1$, zwei Kovariablen, $\mathcal{X} = [0, 10] \times [0, 10]$ und $\boldsymbol{\beta} = (0, -1, -1)^T$.

a) Gesucht ist das A -optimale Design, d. h. es ist $\mathbf{B} = \mathbf{I}_3$. Die Einträge der Matrix \mathbf{S} aus (3.38) sind für $i = 1, 2$ gegeben durch

$$s_{1,1} = \frac{1}{Q(0)} \cdot \left(1 + \frac{1}{x_1^2} + \frac{1}{x_2^2} \right),$$

$$s_{1,i+1} = -\frac{1}{\sqrt{Q(0)Q(-x_i)x_i^2}},$$

$$s_{i+1,i+1} = \frac{1}{Q(-x_i)x_i^2}.$$

In diesem Beispiel sind die zwei Gleichungen des Gleichungssystems (6.8) identisch, woraus unmittelbar $x_1^* = x_2^*$ folgt. Es muss daher nur eine Gleichung gelöst werden. Man erhält $x_1^* = x_2^* = 2,379$. Das A -optimale Design ergibt sich nach Satz 6.5 zu

$$\xi^* = \left\{ \begin{array}{ccc} (2,379; 0) & (0; 2,379) & (0; 0) \\ 0,329 & 0,329 & 0,341 \end{array} \right\}. \quad (6.9)$$

Für das marginale Modell mit einer Kovariablen, $\mathcal{X}_1 = [0, 10]$ und $\boldsymbol{\beta} = (0, -1)^T$ hat das A -optimale Design die Stützpunkte 2,394 und 0 mit Gewichten 0,509 und 0,491. Der Abstand der Stützpunkte auf den Kanten zur Ecke im Modell mit zwei Kovariablen verringert sich in diesem Beispiel unwesentlich im Vergleich zum Abstand der Stützpunkte im marginalen Modell. Es verändert sich jedoch die Gewichtung zwischen den Stützpunkten. Während im marginalen Modell dem Randpunkt 0 ein kleineres Gewicht zugewiesen wird, erhält die Ecke (0; 0) im Modell mit zwei Kovariablen ein größeres Gewicht im Vergleich zu den Stützpunkten auf den Kanten.

Für das aus den optimalen marginalen Designpunkten konstruierte Design mit den Stützpunkten (2,394; 0), (0; 2,394) und (0; 0) ergeben sich die optimalen Gewichte zu $\omega_1^* = \omega_2^* = 0,330$ und $\omega_3^* = 0,341$. Dieses Design weist eine A -Effizienz von 0,99996 auf. Die Beschränkung auf die Optimierung der Gewichte liefert nicht immer Designs mit

so hohen Effizienzen. Beispielsweise hat für den Designraum $\tilde{\mathcal{X}} = [-3, 3] \times [-3, 3]$ das resultierende Design eine A -Effizienz von nur 0,812.

b) Gesucht ist das V -optimale Design. Es sei μ die Gleichverteilung auf \mathcal{X} . Die Matrix \mathbf{B} ist gegeben durch

$$\mathbf{B} = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^T \mu(d\mathbf{x}) = \begin{pmatrix} 1 & 5 & 5 \\ 5 & \frac{100}{3} & 25 \\ 5 & 25 & \frac{100}{3} \end{pmatrix}. \quad (6.10)$$

Das Gleichungssystem (6.8), das sich wieder zu einer Gleichung reduzieren lässt, hat die Lösung $x_1^* = x_2^* = 2,689$. Das V -optimale Design ergibt sich nach Satz 6.5 zu

$$\xi^* = \left\{ \begin{array}{ccc} (2,689; 0) & (0; 2,689) & (0; 0) \\ 0,405 & 0,405 & 0,189 \end{array} \right\}. \quad (6.11)$$

Die beiden Stützpunkte auf den Kanten sind weiter entfernt von der Ecke als beim A -optimalen Design. Die Ecke erhält im Gegensatz zum A -optimalen Design ein geringeres Gewicht als die Stützpunkte auf den Kanten.

Für den Fall, dass einige β_i , $i > 0$, gleich null sind, lassen sich aus numerischen Untersuchungen die gleichen Erkenntnisse wie bei D -Optimalität gewinnen. Optimale Designs lassen sich als Produkt von einem optimalen marginalen Design auf den Ecken für diejenigen Komponenten mit $\beta_k = 0$ und einem Design von der Struktur des L -optimalen Designs aus Satz 6.5 für diejenigen Komponenten mit $\beta_j \neq 0$ konstruieren.

7 ϕ_k -optimale Designs

Dieses Kapitel widmet sich der Bestimmung ϕ_k -optimaler Designs. In Abschnitt 7.1 werden ϕ_k -optimale Designs für $k \in (-\infty, 1)$ berechnet. Da der Äquivalenzsatz 3.12 nicht für $k = -\infty$ gilt, wird dieser Fall, bei dem es sich um E -Optimalität handelt, in Abschnitt 7.2 separat behandelt.

7.1 ϕ_k -optimale Designs für $k \in (-\infty, 1)$

In diesem Abschnitt werden für $k \in (-\infty, 1)$ ϕ_k -optimale Designs für ein Modell mit $p - 1$ Kovariablen, $p \geq 2$, $\mathbf{f}(\mathbf{x}) = (1, \mathbf{x}^T)^T$ und $\mathbf{x} = (x_1, \dots, x_{p-1})^T \in \mathcal{X}$ berechnet. Der Fall $p = 2$ ist also mit eingeschlossen. Die ϕ_k -optimalen Designs lassen sich ähnlich zum Beweis von Satz 6.4 für L -Optimalität bestimmen.

Satz 7.1

Es sei $\mathcal{X} = [u_1, v_1] \times \dots \times [u_{p-1}, v_{p-1}]$ und die Bedingungen (A1)–(A4) seien erfüllt. Für $i = 1, \dots, p - 1$ sei $a_i = v_i$, falls $\beta_i > 0$ und $a_i = u_i$, falls $\beta_i < 0$. Die Matrix \mathbf{S} sei definiert wie in (3.39) als $\mathbf{S} = (\tilde{\mathbf{X}}^T)^{-1} \mathbf{M}(\xi, \boldsymbol{\beta})^{k+1} \tilde{\mathbf{X}}^{-1}$ für ein Design mit Stützpunkten $\mathbf{a} - (x_1/\beta_1)\mathbf{e}_1, \dots, \mathbf{a} - (x_{p-1}/\beta_{p-1})\mathbf{e}_{p-1}$ und $\mathbf{a} = (a_1, \dots, a_{p-1})^T$.

Falls eine Lösung existiert, so seien $x_i^* \in (0, \infty)$ und ω_i^* für $i = 1, \dots, p - 1$ die eindeutigen Lösungen des gemeinsamen Gleichungssystems (7.1) und (3.39):

$$x_i - 2 \cdot \frac{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i)}{Q'(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i)} \cdot \left(1 - \frac{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i)}}{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}} \cdot \frac{s_{1,i+1}}{\sqrt{s_{1,1} s_{i+1,i+1}}} \right) = 0. \quad (7.1)$$

Gilt $x_i^* \leq |\beta_i| (v_i - u_i)$ für $i = 1, \dots, p - 1$, so ist das Design ξ_k^* mit den Stützpunkten $\mathbf{a} - (x_1^*/\beta_1)\mathbf{e}_1, \dots, \mathbf{a} - (x_{p-1}^*/\beta_{p-1})\mathbf{e}_{p-1}$ und \mathbf{a} und zugehörigen Gewichten $\omega_1^*, \dots, \omega_p^*$ das eindeutige ϕ_k -optimale Design.

Der Beweis befindet sich im Anhang. Satz 7.1 unterscheidet sich von Satz 6.5 für L -Optimalität in der Definition der Matrix \mathbf{S} . Da sie von den Gewichten abhängt, beinhaltet das zu lösende Gleichungssystem zusätzlich auch die Gleichungen für die optimalen Gewichte. Dieses gemeinsame Gleichungssystem muss numerisch gelöst werden. Es lässt

sich wie bei L -Optimalität als Fixpunktgleichung schreiben und mit iterativen Verfahren lösen. Bei numerischen Untersuchungen hatte es immer eine Lösung.

Bemerkung 7.2

Satz 7.1 gilt bis auf die Eindeutigkeit des ϕ_k -optimalen Designs auch für $k = 1$. Da in diesem Fall die Kriteriumsfunktion linear und daher nicht streng konkav ist (vgl. Pukelsheim, 1993, S. 138–139), lässt sich die Eindeutigkeit nur für die Stützpunkte, nicht aber für die Gewichte zeigen.

Bemerkung 7.3

Für D -Optimalität gilt $\mathbf{S} = \text{diag}(\omega_1, \dots, \omega_p)$ und somit $s_{1,i+1} = 0$, $i = 1, \dots, p - 1$. Damit vereinfacht sich das Gleichungssystem (7.1) zu $p - 1$ identischen Gleichungen von der Form $\phi_{\mathbf{a}}(x_i) = 0$ mit Lösungen $x_i = \phi_{\mathbf{a}}^{-1}(0)$, wobei die Funktion $\phi_{\mathbf{a}}$ in (4.7) definiert ist. Somit kann Satz 4.9 aus Satz 7.1 gefolgert werden. Insbesondere ist das D -optimale Design aus Satz 4.9 eindeutig.

Beispiel 7.4

Es wird die Situation aus den Beispielen 5.11 und 6.7 betrachtet. Gesucht ist das ϕ_k -optimale Design für $k = -2$. Die Einträge der Matrix \mathbf{S} aus (3.39) sind für $i = 1, 2$ gegeben durch

$$s_{1,1} = \frac{\left(1 + \frac{1}{x_1^2} + \frac{1}{x_2^2}\right)^2}{Q(0)^2 \omega_1} + \frac{1}{Q(0)Q(-x_1)x_1^4 \omega_2} + \frac{1}{Q(0)Q(-x_2)x_2^4 \omega_3},$$

$$s_{1,i+1} = -\frac{1 + \frac{1}{x_1^2} + \frac{1}{x_2^2}}{Q(0)^{\frac{3}{2}}Q(-x_i)^{\frac{1}{2}}x_i^2 \omega_1} - \frac{1}{Q(0)^{\frac{1}{2}}Q(-x_i)^{\frac{3}{2}}x_i^4 \omega_{i+1}},$$

$$s_{i+1,i+1} = \frac{1}{Q(0)Q(-x_i)x_i^4 \omega_1} + \frac{1}{Q(-x_i)^2 x_i^4 \omega_{i+1}}.$$

Das gemeinsame Gleichungssystem (7.1) und (3.39) hat die Lösung $x_1^* = x_2^* = 2,513$, $\omega_1^* = \omega_2^* = 0,327$ und $\omega_3^* = 0,345$. Das ϕ_{-2} -optimale Design ist in Tabelle 7.1 angegeben. Zum Vergleich sind für das gleiche Modell das D -optimale Design, das A -optimale Design (vgl. Beispiel 6.7) und das c -optimale Design für $\mathbf{c} = (0, 1, 1)^T$ (vgl. Beispiel 5.11) in Tabelle 7.1 angegeben. Als c -optimales Design wurde das 3-Punkt-Design mit allen Stützpunkten auf dem Rand des Designraums ausgewählt. Außerdem wurden die Effizienzen der optimalen Designs bezüglich der anderen Optimalitätskriterien berechnet. Es zeigt sich, dass alle Designs relativ hohe Effizienzen bezüglich der anderen Kriterien aufweisen.

Kriterium	Optimales Design	Effizienzen			
		D	c	A	ϕ_{-2}
D -Optimalität (ϕ_0)	$\left\{ \begin{array}{ccc} (2,124; 0) & (0; 2,124) & (0; 0) \\ 1/3 & 1/3 & 1/3 \end{array} \right\}$	1	0,920	0,986	0,965
c -Optimalität	$\left\{ \begin{array}{ccc} (2,723; 0) & (0; 2,723) & (0; 0) \\ 0,380 & 0,380 & 0,241 \end{array} \right\}$	0,930	1	0,943	0,931
A -Optimalität (ϕ_{-1})	$\left\{ \begin{array}{ccc} (2,379; 0) & (0; 2,379) & (0; 0) \\ 0,329 & 0,329 & 0,341 \end{array} \right\}$	0,990	0,956	1	0,996
ϕ_{-2} -Optimalität	$\left\{ \begin{array}{ccc} (2,513; 0) & (0; 2,513) & (0; 0) \\ 0,327 & 0,327 & 0,345 \end{array} \right\}$	0,977	0,960	0,997	1

Tabelle 7.1: Vergleich optimaler Designs für verschiedene Optimalitätskriterien

7.2 E -optimale Designs

In diesem Abschnitt werden E -optimale Designs für das Modell mit einer Kovariablen bestimmt. Es gilt $p = 2$, $\mathbf{f}(x) = (1, x)^T$ und $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$.

Dette und Haines (1994) haben Modelle mit zwei Parametern betrachtet und explizite Formeln für die optimalen Gewichte und die Eigenwerte der Informationsmatrix eines E -optimalen Designs mit zwei Stützpunkten angegeben. Wie im Kapitel über c -optimale Designs sei $\tilde{\mathbf{f}}(x) = \sqrt{Q(\beta_0 + \beta_1 x)} \cdot \mathbf{f}(x)$. Die E -optimalen Gewichte eines Designs mit Stützpunkten x_1 und x_2 lassen sich darstellen als

$$\omega_1^* = \frac{\tilde{\mathbf{f}}(x_2)^T (\tilde{\mathbf{f}}(x_2) - \tilde{\mathbf{f}}(x_1))}{\|\tilde{\mathbf{f}}(x_1) - \tilde{\mathbf{f}}(x_2)\|^2} \quad \text{und} \quad \omega_2^* = \frac{\tilde{\mathbf{f}}(x_1)^T (\tilde{\mathbf{f}}(x_1) - \tilde{\mathbf{f}}(x_2))}{\|\tilde{\mathbf{f}}(x_1) - \tilde{\mathbf{f}}(x_2)\|^2}, \quad (7.2)$$

falls für den Winkel $\alpha = \angle(\tilde{\mathbf{f}}(x_1), \tilde{\mathbf{f}}(x_2)) \in [0^\circ, 180^\circ)$ gilt, dass $\alpha \geq 90^\circ$ ist. Gilt $\alpha < 90^\circ$, so ist $\tilde{\mathbf{f}}(x_2)$ durch $-\tilde{\mathbf{f}}(x_2)$ zu ersetzen. Das folgende Resultat von Dette und Haines (1994) liefert eine geometrische Bedingung, wann das E -optimale Design mit einem c -optimalen Design übereinstimmt. Dabei wird das Konzept des Inkugel-Vektors verwendet, der auf den Berührungspunkt der größten Kugel innerhalb der Elfving-Menge \mathcal{R} mit dem Rand der Elfving-Menge zeigt.

Satz 7.5

Für ein Modell mit zwei Parametern existiert ein Inkugel-Vektor \mathbf{c} , sodass das E -optimale Design mit dem c -optimalen Design für \mathbf{c} übereinstimmt, genau dann, wenn der Rand der Elfving-Menge ein Geradenstück enthält, sodass das mit dem Ursprung gebildete Dreieck einen Winkel größer oder gleich 90° am Ursprung aufweist.

Beträgt der Winkel mehr als 90° , so ist der Inkugel-Vektor bis auf Vorzeichen eindeutig, beträgt er 90° , so kann es nur dann mehrere Inkugel-Vektoren geben, wenn die Elfving-Menge ein Quadrat ist. Im folgenden Satz wird die Funktion ψ_a verwendet, die in (5.2) definiert ist.

Satz 7.6

Es sei $\mathcal{X} = [u, v]$ und die Bedingungen (A1)–(A4) seien erfüllt. Die Funktion g sei definiert als $g(x) := (c_1x - c_2)^2$ für einen Vektor $\mathbf{c} = (c_1, c_2)^T$.

a) Es sei $\beta_1 > 0$ und $x^* = \max\{u, v - \psi_v^{-1}(0)/\beta_1\}$. Gilt $1 + vx^* \geq 0$, so stimmt das eindeutige E -optimale Design mit dem c -optimalen Design für

$$\mathbf{c} = \begin{pmatrix} x^* \sqrt{Q(\beta_0 + \beta_1 x^*)} + v \sqrt{Q(\beta_0 + \beta_1 v)} \\ -\sqrt{Q(\beta_0 + \beta_1 x^*)} - \sqrt{Q(\beta_0 + \beta_1 v)} \end{pmatrix} \quad (7.3)$$

überein:

$$\xi^* = \left\{ \begin{array}{cc} x^* & v \\ \frac{\sqrt{Q(\beta_0 + \beta_1 v)g(v)}}{\sqrt{Q(\beta_0 + \beta_1 x^*)g(x^*)} + \sqrt{Q(\beta_0 + \beta_1 v)g(v)}} & \frac{\sqrt{Q(\beta_0 + \beta_1 x^*)g(x^*)}}{\sqrt{Q(\beta_0 + \beta_1 x^*)g(x^*)} + \sqrt{Q(\beta_0 + \beta_1 v)g(v)}} \end{array} \right\}. \quad (7.4)$$

Gilt $1 + vx^* < 0$, so ist das eindeutige E -optimale Design gegeben durch

$$\xi^* = \left\{ \begin{array}{cc} -1/v & v \\ \frac{v^2 Q(\beta_0 + \beta_1 v)}{Q(\beta_0 - \beta_1/v) + v^2 Q(\beta_0 + \beta_1 v)} & \frac{Q(\beta_0 - \beta_1/v)}{Q(\beta_0 - \beta_1/v) + v^2 Q(\beta_0 + \beta_1 v)} \end{array} \right\}. \quad (7.5)$$

b) Es sei $\beta_1 < 0$ und $x^* = \min\{v, u - \psi_u^{-1}(0)/\beta_1\}$. Gilt $1 + ux^* \geq 0$, so stimmt das eindeutige E -optimale Design mit dem c -optimalen Design für

$$\mathbf{c} = \begin{pmatrix} x^* \sqrt{Q(\beta_0 + \beta_1 x^*)} + u \sqrt{Q(\beta_0 + \beta_1 u)} \\ -\sqrt{Q(\beta_0 + \beta_1 x^*)} - \sqrt{Q(\beta_0 + \beta_1 u)} \end{pmatrix} \quad (7.6)$$

überein:

$$\xi^* = \left\{ \begin{array}{cc} u & x^* \\ \frac{\sqrt{Q(\beta_0 + \beta_1 x^*)g(x^*)}}{\sqrt{Q(\beta_0 + \beta_1 u)g(u)} + \sqrt{Q(\beta_0 + \beta_1 x^*)g(x^*)}} & \frac{\sqrt{Q(\beta_0 + \beta_1 u)g(u)}}{\sqrt{Q(\beta_0 + \beta_1 u)g(u)} + \sqrt{Q(\beta_0 + \beta_1 x^*)g(x^*)}} \end{array} \right\}. \quad (7.7)$$

Gilt $1 + ux^* < 0$, so ist das eindeutige E -optimale Design gegeben durch

$$\xi^* = \left\{ \begin{array}{cc} u & -1/u \\ \frac{Q(\beta_0 - \beta_1/u)}{Q(\beta_0 - \beta_1/u) + u^2 Q(\beta_0 + \beta_1 u)} & \frac{u^2 Q(\beta_0 + \beta_1 u)}{Q(\beta_0 - \beta_1/u) + u^2 Q(\beta_0 + \beta_1 u)} \end{array} \right\}. \quad (7.8)$$

Beweis:

a) Es sei $\beta_1 > 0$. Der Rand der Elfving-Menge wurde im Beweis von Satz 5.2 für das kanonische Modell charakterisiert. Er besteht für $\beta_1 > 0$ aus den Kurven $\pm \tilde{\mathbf{f}}(x)$ für $x \in [x^*, v]$ und aus den Geradenstücken von $-\tilde{\mathbf{f}}(v)$ zu $\tilde{\mathbf{f}}(x^*)$ und von $-\tilde{\mathbf{f}}(x^*)$ zu $\tilde{\mathbf{f}}(v)$, wobei x^* in Satz 7.6 a) definiert ist. Das von den Geradenstücken mit dem Ursprung gebildete Dreieck weist genau dann einen Winkel größer oder gleich 90° am Ursprung auf, wenn $\angle(-\tilde{\mathbf{f}}(v), \tilde{\mathbf{f}}(x^*)) \geq 90^\circ$ gilt. Diese Bedingung ist äquivalent zu $1 + vx^* \geq 0$. Die größte Kugel innerhalb der Elfving-Menge berührt die beiden Geradenstücke an der Stelle mit dem geringsten Abstand zum Ursprung. Der Inkugel-Vektor, der vom Ursprung zu diesem Berührungspunkt zeigt, muss also senkrecht zum Geradenstück sein. Die Richtung des Inkugel-Vektors ergibt sich zu

$$\mathbf{c} = \begin{pmatrix} \tilde{f}_2(x^*) + \tilde{f}_2(v) \\ -\tilde{f}_1(x^*) - \tilde{f}_1(v) \end{pmatrix} = \begin{pmatrix} x^* \sqrt{Q(\beta_0 + \beta_1 x^*)} + v \sqrt{Q(\beta_0 + \beta_1 v)} \\ -\sqrt{Q(\beta_0 + \beta_1 x^*)} - \sqrt{Q(\beta_0 + \beta_1 v)} \end{pmatrix}, \quad (7.9)$$

wobei das Minuszeichen in die zweite Komponente gesetzt wurde. Nach Satz 7.5 stimmt das E -optimale Design mit dem c -optimalen Design für den Vektor \mathbf{c} überein, da c -Optimalität unabhängig von der Skalierung und vom Vorzeichen des Vektors \mathbf{c} ist. Insbesondere erhält man als E -optimale Gewichte die c -optimalen Gewichte, die mit (7.2) übereinstimmen, und die Eindeutigkeit des E -optimalen Designs.

Gilt $1 + vx^* < 0$, so stimmt das E -optimale Design nach Satz 7.5 nicht mit einem c -optimalen Design überein. Die E -optimale Informationsmatrix hat daher einen Eigenwert mit Vielfachheit 2. Nach dem Äquivalenzsatz ist ein Design ξ genau dann E -optimal, wenn eine positiv semidefinite Matrix \mathbf{E} mit $\text{Spur}(\mathbf{E}) = 1$ existiert, sodass

$$h(x) := \mathbf{f}(x)^T \mathbf{E} \mathbf{f}(x) - \frac{\lambda_{\min}(\mathbf{M}(\xi, \boldsymbol{\beta}))}{Q(\beta_0 + \beta_1 x)} \leq 0 \quad (7.10)$$

für alle $x \in [u, v]$ gilt. Aufgrund der Kompaktheit des Designraums und der Stetigkeit der Kriteriumsfunktion existiert ein E -optimales Design, wobei Gleichheit in (7.10) an den Stützpunkten gilt. Wie in den Beweisen von Lemma 3.26 und Satz 6.4 lässt sich zeigen, dass die Funktion $h'(x)$ höchstens zwei Nullstellen hat und $\lim_{x \rightarrow \pm\infty} h(x) = \pm\infty$ gilt. Somit hat das E -optimale Design genau zwei Stützpunkte, wobei sich ein Stützpunkt bei v befinden muss. Nach Lemma 3.3 von Dette und Haines (1994) muss für die E -optimalen Stützpunkte x_1 und v gelten, dass $\tilde{\mathbf{f}}(x_1)$ und $\tilde{\mathbf{f}}(v)$ senkrecht zueinander sind, damit der Eigenwert Vielfachheit 2 besitzt. Daraus folgt $x_1 = -1/v$. Wegen $1 + vx^* < 0$ gilt $x^* < -1/v$. Somit ist der Designraum groß genug, dass $-1/v \in \mathcal{X}$ gilt. Die E -optimalen Gewichte (7.2) berechnen sich zu denen in Satz 7.6.

b) Der Fall $\beta_1 < 0$ kann analog gezeigt werden. □

Bemerkung 7.7

In den Fällen $1 + vx^* \geq 0$ bei a) und $1 + ux^* \geq 0$ bei b) wird die Bedingung (A3) nicht benötigt.

Die Ungleichungen in Bemerkung 7.7 können nur dann nicht erfüllt sein, wenn der Designraum positive und negative Werte enthält. Für Designräume, bei denen alle Werte gleiches Vorzeichen haben, sind diese Ungleichungen also stets erfüllt.

8 Diskussion

Im Mittelpunkt dieser Arbeit steht die Bestimmung optimaler Designs für multiple Regressionsmodelle. Für diese Modelle konnten allgemeine Resultate zur Struktur optimaler Designs gezeigt werden, die für viele Optimalitätskriterien gültig sind. Ist der Designraum ein Polytop, so existiert ein optimales Design, dessen Stützpunkte nur auf den Kanten des Designraums liegen. Bei vielen Kriterien wie z. B. D -, L - und ϕ_k -Optimalität müssen sich die Stützpunkte eines optimalen Designs auf den Kanten des Designraums befinden. Rechteckige Polytope ergeben sich als Designraum, wenn jede Kovariable unabhängig von den anderen Kovariablen Werte aus einem Intervall annehmen kann. Auch für beliebige Designräume wie z. B. eine Kugel lässt sich festhalten, dass die Stützpunkte eines optimalen Designs auf dem Rand liegen.

Für diese Ergebnisse ist es notwendig, dass nur lineare Terme der Kovariablen in das Modell eingehen. Insbesondere dürfen keine Interaktionsterme vorhanden sein. Anhand eines Beispiels wurde gezeigt, dass optimale Designs für Modelle mit einem Interaktionsterm Stützpunkte im Inneren des Designraums haben können.

Diese Strukturresultate sind auch zur numerischen Bestimmung optimaler Designs vorteilhaft, da die Kanten-Restriktion die Anzahl der freien Variablen eines Stützpunktes von $p - 1$ auf 1 reduziert. Ferner konnte gezeigt werden, dass unter geeigneten Bedingungen ein optimales Design nur zwei Stützpunkte je Kante haben kann. Mit Hilfe dieser Ergebnisse wurden optimale Designs für Modelle mit zensierten Daten für verschiedene Optimalitätskriterien hergeleitet.

Für D -Optimalität wurde gezeigt, dass sich optimale Designs für ein multiples Regressionsmodell aus den optimalen Designs in den marginalen Modellen mit einer einzigen Kovariablen konstruieren lassen. Für dieses Ergebnis müssen neben der Existenz eines D -optimalen Designs nur zwei Voraussetzungen erfüllt sein: Die D -optimalen Designs in den marginalen Modellen müssen 2-Punkt-Designs sein und sie müssen einen Randpunkt des Designraums als Stützpunkt haben. Diese Bedingungen sind häufig erfüllt, wenn der Designraum abgeschnitten wird, so dass einer der Stützpunkte des D -optimalen Designs auf dem erweiterten Designraum \mathbb{R} sich außerhalb des abgeschnittenen Designraums befindet. Dies wurde anhand eines Beispiels für das logistische Regressionsmodell demon-

triert, für das D -optimale Designs direkt gefolgert werden konnten. Weitere Beispiele für optimale Designs mit einem Randpunkt des Designraums als Stützpunkt sind bei Biedermann et al. (2006) zu finden. Die zur Konstruktion der D -optimalen Designs im multiplen Regressionsmodell benötigten D -optimalen Designs in den marginalen Modellen sind üblicherweise bekannt oder können zumindest leicht numerisch berechnet werden.

Für die in dieser Arbeit betrachtete Klasse nichtlinearer Modelle sind die genannten Voraussetzungen durch Annahme geeigneter Bedingungen an die Intensitätsfunktion erfüllt. Zu dieser Klasse gehören das Proportional Hazards Modell mit Typ I und zufälliger Zensurierung sowie das Poisson- und Negativ-Binomial-Modell. Für die Zensurierungszeiten sind viele Verteilungen möglich, bei denen die Intensitätsfunktion die Bedingungen (A1)–(A4) erfüllt. Für diese Modellklasse sind die D -optimalen Designs in den marginalen Modellen bekannt. Die Anwendung der bisherigen Resultate lieferte D -optimale Designs für multiple Regression.

Für die additiven Modelle, bei denen beliebige Regressionsfunktionen zugelassen sind, die jeweils nur von einer Kovariablen abhängen, wurde gezeigt, wie sich mit Hilfe der Ergebnisse D -optimale Designs herleiten lassen. Dabei stellte sich heraus, dass die D -optimalen Designs Stützpunkte im Inneren des Designraums haben können.

Ferner wurden L - und ϕ_k -optimale Designs für Modelle mit einer oder mehreren Kovariablen berechnet. Es zeigten sich Gemeinsamkeiten bei der Lage der Stützpunkte optimaler Designs für die verschiedenen Kriterien. Optimale Designs haben die Ecke des Designraums als Stützpunkt, bei der die Intensitätsfunktion maximal wird. Die weiteren Stützpunkte liegen auf den zu dieser Ecke inzidenten Kanten des Designraums. Insbesondere besitzen die optimalen Designs immer einen minimalen Träger. Dies ist speziell bei einer großen Anzahl von Kovariablen vorteilhaft, da diese Designs in der Praxis einfacher umgesetzt werden können als Designs mit mehr Stützpunkten, wie z. B. das 2^{p-1} vollfaktorielle Design auf Basis der optimalen marginalen Designpunkte. Es ist zudem einfacher das Design zu einem exakten Design zu runden, falls dies notwendig ist.

Während bei D -Optimalität nur die eindeutige Lösung einer nichtlinearen Gleichung zu finden ist, muss bei L -Optimalität ein nichtlineares Gleichungssystem für die Stützpunkte und bei ϕ_k -Optimalität ein nichtlineares Gleichungssystem für die Stützpunkte und Gewichte gelöst werden. Für L -Optimalität wurde die Existenz und Eindeutigkeit der Lösung des Gleichungssystems bewiesen. Für ϕ_k -Optimalität ist die Existenz einer Lösung nicht bewiesen, jedoch die Eindeutigkeit, falls eine Lösung existiert. Obwohl das Gleichungssystem insbesondere bei ϕ_k -Optimalität eine Schwierigkeit bei der Berechnung eines optimalen Designs darstellt und letztendlich numerisch gelöst werden muss,

bieten diese Resultate Vorteile gegenüber einer direkten numerischen Herangehensweise zur Berechnung optimaler Designs. So hat man die Sicherheit, dass es sich tatsächlich um ein optimales Design handelt, wenn eine Lösung des Gleichungssystems gefunden wurde. Bei numerischen Optimierungsverfahren besteht immer die Problematik, dass nur ein lokales Maximum oder Minimum gefunden werden kann. Zudem sind diese Optimierungsverfahren in der Regel deutlich aufwendiger.

Bei der Betrachtung der ϕ_k -Kriterien wurde der Fall $k = -\infty$ für E -Optimalität zunächst ausgeschlossen, da die Kriteriumsfunktion nicht differenzierbar sein muss und ein separater Äquivalenzsatz gilt. Für das Modell mit einer Kovariablen konnten die E -optimalen Designs bestimmt werden. Die Berechnung E -optimaler Designs für mehrere Kovariablen ist vermutlich sehr schwierig und ein Thema für weitere Forschung.

Auch bei der Berechnung D - und L -optimaler Designs für Modelle mit einer Kovariablen und einem diskreten Designraum lassen sich Ähnlichkeiten feststellen. So gibt es drei Typen von möglichen D -optimalen Designs, die auch bei L -Optimalität vorkommen. Zusätzlich gibt es bei L -Optimalität jedoch einen vierten Typ von Design, der nicht als D -optimales Design möglich ist.

Bei c -Optimalität gelang für Modelle mit einer Kovariablen eine umfassende Charakterisierung c -optimaler Designs für einen stetigen und diskreten Designraum. Bisher waren nur für spezielle Modelle aus der betrachteten Modellklasse oder bestimmte \mathbf{c} -Vektoren c -optimale Designs bekannt.

Für multiple Regressionsmodelle wurden c -optimale Designs für bestimmte \mathbf{c} -Vektoren hergeleitet. Ihre Struktur weicht von der für die anderen Optimalitätskriterien ab. So existieren c -optimale Designs mit Stützpunkten im Inneren des Designraums. Als Ursache dafür konnte die fehlende Regularität der im Äquivalenzsatz für c -Optimalität auftretenden Matrix identifiziert werden. In Übereinstimmung mit den Ergebnissen für multiple Regressionsmodelle existiert auch ein c -optimales Design mit Stützpunkten ausschließlich auf den Kanten des Designraums. Eine natürliche Erweiterung der vorliegenden Resultate ist die Bestimmung c -optimaler Designs für die bisher nicht betrachteten Vektoren \mathbf{c} .

Bei der Berechnung D -optimaler Designs für Modelle mit zwei Kovariablen und einem Interaktionsterm konnten für bestimmte Parameter- und Designraum-Kombinationen Designs gefunden werden, die in einer gewissen Klasse optimal sind. Die Verifikation der bisher nur aufgrund numerischer Berechnungen vermuteten D -Optimalität dieser Designs ist ein Thema zukünftiger Forschung.

Zur Lösung der Parameterabhängigkeit lokal optimaler Designs gibt es mehrere Ansätze. In dieser Arbeit wurden standardisiert Maximin D -optimale Designs und Bayessche D - und c -optimale Designs bestimmt.

Ein Resultat von Konstantinou et al. (2014) zum standardisiert Maximin D -optimalen Design für Modelle mit einer Kovariablen konnte auf den mehrdimensionalen Fall innerhalb einer speziellen Designklasse erweitert werden. In Beispielen wurden standardisiert Maximin D -optimale Designs und Bayessche D -optimale Designs für multiple Regressionsmodelle aus den optimalen Designs in den marginalen Modellen konstruiert. Dabei sind die optimalen Designs in den marginalen Modellen 2-Punkt-Designs, die den Randpunkt 0 des Designraums als Stützpunkt haben. Ob dieses Vorgehen immer optimale Designs liefert, ist ein Thema für weitere Untersuchungen.

Für die Bayesschen Kriterien konnten unter Verwendung von Quadraturformeln zur näherungsweise Berechnung der Integrale numerisch optimale Designs für das Proportional Hazards Modell mit Typ I Zensierung bestimmt werden. Für das Proportional Hazards Modell mit exponentialverteilten Zensierungszeiten bzw. das Negativ-Binomial-Modell und einer Kovariablen wurden für verschiedene Gewichtungen analytische Resultate hergeleitet, wobei die gefundenen Designs optimal in der Klasse der 2-Punkt-Designs sind.

Die vorgestellten Methoden und Resultate zur Bestimmung optimaler Designs für multiple Regression lassen sich auf weitere Modelle und Optimalitätskriterien anwenden. Bisher waren für nichtlineare multiple Regressionsmodelle nur wenige Ergebnisse sowohl für D -Optimalität als auch insbesondere für c -, L - und ϕ_k -Optimalität bekannt. Diese Arbeit liefert somit einen wichtigen Beitrag zur weiteren Entwicklung der optimalen Versuchsplanung für nichtlineare multiple Regressionsmodelle.

A Anhang

Beweise zu Kapitel 4

Beweis von Satz 4.19:

Die Funktion Q berechnet sich zu $Q(\theta) = e^\theta / (e^\theta + \lambda) = 1 / (1 + \lambda e^{-\theta})$. Für ein 2-Punkt-Design ξ mit Stützpunkten x_1 und x_2 gilt

$$\det(\mathbf{M}(\xi, \beta))^{-1} = \frac{(1 + \lambda e^{-\beta_0 - \beta_1 x_1})(1 + \lambda e^{-\beta_0 - \beta_1 x_2})}{\omega_1 \omega_2 (x_2 - x_1)^2}.$$

Zur Berechnung der Kriteriumsfunktion ist das folgende Integral hilfreich:

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-\beta x} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(\beta-\mu)^2}{2\sigma^2}} d\beta &= e^{-x\mu + \frac{x^2\sigma^2}{2}} \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(\beta-\mu+x\sigma^2)^2}{2\sigma^2}} d\beta \\ &= e^{-x\mu + \frac{x^2\sigma^2}{2}}. \end{aligned} \quad (\text{A.1})$$

Die Kriteriumsfunktion ergibt sich ohne den Logarithmus, der nur eine monotone Transformation darstellt, zu

$$E_\beta \left[\det(\mathbf{M}(\xi, \beta))^{-1} \right] = \frac{1 + \lambda e^{\frac{x_1^2\sigma_1^2}{2} - \mu_0 + \frac{\sigma_0^2}{2}} + \lambda e^{\frac{x_2^2\sigma_1^2}{2} - \mu_0 + \frac{\sigma_0^2}{2}} + \lambda^2 e^{\frac{\sigma_1^2}{2}(x_1+x_2)^2 - 2\mu_0 + 2\sigma_0^2}}{\omega_1 \omega_2 (x_2 - x_1)^2}.$$

Die optimalen Gewichte sind $\omega_1^* = \omega_2^* = 1/2$. Es seien C_1 und C_2 definiert wie in Satz 4.19 und $r(z) := e^{\frac{1}{2}\sigma_1^2 z^2}$. Dann kann die Kriteriumsfunktion geschrieben werden als

$$h(x_1, x_2) := \frac{1 + C_1 \cdot (r(x_1) + r(x_2)) + C_2 \cdot r(x_1 + x_2)}{\frac{1}{4}(x_2 - x_1)^2}.$$

Es gelte $x_2 > x_1$. Der Punkt $(-x, x)$ liegt auf der durch $x_1 = -x_2$ definierten Ebene. Wegen $h(-x + y, x + y) = h(-x - y, x - y)$ für alle $x > 0$ und $y \geq 0$ ist die Funktion h symmetrisch zur Ebene, die durch $x_1 = -x_2$ gegeben ist. Es wird die Funktion h senkrecht zu dieser Symmetrieebene betrachtet. Dazu sei $\tilde{h}(y) := h(-x + y, x + y)$. Die

Ableitung von $\tilde{h}(y)$ berechnet sich für $x > 0$ und $y \geq 0$ zu

$$\tilde{h}'(y) = \frac{C_1 \cdot (r'(-x+y) + r'(x+y)) + C_2 \cdot 2r'(2y)}{x^2}.$$

Dabei ist die Ableitung der Funktion r gegeben durch $r'(z) = \sigma_1^2 z e^{\frac{1}{2}\sigma_1^2 z^2}$. Es gelten $r'(-z) = -r'(z)$ und $r'(0) = 0$ und somit folgt $\tilde{h}'(0) = 0$. Die zweite Ableitung von r berechnet sich zu $r''(z) = (\sigma_1^2 + \sigma_1^4 z^2) e^{\frac{1}{2}\sigma_1^2 z^2} > 0$. Die Funktion r' ist also streng monoton wachsend. Dann ist die Funktion \tilde{h}' als Summe streng monoton wachsender Funktionen ebenfalls streng monoton wachsend. Es folgt $\tilde{h}'(y) > 0$ für $y > 0$. Die Funktion h wächst zu beiden Seiten um die Symmetrieebene und wird folglich für $x_1 = -x_2$ minimiert. Es wird daher die Funktion

$$\hat{h}(x) := h(-x, x) = \frac{1 + C_1 \cdot 2r(x) + C_2}{x^2}$$

bezüglich $x > 0$ minimiert. Die Ableitung von \hat{h} ist gegeben durch

$$\hat{h}'(x) = \frac{C_1 \cdot 2\sigma_1^2 x^3 r(x) - (1 + C_1 \cdot 2r(x) + C_2) \cdot 2x}{x^4}.$$

Sie ist für $x > 0$ genau dann gleich null, wenn gilt:

$$\begin{aligned} 0 &= C_1 r(x) \cdot (\sigma_1^2 x^2 - 2) - 1 - C_2 \\ \Leftrightarrow \quad \frac{1 + C_2}{2C_1 e} &= e^{\frac{1}{2}\sigma_1^2 x^2 - 1} \cdot \left(\frac{1}{2}\sigma_1^2 x^2 - 1 \right) \\ \Leftrightarrow \quad W\left(\frac{1 + C_2}{2C_1 e}\right) &= \frac{1}{2}\sigma_1^2 x^2 - 1 \\ \Leftrightarrow \quad x &= \pm \frac{1}{\sigma_1} \cdot \sqrt{2 + 2W\left(\frac{1 + C_2}{2C_1 e}\right)}. \end{aligned}$$

Da $\hat{h}(x)$ für $x \rightarrow 0$ und $x \rightarrow \infty$ gegen unendlich strebt, liegt bei der positiven Lösung, die mit x^* bezeichnet wird, ein Minimum vor. Es folgt $x_1^* = -x^*$ und $x_2^* = x^*$.

Nun wird der Fall betrachtet, dass $-x^*$ außerhalb des Designraums liegt, d. h. es gilt $-x^* < u$. Die Funktion $h(x_1, x_2)$ wächst zu beiden Seiten senkrecht zur Symmetrieebene und fällt auf der Symmetrieebene in Richtung $(-x^*, x^*)$. Sie wird daher für $u + v \geq 0$ bei $x_1^* = u$ minimiert und für $u + v \leq 0$ bei $x_2^* = v$ (vgl. Abbildung A.1).

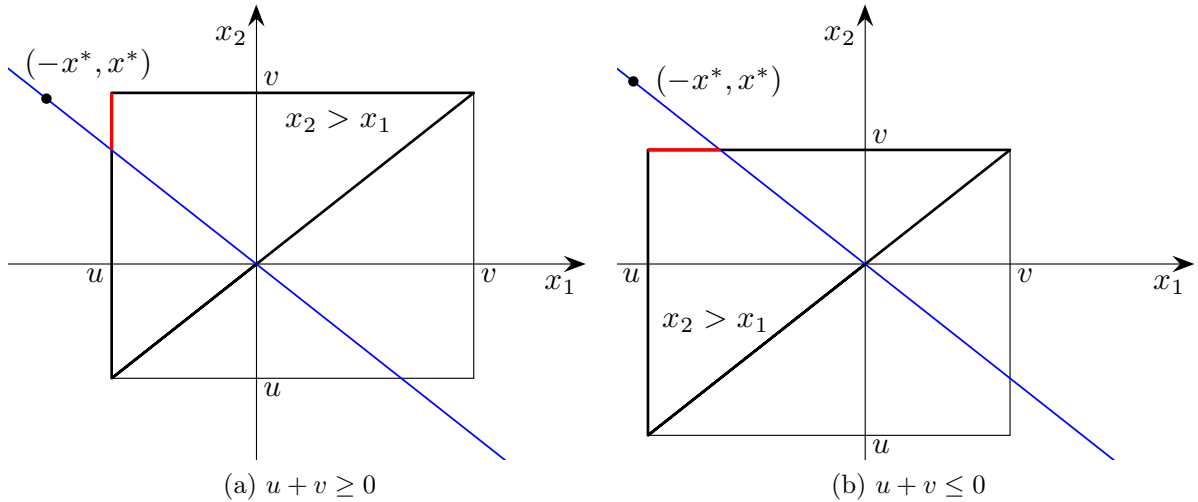


Abbildung A.1: Beweisskizze: Die Symmetrieebene ist durch die blaue Gerade dargestellt. Die Funktion $h(x_1, x_2)$ wird im Dreieck mit $x_2 > x_1$ auf der roten Linie minimiert.

Im ersten Fall $u + v \geq 0$ ist die Funktion

$$k(x_2) := h(u, x_2) = \frac{1 + C_1 \cdot (r(u) + r(x_2)) + C_2 \cdot r(u + x_2)}{\frac{1}{4}(x_2 - u)^2}$$

für $x_2 > u$ zu minimieren. Ihre Ableitung berechnet sich zu

$$k'(x_2) = \frac{4}{(x_2 - u)^3} \cdot \left[(C_1 \cdot \sigma_1^2 x_2 r(x_2) + C_2 \cdot \sigma_1^2 (u + x_2) r(u + x_2)) \cdot (x_2 - u) - 2 \cdot (1 + C_1 \cdot (r(u) + r(x_2)) + C_2 \cdot r(u + x_2)) \right].$$

Diese Ableitung ist genau dann gleich null, wenn der Term in eckigen Klammern gleich null ist. Dieser kann geschrieben werden als

$$D_u(x_2) := C_1 r(x_2) (\sigma_1^2 x_2 (x_2 - u) - 2) + C_2 r(u + x_2) (\sigma_1^2 (u + x_2) (x_2 - u) - 2) - 2 - 2C_1 r(u).$$

Er ist genau dann gleich null, wenn $l(x_2) := D_u(x_2)/(x_2 - u) = 0$ gilt. Nun wird gezeigt, dass die Funktion $l(x_2)$ streng monoton wachsend ist für $x_2 > u$ und $k'(x_2)$ daher höchstens eine Nullstelle im Intervall (u, ∞) hat. Dazu wird zunächst die Funktion $l_1(x_2) := C_1 r(x_2) (\sigma_1^2 x_2 (x_2 - u) - 2)/(x_2 - u)$ für $x_2 > u$ betrachtet. Ihre Ableitung ist

gegeben durch

$$\begin{aligned}
 l'_1(x_2) &= \frac{1}{(x_2 - u)^2} \cdot \left[C_1 r(x_2) \cdot [\sigma_1^2 x_2 (\sigma_1^2 x_2 (x_2 - u) - 2) + \sigma_1^2 (2x_2 - u)] \cdot (x_2 - u) \right. \\
 &\quad \left. - C_1 r(x_2) (\sigma_1^2 x_2 (x_2 - u) - 2) \right] \\
 &= \frac{C_1 r(x_2) \cdot [\sigma_1^2 (x_2 - u) (\sigma_1^2 x_2^2 (x_2 - u) - u - x_2) + 2]}{(x_2 - u)^2}.
 \end{aligned}$$

Diese Ableitung ist positiv, wenn $m(x_2) := \sigma_1^2 (x_2 - u) (\sigma_1^2 x_2^2 (x_2 - u) - u - x_2) + 2$ positiv ist. Da $m(x_2)$ für $x_2 \rightarrow \infty$ gegen unendlich strebt und $m(u) = 2 > 0$ gilt, genügt es $m(x_2)$ an den Extrema auszuwerten, um zu zeigen, dass $m(x_2)$ für $x_2 > u$ positiv ist. Die Ableitung von $m(x_2)$ berechnet sich zu

$$\begin{aligned}
 m'(x_2) &= \sigma_1^2 (\sigma_1^2 x_2^2 (x_2 - u) - u - x_2) + \sigma_1^2 (x_2 - u) (\sigma_1^2 (3x_2^2 - 2x_2 u) - 1) \\
 &= \sigma_1^2 x_2 (\sigma_1^2 (x_2 - u) (4x_2 - 2u) - 2).
 \end{aligned}$$

Sie hat die drei Nullstellen

$$x_2 = 0, \quad x_2 = \frac{3}{4}u + \sqrt{\frac{u^2}{16} + \frac{1}{2\sigma_1^2}}, \quad x_2 = \frac{3}{4}u - \sqrt{\frac{u^2}{16} + \frac{1}{2\sigma_1^2}}.$$

Da die dritte Nullstelle kleiner als u ist, sind nur die ersten beiden Nullstellen von Interesse. Es gilt $m(0) = \sigma_1^2 u^2 + 2 > 0$. Mit $S := \sqrt{\frac{u^2}{16} + \frac{1}{2\sigma_1^2}}$ gilt:

$$\begin{aligned}
 m\left(\frac{3}{4}u + S\right) &= \sigma_1^4 \left(-\frac{1}{4}u + S\right)^2 \left(\frac{3}{4}u + S\right)^2 + \sigma_1^2 \left(-\frac{1}{4}u + S\right) \left(-\frac{7}{4}u - S\right) + 2 \\
 &= \sigma_1^4 \left(-\frac{3}{16}u^2 + \frac{1}{2}uS + \frac{u^2}{16} + \frac{1}{2\sigma_1^2}\right)^2 + \sigma_1^2 \left(\frac{7}{16}u^2 - \frac{3}{2}uS - \frac{u^2}{16} - \frac{1}{2\sigma_1^2}\right) + 2 \\
 &= \sigma_1^4 \left(\frac{1}{64}u^4 - \frac{1}{8}u^3S - \frac{u^2}{8\sigma_1^2} + \frac{1}{4}u^2 \left(\frac{u^2}{16} + \frac{1}{2\sigma_1^2}\right) + \frac{uS}{2\sigma_1^2} + \frac{1}{4\sigma_1^4}\right) \\
 &\quad + \sigma_1^2 \left(\frac{3}{8}u^2 - \frac{3}{2}uS\right) + \frac{3}{2} \\
 &= \frac{1}{32}u^4 \sigma_1^4 - \frac{1}{8}u^3 \sigma_1^4 S + \frac{3}{8}u^2 \sigma_1^2 - u \sigma_1^2 S + \frac{7}{4}.
 \end{aligned}$$

Zur Bestimmung der Werte von u , für die der obige Ausdruck gleich null ist, muss die folgende Gleichung nach u aufgelöst werden:

$$\left(\frac{1}{8}u^3\sigma_1^4 + u\sigma_1^2\right) \cdot S = \frac{1}{32}u^4\sigma_1^4 + \frac{3}{8}u^2\sigma_1^2 + \frac{7}{4}.$$

Quadrieren liefert:

$$\begin{aligned} \left(\frac{1}{64}u^6\sigma_1^8 + \frac{1}{4}u^4\sigma_1^6 + u^2\sigma_1^4\right) \cdot \left(\frac{u^2}{16} + \frac{1}{2\sigma_1^2}\right) &= \frac{u^8\sigma_1^8}{1024} + \frac{3}{128}u^6\sigma_1^6 + \frac{1}{4}u^4\sigma_1^4 + \frac{21}{16}u^2\sigma_1^2 + \frac{49}{16} \\ \Leftrightarrow 0 &= \frac{1}{16}u^4\sigma_1^4 + \frac{13}{16}u^2\sigma_1^2 + \frac{49}{16}. \end{aligned}$$

Mit der Substitution $z = u^2$ erhält man die Gleichung $0 = \frac{1}{16}z^2\sigma_1^4 + \frac{13}{16}z\sigma_1^2 + \frac{49}{16}$ mit den imaginären Lösungen $z = -\frac{13}{2\sigma_1^2} \pm \sqrt{-\frac{27}{4\sigma_1^4}}$. Daher sind auch die Lösungen $u = \pm\sqrt{z}$ imaginär. Der Ausdruck $m\left(\frac{3}{4}u + S\right)$ wird also für kein reelles u gleich null und ist somit entweder für alle u positiv oder negativ. Für $u = 0$ gilt $m(S) = 7/4 > 0$ und es folgt, dass $m(x_2)$ positiv ist für $x_2 > u$ und $l_1(x_2)$ somit streng monoton wächst.

Analog folgt, dass die Funktion $l_2(x_2) := C_2r(u + x_2)(\sigma_1^2(u + x_2)(x_2 - u) - 2)/(x_2 - u)$ streng monoton wachsend ist für $x_2 > u$. Die Funktion $l_3(x_2) := (-2 - 2C_1r(u))/(x_2 - u)$ ist auch streng monoton wachsend für $x_2 > u$. Wegen $l(x_2) = l_1(x_2) + l_2(x_2) + l_3(x_2)$ ist $l(x_2)$ streng monoton wachsend für $x_2 > u$. Folglich hat $k'(x_2)$ höchstens eine Nullstelle im Intervall (u, ∞) . Da $k(x_2)$ für $x_2 \rightarrow u$ und $x_2 \rightarrow \infty$ gegen unendlich strebt, existiert genau ein Minimum. Liegt das Minimum außerhalb des Designraums, dann ist $x_2^* = v$ der zweite Stützpunkt.

Die Fälle $u + v \leq 0$ und $x^* > v$ können analog gezeigt werden. \square

Beweis von Satz 4.23:

Es wird der Fall $\beta_{12} < 0$ betrachtet. Für $\beta_{12} = 0$ folgt der Beweis analog. Die Determinante der Informationsmatrix von ξ aus (4.27) berechnet sich mit Hilfe von (3.35) zu

$$\det(\mathbf{M}(\xi, \beta)) = \frac{1}{256} \prod_{i=1}^4 Q(\mathbf{f}(\mathbf{x}_i)^T \beta) \cdot \det(\mathbf{X})^2$$

mit

$$\det(\mathbf{X}) = \det \begin{pmatrix} 1 & v_1 & v_2 & v_1v_2 \\ 1 & v_1 - x_1 & v_2 & (v_1 - x_1)v_2 \\ 1 & v_1 & v_2 - x_2 & v_1(v_2 - x_2) \\ 1 & v_1 - x_3 & v_2 - x_4 & (v_1 - x_3)(v_2 - x_4) \end{pmatrix}.$$

Subtrahieren der ersten Zeile von der zweiten, dritten und vierten Zeile liefert

$$\det(\mathbf{X}) = \det \begin{pmatrix} 1 & v_1 & v_2 & v_1v_2 \\ 0 & -x_1 & 0 & -x_1v_2 \\ 0 & 0 & -x_2 & -v_1x_2 \\ 0 & -x_3 & -x_4 & x_3x_4 - v_2x_3 - v_1x_4 \end{pmatrix}.$$

Subtrahiert man nun das (x_3/x_1) -fache der zweiten Zeile und das (x_4/x_2) -fache der dritten Zeile von der vierten Zeile, so ergibt sich

$$\det(\mathbf{X}) = \det \begin{pmatrix} 1 & v_1 & v_2 & v_1v_2 \\ 0 & -x_1 & 0 & -x_1v_2 \\ 0 & 0 & -x_2 & -v_1x_2 \\ 0 & 0 & 0 & x_3x_4 \end{pmatrix} = x_1x_2x_3x_4.$$

Die Determinante der Informationsmatrix von ξ ist ein Produkt verschiedener Faktoren, die unabhängig voneinander maximiert werden können. Zur Maximierung bezüglich x_1 ist die folgende Funktion zu maximieren:

$$\begin{aligned} g_1(x_1) &= Q(\beta_0 + \beta_1(v_1 - x_1) + \beta_2v_2 + \beta_{12}(v_1 - x_1)v_2) \cdot x_1^2 \\ &= Q(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - (\beta_1 + \beta_{12}v_2)x_1) \cdot x_1^2. \end{aligned}$$

Ihre Ableitung ist gegeben durch

$$\begin{aligned} g_1'(x_1) &= -(\beta_1 + \beta_{12}v_2)Q'(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - (\beta_1 + \beta_{12}v_2)x_1)x_1^2 \\ &\quad + 2x_1Q(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - (\beta_1 + \beta_{12}v_2)x_1). \end{aligned}$$

Diese Ableitung hat eine Nullstelle bei $x_1 = 0$. Es handelt sich um ein Minimum von g_1 , da $g_1(0) = 0$ und $g_1(x_1) > 0$ für $x_1 \neq 0$. Eine weitere Nullstelle von g_1' ergibt sich als Lösung der folgenden Gleichung:

$$\begin{aligned} 0 &= -(\beta_1 + \beta_{12}v_2)x_1 + 2 \cdot \frac{Q(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - (\beta_1 + \beta_{12}v_2)x_1)}{Q'(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - (\beta_1 + \beta_{12}v_2)x_1)} \\ &= -\phi_{\mathbf{v}}((\beta_1 + \beta_{12}v_2)x_1). \end{aligned}$$

Die eindeutige Lösung $x_1^* = \phi_{\mathbf{v}}^{-1}(0)/(\beta_1 + \beta_{12}v_2)$ befindet sich im Intervall $(0, \infty)$, da die Funktionen Q und Q' positiv sind und $\beta_1 + \beta_{12}v_2 > 0$ gilt. Es handelt sich um ein Maximum von g_1 , da die Ableitung an dieser Stelle das Vorzeichen von positiv zu negativ

wechselt. Gilt $\phi_{\mathbf{v}}^{-1}(0)/(\beta_1 + \beta_{12}v_2) > v_1 - u_1$, so wird die Determinante bei $x_1^* = v_1 - u_1$ maximiert.

Zur Maximierung der Determinante bezüglich x_2 ist folgende Funktion zu maximieren:

$$\begin{aligned} g_2(x_2) &= Q(\beta_0 + \beta_1v_1 + \beta_2(v_2 - x_2) + \beta_{12}v_1(v_2 - x_2)) \cdot x_2^2 \\ &= Q(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - (\beta_2 + \beta_{12}v_1)x_2) \cdot x_2^2. \end{aligned}$$

Analog wie bei der Berechnung von x_1^* folgt $x_2^* = \phi_{\mathbf{v}}^{-1}(0)/(\beta_2 + \beta_{12}v_1)$. Für den Fall $\phi_{\mathbf{v}}^{-1}(0)/(\beta_2 + \beta_{12}v_1) > v_2 - u_2$ wird die Determinante bei $x_2^* = v_2 - u_2$ maximiert.

Zur Bestimmung der optimalen Werte von x_3 und x_4 ist die folgende Funktion zu maximieren:

$$g(x_3, x_4) := Q(\beta_0 + \beta_1(v_1 - x_3) + \beta_2(v_2 - x_4) + \beta_{12}(v_1 - x_3)(v_2 - x_4)) \cdot (x_3x_4)^2.$$

Sie wird auf den Kurven betrachtet, bei denen das Argument der Funktion Q konstant ist, d. h.

$$\begin{aligned} -\beta_1x_3 - \beta_2x_4 - \beta_{12}x_3v_2 - \beta_{12}v_1x_4 + \beta_{12}x_3x_4 &= \eta \\ \iff x_3 = x_3(x_4, \eta) &= -\frac{\eta + (\beta_2 + \beta_{12}v_1)x_4}{\beta_1 + \beta_{12}v_2 - \beta_{12}x_4} \end{aligned} \quad (\text{A.2})$$

mit $\eta < 0$. Dabei handelt es sich um streng monoton fallende Kurven, denn es gilt

$$\begin{aligned} \frac{dx_3(x_4, \eta)}{dx_4} &= -\frac{(\beta_2 + \beta_{12}v_1)(\beta_1 + \beta_{12}v_2 - \beta_{12}x_4) + (\eta + (\beta_2 + \beta_{12}v_1)x_4)\beta_{12}}{(\beta_1 + \beta_{12}v_2 - \beta_{12}x_4)^2} \\ &= -\frac{(\beta_2 + \beta_{12}v_1)(\beta_1 + \beta_{12}v_2) + \eta\beta_{12}}{(\beta_1 + \beta_{12}v_2 - \beta_{12}x_4)^2} < 0. \end{aligned}$$

Es muss $x_3 > 0$ gelten und daher gilt auf diesen Kurven $x_4 < -\eta/(\beta_2 + \beta_{12}v_1)$. Somit ist die folgende Funktion bezüglich $x_4 \in (0, -\frac{\eta}{\beta_2 + \beta_{12}v_1})$ zu maximieren:

$$g_\eta(x_4) := g(x_3(x_4, \eta), x_4) = Q(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} + \eta) \cdot \left(\frac{\eta + (\beta_2 + \beta_{12}v_1)x_4}{\beta_1 + \beta_{12}v_2 - \beta_{12}x_4} \right)^2 \cdot x_4^2.$$

Ihre Ableitung ergibt sich zu

$$\begin{aligned} g'_\eta(x_4) &= Q(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} + \eta) \cdot 2 \cdot \frac{\eta + (\beta_2 + \beta_{12}v_1)x_4}{(\beta_1 + \beta_{12}v_2 - \beta_{12}x_4)^3} \cdot x_4 \\ &\quad \cdot \left[(\eta + 2(\beta_2 + \beta_{12}v_1)x_4) \cdot (\beta_1 + \beta_{12}v_2 - \beta_{12}x_4) + \beta_{12}(\eta + (\beta_2 + \beta_{12}v_1)x_4)x_4 \right]. \end{aligned}$$

Diese Ableitung hat vier Nullstellen, eine bei $x_4 = 0$, eine bei $x_4 = -\eta/(\beta_2 + \beta_{12}v_1) > 0$ und zwei weitere bei den Nullstellen des quadratischen Polynoms in eckigen Klammern. Bei den ersten beiden Nullstellen handelt es sich jeweils um ein Minimum der Funktion g_η , weil sie an diesen Stellen gleich null ist und ansonsten nichtnegativ. Da die Polstelle von g_η bei $x_4 = (\beta_1 + \beta_{12}v_2)/\beta_{12} < 0$ liegt, muss zwischen den Minima ein Maximum existieren. Wegen

$$\lim_{x_4 \rightarrow -\infty} g_\eta(x_4) = \infty \quad \text{und} \quad \lim_{x_4 \rightarrow (\beta_1 + \beta_{12}v_2)/\beta_{12}} g_\eta(x_4) = \infty$$

muss es außerdem ein Minimum im Intervall $(-\infty, \frac{\beta_1 + \beta_{12}v_2}{\beta_{12}})$ geben. Damit sind alle Extrema von g_η klassifiziert. Zur Berechnung der Ortskurve, auf der die Maxima liegen, wird das Polynom in eckigen Klammern nach η aufgelöst:

$$\eta = -\frac{x_4(\beta_2 + \beta_{12}v_1)(2\beta_1 + 2\beta_{12}v_2 - \beta_{12}x_4)}{\beta_1 + \beta_{12}v_2}.$$

Einsetzen von η in (A.2) liefert mit $\tilde{C} := (\beta_2 + \beta_{12}v_1)/(\beta_1 + \beta_{12}v_2)$

$$x_3 = -\frac{-x_4\tilde{C}(2\beta_1 + 2\beta_{12}v_2 - \beta_{12}x_4) + \tilde{C}(\beta_1 + \beta_{12}v_2)x_4}{\beta_1 + \beta_{12}v_2 - \beta_{12}x_4} = \tilde{C}x_4.$$

Auf dieser Geraden liegen die Maxima der Funktion $g_\eta(x_4)$ für verschiedene η . Nun wird die Funktion $g(x_3, x_4)$ auf dieser Geraden betrachtet:

$$\tilde{g}(x_4) := g(\tilde{C}x_4, x_4) = Q(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - x_4\tilde{C}(2\beta_1 + 2\beta_{12}v_2 - \beta_{12}x_4)) \cdot \tilde{C}^2 \cdot x_4^4.$$

Die Ableitung von \tilde{g} berechnet sich zu

$$\begin{aligned} \tilde{g}'(x_4) &= Q(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - x_4\tilde{C}(2\beta_1 + 2\beta_{12}v_2 - \beta_{12}x_4)) \cdot \tilde{C}^2 \cdot 4x_4^3 \\ &\quad - Q'(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - x_4\tilde{C}(2\beta_1 + 2\beta_{12}v_2 - \beta_{12}x_4)) \cdot \tilde{C}^3 \cdot x_4^4 \cdot 2(\beta_1 + \beta_{12}v_2 - \beta_{12}x_4). \end{aligned}$$

Diese Ableitung hat eine dreifache Nullstelle bei $x_4 = 0$, bei der es sich um ein Minimum von \tilde{g} handelt, da $\tilde{g}(0) = 0$ und $\tilde{g}(x_4) > 0$ für $x_4 \neq 0$. Eine weitere Nullstelle von \tilde{g}' ergibt sich als Lösung der folgenden Gleichung:

$$2 \cdot \frac{Q(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - x_4\tilde{C}(2\beta_1 + 2\beta_{12}v_2 - \beta_{12}x_4))}{Q'(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - x_4\tilde{C}(2\beta_1 + 2\beta_{12}v_2 - \beta_{12}x_4))} = x_4\tilde{C}(\beta_1 + \beta_{12}v_2 - \beta_{12}x_4).$$

Das Argument der Funktion Q/Q' auf der linken Seite ist eine quadratische Funktion in x_4 mit einem Maximum bei $x_4 = (\beta_1 + \beta_{12}v_2)/\beta_{12} < 0$. Sie ist daher für $x_4 \geq 0$ monoton

fallend und wegen (A4) gilt dies für die gesamte linke Seite. Die quadratische Funktion auf der rechten Seite hat ein Minimum bei $x_4 = (\beta_1 + \beta_{12}v_2)/(2\beta_{12}) < 0$ und ist somit streng monoton wachsend für $x_4 \geq 0$. Da die linke Seite positiv ist und die rechte Seite für $x_4 = 0$ gleich null ist, existiert genau eine Lösung \hat{x} im Intervall $(0, \infty)$. Es handelt sich um ein Maximum von \tilde{g} , da $\tilde{g}(0) = 0$ und $\lim_{x_4 \rightarrow \infty} \tilde{g}(x_4) = 0$ nach Lemma 2.3.

Mit der Skalierung $x_4 = x/(\beta_2 + \beta_{12}v_1)$ ergibt sich die Gleichung zu

$$2 \cdot \frac{Q(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - x \cdot (2 - \frac{\beta_{12}x}{C}))}{Q'(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - x \cdot (2 - \frac{\beta_{12}x}{C}))} = x \cdot \left(1 - \frac{\beta_{12}x}{C}\right)$$

mit $C = (\beta_2 + \beta_{12}v_1) \cdot (\beta_1 + \beta_{12}v_2)$. Sie hat die eindeutige Lösung $\tilde{x} = \hat{x} \cdot (\beta_2 + \beta_{12}v_1)$ im Intervall $(0, \infty)$. Es folgt $x_4^* = \tilde{x}/(\beta_2 + \beta_{12}v_1)$ und $x_3^* = \tilde{x}/(\beta_1 + \beta_{12}v_2)$.

Im Folgenden wird der Fall betrachtet, dass $\tilde{x}/(\beta_2 + \beta_{12}v_1) > v_2 - u_2$ gilt und das berechnete Maximum somit außerhalb des Designraums liegt. Die Funktion $g_\eta(x_4)$ ist auf fallenden Kurven definiert und wird für jedes $\eta < 0$ auf der Geraden $x_4 = x_3/\tilde{C}$ maximiert. Die Funktion $g(x_3, x_4)$ wächst auf dieser Geraden streng monoton bis zum berechneten Maximum. Daher wird $g(x_3, x_4)$ für $(v_1 - u_1)/\tilde{C} \geq v_2 - u_2$ bei $x_4 = v_2 - u_2$ maximiert und für $(v_1 - u_1)/\tilde{C} \leq v_2 - u_2$ bei $x_3 = v_1 - u_1$ (vgl. Abbildung A.2).

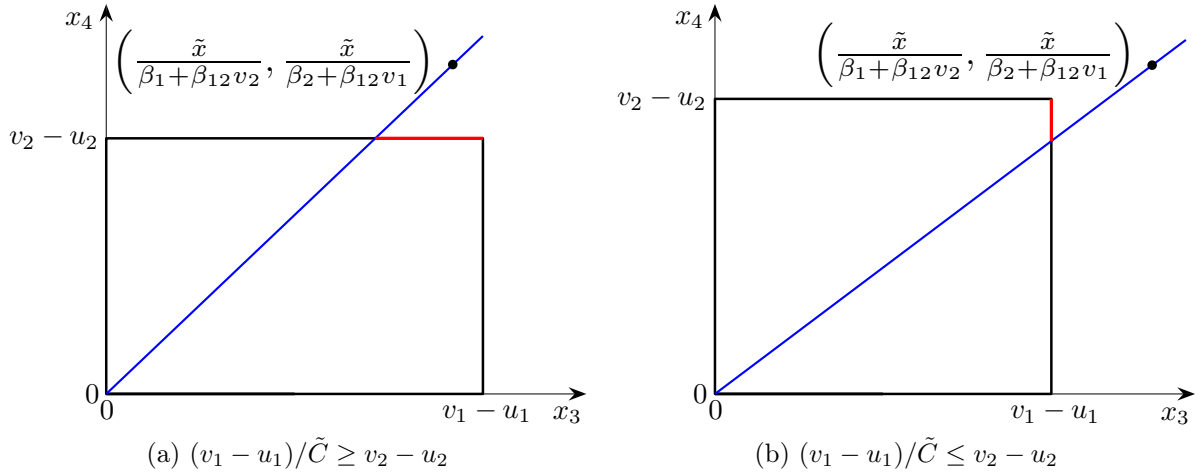


Abbildung A.2: Beweisskizze: Die blaue Gerade stellt die Gerade $x_4 = x_3/\tilde{C}$ dar. Verläuft sie bei $x_3 = v_1 - u_1$ oberhalb von $v_2 - u_2$, so wird die Funktion $g(x_3, x_4)$ auf der roten Linie bei $x_4 = v_2 - u_2$ maximiert. Verläuft sie bei $x_3 = v_1 - u_1$ unterhalb von $v_2 - u_2$, so wird die Funktion $g(x_3, x_4)$ auf der roten Linie bei $x_3 = v_1 - u_1$ maximiert.

Im ersten Fall $(v_1 - u_1)/\tilde{C} \geq v_2 - u_2$ ist die Funktion $g(x_3, v_2 - u_2)$ bezüglich x_3 zu maximieren und analog zur Berechnung von x_1^* folgt $x_3^* = \phi_{(v_1, u_2)}^{-1}(0)/(\beta_1 + \beta_{12}u_2)$.

Im zweiten Fall $(v_1 - u_1)/\tilde{C} \leq v_2 - u_2$ ist die Funktion $g(v_1 - u_1, x_4)$ bezüglich x_4 zu

maximieren. Das Maximum ergibt sich analog zu vorherigen Berechnungen als Lösung der Gleichung $\phi_{(u_1, v_2)^T}((\beta_2 + \beta_{12}u_1)x_4) = 0$. Es wird nun gezeigt, dass die Lösung dieser Gleichung größer als $v_2 - u_2$ ist. Nach Voraussetzung gilt

$$\tilde{x} > (v_2 - u_2)(\beta_2 + \beta_{12}v_1) \geq \frac{(v_1 - u_1)(\beta_2 + \beta_{12}v_1)}{\bar{C}} = (v_1 - u_1)(\beta_1 + \beta_{12}v_2)$$

und somit folgt $u_1 > v_1 - \tilde{x}/(\beta_1 + \beta_{12}v_2)$. Insbesondere gilt $\mathbf{f}(u_1, v_2)^T \boldsymbol{\beta} > \mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - \tilde{x}$. Unter Verwendung dieser beiden Abschätzungen für die erste Ungleichung und der Abschätzung $x_4 < \tilde{x}/(\beta_2 + \beta_{12}v_1)$ für die zweite Ungleichung folgt für alle $x_4 \in (0, v_2 - u_2]$

$$\begin{aligned} \phi_{(u_1, v_2)^T}((\beta_2 + \beta_{12}u_1)x_4) &= (\beta_2 + \beta_{12}u_1)x_4 - 2 \cdot \frac{Q(\mathbf{f}(u_1, v_2)^T \boldsymbol{\beta} - (\beta_2 + \beta_{12}u_1)x_4)}{Q'(\mathbf{f}(u_1, v_2)^T \boldsymbol{\beta} - (\beta_2 + \beta_{12}u_1)x_4)} \\ &\leq \left(\beta_2 + \beta_{12}v_1 - \frac{\beta_{12}\tilde{x}}{\beta_1 + \beta_{12}v_2} \right) x_4 - 2 \cdot \frac{Q(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - \tilde{x} - \left(\beta_2 + \beta_{12}v_1 - \frac{\beta_{12}\tilde{x}}{\beta_1 + \beta_{12}v_2} \right) x_4)}{Q'(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - \tilde{x} - \left(\beta_2 + \beta_{12}v_1 - \frac{\beta_{12}\tilde{x}}{\beta_1 + \beta_{12}v_2} \right) x_4)} \\ &< \left(1 - \frac{\beta_{12}\tilde{x}}{C} \right) \tilde{x} - 2 \cdot \frac{Q(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - (2 - \frac{\beta_{12}\tilde{x}}{C})\tilde{x})}{Q'(\mathbf{f}(\mathbf{v})^T \boldsymbol{\beta} - (2 - \frac{\beta_{12}\tilde{x}}{C})\tilde{x})} = 0 \end{aligned}$$

mit $C = (\beta_2 + \beta_{12}v_1)(\beta_1 + \beta_{12}v_2)$ und daher wird $g(v_1 - u_1, x_4)$ bei $x_4^* = v_2 - u_2$ maximiert. Der Fall $\tilde{x}/(\beta_1 + \beta_{12}v_2) > v_1 - u_1$ kann analog behandelt werden. \square

Beweise zu Kapitel 5

Beweis von Satz 5.12:

Für ein 2-Punkt-Design ξ mit Stützpunkten x_1 und x_2 gilt mit $\theta_i = \beta_0 + \beta_1 x_i$ für $i = 1, 2$

$$\mathbf{c}^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{c} = \frac{\omega_1 Q(\theta_1) + \omega_2 Q(\theta_2)}{\omega_1 \omega_2 Q(\theta_1) Q(\theta_2) (x_2 - x_1)^2} = \frac{\omega_1 (1 + \lambda e^{-\theta_2}) + \omega_2 (1 + \lambda e^{-\theta_1})}{\omega_1 \omega_2 (x_2 - x_1)^2} \quad (\text{A.3})$$

und die Kriteriumsfunktion berechnet sich mit Hilfe von (A.1) zu

$$\begin{aligned} E_{\boldsymbol{\beta}}[\mathbf{c}^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{c}] \\ = \frac{\omega_1 \left(1 + \lambda \exp\left(-\mu_0 + \frac{\sigma_0^2}{2} - x_2 \mu_1 + \frac{x_2^2 \sigma_1^2}{2}\right) \right) + \omega_2 \left(1 + \lambda \exp\left(-\mu_0 + \frac{\sigma_0^2}{2} - x_1 \mu_1 + \frac{x_1^2 \sigma_1^2}{2}\right) \right)}{\omega_1 \omega_2 (x_2 - x_1)^2}. \end{aligned}$$

Sie ist von der Form (A.3) wie die Kriteriumsfunktion für c -Optimalität eines 2-Punkt-Designs, für das die c -optimalen Gewichte in (5.1) angegeben sind. Mit der in (5.12)

definierten Funktion w folgen die optimalen Gewichte

$$\omega_1^* = \frac{\frac{1}{w(x_2)}}{\frac{1}{w(x_1)} + \frac{1}{w(x_2)}} = \frac{w(x_1)}{w(x_1) + w(x_2)}, \quad (\text{A.4a})$$

$$\omega_2^* = \frac{\frac{1}{w(x_1)}}{\frac{1}{w(x_1)} + \frac{1}{w(x_2)}} = \frac{w(x_2)}{w(x_1) + w(x_2)}. \quad (\text{A.4b})$$

Einsetzen der optimalen Gewichte in die Kriteriumsfunction liefert

$$E_{\beta}[\mathbf{c}^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{c}] = \frac{(w(x_1) + w(x_2))^2}{(x_2 - x_1)^2} =: h(x_1, x_2).$$

Zur kürzeren Schreibweise sei $C := \exp(-\mu_0 + \frac{\sigma_0^2}{2} - \frac{\mu_1^2}{2\sigma_1^2})$ und die Funktion r sei definiert als

$$r(z) := \lambda \exp\left(-\mu_0 + \frac{\sigma_0^2}{2} - z\mu_1 + \frac{z^2\sigma_1^2}{2}\right) = \lambda C \exp\left(\frac{\sigma_1^2}{2} \left(z - \frac{\mu_1}{\sigma_1^2}\right)^2\right).$$

Es gilt der Zusammenhang $w(z) = \sqrt{1 + r(z)}$ und $w(\mu_1/\sigma_1^2 - z) = w(\mu_1/\sigma_1^2 + z)$.

Es gelte $x_2 > x_1$. Der Punkt $(\mu_1/\sigma_1^2 - x, \mu_1/\sigma_1^2 + x)$ liegt auf der durch $x_1 + x_2 = 2\mu_1/\sigma_1^2$ definierten Ebene. Es gilt $h(\mu_1/\sigma_1^2 - x + y, \mu_1/\sigma_1^2 + x + y) = h(\mu_1/\sigma_1^2 - x - y, \mu_1/\sigma_1^2 + x - y)$ für alle $x > 0$ und $y \geq 0$ und daher ist die Funktion h symmetrisch zur Ebene, die durch $x_1 + x_2 = 2\mu_1/\sigma_1^2$ gegeben ist. Es wird die Funktion $\tilde{h}(y) := h(\mu_1/\sigma_1^2 - x + y, \mu_1/\sigma_1^2 + x + y)$ senkrecht zur Symmetrieebene betrachtet. Ihre Ableitung berechnet sich für $x > 0$ und $y \geq 0$ zu

$$\tilde{h}'(y) = 2 \cdot \frac{w\left(\frac{\mu_1}{\sigma_1^2} - x + y\right) + w\left(\frac{\mu_1}{\sigma_1^2} + x + y\right)}{4x^2} \cdot \left[w'\left(\frac{\mu_1}{\sigma_1^2} - x + y\right) + w'\left(\frac{\mu_1}{\sigma_1^2} + x + y\right) \right].$$

Die Ableitung der Funktion $w(z) = \sqrt{1 + r(z)}$ ist gegeben durch

$$w'(z) = \frac{1}{2} \cdot \frac{r'(z)}{\sqrt{1 + r(z)}} = \frac{\sigma_1^2}{2} \cdot \left(z - \frac{\mu_1}{\sigma_1^2}\right) \cdot \frac{r(z)}{\sqrt{1 + r(z)}}.$$

Aus $r(\mu_1/\sigma_1^2 - z) = r(\mu_1/\sigma_1^2 + z)$ folgt $w'(\mu_1/\sigma_1^2 - z) = -w'(\mu_1/\sigma_1^2 + z)$ und damit $\tilde{h}'(0) = 0$. Die zweite Ableitung von $w(z)$ berechnet sich zu

$$w''(z) = \frac{\sigma_1^2}{2} \cdot \frac{\left(\sigma_1^2 \left(z - \frac{\mu_1}{\sigma_1^2}\right)^2 + 1\right) r(z) \sqrt{1 + r(z)} - \frac{\sigma_1^2}{2} \cdot \left(z - \frac{\mu_1}{\sigma_1^2}\right)^2 \cdot \frac{r(z)^2}{\sqrt{1 + r(z)}}}{1 + r(z)}$$

$$= \frac{\sigma_1^2}{2} \cdot r(z) \cdot \frac{\sigma_1^2 \left(z - \frac{\mu_1}{\sigma_1^2}\right)^2 + 1 + r(z) + \frac{\sigma_1^2}{2} \cdot \left(z - \frac{\mu_1}{\sigma_1^2}\right)^2 \cdot r(z)}{(1 + r(z))^{\frac{3}{2}}}.$$

Es gilt $w''(z) > 0$ für alle z und somit ist w' eine streng monoton wachsende Funktion. Insbesondere folgt für $y > 0$, dass $w'(\mu_1/\sigma_1^2 - x + y) + w'(\mu_1/\sigma_1^2 + x + y) > 0$ ist und damit $\tilde{h}'(y) > 0$ gilt. Die Funktion h wächst zu beiden Seiten um die Symmetrieebene und wird für $x_1 = 2\mu_1/\sigma_1^2 - x_2$ minimal. Es ist also die Funktion

$$\hat{h}(x) := h(2\mu_1/\sigma_1^2 - x, x) = \frac{1 + r(x)}{\left(x - \frac{\mu_1}{\sigma_1^2}\right)^2}$$

für $x > \mu_1/\sigma_1^2$ zu minimieren. Die Ableitung von \hat{h} berechnet sich zu

$$\hat{h}'(x) = \frac{\sigma_1^2 \left(x - \frac{\mu_1}{\sigma_1^2}\right)^2 r(x) - 2 \cdot (1 + r(x))}{\left(x - \frac{\mu_1}{\sigma_1^2}\right)^3}.$$

Sie ist genau dann gleich null, wenn gilt:

$$\begin{aligned} 0 &= \sigma_1^2 \left(x - \frac{\mu_1}{\sigma_1^2}\right)^2 r(x) - 2 - 2r(x) \\ \iff 2 &= \left(\sigma_1^2 \left(x - \frac{\mu_1}{\sigma_1^2}\right)^2 - 2\right) r(x) \\ \iff \frac{1}{\lambda C e} &= \left(\frac{\sigma_1^2}{2} \left(x - \frac{\mu_1}{\sigma_1^2}\right)^2 - 1\right) \exp\left(\frac{\sigma_1^2}{2} \left(x - \frac{\mu_1}{\sigma_1^2}\right)^2 - 1\right) \\ \iff W\left(\frac{1}{\lambda C e}\right) &= \frac{\sigma_1^2}{2} \left(x - \frac{\mu_1}{\sigma_1^2}\right)^2 - 1 \\ \iff x &= \frac{\mu_1}{\sigma_1^2} \pm \frac{1}{\sigma_1} \cdot \sqrt{2 + 2W\left(\frac{1}{\lambda C e}\right)}. \end{aligned}$$

Es sei x^* definiert wie in (5.9). Da die Funktion $\hat{h}(x)$ für $x \rightarrow \mu_1/\sigma_1^2$ und $x \rightarrow \infty$ gegen unendlich strebt, liegt bei $x = \mu_1/\sigma_1^2 + x^*$ ein Minimum vor. Die Funktion $h(x_1, x_2)$ wird daher für $x_2^* = \mu_1/\sigma_1^2 + x^*$ und $x_1^* = \mu_1/\sigma_1^2 - x^*$ minimiert. Da die optimalen Stützpunkte symmetrisch um μ_1/σ_1^2 liegen, ergeben sich die optimalen Gewichte in (A.4) zu $\omega_1^* = \omega_2^* = \frac{1}{2}$.

Nun wird der Fall $\mu_1/\sigma_1^2 - x^* < u$ betrachtet, bei dem das Minimum außerhalb des Designraums liegt. Da die Funktion $h(x_1, x_2)$ zu beiden Seiten senkrecht zur Symmetrieebene wächst und auf der Symmetrieebene in Richtung $(\mu_1/\sigma_1^2 - x^*, \mu_1/\sigma_1^2 + x^*)$ fällt, kann wie im Beweis von Satz 4.19 gefolgert werden, dass sie für $u + v \geq 2\mu_1/\sigma_1^2$ bei $x_1^* = u$ und für $u + v \leq 2\mu_1/\sigma_1^2$ bei $x_2^* = v$ minimiert wird (vgl. Abbildung A.1). Im ersten Fall $u + v \geq 2\mu_1/\sigma_1^2$ ist also die Funktion

$$k(x_2) := h(u, x_2) = \frac{(w(u) + w(x_2))^2}{(x_2 - u)^2}$$

für $x_2 > u$ zu minimieren. Ihre Ableitung berechnet sich zu

$$k'(x_2) = 2 \cdot \frac{w(u) + w(x_2)}{x_2 - u} \cdot \frac{w'(x_2)(x_2 - u) - w(u) - w(x_2)}{(x_2 - u)^2}.$$

Die Funktion $l(z_1, z_2)$ sei definiert als

$$l(z_1, z_2) := w'(z_1)(z_1 - z_2) - w(z_1) - w(z_2).$$

Die Ableitung $k'(x_2)$ ist für $x_2 > u$ genau dann gleich null, wenn $\tilde{l}_u(x_2) := l(x_2, u)$ gleich null ist. Die Ableitung von $\tilde{l}_u(x_2)$ ist gegeben durch

$$\tilde{l}'_u(x_2) = w''(x_2)(x_2 - u) + w'(x_2) - w'(x_2) = w''(x_2)(x_2 - u) > 0.$$

Die Funktion $\tilde{l}_u(x_2)$ ist für $x_2 > u$ streng monoton wachsend und kann daher höchstens eine Nullstelle haben. Folglich kann die Funktion $k(x_2)$ höchstens ein Extremum im Intervall (u, ∞) haben. Da $k(x_2)$ für $x_2 \rightarrow u$ und $x_2 \rightarrow \infty$ gegen unendlich strebt, existiert genau ein Minimum, das sich bei der eindeutigen Lösung der Gleichung $\tilde{l}_u(x_2)$ im Intervall (u, ∞) befindet. Liegt das Minimum außerhalb des Designraums, so ist $x_2^* = v$ der zweite Stützpunkt.

Für den zweiten Fall $u + v \leq 2\mu_1/\sigma_1^2$ ist die Funktion $h(x_1, v)$ für $x_1 < v$ zu maximieren. Es soll nun gezeigt werden, dass der Stützpunkt bei $x_1 = u$ liegt. Analog zu obiger Herleitung folgt, dass die Funktion $h(x_1, v)$ genau ein Minimum hat, das sich als eindeutige Lösung der Gleichung $l(x_1, v) = 0$ im Intervall $(-\infty, v)$ berechnet. Ebenfalls folgt, dass die Funktion $\tilde{l}_v(x_1) := l(x_1, v)$ für $x_1 < v$ streng monoton fallend ist. Nun soll für alle $x_1 \in [u, v]$ die folgende Ungleichungskette gezeigt werden:

$$\tilde{l}_v(x_1) = l(x_1, v) \leq l(u, v) < l(\mu_1/\sigma_1^2 - x^*, v) < l(\mu_1/\sigma_1^2 - x^*, \mu_1/\sigma_1^2 + x^*) = 0. \quad (\text{A.5})$$

Für die eindeutige Lösung \tilde{x} der Gleichung $l(x_1, v) = 0$ im Intervall $(-\infty, v)$ gilt dann $\tilde{x} < u$. Das Minimum liegt also außerhalb des Designraums und $h(x_1, v)$ wird bei $x_1^* = u$ minimiert.

Die ersten zwei Ungleichungen gelten, da $\tilde{l}_v(x_1)$ für $x_1 < v$ streng monoton fallend ist und da nach Voraussetzung $\mu_1/\sigma_1^2 - x^* < u \leq x_1$ gilt. Für die dritte Ungleichung wird zunächst festgehalten, dass aus $\mu_1/\sigma_1^2 - x^* < u \leq 2\mu_1/\sigma_1^2 - v$ die Ungleichung $v < \mu_1/\sigma_1^2 + x^*$ folgt. Die Funktion \hat{l} sei definiert als $\hat{l}(z) := l(\mu_1/\sigma_1^2 - x^*, z)$, wobei $z < \mu_1/\sigma_1^2 + x^*$. Für ihre Ableitung gilt

$$\hat{l}'(z) = -w'(\mu_1/\sigma_1^2 - x^*) - w'(z) > -w'(\mu_1/\sigma_1^2 - x^*) - w'(\mu_1/\sigma_1^2 + x^*) = 0,$$

da die Funktion w' streng monoton wachsend ist. Somit ist $\hat{l}(z)$ streng monoton wachsend und wegen $v < \mu_1/\sigma_1^2 + x^*$ folgt die Ungleichungskette (A.5).

Der Fall $\mu_1/\sigma_1^2 + x^* > v$ kann analog gezeigt werden. \square

Beweis von Satz 5.13:

Aus $u > -\vartheta$ folgt $x + \vartheta > 0$ und somit gilt

$$\begin{aligned} \int_0^\infty e^{-\beta_0 - \beta_1 x} \cdot \frac{b^a}{\Gamma(a)} \beta_0^{a-1} e^{-b\beta_0} d\beta_0 &= \frac{b^a e^{-\beta_1 x}}{(b+1)^a} \cdot \int_0^\infty \frac{(b+1)^a}{\Gamma(a)} \beta_0^{a-1} e^{-(b+1)\beta_0} d\beta_0 = \frac{b^a e^{-\beta_1 x}}{(b+1)^a} \\ \int_0^\infty \frac{b^a e^{-\beta_1 x}}{(b+1)^a} \cdot \vartheta e^{-\vartheta\beta_1} d\beta_1 &= \frac{b^a \vartheta}{(b+1)^a (\vartheta + x)} \cdot \int_0^\infty (\vartheta + x) e^{-(\vartheta+x)\beta_1} d\beta_1 = \frac{b^a \vartheta}{(b+1)^a (\vartheta + x)} \end{aligned}$$

für alle $x \in \mathcal{X}$. Die Kriteriumsfunktion eines 2-Punkt-Designs ξ mit Stützpunkten x_1 und x_2 ergibt sich mit (A.3) zu

$$E_{\beta}[\mathbf{c}^T \mathbf{M}(\xi, \beta)^{-1} \mathbf{c}] = \frac{\omega_1 \left(1 + \frac{\lambda b^a \vartheta}{(b+1)^a (\vartheta + x_2)} \right) + \omega_2 \left(1 + \frac{\lambda b^a \vartheta}{(b+1)^a (\vartheta + x_1)} \right)}{\omega_1 \omega_2 (x_2 - x_1)^2}.$$

Mit der in (5.14) definierten Funktion w folgen analog zum Beweis von Satz 5.12 die optimalen Gewichte

$$\omega_1^* = \frac{w(x_1)}{w(x_1) + w(x_2)} \quad \text{und} \quad \omega_2^* = \frac{w(x_2)}{w(x_1) + w(x_2)}.$$

Setzt man die optimalen Gewichte in die Kriteriumsfunktion ein, so ergibt sie sich zu

$$E_{\beta}[\mathbf{c}^T \mathbf{M}(\xi, \beta)^{-1} \mathbf{c}] = \frac{(w(x_1) + w(x_2))^2}{(x_2 - x_1)^2}.$$

Sie ist für $x_2 > x_1$ streng monoton fallend mit x_2 und wird daher bei $x_2^* = v$ minimiert.

Zur Bestimmung des zweiten Stützpunktes muss die Funktion

$$h(x_1) := \frac{(w(x_1) + w(v))^2}{(v - x_1)^2}$$

minimiert werden. Ihre Ableitung berechnet sich zu

$$h'(x_1) = 2 \cdot \frac{w(x_1) + w(v)}{v - x_1} \cdot \frac{w'(x_1)(v - x_1) + w(x_1) + w(v)}{(v - x_1)^2}.$$

Zur kürzeren Schreibweise sei $C := \lambda b^a \vartheta / (b + 1)^a$ und die Funktion r sei definiert als $r(x) := C / (\vartheta + x)$. Die Ableitung $h'(x_1)$ ist genau dann gleich null, wenn gilt:

$$\begin{aligned} 0 &= w'(x_1)(v - x_1) + w(x_1) + w(v) \\ &= -\frac{C}{2} \cdot \frac{(v - x_1)}{\sqrt{1 + r(x_1)}(\vartheta + x_1)^2} + \sqrt{1 + r(x_1)} + \sqrt{1 + r(v)} \\ \Leftrightarrow \sqrt{1 + r(v)}\sqrt{1 + r(x_1)} &= \frac{C}{2} \cdot \frac{(v - x_1)}{(\vartheta + x_1)^2} - 1 - r(x_1). \end{aligned}$$

Quadrieren beider Seiten liefert:

$$\begin{aligned} (1 + r(v))(1 + r(x_1)) &= \frac{\left(\frac{C}{2} \cdot (v - x_1) - (\vartheta + x_1)^2 - C \cdot (\vartheta + x_1)\right)^2}{(\vartheta + x_1)^4} \\ \Leftrightarrow \frac{(\vartheta + v + C)(\vartheta + x_1 + C)}{(\vartheta + v)(\vartheta + x_1)} &= \frac{\left(\frac{C}{2} \cdot (v - x_1) - (\vartheta + x_1)(\vartheta + x_1 + C)\right)^2}{(\vartheta + x_1)^4}. \end{aligned}$$

Diese Gleichung ist äquivalent zu:

$$\begin{aligned} 0 &= \left(\frac{C(v - x_1)}{2} - (\vartheta + x_1)(\vartheta + x_1 + C)\right)^2 (\vartheta + v) - (\vartheta + v + C)(\vartheta + x_1 + C)(\vartheta + x_1)^3 \\ &= \left(\frac{C^2(v - x_1)^2}{4} - C(v - x_1)(\vartheta + x_1)(\vartheta + x_1 + C)\right) (\vartheta + v) \\ &\quad + C(\vartheta + x_1)^2(\vartheta + x_1 + C)(v - x_1) \\ &= \frac{C^2(v - x_1)^2(\vartheta + v)}{4} - C(v - x_1)^2(\vartheta + x_1)(\vartheta + x_1 + C) \\ &= C(v - x_1)^2 \cdot \left[\frac{C}{4}(\vartheta + v) - (\vartheta + x_1)(\vartheta + x_1 + C)\right]. \end{aligned}$$

Es müssen die Nullstellen der quadratischen Funktion in eckigen Klammern bestimmt

werden. Diese berechnen sich zu:

$$x_1 = -\vartheta - \frac{C}{2} \pm \frac{1}{2} \sqrt{C^2 + C(\vartheta + v)}.$$

Wegen $-\vartheta - \frac{C}{2} - \frac{1}{2} \sqrt{C^2 + C(\vartheta + v)} < -\vartheta$ ist nur die andere Lösung zu betrachten, die größer als $-\vartheta$ ist. Da $h(x_1)$ für $x_1 \rightarrow -\vartheta$ und $x_1 \rightarrow v$ gegen unendlich strebt, handelt es sich um ein Minimum. Liegt das Minimum außerhalb des Designraums, so wird $h(x_1)$ bei $x_1^* = u$ minimiert. \square

Beweise zu Kapitel 6 und Hilfssatz

Beweis von Satz 6.2:

a) Es sei $\beta_1 > 0$ und x^* sei die eindeutige Lösung der Gleichung $\gamma_{x_k}(x) = 0$ im Intervall $(-\infty, x_k)$. Die Funktion h sei definiert als

$$h(x) := \mathbf{f}(x)^T \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{B} \mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{f}(x) - \frac{\text{Spur}(\mathbf{M}(\xi, \boldsymbol{\beta})^{-1} \mathbf{B})}{Q(\mathbf{f}(x)^T \boldsymbol{\beta})}. \quad (\text{A.6})$$

Nach dem Äquivalenzsatz ist ein Design ξ genau dann L -optimal, wenn $h(x_i) \leq 0$ für alle $x_i \in \mathcal{X}$ gilt mit Gleichheit an den Stützpunkten. Im Beweis von Satz 6.4 wird gezeigt, dass die Funktion h' höchstens zwei Nullstellen hat und $\lim_{x \rightarrow \pm\infty} h(x) = \pm\infty$. Die Funktion h kann daher höchstens drei Nullstellen $a_1 < a_2 < a_3$ besitzen. In Abbildung A.3 ist der Verlauf dieser Funktion graphisch dargestellt.

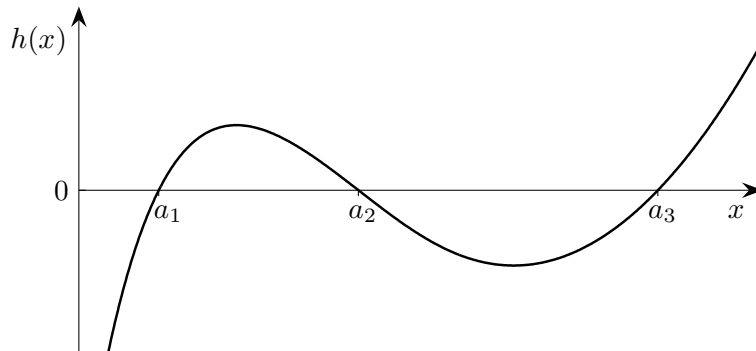


Abbildung A.3: Beweisskizze: Verlauf der Funktion $h(x)$

Gilt $x^* \leq x_1$, so hat das L -optimale Design auf dem stetigen Designraum $[x_1, x_k]$ nach Satz 6.1 die Stützpunkte x_1 und x_k . Dieses Design ist auf dem diskreten Designraum somit ebenfalls L -optimal. Sei nun $x^* > x_1$. Zunächst werden 2-Punkt-Designs betrachtet.

i) Ein L -optimales Design mit Stützpunkten $x_i < x_j$ muss einen Stützpunkt haben, der

größer als x^* ist. Angenommen, es gilt $x_i < x_j \leq x^*$. Ist $x_j = a_3$, so folgt $h(x) > 0$ für alle $x > x_j$ und insbesondere $h(x_k) > 0$. Es muss also $x_j = a_2$ sein. Dann gilt allerdings wegen der L -Optimalität $h(x_k) \leq 0$ und somit $h(x^*) \leq 0$ und das Design wäre auch auf $\mathcal{X} \cup \{x^*\}$ L -optimal. Dies widerspricht Satz 6.1.

ii) Hat ein L -optimales Design zwei Stützpunkte $x_i < x_j$ mit $x^* \leq x_i < x_j$, so muss $x_j = x_k$ gelten. Angenommen, es gilt $x_j < x_k$. Wie bei i) folgt $x_j = a_2$ und somit $x_i = a_1$. Da $h(x) \leq 0$ für alle $x \leq a_1 = x_i$, gilt insbesondere $h(x^*) \leq 0$. Es folgt wieder der Widerspruch, dass das Design auch auf $\mathcal{X} \cup \{x^*\}$ L -optimal ist.

iii) Hat ein L -optimales Design zwei Stützpunkte $x_i < x_k$ mit $x^* \leq x_i < x_k$, so muss $x_i = x^+(x^*)$ gelten. Angenommen, es gilt $x_i > x^+(x^*)$. Ist $x_i = a_1$, so folgt $h(x) < 0$ für alle $x < x_i$ und insbesondere $h(x^*) < 0$. Es folgt der Widerspruch, dass das Design auch auf $\mathcal{X} \cup \{x^*\}$ L -optimal ist. Also gilt $x_i = a_2$. Wegen der L -Optimalität muss $h(x^*) > 0$ gelten, d. h. $x^* \in (a_1, x_i)$. Da $x^+(x^*) \in [x^*, x_i)$, folgt $h(x^+(x^*)) > 0$ und damit der Widerspruch, dass das Design nicht L -optimal ist.

iv) Hat ein L -optimales Design zwei Stützpunkte $x_i < x_j$ mit $x_i < x^* < x_j$, so muss $x_i = x^-(x^*)$ gelten. Angenommen, es gilt $x_i < x^-(x^*)$. Ist $x_i = a_2$, so folgt $x_j = a_3$ und somit $h(x^*) < 0$ und das Design wäre auch auf $\mathcal{X} \cup \{x^*\}$ L -optimal. Also muss $x_i = a_1$ sein und $h(x^*) > 0$, d. h. $x^* \in (x_i, a_2)$. Da $x^-(x^*) \in (x_i, x^*)$, gilt $h(x^-(x^*)) > 0$ und es folgt der Widerspruch, dass das Design nicht L -optimal ist.

v) Hat ein L -optimales Design zwei Stützpunkte $x_i < x_j$ mit $x_i < x^* < x_j$, so muss $x_j = x^+(x^*)$ oder $x_j = x_k$ gelten. Angenommen, es gilt $x_j \notin \{x^+(x^*), x_k\}$. Wie bei i) folgt $x_j = a_2$ und es muss $h(x^*) > 0$ gelten, d. h. $x^* \in (a_1, x_j)$. Da $x^+(x^*) \in [x^*, x_j)$, gilt $h(x^+(x^*)) > 0$ und somit ist das Design nicht L -optimal.

Ist ein 3-Punkt-Design mit Stützpunkten $x_i < x_j < x_l$ L -optimal, so gilt $x_i = a_1$, $x_j = a_2$ und $x_l = a_3$. Da $h(x) > 0$ für alle $x > x_l$, muss für ein L -optimales Design $x_l = x_k$ sein. Außerdem muss $h(x^*) > 0$ gelten, d. h. $x^* \in (x_i, x_j)$. Es folgt $x_i = x^-(x^*)$ und $x_j = x^+(x^*)$.

b) Der Fall $\beta_1 < 0$ folgt analog.

c) Der Fall $\beta_1 = 0$ folgt aus Resultaten für das lineare Modell. □

Beweis von Bemerkung 6.6:

Die Matrix \mathbf{S} lässt sich mit Hilfe von Lemma 3.19 darstellen als

$$\mathbf{S} = (\tilde{\mathbf{X}}^T)^{-1} \mathbf{B} \tilde{\mathbf{X}}^{-1} = \mathbf{Q}^{-\frac{1}{2}} (\mathbf{X}^T)^{-1} \mathbf{B} \mathbf{X}^{-1} \mathbf{Q}^{-\frac{1}{2}}.$$

Die Matrix \mathbf{X} ist gegeben durch

$$\mathbf{X} = \begin{pmatrix} 1 & \mathbf{a}^T \\ \mathbf{1}_{p-1} & \mathbf{1}_{p-1}\mathbf{a}^T - \text{diag}(\mathbf{x}^*) \end{pmatrix}$$

mit Vektoren $\mathbf{x}^* = (x_1^*/\beta_1, \dots, x_{p-1}^*/\beta_{p-1})^T$ und $\mathbf{1}_{p-1} = (1, \dots, 1)^T$. Die Inverse der Matrix \mathbf{X} berechnet sich zu

$$\mathbf{X}^{-1} = \begin{pmatrix} 1 - \mathbf{a}^T \mathbf{y}^* & (\mathbf{z}^*)^T \\ \mathbf{y}^* & -\text{diag}(\mathbf{y}^*) \end{pmatrix}$$

mit $\mathbf{y}^* = (\beta_1/x_1^*, \dots, \beta_{p-1}/x_{p-1}^*)^T$ und $\mathbf{z}^* = (a_1\beta_1/x_1^*, \dots, a_{p-1}\beta_{p-1}/x_{p-1}^*)^T$.

Die Einträge $r_{i,j}$ der Matrix $\mathbf{R} := \mathbf{B}\mathbf{X}^{-1}$ sind für $i = 1, \dots, p$ und $j = 2, \dots, p$ gegeben durch

$$r_{i,1} = b_{i,1}(1 - \mathbf{a}^T \mathbf{y}^*) + \sum_{k=2}^p b_{i,k} \frac{\beta_{k-1}}{x_{k-1}^*},$$

$$r_{i,j} = \frac{\beta_{j-1}}{x_{j-1}^*} \cdot (b_{i,1}a_{j-1} - b_{i,j}).$$

Nun sind die Einträge $d_{i,j}$ der symmetrischen Matrix $\mathbf{D} := (\mathbf{X}^T)^{-1}\mathbf{R} = (\mathbf{X}^T)^{-1}\mathbf{B}\mathbf{X}^{-1}$ zu bestimmen. Der Eintrag $d_{1,1}$ ergibt sich zu

$$\begin{aligned} d_{1,1} &= (1 - \mathbf{a}^T \mathbf{y}^*)r_{1,1} + \sum_{l=2}^p y_{l-1}^* r_{l,1} \\ &= (1 - \mathbf{a}^T \mathbf{y}^*) \left(b_{1,1}(1 - \mathbf{a}^T \mathbf{y}^*) + \sum_{k=2}^p b_{1,k} \frac{\beta_{k-1}}{x_{k-1}^*} \right) \\ &\quad + \sum_{l=2}^p \frac{\beta_{l-1}}{x_{l-1}^*} \left(b_{l,1}(1 - \mathbf{a}^T \mathbf{y}^*) + \sum_{k=2}^p b_{l,k} \frac{\beta_{k-1}}{x_{k-1}^*} \right) \\ &= b_{1,1}(1 - \mathbf{a}^T \mathbf{y}^*)^2 + 2(1 - \mathbf{a}^T \mathbf{y}^*) \sum_{k=2}^p b_{1,k} \frac{\beta_{k-1}}{x_{k-1}^*} + \sum_{l=2}^p \sum_{k=2}^p b_{l,k} \frac{\beta_{l-1}}{x_{l-1}^*} \frac{\beta_{k-1}}{x_{k-1}^*}. \end{aligned}$$

Die restlichen Einträge in der ersten Zeile bzw. Spalte berechnen sich für $j = 2, \dots, p$ zu

$$\begin{aligned} d_{1,j} = d_{j,1} &= (1 - \mathbf{a}^T \mathbf{y}^*)r_{1,j} + \sum_{k=2}^p y_{k-1}^* r_{k,j} \\ &= \frac{\beta_{j-1}}{x_{j-1}^*} \cdot \left((1 - \mathbf{a}^T \mathbf{y}^*)(b_{1,1}a_{j-1} - b_{1,j}) + \sum_{k=2}^p \frac{\beta_{k-1}}{x_{k-1}^*} (b_{k,1}a_{j-1} - b_{k,j}) \right). \end{aligned}$$

Die übrigen Einträge ergeben sich für $i, j = 2, \dots, p$ zu

$$\begin{aligned}
d_{i,j} &= z_{i-1}^* r_{1,j} + (-y_{i-1}^*) r_{i,j} \\
&= \frac{a_{i-1} \beta_{i-1}}{x_{i-1}^*} \cdot \frac{\beta_{j-1}}{x_{j-1}^*} \cdot (b_{1,1} a_{j-1} - b_{1,j}) - \frac{\beta_{i-1}}{x_{i-1}^*} \cdot \frac{\beta_{j-1}}{x_{j-1}^*} \cdot (b_{i,1} a_{j-1} - b_{i,j}) \\
&= \frac{\beta_{i-1} \beta_{j-1}}{x_{i-1}^* x_{j-1}^*} \cdot (b_{1,1} a_{i-1} a_{j-1} - b_{1,j} a_{i-1} - b_{i,1} a_{j-1} + b_{i,j}).
\end{aligned}$$

Es gilt $\mathbf{S} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{D} \mathbf{Q}^{-\frac{1}{2}}$. Die beiden Diagonalmatrizen $\mathbf{Q}^{-\frac{1}{2}}$ bewirken eine Multiplikation des Eintrags $d_{i,j}$ der Matrix \mathbf{D} mit dem Faktor $1/\sqrt{q_{i,i} q_{j,j}}$, wobei $q_{k,k}$ den Diagonaleintrag (k, k) der Matrix $\mathbf{Q} = \text{diag}(Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta}), Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_1^*), \dots, Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_{p-1}^*))$ bezeichnet. \square

Lemma A.1

Die Matrix $\mathbf{S} = (s_{i,j})_{i,j=1,\dots,p}$ sei definiert wie in (3.38) als $\mathbf{S} = (\tilde{\mathbf{X}}^T)^{-1} \mathbf{B} \tilde{\mathbf{X}}^{-1}$ für ein Design mit Stützpunkten $x_1 \mathbf{e}_1, \dots, x_{p-1} \mathbf{e}_{p-1}$ und $\mathbf{0}_{p-1} = (0, \dots, 0)^T$, wobei $x_i < 0$ für $i = 1, \dots, p-1$ gilt. Dabei ist $\mathbf{B} = (b_{i,j})_{i,j=1,\dots,p}$ eine symmetrische positiv definite Matrix. Dann gilt für $i = 1, \dots, p-1$:

$$\lim_{x_i \rightarrow 0} \frac{s_{1,i+1}}{\sqrt{s_{1,1} s_{i+1,i+1}}} = -1.$$

Beweis:

Aus dem Beweis von Bemerkung 6.6 folgen die Einträge $d_{i,j}$ der symmetrischen Matrix $\mathbf{D} := (\mathbf{X}^T)^{-1} \mathbf{B} \mathbf{X}^{-1}$ für $i, j = 2, \dots, p$:

$$\begin{aligned}
d_{1,1} &= b_{1,1} - 2 \sum_{k=2}^p \frac{b_{1,k}}{x_{k-1}} + \sum_{l=2}^p \sum_{k=2}^p \frac{1}{x_{l-1}} \frac{b_{l,k}}{x_{k-1}}, \\
d_{1,j} = d_{j,1} &= \frac{b_{1,j}}{x_{j-1}} - \sum_{l=2}^p \frac{1}{x_{l-1}} \frac{b_{l,j}}{x_{j-1}}, \\
d_{i,j} &= \frac{b_{i,j}}{x_{i-1} x_{j-1}}.
\end{aligned}$$

Außerdem folgt $\mathbf{S} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{D} \mathbf{Q}^{-\frac{1}{2}}$ und damit $s_{i,j} = d_{i,j} / \sqrt{q_{i,i} q_{j,j}}$, wobei $q_{k,k}$ den Diagonaleintrag (k, k) der Matrix \mathbf{Q} bezeichnet. Es gilt $s_{1,i+1} / \sqrt{s_{1,1} s_{i+1,i+1}} = d_{1,i+1} / \sqrt{d_{1,1} d_{i+1,i+1}}$. Aufgrund der positiven Definitheit sind die Diagonaleinträge der Matrizen \mathbf{S} , \mathbf{D} und \mathbf{B}

positiv. Es folgt für $i = 1, \dots, p-1$:

$$\begin{aligned} \lim_{x_i \rightarrow 0} \frac{s_{1,i+1}}{\sqrt{s_{1,1} s_{i+1,i+1}}} &= \lim_{x_i \rightarrow 0} \frac{d_{1,i+1}}{\sqrt{d_{1,1} d_{i+1,i+1}}} = \lim_{x_i \rightarrow 0} \frac{x_i^2 d_{1,i+1}}{\sqrt{x_i^2 d_{1,1}} \sqrt{x_i^2 d_{i+1,i+1}}} \\ &= \frac{-b_{i+1,i+1}}{\sqrt{b_{i+1,i+1}} \sqrt{b_{i+1,i+1}}} = -1. \end{aligned} \quad \square$$

Beweis zu Kapitel 7

Beweis von Satz 7.1:

Es wird gezeigt, dass das Design ξ_k^* auch auf dem unbeschränkten Designraum mit $x_i \in (-\infty, v_i]$ für $\beta_i > 0$ und $x_i \in [u_i, \infty)$ für $\beta_i < 0$ ϕ_k -optimal ist. Für $i = 1, \dots, p-1$ seien

$$p_i(x) := \mathbf{f}(\mathbf{a} + (x - a_i)\mathbf{e}_i)^T \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{k-1} \mathbf{f}(\mathbf{a} + (x - a_i)\mathbf{e}_i).$$

Das Design ξ_k^* ist nach Satz 3.12, Korollar 3.23 und Bemerkung 4.3 genau dann ϕ_k -optimal, wenn

$$Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} + \beta_i(x_i - a_i)) \cdot p_i(x_i) - \text{Spur}(\mathbf{M}(\xi_k^*, \boldsymbol{\beta})^k) \leq 0$$

für $i = 1, \dots, p-1$ gilt. Diese Bedingung ist äquivalent zu

$$h_i(x_i) := p_i(x_i) - \frac{\text{Spur}(\mathbf{M}(\xi_k^*, \boldsymbol{\beta})^k)}{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} + \beta_i(x_i - a_i))} \leq 0.$$

Die Informationsmatrix des Designs ξ_k^* lässt sich gemäß Lemma 3.19 in der Form

$$\mathbf{M}(\xi_k^*, \boldsymbol{\beta}) = \tilde{\mathbf{X}}^T \mathbf{D}_\omega \tilde{\mathbf{X}} = \mathbf{X}^T \mathbf{Q}^{\frac{1}{2}} \mathbf{D}_\omega \mathbf{Q}^{\frac{1}{2}} \mathbf{X}$$

darstellen. Die Matrix \mathbf{X} und die Inverse \mathbf{X}^{-1} sind wie im Beweis von Bemerkung 6.6 gegeben durch

$$\mathbf{X} = \begin{pmatrix} 1 & \mathbf{a}^T \\ \mathbf{1}_{p-1} & \mathbf{1}_{p-1} \mathbf{a}^T - \text{diag}(\mathbf{x}^*) \end{pmatrix} \quad \text{und} \quad \mathbf{X}^{-1} = \begin{pmatrix} 1 - \mathbf{a}^T \mathbf{y}^* & (\mathbf{z}^*)^T \\ \mathbf{y}^* & -\text{diag}(\mathbf{y}^*) \end{pmatrix},$$

wobei $\mathbf{1}_{p-1} = (1, \dots, 1)^T$, $\mathbf{x}^* = (x_1^*/\beta_1, \dots, x_{p-1}^*/\beta_{p-1})^T$, $\mathbf{y}^* = (\beta_1/x_1^*, \dots, \beta_{p-1}/x_{p-1}^*)^T$ und $\mathbf{z}^* = (a_1 \beta_1/x_1^*, \dots, a_{p-1} \beta_{p-1}/x_{p-1}^*)^T$.

Unter Verwendung der Matrix $\mathbf{S} = (\tilde{\mathbf{X}}^T)^{-1} \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{k+1} \tilde{\mathbf{X}}^{-1}$ aus (3.39) folgt

$$\begin{aligned} \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{k-1} &= \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{-1} \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{k+1} \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{-1} \\ &= \tilde{\mathbf{X}}^{-1} \mathbf{D}_\omega^{-1} (\tilde{\mathbf{X}}^T)^{-1} \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{k+1} \tilde{\mathbf{X}}^{-1} \mathbf{D}_\omega^{-1} (\tilde{\mathbf{X}}^T)^{-1} \\ &= \mathbf{X}^{-1} \mathbf{Q}^{-\frac{1}{2}} \mathbf{D}_\omega^{-1} \mathbf{S} \mathbf{D}_\omega^{-1} \mathbf{Q}^{-\frac{1}{2}} (\mathbf{X}^T)^{-1} \end{aligned}$$

und

$$\begin{aligned} \text{Spur}(\mathbf{M}(\xi_k^*, \boldsymbol{\beta})^k) &= \text{Spur}(\mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{-1} \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{k+1}) \\ &= \text{Spur}(\tilde{\mathbf{X}}^{-1} \mathbf{D}_\omega^{-1} (\tilde{\mathbf{X}}^T)^{-1} \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{k+1}) \\ &= \text{Spur}(\mathbf{D}_\omega^{-1} (\tilde{\mathbf{X}}^T)^{-1} \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{k+1} \tilde{\mathbf{X}}^{-1}) = \text{Spur}(\mathbf{D}_\omega^{-1} \mathbf{S}) \\ &= \sum_{i=1}^p \frac{s_{ii}}{\omega_i^*} = \sum_{j=1}^p \sqrt{s_{jj}} \cdot \sum_{i=1}^p \sqrt{s_{ii}} = \left(\sum_{j=1}^p \sqrt{s_{jj}} \right)^2. \end{aligned}$$

Es bezeichne $\mathbf{e}_{i,p}$ den i -ten Einheitsvektor im \mathbb{R}^p . Es gilt $(\mathbf{X}^T)^{-1} \mathbf{f}(\mathbf{a}) = \mathbf{e}_{1,p}$ und $(\mathbf{X}^T)^{-1} \mathbf{f}(\mathbf{a} - (x_i^*/\beta_i) \mathbf{e}_i) = \mathbf{e}_{i+1,p}$, $i = 1, \dots, p-1$. Damit folgt

$$\begin{aligned} p_i(a_i) &= \mathbf{f}(\mathbf{a})^T \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{k-1} \mathbf{f}(\mathbf{a}) = \mathbf{f}(\mathbf{a})^T \mathbf{X}^{-1} \mathbf{Q}^{-\frac{1}{2}} \mathbf{D}_\omega^{-1} \mathbf{S} \mathbf{D}_\omega^{-1} \mathbf{Q}^{-\frac{1}{2}} (\mathbf{X}^T)^{-1} \mathbf{f}(\mathbf{a}) \\ &= \mathbf{e}_{1,p}^T \mathbf{Q}^{-\frac{1}{2}} \mathbf{D}_\omega^{-1} \mathbf{S} \mathbf{D}_\omega^{-1} \mathbf{Q}^{-\frac{1}{2}} \mathbf{e}_{1,p} = \frac{s_{1,1}}{(\omega_1^*)^2 Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})} = \frac{\left(\sum_{j=1}^p \sqrt{s_{jj}} \right)^2}{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})} \end{aligned}$$

und analog

$$\begin{aligned} p_i(a_i - x_i^*/\beta_i) &= \mathbf{e}_{i+1,p}^T \mathbf{Q}^{-\frac{1}{2}} \mathbf{D}_\omega^{-1} \mathbf{S} \mathbf{D}_\omega^{-1} \mathbf{Q}^{-\frac{1}{2}} \mathbf{e}_{i+1,p} = \frac{s_{i+1,i+1}}{(\omega_{i+1}^*)^2 Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i^*)} \\ &= \frac{\left(\sum_{j=1}^p \sqrt{s_{jj}} \right)^2}{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i^*)}. \end{aligned}$$

Es gilt $h_i(a_i) = 0$ und $h_i(a_i - x_i^*/\beta_i) = 0$. Wie in den Beweisen von Lemma 3.26 und Satz 6.4 kann gezeigt werden, dass die Funktionen $h_i'(x_i)$ höchstens zwei Nullstellen haben und $\lim_{x_i \rightarrow \pm\infty} h_i(x_i) = \pm\infty$ für $\beta_i > 0$ und $\lim_{x_i \rightarrow \pm\infty} h_i(x_i) = \mp\infty$ für $\beta_i < 0$. Es muss gezeigt werden, dass die Funktionen $h_i(x_i)$ bei $x_i = a_i - x_i^*/\beta_i$ ein Maximum haben. Wegen der Grenzwertbetrachtungen genügt es zu zeigen, dass die Gleichungen

$h'_i(a_i - x_i^*/\beta_i) = 0$ für $i = 1, \dots, p-1$ erfüllt sind. Es gilt

$$p'_i(x) = 2 \cdot \mathbf{e}_{i+1,p}^T \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{k-1} \mathbf{f}(\mathbf{a} + (x - a_i) \mathbf{e}_i)$$

und mit $\mathbf{e}_{i+1,p}^T \mathbf{X}^{-1} = (\beta_i/x_i^*, -(\beta_i/x_i^*) \mathbf{e}_i^T)$ folgt

$$\begin{aligned} p'_i(a_i - x_i^*/\beta_i) &= 2 \cdot \mathbf{e}_{i+1,p}^T \mathbf{M}(\xi_k^*, \boldsymbol{\beta})^{k-1} \mathbf{f}(\mathbf{a} - (x_i^*/\beta_i) \mathbf{e}_i) \\ &= 2 \cdot \mathbf{e}_{i+1,p}^T \mathbf{X}^{-1} \mathbf{Q}^{-\frac{1}{2}} \mathbf{D}_\omega^{-1} \mathbf{S} \mathbf{D}_\omega^{-1} \mathbf{Q}^{-\frac{1}{2}} (\mathbf{X}^T)^{-1} \mathbf{f}(\mathbf{a} - (x_i^*/\beta_i) \mathbf{e}_i) \\ &= 2(\beta_i/x_i^*) \cdot (1, -\mathbf{e}_i^T) \mathbf{Q}^{-\frac{1}{2}} \mathbf{D}_\omega^{-1} \mathbf{S} \mathbf{D}_\omega^{-1} \mathbf{Q}^{-\frac{1}{2}} \mathbf{e}_{i+1,p} \\ &= \frac{2 \cdot \frac{\beta_i}{x_i^*} \cdot s_{1,i+1}}{\omega_1^* \omega_{i+1}^* \sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})} \sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i^*)}} - \frac{2 \cdot \frac{\beta_i}{x_i^*} \cdot s_{i+1,i+1}}{(\omega_{i+1}^*)^2 Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i^*)}. \end{aligned}$$

Daher gilt

$$\begin{aligned} h'_i\left(a_i - \frac{x_i^*}{\beta_i}\right) &= \left(\sum_{j=1}^p \sqrt{s_{jj}}\right)^2 \cdot \left[\frac{2 \cdot \frac{\beta_i}{x_i^*} \cdot \frac{s_{1,i+1}}{\sqrt{s_{1,1} s_{i+1,i+1}}}}{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})} \sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i^*)}} - \frac{2 \cdot \frac{\beta_i}{x_i^*}}{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i^*)} \right. \\ &\quad \left. + \frac{\beta_i Q'(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i^*)}{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i^*)^2} \right]. \end{aligned}$$

Die Gleichungen $h'_i(a_i - x_i^*/\beta_i) = 0$ sind äquivalent zu den Gleichungen

$$2 \cdot \frac{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i^*)}{Q'(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i^*)} \cdot \left(\frac{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta} - x_i^*)}}{\sqrt{Q(\mathbf{f}(\mathbf{a})^T \boldsymbol{\beta})}} \cdot \frac{s_{1,i+1}}{\sqrt{s_{1,1} s_{i+1,i+1}}} - 1 \right) + x_i^* = 0.$$

Lösen die x_i^* und ω_i^* dieses Gleichungssystem gemeinsam mit (3.39), so ist das Design ξ_k^* ϕ_k -optimal. Mit gleicher Argumentation wie im Beweis von Satz 6.4 kann gefolgert werden, dass eine Lösung dieses gemeinsamen Gleichungssystems für die Stützpunkte eindeutig sein muss.

Angenommen, es gäbe zwei ϕ_k -optimale Designs ξ_1 und ξ_2 mit gleichen Stützpunkten aber unterschiedlichen Gewichten $\omega_i^{(1)}$ bzw. $\omega_i^{(2)}$, $i = 1, \dots, p$. Da die optimale Informationsmatrix eindeutig ist, gilt $\mathbf{0} = \mathbf{M}(\xi_1, \boldsymbol{\beta}) - \mathbf{M}(\xi_2, \boldsymbol{\beta}) = \tilde{\mathbf{X}}^T (\mathbf{D}_\omega^{(1)} - \mathbf{D}_\omega^{(2)}) \tilde{\mathbf{X}}$, wobei $\mathbf{D}_\omega^{(j)} = \text{diag}(\omega_1^{(j)}, \dots, \omega_p^{(j)})$, $j = 1, 2$. Da die Matrix $\tilde{\mathbf{X}}$ regulär ist, folgt $\mathbf{D}_\omega^{(1)} - \mathbf{D}_\omega^{(2)} = \mathbf{0}$ und damit $\omega_i^{(1)} = \omega_i^{(2)}$ für $i = 1, \dots, p$. Es folgt $\xi_1 = \xi_2$ und damit die Eindeutigkeit der Lösung des gemeinsamen Gleichungssystems. \square

Abbildungsverzeichnis

4.1	Graphische Darstellung der Funktion $d(\xi_1^*, \frac{z}{z+1})$	47
4.2	Graphische Darstellung der Funktion $d(\xi^*, \mathbf{x})$	47
4.3	Graphische Darstellung der Funktion $d(\xi, \mathbf{x})$	48
4.4	Fréchet-Ableitung für die optimalen Designs aus Tabelle 4.2	50
4.5	Fréchet-Ableitung für das Design ξ^* in (4.21)	51
4.6	Sensitivitätsfunktion $d(\xi_{\alpha_1, \alpha_2}^*, \mathbf{x})$ für die Designs $\xi_{\alpha_1, \alpha_2}^*$ in (4.25)	54
4.7	Sensitivitätsfunktion $d(\xi^*, \mathbf{x})$ für das Design ξ^* in (4.26)	55
4.8	Sensitivitätsfunktion $d(\xi^*, \mathbf{x})$ für das Design ξ^* in (4.37)	59
5.1	Elfving-Menge	63
5.2	Struktur der c -optimalen Designs	71
5.3	Fréchet-Ableitung für die optimalen Designs aus Tabelle 5.1	73
6.1	Fréchet-Ableitung für die Designs in (6.3) und (6.4)	80
A.1	Beweisskizze zu Satz 4.19	101
A.2	Beweisskizze zu Satz 4.23	107
A.3	Beweisskizze zu Satz 6.2	114

Symbolverzeichnis

$\mathbf{0}_p = (0, \dots, 0)^T$	p -dimensionaler Nullvektor
$\mathbf{1}_p = (1, \dots, 1)^T$	p -dimensionaler Einsvektor
$\ \cdot\ $	Euklidische Norm
$\sphericalangle(\mathbf{v}_1, \mathbf{v}_2)$	von den Vektoren \mathbf{v}_1 und \mathbf{v}_2 eingeschlossener Winkel
$\xi_1 \otimes \xi_2$	Produktdesign von ξ_1 und ξ_2
$\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^T$	unbekannter Parametervektor
c	Experimentdauer
C_i	Zensierungszeitpunkt des i -ten Individuums
$\text{conv}(S)$	Konvexe Hülle der Menge S
δ_i	Zensierungsindikator des i -ten Individuums
$\text{diag}(x_1, \dots, x_p)$	Diagonalmatrix mit Einträgen x_1, \dots, x_p
\mathbf{e}_i	i -ter Einheitsvektor im \mathbb{R}^{p-1}
$\mathbf{e}_{i,k}$	i -ter Einheitsvektor im \mathbb{R}^k
$\text{eff}_c(\xi, \boldsymbol{\beta})$	c -Effizienz des Designs ξ
$\text{eff}_D(\xi, \boldsymbol{\beta})$	D -Effizienz des Designs ξ
$\text{eff}_L(\xi, \boldsymbol{\beta})$	L -Effizienz des Designs ξ
$\text{eff}_{\phi_k}(\xi, \boldsymbol{\beta})$	ϕ_k -Effizienz des Designs ξ
$\text{Exp}(\lambda)$	Exponentialverteilung mit Parameter λ
$\mathbf{f}(\cdot)$	Vektor der Regressionsfunktionen
$F_\Phi(\mathbf{M}_1, \mathbf{M}_2)$	Fréchet-Ableitung von Φ bei \mathbf{M}_1 in Richtung von \mathbf{M}_2
$F_Y(\cdot)$	Verteilungsfunktion von Y
$f_Y(\cdot)$	Dichtefunktion von Y
$\Gamma(\cdot)$	Gammafunktion
$\text{Gamma}(a, b)$	Gammaverteilung mit Parametern a und b
$H_Y(\cdot)$	Kumulierte Hazardfunktion von Y
$\mathbf{I}(\mathbf{x}, \boldsymbol{\beta})$	Fisher-Informationsmatrix für eine Beobachtung bei \mathbf{x}
$\mathbf{I}(\xi_n, \boldsymbol{\beta})$	Fisher-Informationsmatrix für ein exaktes Design ξ_n
\mathbf{I}_p	Einheitsmatrix der Größe $p \times p$
$L(\boldsymbol{\beta})$	Likelihood-Funktion
$\lambda_0(\cdot)$	Baseline-Hazardfunktion

$\lambda_{\min}(\mathbf{M})$	Kleinster Eigenwert der Matrix \mathbf{M}
$\lambda_Y(\cdot)$	Hazardfunktion von Y
$\mathbf{M}(\xi, \boldsymbol{\beta})$	Informationsmatrix des Designs ξ
\mathbf{M}^-	Verallgemeinerte Inverse von \mathbf{M}
$\mathcal{M}_{\boldsymbol{\beta}}$	Menge aller Informationsmatrizen für festes $\boldsymbol{\beta}$
$N(\mu, \sigma^2)$	Normalverteilung mit Parametern μ und σ^2
$\mathcal{N}(\xi)$	Menge aller Parameterwerte $\boldsymbol{\beta}$, bei denen die D -Effizienz des Designs ξ minimal ist
$\mathcal{N}_{\mathbf{A}}$	Matrizen in $\text{PSD}(p)$, die die Identifizierbarkeitsbedingung erfüllen
ω_i	Gewicht der Versuchseinstellung \mathbf{x}_i
p	Anzahl der Parameter / Dimension des Parametervektors $\boldsymbol{\beta}$
$\text{PD}(p)$	Menge der symmetrischen positiv definiten $p \times p$ -Matrizen
$\Phi(\cdot)$	Kriteriumsfunktion
$\text{PSD}(p)$	Menge der symmetrischen positiv semidefiniten $p \times p$ -Matrizen
$Q(\cdot)$	Intensitätsfunktion
\mathcal{R}	Elfving-Menge
$S_Y(\cdot)$	Überlebensfunktion von Y
T_i	Beobachtungszeitpunkt des i -ten Individuums
Θ	Wertebereich für die Parameter
$W(\cdot)$	Lambertsche W -Funktion
\mathcal{X}	Designraum
$x^+(\cdot)$	Nächstgrößeres Element des Designraums, Gleichheit zugelassen
$x^-(\cdot)$	Nächstkleineres Element des Designraums
\mathbf{x}_i	Vektor der Kovariablen / Versuchseinstellungen des i -ten Individuums
Ξ	Menge aller Wahrscheinlichkeitsmaße auf \mathcal{X}
ξ	Design
$\xi_{\boldsymbol{\beta}}^*$	lokal optimales Design für Parametervektor $\boldsymbol{\beta}$
Y_i	Lebensdauer des i -ten Individuums

Literaturverzeichnis

- [1] Atkinson, A. C., Donev, A. N. und Tobias, R. D. (2007). *Optimum Experimental Designs, with SAS*. Oxford University Press, Oxford.
- [2] Atwood, C. L. (1969). *Optimal and Efficient Designs of Experiments*. Ann. Math. Statist., 40, 1570–1602.
- [3] Atwood, C. L. (1973). *Sequences Converging to D-Optimal Designs of Experiments*. Ann. Statist., 1, 342–352.
- [4] Bagnoli, M. und Bergstrom, T. (2005). *Log-concave probability and its applications*. Economic Theory, 26, 445–469.
- [5] Biedermann, S., Dette, H. und Zhu, W. (2006). *Optimal Designs for Dose-Response Models with Restricted Design Spaces*. J. Amer. Statist. Assoc., 101, 747–759.
- [6] Braess, D. und Dette, H. (2007). *On the number of support points of maximin and Bayesian optimal designs*. Ann. Statist., 35, 772–792.
- [7] Chaloner, K. und Larntz, K. (1989). *Optimal Bayesian design applied to logistic regression experiments*. J. Statist. Plann. Inference, 21, 191–208.
- [8] Chaloner, K. und Verdinelli, I. (1995). *Bayesian Experimental Design: A Review*. Statistical Science, 10, 273–304.
- [9] Chernoff, H. (1953). *Locally optimal designs for estimating parameters*. Ann. Math. Statist., 24, 586–602.
- [10] Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J. und Knuth, D. E. (1996). *On the Lambert W Function*. Advances in Computational Mathematics, 5, 329–359.
- [11] Cox, D. R. (1972). *Regression Models and Life-Tables*. J. R. Statist. Soc. B, 34, 187–220.
- [12] Cox, D. R. und Oakes, D. (1984). *Analysis of Survival Data*. Chapman and Hall, London.

- [13] Dette, H. (1996). *Lower bounds for efficiencies with applications*. Research Developments in Probability and Statistics: Festschrift in Honor of Madan L. Puri on the Occasion of his 65th Birthday (Hrsg. E. Brunner und M. Denker), 111–124, VSP, Utrecht.
- [14] Dette, H. (1997). *Designing Experiments with Respect to ‘Standardized’ Optimality Criteria*. J. R. Statist. Soc. B, 59, 97–110.
- [15] Dette, H. und Haines, L. M. (1994). *E-optimal designs for linear and nonlinear models with two parameters*. Biometrika, 81, 739–754.
- [16] Dette, H., Haines, L. M. und Imhof, L. A. (2007). *Maximin and Bayesian optimal designs for regression models*. Statistica Sinica, 17, 463–480.
- [17] Dette, H. und Studden, W. J. (1993). *Geometry of E-Optimality*. Ann. Statist., 21, 416–433.
- [18] Duchateau, L. und Janssen, P. (2008). *The Frailty Model*. Springer, New York.
- [19] Elfving, G. (1952). *Optimum Allocation in Linear Regression Theory*. Ann. Math. Statist., 23, 255–262.
- [20] Fedorov, V. V. (1972). *Theory of Optimal Experiments*. Academic Press, New York.
- [21] Firth, D. und Hinde, J. P. (1997). *On Bayesian D-optimum Design Criteria and the Equivalence Theorem in Non-linear Models*. J. R. Statist. Soc. B, 59, 793–797.
- [22] Ford, I., Torsney, B. und Wu, C. F. J. (1992). *The Use of a Canonical Form in the Construction of Locally Optimal Designs for Non-Linear Problems*. J. R. Statist. Soc. B, 54, 569–583.
- [23] Graßhoff, U., Holling, H. und Schwabe, R. (2016). *Optimal Design for the Rasch Poisson-Gamma Model*. In mODa11 - Advances in Model-Oriented Design and Analysis (Herausgegeben von J. Kunert, C. H. Müller und A. C. Atkinson), 133–141. Springer.
- [24] Kabera, G. M., Haines, L. M. und Ndlovu, P. (2015). *The analytic construction of D-optimal designs for the two-variable binary logistic regression model without interaction*. Statistics, 49, 1169–1186.
- [25] Kiefer, J. (1959). *Optimum Experimental Designs*. J. R. Statist. Soc. B, 21, 272–319.
- [26] Kiefer, J. (1974). *General Equivalence Theory for Optimum Designs (Approximate Theory)*. Ann. Statist., 2, 849–879.

- [27] Klein, J. P. und Moeschberger, M. L. (2003). *Survival analysis: Techniques for Censored and Truncated Data*. 2nd ed. Springer, New York.
- [28] Könnner, D. (2011). *Optimale Designs für Poisson-Regression*. Universität Magdeburg, Fakultät für Mathematik, Diplomarbeit.
- [29] Konstantinou, M. (2013). *Locally optimal and robust designs for two-parameter nonlinear models with application to survival models*. University of Southampton, Mathematics, Doctoral Thesis.
- [30] Konstantinou, M., Biedermann, S. und Kimber, A. (2014). *Optimal designs for two-parameter nonlinear models with application to survival models*. *Statistica Sinica*, 24, 415–428.
- [31] Kulpa, W. (1997). *The Poincaré-Miranda Theorem*. *Amer. Math. Monthly*, 104, 545–550.
- [32] Müller, C. H. (1995). *Maximin efficient designs for estimating nonlinear aspects in linear models*. *J. Statist. Plann. Inference*, 44, 117–132.
- [33] Prékopa, A. (1973). *On logarithmic concave measures and functions*. *Acta Scientiarum Mathematicarum*, 34, 335–343.
- [34] Pronzato, L. und Pázman, A. (2013). *Design of Experiments in Nonlinear Models: Asymptotic Normality, Optimality Criteria and Small-Sample Properties*. Springer, New York.
- [35] Pukelsheim, F. (1993). *Optimal Design of Experiments*. Wiley, New York.
- [36] Pukelsheim, F. und Rieder, B. (1992). *Efficient rounding of approximate designs*. *Biometrika*, 79, 763–770.
- [37] Pukelsheim, F. und Studden, W. J. (1993). *E-Optimal Designs for Polynomial Regression*. *Ann. Statist.*, 21, 402–415.
- [38] Pukelsheim, F. und Titterington, D. M. (1983). *General Differential and Lagrangian Theory for Optimal Experimental Design*. *Ann. Statist.*, 11, 1060–1068.
- [39] Pukelsheim, F. und Torsney, B. (1991). *Optimal Weights for Experimental Designs on Linearly Independent Support Points*. *Ann. Statist.*, 19, 1614–1625.
- [40] Rodríguez-Torreblanca, C. und Rodríguez-Díaz, J. M. (2007). *Locally D- and c-optimal designs for Poisson and negative binomial regression models*. *Metrika*, 66, 161–172.

- [41] Russell, K. G., Woods, D. C., Lewis, S. M. und Eccleston, E. C. (2009). *D-optimal designs for Poisson regression models*. *Statistica Sinica*, 19, 721–730.
- [42] Schmidt, D. (2013). *Optimales Design für zensierte Daten*. Universität Magdeburg, Fakultät für Mathematik, Diplomarbeit.
- [43] Schmidt, D. und Schwabe, R. (2015). *On optimal designs for censored data*. *Metrika*, 78, 237–257, Erratum, 259.
- [44] Schmidt, D. und Schwabe, R. (2017). *Optimal Design for Multiple Regression with Information Driven by the Linear Predictor*. *Statistica Sinica*, 27, 1371–1384.
- [45] Schwabe, R. (1996). *Optimum Designs for Multi-Factor Models*. Springer, New York.
- [46] Sebastiani, P. und Settimi, R. (1997). *A note on D-optimal designs for a logistic regression model*. *J. Statist. Plann. Inference*, 59, 359–368.
- [47] Silvey, S. D. (1978). *Optimal design measures with singular information matrices*. *Biometrika*, 65, 553–559.
- [48] Silvey, S. D. (1980). *Optimal Design*. Chapman and Hall, London.
- [49] Sitter, R. R. und Torsney, R. (1995). *Optimal designs for binary response experiments with two design variables*. *Statistica Sinica*, 5, 405–419.
- [50] Wang, Y., Myers, R. H., Smith, E. P. und Ye, K. (2006). *D-optimal designs for Poisson regression models*. *J. Statist. Plann. Inference*, 136, 2831–2845.