Lineare Spiele in der Stichprobentheorie

Inauguraldissertation
zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften
der Universität Mannheim

vorgelegt von
Jochen Schmidt

Mannheim 2002

Dekan:	Professor Dr. C. Buchheim
	Universität Mannheim
Referent:	Professor Dr. H. Stenger
	Universität Mannheim
Korreferent:	PD Dr. S. Gabler
	ZUMA
Tag der mündlichen Prüfung:	13.06.2002

Inhaltsverzeichnis

Einleitung

Bei der Auswertung einer Umfrage muss der Statistiker ausgehend von einer Stichprobe Rückschlüsse auf die Gesamtheit ziehen. Es stellt sich die Frage, wie die Stichprobenelemente auszuwählen sind und welcher Schätzer zur Hochrechnung der erhobenen Daten verwendet werden soll. Die Präzision der Schätzung wird erhöht, wenn bekannte Informationen (so genannte a-priori-Informationen) einbezogen werden. Anhand der a-priori-Informationen lassen sich meist gewisse Parameter ausschließen, und die Strategie soll auf der Menge der möglichen Parameter, dem Parameterraum, effizient sein. Wie dieser festzulegen ist, werden wir im ersten Kapitel diskutieren. Grob lassen sich die von uns betrachteten Parameterräume in das Ellipsoid und den Quader einteilen, was auch den Aufbau der vorliegenden Arbeit festlegt: Während die Kapitel 1 und 2 allgemein gültige Aussagen beinhalten, beziehen sich Kapitel 3 bis 6 auf das Ellipsoid und Kapitel 7 und 8 auf den Quader.

Im Allgemeinen wird man keine (in Hinblick auf den quadratischen Verlust) beste Strategie finden, so dass andere Optimalitätskriterien in den Vordergrund treten: Wir suchen eine Strategie, die im schlimmsten Fall am effizientesten ist. Dass solche – sogar zulässige – Minimax-Strategien für die von uns betrachteten Parameterräume immer existieren, wird für das Ellipsoid in Kapitel 3 und für den Quader in Kapitel 7 bewiesen. Die Beweise basieren auf der Idee von Prof. Stenger, das Auswertungsproblem als ein lineares Spiel aufzufassen:

> Ein lineares Spiel besteht aus zwei Teilmengen eines Hilbertraums, wobei die Teilmengen stellvertretend für die Strategien von zwei Spielern, in unserem Fall Natur und Statistiker, stehen. Beide wählen unabhängig voneinander jeweils eine Strategie, also ein Element der ihnen zugeordneten Teilmenge. Der Wert des Skalarprodukts ist als Verlust des Statistikers und Gewinn der Natur aufzufassen.

Genauer betrachten wir den Vektorraum der symmetrischen Matrizen mit dem durch den Spur-Operator festgelegten Skalarprodukt. In Kapitel 2 werden die Strategienräume der Spieler als Teilmengen dieses Vektorraums dargestellt. Wir werden sehen, dass jede Strategie des Statistikers durch eine zulässige Strategie dominiert wird.

In Kapitel 3 beweisen wir neben der Existenz von zulässigen Minimax-Strategien, dass auch die Natur eine Sicherheitsstrategie besitzt. Tatsächlich ist es zur Berechnung von Minimax-Strategien nützlich, der Natur ein vernunftmäßiges Verhalten zu unterstellen. Wir erläutern das Prinzip der Indifferenz (eine Anleihe aus der Spieltheorie), das für die Lösung des Minimax-Problems in vielen Fällen sehr hilfreich ist.

Es ist nicht einfach, Minimax-Strategien zu berechnen. Die Komplexität des Problems zeigt sich schon im Fall Zwei aus Drei, der für das Ellipsoid in Kapitel 4 und für den Quader in Kapitel 7 ausführlich behandelt wird. Auch wenn dieser Fall für die Praxis wenig relevant ist, hoffen wir, unser Verständnis für das Spiel zu verbessern.

Kapitel 5 befasst sich mit Stichproben vom Umfang Zwei aus einer beliebigen Gesamtheit. Wir suchen Existenzbedingungen für Indifferenz-Strategien der Natur und beschäftigen uns mit der Frage, ob es sinnvoll ist, gewisse Einheiten bewusst auszuwählen?

Häufig zerfällt die Gesamtheit auf natürliche Weise in Schichten, was die Verwendung von Schichtungsstrategien nahe legt. Welche Aufteilungen (in Abhängigkeit weiterer a-priori-Informationen) dabei das Minimax-Prinzip erfüllen, werden wir in Kapitel 6 sehen.

Die Kapitel 7 und 8 sind ganz dem Quader gewidmet. Er ist schwieriger zu handhaben, da sich z.B. das Indifferenz-Prinzip nur in Sonderfällen anwenden lässt. Trotzdem werden wir das Minimax-Problem in Spezialfällen lösen, und – wie zu Beginn bemerkt – die Existenz von zulässigen Minimax-Strategien beweisen, wobei auch hier die Natur eine Sicherheitsstrategie besitzt.

Der Anhang besteht aus den für unsere Anwendungen notwendigen mathematischen Ergänzungen. Um die Eigenständigkeit der Kapitel zu gewährleisten, kann es zu Wiederholungen kommen.

Es ist kein Anliegen dieser Arbeit, eine chronologische Übersicht der bisherigen Ergebnisse zu liefern. Sie ist eine Mischung aus bekannten und neuen Resultaten. Die bekannten Resultate, die mit entsprechenden Literaturangaben versehen sind, werden im Rahmen eines linearen Spiels diskutiert und alternativ bewiesen. Speziell lenken wir unsere Aufmerksamkeit auf die Sicherheitsstrategie der Natur, die so genannte Maximin-Strategie: Oftmals helfen Kenntnisse über die Maximin-Strategie bei der Berechnung der Minimax-Strategie. Um die Struktur des Spiels zu erfassen, werden wir viele Beispiele behandeln.

An dieser Stelle möchte ich all jenen danken, die zum Gelingen dieser Arbeit beigetragen haben. Mein besonderer Dank gilt Herrn Prof. Dr. Horst Stenger für die Anregung zu dieser Arbeit und die stetige Betreuung während ihrer Verfassung. Weiter bedanke ich mich herzlich bei Herrn PD Dr. Siegfried Gabler für die Übernahme des Korreferats.

Kapitel 1

Einführung der Grundbegriffe

Das Auswertungsproblem besteht darin, die Summe der Werte (oder deren arithmetisches Mittel) eines uns interessierenden Merkmales einer Gesamtheit anhand einer Teilerhebung zu schätzen: Angenommen, man interessiert sich für den Umsatz der Handwerkbetriebe einer Region, aus Kostengründen können aber nicht alle Betriebe befragt werden. Welche Betriebe sollen ausgewählt werden, und wie schätzt man mittels der erhobenen Daten den Gesamtumsatz möglichst effizient?

Wir werden in diesem Kapitel das Auswertungsproblem als lineares Spiel formulieren, wobei wir allgemein davon ausgehen, dass bei der Erhebung weder Antwortausfälle (eine ausgewählte Einheit verweigert die Auskunft) noch Antwortfehler (eine ausgewählte Einheit gibt eine falsche Antwort) auftreten.

1.1 Grundbegriffe der Stichprobentheorie

Gegeben sei eine Gesamtheit $\mathcal{U}$, bestehend aus den Einheiten $1, \ldots, N$, wobei jeder Einheit ein uns nicht bekannter Wert $y_i \in \mathbb{R}$ eines interessierenden Merkmales zugeordnet ist. Der resultierende Vektor $\underline{y} := (y_1, \ldots, y_N)^t \in \mathbb{R}^N$ heißt *Parameter* der Gesamtheit. Um die Merkmalsumme $y := \sum_{i=1}^N y_i$ oder deren arithmetisches Mittel $\overline{y} = \frac{1}{N} y$ zu schätzen, gehen wir wie folgt vor: Wir wählen einige Einheiten aus, ermitteln die zugehörigen y-Werte und schätzen mit deren Hilfe die uns interessierende Größe. Welche Einheiten wir auswählen und wie wir schätzen, wird stark von unseren a-priori-Informationen abhängen. Oftmals kennt man die Werte eines oder mehrerer Hilfsmerkmale, die in einem Zusammenhang zu den Werten des uns interessierenden Merkmales stehen. Angenommen, man kennt die Mitarbeiterzahlen von N Betrieben einer Branche und interessiert sich für die

Krankmeldungen y an einem bestimmten Tag. Sei x_i die Anzahl der Mitarbeiter des i-ten Betriebs und $\underline{x} = (x_1, \ldots, x_N)^t \in \mathbb{R}^N$. Das Hilfsmerkmal grenzt die Menge der a-priori möglichen Parameter ein. Für den Parameter $\underline{y} \in \mathbb{R}^N$ gilt

$$0 \leq y_i \leq x_i \quad \forall i = 1, \ldots, N.$$

Interessiert man sich für den Umsatz y der N Betriebe, so lässt sich vermuten, dass für den Parameter $\underline{y}$ in etwa gilt

$$\underline{y} = \beta \underline{x}$$

für ein (unbekanntes) $\beta \in \mathbb{R}$. Allgemeiner gehen wir von der folgenden Situation aus: Seien $\underline{x}_1, \ldots, \underline{x}_K \in \mathbb{R}^N$, $1 \leq K < N$, linear unabhängige Vektoren. Wir nehmen an, dass der Parameter $\underline{y}$ in der Nähe der Unterraums

$$\text{Bild}(X) = \{X\underline{\beta}; \underline{\beta} \in \mathbb{R}^K\}$$

liegt, wobei X die $N \times K$-Matrix mit den Spalten $\underline{x}_1, \ldots, \underline{x}_K$ bezeichnet. Wir nennen eine $N \times K$-Matrix mit vollem Spaltenrang eine *Regressionsmatrix*. Der Begriff der Nähe wird modelliert durch eine symmetrische positiv semidefinite Matrix U vom Rang $N - K$ mit $UX = 0$: $\underline{y}$ liegt in der Nähe von Bild(X), falls $\underline{y}^t U \underline{y}$ klein ist. Eine genauere Motivation für die Matrix U finden wir in Abschnitt 1.3. Falls die Regressionsmatrix durch einen Vektor $\underline{x} \in \mathbb{R}^N$ gegeben ist, fordern wir $x_i > 0$ für $i = 1, \ldots, N$,

$$x := \sum_i x_i = 1$$

und nennen $\underline{x}$ einen *Regressionsvektor*.

Die Menge der a-priori möglichen Parameter heißt *Parameterraum* und wird mit

$$\Theta$$

bezeichnet. Wir werden in dieser Arbeit immer von der Existenz einer Regressionsmatrix X bzw. eines Regressionsvektors $\underline{x}$ ausgehen. Als Parameterräume betrachten wir

$$\Theta_Q := \{\underline{y} \in \mathbb{R}^N; 0 \leq y_i \leq x_i \ \forall i = 1, \ldots, N\}$$

oder

$$\Theta_U := \{\underline{y} \in \mathbb{R}^N; \underline{y}^t U \underline{y} \leq 1\},$$

wobei U eine symmetrische positiv semidefinite $N \times N$-Matrix mit Rang $N-K$ und $UX = 0$ bezeichnet.

Jede Teilmenge s der Gesamtheit heißt *Stichprobe*, die Anzahl der Elemente $|s|$ in s *Stichprobenumfang*. Ein *Auswahlplan* p ist eine Verteilung auf der Menge aller Stichproben, also auf der Potenzmenge $2^{\mathcal{U}}$ von $\mathcal{U}$. Der *Träger* $S(p)$ von p ist definiert durch

$$S(p) := \{s \in 2^{\mathcal{U}}; p(s) > 0\}.$$

Falls der Träger nur Stichproben vom Umfang n enthält, nennen wir p einen Auswahlplan *vom Umfang* n. In Hinblick auf spätere Anwendungen definieren wir

$$S(n) := \{s \in 2^{\mathcal{U}}; |s| = n\} \quad \text{für } n \in \mathbb{N}.$$

Für einen beliebigen Auswahlplan p sind die *Inklusionswahrscheinlichkeiten* $\pi_{ij}(p)$ von p definiert durch

$$\pi_{ij}(p) := \sum_{s; i,j \in s} p(s) \quad \forall i, j = 1, \dots, N.$$

Dabei ist $\pi_{ij}(p)$ die Wahrscheinlichkeit, mit der die Einheiten i und j in die Auswahl gelangen. Falls $i = j$, schreiben wir statt $\pi_{ii}(p)$ auch $\pi_i(p)$ für $i = 1, \dots, N$. Man nennt $\pi_i(p)$ Inklusionswahrscheinlichkeit *erster Ordnung* und $\pi_{ij}(p)$ für $i \neq j$ Inklusionswahrscheinlichkeit *zweiter Ordnung*.

Eine Abbildung $t : 2^{\mathcal{U}} \times \Theta \to \mathbb{R}$, $(s, \underline{y}) \mapsto t(s, \underline{y})$, die (bei festem s) nur von den y_i-Werten für $i \in s$ abhängt, heißt *Schätzer*, $t(s, \underline{y})$ *Schätzung*. Falls p ein Auswahlplan ist und t ein Schätzer (der nur für $s \in S(p)$ definiert sein muss), bezeichnen wir (p, t) als *Strategie* des Statistikers. Für $\underline{y} \in \Theta$ sei

$$\mathrm{E}_p t(\,.\,, \underline{y}) := \sum_s p(s) t(s, \underline{y})$$

der Erwartungswert von (p, t) (bzgl. $\underline{y}$) und

$$\mathrm{var}_p t(\,.\,, \underline{y}) := \sum_s p(s)[t(s, \underline{y}) - \mathrm{E}_p t(\,.\,, \underline{y})]^2$$

die Varianz von (p, t) (bzgl. $\underline{y}$). Falls

$$\mathrm{E}_p t(\,.\,, \underline{y}) = y \quad \forall \underline{y} \in \Theta,$$

heißt die Strategie (p, t) *unverzerrt* für y. Als Maß für die Abweichung verwenden wir den *mittleren quadratischen Verlust* (MSE)

$$\mathrm{MSE}(\underline{y}, (p, t)) := \mathrm{E}_p(t(\,.\,, \underline{y}) - y)^2 = \sum_s p(s)(t(s, \underline{y}) - y)^2,$$

der auch als *Risiko* der Strategie (p,t) (bzgl. $\underline{y}$) bezeichnet wird. Falls eine Strategie (p,t) unverzerrt für y ist, gilt offensichtlich

$$\mathrm{MSE}(\underline{y},(p,t)) = \mathrm{var}_p t(\,.\,,\underline{y}) \quad \forall \underline{y} \in \Theta.$$

1.2 Beispiele für Strategien des Statistikers

Eine Strategie (p,t) (bzw. der Schätzer t) heißt *affin-linear*, falls

$$t(s,\underline{y}) = \sum_{i \in s} a_{si} y_i + b_s \quad \forall s \in S(p), \forall \underline{y} \in \mathbb{R}^N,$$

wobei $a_{si}, b_s \in \mathbb{R}$. Falls $b_s = 0$ für alle $s \in S(p)$, heißt die Strategie (bzw. der Schätzer) *linear*. Ist p ein Auswahlplan vom Umfang n, so nennen wir (p,t) (affin-)lineare Strategie vom Umfang n. Lineare Strategien vom Umfang n lassen sich wie folgt charakterisieren: Sei (p,t) eine lineare Strategie mit

$$t(s,\underline{y}) = \sum_{i \in s} a_{si} y_i \quad \forall s \in S(p), \forall \underline{y} \in \mathbb{R}^N.$$

Definieren wir $a_{si} := 0$ für $i \notin s$ und $\underline{a}_s := (a_{s1}, \ldots, a_{sN})^t \in \mathbb{R}^N$, so ist

$$t(s,\underline{y}) = \underline{y}^t \underline{a}_s = (\underline{y}, \underline{a}_s)_2,$$

wobei $(\,.\,,\,.\,)_2$ das Euklidische Skalarprodukt bezeichnet.

Definition 1.1 *Ein Vektor* $\underline{a} \in \mathbb{R}^N$ *heißt* Schätzvektor *(vom Umfang n), falls*

$$|\{i; a_i = 0\}| \geq N - n.$$

Ein Schätzvektor $\underline{a}_s$ heißt Schätzvektor bzgl. der Stichprobe s, falls $\underline{a}_s$ ein Schätzvektor ist und

$$a_{sj} = 0 \quad \forall j \notin s.$$

Ein linearer Schätzer t entspricht einer Menge $\{\underline{a}_s; s \in S(p)\}$ von Schätzvektoren. Andererseits definiert eine beliebige Menge von Schätzvektoren $\{\underline{a}_s; s \in S(p)\}$ in offensichtlicher Weise einen linearen Schätzer. Statt (p,t) schreiben wir auch $(p, \{\underline{a}_s\})$.

Sei X eine Regressionsmatrix mit vollem Spaltenrang K. Eine lineare Strategie $(p, \{\underline{a}_s\})$ heißt *repräsentativ* bzgl. X (siehe Hajek [12]), falls

$$X^t(\underline{a}_s - \underline{1}) = \underline{0} \quad \forall s \in S(p),$$

wobei $\underline{1} = (1, \ldots, 1)^t \in \mathbb{R}^N$ und $\underline{0} = (0, \ldots, 0)^t \in \mathbb{R}^K$.

Bemerkung 1.2 *Die Schätzung mittels einer repräsentativen Strategie ist für alle Parameter aus Bild(X) exakt, d.h. das Risiko ist Null. Unter der Annahme, dass der Parameter in der Nähe des Unterraums Bild(X) liegt, ist die Verwendung von repräsentativen Strategien naheliegend.*

In Hinblick auf spätere Anwendungen definieren wir:

Definition 1.3 *Ein Schätzvektor $\underline{a}$ heißt* repräsentativ *(bezüglich X), falls*

$$X^t(\underline{a} - \underline{1}) = \underline{0}.$$

Weiter sei

$$\begin{aligned} A &:= \{\underline{a} \in \mathbb{R}^N; \underline{a} \text{ ist repräsentativer Schätzvektor}\}, \\ A_s &:= \{\underline{a}_s \in A; \underline{a}_s \text{ ist repräsentativer Schätzvektor bzgl. } s\}. \end{aligned}$$

Weitere Beispiele für Strategien sind:
1) Die *Standardstrategie* (p_0, t_0): Sei $n \in \mathbb{N}$, $n \leq N$. Mit

$$p_0(s) := \begin{cases} \binom{N}{n}^{-1} & : \ |s| = n \\ 0 & : \ \text{sonst} \end{cases}$$

ist p_0 ein Auswahlplan vom Umfang n, auch bekannt als *einfache Zufallsauswahl.* Der *Standardschätzer* t_0 ist definiert durch

$$t_0(s, \underline{y}) := \frac{N}{n} \sum_{i \in s} y_i \quad \forall s \in S(p), \forall \underline{y} \in \mathbb{R}^N,$$

und die Strategie (p_0, t_0) ist unverzerrt für y mit

$$\text{var}_{p_0} t_0(\,.\,, \underline{y}) = \frac{(N-n)N}{n} s_{yy}$$

wobei $s_{yy} = \frac{1}{N-1} \sum_i (y_i - \overline{y})^2$ die *korrigierte Varianz* von $\underline{y}$ bezeichnet.
2) Sei p ein beliebiger Auswahlplan mit den Inklusionswahrscheinlichkeiten π_{ij} und $\pi_i > 0$ für $i = 1, \ldots, N$. Der durch

$$t_{HT}(s, \underline{y}) := \sum_{i \in s} \frac{y_i}{\pi_i} \quad \forall s \in S(p), \forall \underline{y} \in \mathbb{R}^N$$

definierte Schätzer heißt *Horvitz-Thompson-Schätzer.* Die Strategie (p, t_{HT}) heißt *Horvitz-Thompson-Strategie* (HT-Strategie; vgl. Horvitz und Thompson [14]) und ist unverzerrt für y mit

$$\text{var}_p t_{HT}(\,.\,, \underline{y}) = \sum_{i,j} \frac{y_i y_j}{\pi_i \pi_j} (\pi_{ij} - \pi_i \pi_j).$$

Im Fall $p = p_0$ entspricht der Horvitz-Thompson-Schätzer dem Standardschätzer.
3) Die *Rao-Hartley-Cochran-Strategie* (RHC-Strategie; vgl. Rao, Hartley und Cochran [18]) (p_{RHC}, t_{RHC}) : Sei $\underline{x} \in \mathbb{R}^N$ ein Regressionsvektor. Der Einfachheit halber nehmen wir an, dass $\frac{N}{n}$ eine natürliche Zahl ist, und zerlegen $\mathcal{U}$ in n Schichten $\mathcal{U}(1), \ldots, \mathcal{U}(n)$ der Größe $\frac{N}{n}$. Aus jeder Schicht wählen wir eine Einheit wie folgt zufällig aus: Für $i \in \mathcal{U}(h)$ ist die Wahrscheinlichkeit, die i-te Einheit auszuwählen gegeben durch

$$\frac{x_i}{\sum_{j \in \mathcal{U}(h)} x_j}.$$

Wir erhalten eine Stichprobe s vom Umfang n und benutzen für die Schätzung von y den Schätzer

$$t_{RHC}(s, \underline{y}) := \sum_{i \in s} \frac{y_i}{x_i} \sum_{\substack{j \in \mathcal{U}(h): \\ i \in \mathcal{U}(h)}} x_j.$$

Für eine Stichprobe s ist $t_{RHC}(s, \,.\,)$ eine Zufallsvariable, abhängig von der Zerlegung von $\mathcal{U}$. Bei vorgegebener Zerlegung ist die RHC-Strategie eine repräsentative Strategie vom Umfang n. Es gilt

$$\mathrm{E}_{p_{RHC}} t_{RHC}(\,.\,, \underline{y}) = y,$$

$$\mathrm{var}_{p_{RHC}} t_{RHC}(\,.\,, \underline{y}) = \frac{1}{n} \frac{N-n}{N-1} \sum_{i=1}^{N} x_i \Big(\frac{y_i}{x_i} - y\Big)^2$$

für alle $\underline{y} \in \mathbb{R}^N$.
4) Die *Hansen-Hurwitz-Strategie* (HH-Strategie; vgl. Hansen und Hurwitz [13]) $(\tilde{p}_{HH}, t_{HH})$: Sei $\underline{x} \in \mathbb{R}^N$ ein Regressionsvektor. $\tilde{p}_{HH}$ bezeichnet den geordneten Auswahlplan, der jeder geordneten Stichprobe $\tilde{s} = (i(1), \ldots, i(n))$ mit $1 \leq i(1), \ldots, i(N) \leq N$ die Wahrscheinlichkeit

$$\tilde{p}_{HH}(\tilde{s}) = x_{i(1)} \cdots x_{i(n)}$$

zuordnet. Der Hansen-Hurwitz-Schätzer (HH-Schätzer) ist definiert durch

$$t_{HH}(\tilde{s}, \underline{y}) := \frac{1}{n} \sum_{i=1}^{N} L_i \frac{y_i}{x_i},$$

wobei die Zufallsvariable L_i angibt, wie oft die Einheit i in der geordneten Stichprobe $\tilde{s}$ vorkommt. Es ist $\mathrm{E}_{\tilde{p}_{HH}} L_i = n x_i$ für $i = 1, \ldots, N$ und somit

$$\mathrm{E}_{\tilde{p}_{HH}} t_{HH}(\,.\,, \underline{y}) = y \quad \forall \underline{y} \in \mathbb{R}^N.$$

Wegen $\mathrm{var}_{\tilde{p}_{HH}} L_i = nx_i(1-x_i)$ für $i = 1, \ldots, N$ und $\mathrm{cov}_{\tilde{p}_{HH}}(L_i, L_j) = -nx_i x_j$ für $i \neq j$ gilt

$$\mathrm{var}_{\tilde{p}_{HH}} t_{HH}(\,.\,, \underline{y}) = \frac{1}{n}\sum_{i=1}^{N} x_i\Big(\frac{y_i}{x_i} - y\Big)^2 \quad \forall \underline{y} \in \mathbb{R}^N.$$

1.3 Superpopulationsmodelle

Bei gegebener Regressionsmatrix X wird man oftmals (aufgrund weiterer Vorkenntnisse) von einem *Superpopulationsmodell* $M(\Omega)$ ausgehen: Wir interpretieren $\underline{y}$ als Realisation eines Zufallsvektors $\underline{Y} = (Y_1, \ldots, Y_N)^t$ mit

$$\underline{Y} = X\underline{\beta} + \underline{\epsilon},$$

wobei $\underline{\beta} \in \mathbb{R}^K$ unbekannt ist und $\underline{\epsilon} = (\epsilon_1, \ldots, \epsilon_N)^t$ einen Zufallsvektor bezeichnet mit

$$\mathrm{E}(\underline{\epsilon}) = \underline{0}, \quad \mathrm{var}(\underline{\epsilon}) = \Omega.$$

Ausgehend von einer Stichprobe s (vom Umfang n) ist eine lineare Prognose für $\sum Y_i$ gegeben durch

$$\underline{a}_s^t \underline{Y},$$

wobei $\underline{a}_s$ einen Schätzvektor bzgl. s bezeichnet. Diese Prognose ist genau dann unverzerrt, wenn

$$X^t(\underline{a}_s - \underline{1}) = \underline{0},$$

d.h. falls $\underline{a}_s$ repräsentativ ist. Unter diesen Voraussetzungen wird

$$\mathrm{var}((\underline{a}_s - \underline{1})^t \underline{Y}) = (\underline{a}_s - \underline{1})^t \Omega (\underline{a}_s - \underline{1})$$

minimiert durch eine so genannte beste lineare unverzerrte (BLU) Prognose. Allgemein wird man nicht davon ausgehen, dass das Superpopulationsmodell $M(\Omega)$ in Anwendungssituationen exakt erfüllt ist, was die Verwendung eines Auswahlplans motiviert: Durch die zufällige Auswahl der Stichprobe wird eine gewisse Robustheit sichergestellt (vgl. Stenger [26]).

Die zu verwendende Strategie $(p, \{\underline{a}_s\})$ sollte in einer Umgebung von Bild(X) zuverlässig sein. Wir nehmen im Folgenden an, dass $\underline{\epsilon}$ normalverteilt ist mit regulärer Kovarianzmatrix Ω. Dann ist die Dichtefunktion $\phi(.|\underline{\beta})$ von $\underline{Y}$ (bei Vorgabe von $\underline{\beta}$) gegeben durch

$$\phi(\underline{y}|\underline{\beta}) = ((2\pi)^N \det(\Omega))^{-\frac{1}{2}} \exp[-\frac{(\underline{y} - X\underline{\beta})^t \Omega^{-1} (\underline{y} - X\underline{\beta})}{2}],$$

und die Wahrscheinlichkeit, dass $\underline{Y}$ einen Vektor aus der Menge

$$\{\underline{y} \in \mathbb{R}^N; \max_{\underline{\beta}} \phi(\underline{y}|\underline{\beta}) \geq \tilde{c}^2\}$$

realisiert, ist sehr groß, falls $\tilde{c} \in \mathbb{R}_+$ nicht zu groß ist. Obige Menge entspricht

$$\{\underline{y} \in \mathbb{R}^N; \min_{\underline{\beta}} (\underline{y} - X\underline{\beta})^t \Omega^{-1} (\underline{y} - X\underline{\beta}) \leq c^2\}$$

für ein entsprechend gewähltes $c \in \mathbb{R}$, und ein natürlicher Abstandbegriff ist definiert durch

$$\min_{\underline{\beta} \in \mathbb{R}^K} (\underline{y} - X\underline{\beta})^t \Omega^{-1} (\underline{y} - X\underline{\beta}) = \underline{y}^t (\Omega^{-1} - \Omega^{-1} X (X^t \Omega^{-1} X)^{-1} X^t \Omega^{-1}) \underline{y},$$

da das Minimum angenommen wird in

$$\underline{\beta}_0 := (X^t \Omega^{-1} X)^{-1} X^t \Omega^{-1} \underline{y}.$$

Dabei ist

$$U := \Omega^{-1} - \Omega^{-1} X (X^t \Omega^{-1} X)^{-1} X^t \Omega^{-1}$$

eine positiv semidefinite $N \times N$-Matrix mit Rang $N - K$ und $UX = 0$.

Beispiel 1.4 (Der HH-Raum) *Sei $\underline{x} \in \mathbb{R}^N$ ein Regressionsvektor und $\Omega :=$ diag($\underline{x}$) die Diagonalmatrix, deren Einträge die Komponenten von $\underline{x}$ sind. Dann ist*

$$U = diag^{-1}(\underline{x}) - \underline{1}\,\underline{1}^t$$

und

$$\Theta_U = \{\underline{y} \in \mathbb{R}^N; \sum_{i=1}^N x_i (\frac{y_i}{x_i} - y)^2 \leq 1\} = \{\underline{y} \in \mathbb{R}^N; var_{\tilde{p}_{HH}} t_{HH}(\,.\,, \underline{y}) \leq \frac{1}{n}\}.$$

In diesem Fall bezeichnet man Θ_U als HH-Raum. Viele Beispiele der vorliegenden Arbeit werden sich auf diesen Parameterraum beziehen. Man verwendet den HH-Raum auch, um einen Effizienzvergleich zwischen der HH-Strategie und der HT-Strategie anzustellen (siehe z.B. Gabler [6], S. 62).

1.4 Minimax-Strategien

Sei Θ ein Parameterraum und D eine Menge von Strategien des Statistikers. Die Frage, welche Strategie (aus D) der Statistiker zur Schätzung verwenden sollte, hängt von den Optimalitätskriterien ab, die man zugrunde legt. Godambe [11] zeigte, dass für fast alle Auswahlverfahren p kein linearer, p-unverzerrter Schätzer existiert, der die Varianz bzgl. aller $\underline{y} \in \mathbb{R}^N$ minimiert. Wir werden also im Allgemeinen keine Strategie $(p^*, t^*) \in D$ finden mit

$$\mathrm{MSE}(\underline{y}, (p^*, t^*)) \leq \mathrm{MSE}(\underline{y}, (p, t)) \quad \forall \underline{y} \in \Theta\, \forall (p, t) \in D.$$

Ein gängiges Optimalitätskriterium ist die *Zulässigkeit*: Eine Strategie (p^*, t^*) *dominiert* eine Strategie (p, t), falls

$$\mathrm{MSE}(\underline{y}, (p^*, t^*)) \leq \mathrm{MSE}(\underline{y}, (p, t)) \quad \forall \underline{y} \in \Theta.$$

Ist für mindestens ein $\underline{y} \in \Theta$ die strenge Ungleichheit erfüllt, so *dominiert* (p^*, t^*) die Strategie (p, t) *strikt*. Eine Strategie aus D heißt *zulässig*, wenn sie von keiner Strategie aus D strikt dominiert wird.

In dieser Arbeit verwenden wir als Optimalitätskriterium das *Minimax-Prinzip*: Eine Strategie $(p^*, t^*) \in D$ heißt *Minimax-Strategie* oder auch *Sicherheitsstrategie* (bzgl. (Θ, D)), falls sie das maximale Risiko auf Θ minimiert bzgl. aller Strategien aus D:

$$\sup_{\underline{y} \in \Theta} \mathrm{MSE}(\underline{y}, (p^*, t^*)) = \min_{(p,t) \in D} \sup_{\underline{y} \in \Theta} \mathrm{MSE}(\underline{y}, (p, t)).$$

Damit das Minimax-Problem sinnvoll gestellt ist, fordert man, dass das *Minimax-Risiko*

$$\sup_{\underline{y} \in \Theta} \mathrm{MSE}(\underline{y}, (p^*, t^*))$$

endlich ist.

Minimax-Strategien sind robuste Lösungen, die uns gegen den "unerwünschten" Parameter aus Θ absichern; sie hängen stark von der Wahl des Parameterraums Θ und der Menge D von Strategien des Statistikers ab. Wir beschäftigen uns in dieser Arbeit mit der Menge der repräsentativen Strategien vom Umfang n und wählen für Θ die in Abschnitt 1.1 eingeführten Parameterräume Θ_U oder Θ_Q.

Im Fall $\Theta = \Theta_U$ ist das maximale Risiko einer linearen Strategie $(p, \{\underline{a}_s\})$

$$\sup_{\underline{y} \in \Theta_U} \mathrm{MSE}(\underline{y}, (p, \{\underline{a}_s\}))$$

genau dann endlich, wenn die Strategie repräsentativ ist. Somit ist eine Minimax-Strategie bzgl. der Menge der repräsentativen Strategien vom Umfang n auch eine Minimax-Strategie bzgl. der Menge der linearen Strategien vom Umfang n. Dies gilt nicht für den Parameterraum Θ_Q (siehe z.B. Gabler [6], Bsp. 1 auf S. 27). Hier beschränken wir uns wegen Bemerkung 1.2 auf repräsentative Strategien.

Unter gewissen Symmetriebedingungen des Parameterraums Θ lässt sich die Menge der potentiellen Minimax-Strategien einschränken (vgl. Stenger [20]): Sei Γ eine Permutation von $\mathcal{U} = \{1, \ldots, N\}$. Für $s \subseteq \mathcal{U}$ definieren wir

$$\Gamma s := \{\Gamma i; i \in s\},$$

und für $\underline{y} \in \mathbb{R}^N$ definieren wir $\Gamma \underline{y} \in \mathbb{R}^N$ durch

$$(\Gamma \underline{y})_{\Gamma i} := y_i.$$

Sei G eine Untergruppe der Permutationsgruppe bzgl. $\mathcal{U}$ mit

$$\Gamma\Theta = \Theta,$$

wobei

$$\Gamma\Theta = \{\Gamma\underline{y}; \underline{y} \in \Theta\}.$$

Für eine Strategie (p, t) und eine Permutation $\Gamma \in G$ definieren wir

$$p_\Gamma(\Gamma s) := p(s),$$

$$t_\Gamma(\Gamma s, \Gamma \underline{y}) := t(s, \underline{y})$$

und

$$\overline{p}(s) := \frac{1}{|G|} \sum_{\Gamma \in G} p_\Gamma(s),$$

$$\overline{t}(s, \underline{y}) := \begin{cases} \frac{1}{|G|\overline{p}(s)} \sum_{\Gamma \in G} t_\Gamma(s, \underline{y}) p_\Gamma(s) & : \quad \overline{p}(s) > 0, \\ 0 & : \quad \text{sonst} \end{cases}.$$

Lemma 1.5 *Sei (p, t) eine Strategie des Statistikers. Dann ist*

$$\max_{\underline{y} \in \Theta} \mathrm{MSE}(\underline{y}, (\overline{p}, \overline{t})) \leq \max_{\underline{y} \in \Theta} \mathrm{MSE}(\underline{y}, (p, t)).$$

Den Beweis findet man bei Gabler [6], S. 15, oder Stenger [21].

Beispiel 1.6 (Die Standardstrategie als Minimax-Strategie) *Sei* $\underline{x} = \frac{1}{N}\underline{1}$ *und* $U := diag^{-1}(\underline{x}) - \underline{1}\,\underline{1}^t = NI - \underline{1}\,\underline{1}^t$, *wobei* I *die* $N \times N$*-Einheitsmatrix bezeichnet. Dann ist*

$$\Theta_U = \{\underline{y} \in \mathbb{R}^N; s_{yy} \leq \frac{1}{N(N-1)}\}.$$

Aufgrund der symmetrischen Bedingungen ist die Standardstrategie (p_0, t_0) *eine Minimax-Strategie (bzgl. aller linearen Strategien), und es gilt*

$$MSE(\underline{y}, (p_0, t_0)) = \frac{(N-n)N}{n} s_{yy} \leq \frac{N-n}{n(N-1)} \quad \forall \underline{y} \in \Theta_U.$$

1.5 Ein lineares Spiel

Der mittlere quadratische Verlust von $\underline{y} \in \mathbb{R}^N$ und einer linearen Strategie $(p, \{\underline{a}_s\})$ vom Umfang n ist gegeben durch

$$\begin{aligned} \mathrm{MSE}(\underline{y}, (p, \{\underline{a}_s\})) &= \sum_s p(s)(\underline{y}^t \underline{a}_s - y)^2 = \sum_s p(s)(\underline{y}^t(\underline{a}_s - \underline{1}))^2 \\ &= \sum_s p(s)(\underline{a}_s - \underline{1})^t \underline{y}\,\underline{y}^t(\underline{a}_s - \underline{1}) \\ &= \mathrm{tr}[\underline{y}\,\underline{y}^t \sum_s p(s)(\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t], \end{aligned}$$

wobei tr den Spuroperator bezeichnet. Offensichtlich sind

$$\underline{y}\,\underline{y}^t$$

und

$$\sum_s p(s)(\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t$$

symmetrische positiv semidefinite $N \times N$-Matrizen. Falls die Schätzvektoren repräsentativ sind, heißt $\sum_s p(s)(\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t$ *risikoerzeugende Matrix*, und wir bezeichnen mit

$$\mathcal{A}_X$$

die Menge aller risikoerzeugenden Matrizen. (Wir verschärfen die von Cheng und Li [5] eingeführte Definition einer risikoerzeugenden Matrix durch die Forderung der Repräsentativität.) Für einen Parameterraum Θ sei

$$\mathcal{Y}$$

die konvexe Hülle aller $\underline{y}\,\underline{y}^t$ mit $\underline{y} \in \Theta$. Falls $\Theta = \Theta_U$ bzw. $\Theta = \Theta_Q$, schreiben wir $\mathcal{Y}_U$ bzw. $\mathcal{Y}_Q$. Bezeichnen wir mit

$$\mathcal{Q}$$

den Vektorraum aller symmetrischen $N \times N$-Matrizen und definieren

$$\mathcal{Q}_0 := \{Q \in \mathcal{Q}; Q \text{ ist positiv semidefinit}\},$$

so gilt

$$\mathcal{Y},\ \mathcal{A}_X \subseteq \mathcal{Q}_0 \subseteq \mathcal{Q}.$$

Durch

$$\langle Q_1, Q_2 \rangle := \operatorname{tr}(Q_1 Q_2) \quad \forall Q_1, Q_2 \in \mathcal{Q}$$

ist ein Skalarprodukt auf $\mathcal{Q}$ definiert, und $(\mathcal{Q}, \langle\,.\,,\,.\,\rangle)$ ist ein Hilbertraum. Wir bezeichnen mit $|\,.\,|_{tr}$ die durch $\langle\,.\,,\,.\,\rangle$ induzierte Norm, d.h.

$$|Q|_{tr} := \sqrt{\langle Q, Q \rangle} \quad \forall Q \in \mathcal{Q}.$$

Das Auswertungsproblem lässt sich als lineares Spiel auffassen: $\mathcal{Y}$ ist der Strategienraum der Natur und $\mathcal{A}_X$ der Strategienraum des Statistikers. Die Natur wählt eine Matrix $V \in \mathcal{Y}$ und der Statistiker eine Matrix $W \in \mathcal{A}_X$. Der Verlust des Statistikers (bzw. der Gewinn der Natur) ist gegeben durch

$$\operatorname{tr}(VW) = \langle V, W \rangle.$$

Eine Strategie W^* des Statistikers heißt (in Analogie zu Abschnitt 1.4) *Minimax-Strategie* (oder auch *Sicherheitsstrategie*), falls

$$\infty > \sup_V \langle V, W^* \rangle = \min_W \sup_V \langle V, W \rangle.$$

Entsprechend heißt eine Strategie V^* der Natur *Maximin-Strategie* (oder auch *unerwünschte Strategie* bzw. *Sicherheitsstrategie* der Natur), falls

$$-\infty < \inf_W \langle V^*, W \rangle = \max_V \inf_W \langle V, W \rangle.$$

Wenn eine Maximin-Strategie V^* und eine Minimax-Strategie W^* existieren, heißt das Spiel $(\mathcal{Y}, \mathcal{A}_X)$ *lösbar*, und wir bezeichnen mit

$$\langle V^*, W^* \rangle$$

den *Wert des Spiels*. Das Tupel (V^*, W^*) nennen wir ein *Gleichgewicht*.

Beispiel 1.7 *Sei* $\underline{x} = \frac{1}{N}\underline{1}$, $U := diag^{-1}(\underline{x}) - \underline{1}\,\underline{1}^t = NI - \underline{1}\,\underline{1}^t$ *und* $\Theta = \Theta_U$. *Dann ist die durch die Standardstrategie definierte risikoerzeugende Matrix*

$$W^* = \frac{N-n}{(N-1)n}U$$

eine Minimax-Strategie (vgl. Beispiel 1.6). Aufgrund der symmetrischen Bedingungen ist die Maximin-Strategie V^* *gegeben durch*

$$V^* = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{N-1}\underline{e}_i\underline{e}_i^t = \frac{1}{N(N-1)}I,$$

und der Wert des Spiels beträgt $\frac{N-n}{(N-1)n}$. *Offensichtlich hängt die Maximin-Strategie nicht vom Stichprobenumfang* n *ab, was sich allerdings nicht verallgemeinern lässt (vgl. Beispiel 4.8 in Abschnitt 4.6).*

Eine Strategie $V^* \in \mathcal{Y}$ heißt *beste Antwort* auf $W \in \mathcal{A}_X$, falls

$$V^* \in \text{argmax}_{V\in\mathcal{Y}}\langle V, W\rangle.$$

Die Menge aller besten Antworten auf W bezeichnen wir mit $\mathcal{BA}(W)$. Analog heißt eine Strategie $W^* \in \mathcal{A}_X$ *beste Antwort* auf $V \in \mathcal{Y}$, falls

$$W^* \in \text{argmin}_{W\in\mathcal{A}_X}\langle V, W\rangle,$$

und $\mathcal{BA}(V)$ bezeichnet die Menge aller besten Antworten auf V. Offenbar ist ein Tupel (V^*, W^*) genau dann ein Gleichgewicht des Spiels $(\mathcal{Y}, \mathcal{A}_X)$, wenn

$$V^* \in \mathcal{BA}(W^*)$$

und

$$W^* \in \mathcal{BA}(V^*).$$

Ein Gleichgewicht (V^*, W^*) ist also ein *Nash-Gleichgewicht* im Sinne der Spieltheorie (vgl. z.B. Osborne und Rubinstein [17], Prop. 22.2 auf S. 22). Falls $W^* = \sum_s p^*(s)(\underline{a}_s^* - \underline{1})(\underline{a}_s^* - \underline{1})^t$ eine Minimax-Strategie ist, ist $(p^*, \{\underline{a}_s^*\})$ eine Minimax-Strategie im Sinne von Abschnitt 1.4, und es gilt

$$(\underline{a}_s^* - \underline{1})(\underline{a}_s^* - \underline{1})^t \in \mathcal{BA}(V^*) \quad \forall s \in S(p^*). \tag{1.1}$$

Speziell sind die Schätzvektoren der Minimax-Strategie BLU-Schätzer im Superpopulationsmodell $M(V^*)$.

Wir bezeichnen in dieser Arbeit sowohl eine risikoerzeugende Matrix als Strategie als auch ein Tupel (p, t), was nicht zu Missverständnissen führen sollte. Hauptsächlich sind wir daran interessiert, die Minimax-Strategie (p^*, t^*) zu bestimmen.

Dass sich eine Minimax-Strategie $W^* \in \mathcal{A}_X$ i.A. nicht eindeutig in einen Auswahlplan p^* und einen Schätzer t^* zerlegen lässt, zeigt das folgende Beispiel:

Beispiel 1.8 *Sei* $n = 2$, $N = 4$, $\underline{x} = \frac{1}{4}\underline{1} \in \mathbb{R}^4$, $U = diag^{-1}(\underline{x}) - \underline{1}\,\underline{1}^t$ *und* $\Theta = \Theta_U$. *Dann ist die Standardstrategie mit risikoerzeugender Matrix*

$$W^* = \frac{1}{3}U$$

eine Minimax-Strategie (vgl. Beispiel 1.7). Definieren wir

$$p^*(s) := \begin{cases} \frac{1}{3} & : \quad 4 \in s \\ 0 & : \quad sonst \end{cases}$$

und

$$\underline{a}^*_{\{i,4\}} := 2(\underline{e}_i + \underline{e}_4) \quad \forall i = 1, 2, 3,$$

so gilt

$$\sum_s p^*(s)(\underline{a}^*_s - \underline{1})(\underline{a}^*_s - \underline{1})^t = \frac{1}{3}U = W^*,$$

d.h. $(p^*, \{\underline{a}^*_s\})$ *ist ebenfalls eine Minimax-Strategie und besitzt dieselbe risikoerzeugende Matrix wie die Standardstrategie.*

1.6 Einige bekannte Resultate

Die folgenden Aussagen beziehen sich auf lineare Strategien vom Umfang n. In Beispiel 1.6 haben wir gesehen, dass die Standardstrategie für den Parameterraum

$$\{\underline{y} \in \mathbb{R}^N; \sum (y_i - \overline{y})^2 \leq 1\} \tag{1.2}$$

eine Minimax-Strategie ist. Dass sie die Minimax-Eigenschaft für größere Mengen von Strategien beibehält, lässt sich z.B. bei Bickel und Lehmann [2] oder Gabler [6], S. 121, nachlesen.

Stenger und Gabler [25] diskutierten den allgemeineren Parameterraum

$$\{\underline{y} \in \mathbb{R}^N; \sum \sum d_{ij}(y_i - \overline{y})(y_j - \overline{y}) \leq 1\}, \tag{1.3}$$

wobei (d_{ij}) eine positiv definite $N \times N$ Matrix bezeichnet. Auf die Ergebnisse werden wir in Abschnitt 3.3 eingehen.

$$\{\underline{y} \in \mathbb{R}^N; \sum (y_i - \frac{y}{x} x_i)^2 \leq 1\} \tag{1.4}$$

ist ein Beispiel für einen Parameterraum, der von einem Vektor $\underline{x} \in \mathbb{R}^N$ abhängt. Für diese Menge zeigte Stenger [24], dass die Verhältnisstrategie in einem asymptotischen Sinn minimax ist.

Minimax-Strategien für den Parameterraum

$$\{\underline{y} \in \mathbb{R}^N; \sum \sum d_{ij}(y_i - \frac{y}{x} x_i)(y_j - \frac{y}{x} x_j) \leq 1\} \tag{1.5}$$

wurden hergeleitet von Gabler und Stenger [9] unter den Voraussetzungen:

$$\underline{x} \text{ nahe bei } \underline{1}$$

und

$$(d_{ij}) \text{ in der Nähe der Einheitsmatrix.}$$

Die obigen Parameterräume sind alle von der Gestalt

$$\{\underline{y} \in \mathbb{R}^N; \underline{y}^t U \underline{y} \leq 1\},$$

wobei U eine positiv semidefinite Matrix bezeichnet mit $U\underline{1} = \underline{0}$ im Fall (1.2), (1.3) und $U\underline{x} = \underline{0}$ im Fall (1.4), (1.5). Allgemein wurde für eine Regressionsmatrix X und eine symmetrische positiv semidefinite $N \times N$-Matrix U mit $UX = 0$ für den Parameterraum Θ_U von Stenger, Gabler und Schmidt [27] die Existenz von Minimax-Strategien geklärt. Auf den zugrunde liegenden Beweis werden wir in Kapitel 3 eingehen.

Kapitel 2

Die Strategien-Räume der Spieler

Der erste Abschnitt dieses Kapitels enthält die – in Hinblick auf unsere Anwendungen – wichtigsten Aussagen über den Vektorraum $\mathcal{Q}$ der symmetrischen $N \times N$-Matrizen. Wir diskutieren in den Abschnitten 2.3 und 2.4 die Strategien-Räume des Statistikers und der Natur als Teilmengen von $\mathcal{Q}_0$. Speziell werden wir sehen, dass zu einer beliebigen Strategie des Statistikers immer eine zulässige Strategie existiert, die diese dominiert. Dabei sei $N > K \geq 1$, $U \in \mathcal{Q}_0$ mit Rang $N - K$ und X eine $N \times K$-Matrix mit vollem Rang (eine Regressionsmatrix), so dass

$$UX = 0.$$

2.1 Die Löwner-Halbordnung

Wie bereits bemerkt, ist $(\mathcal{Q}, \langle\, .\, ,\, .\, \rangle)$ ein Hilbertraum, und es gilt (vgl. Anhang, Lemma A.1)

$$\langle P, Q \rangle \geq 0 \quad \forall P, Q \in \mathcal{Q}_0. \tag{2.1}$$

Die Teilmenge $\mathcal{Q}_0$ ist abgeschlossen (vgl. Anhang, Satz A.6), und durch sie ist die so genannte Löwner-Halbordnung $\succeq$ auf $\mathcal{Q}$ definiert (siehe z.B. Marshall und Olkin [16], S. 462): Für $P, Q \in \mathcal{Q}$ mit $P - Q \in \mathcal{Q}_0$ schreiben wir

$$P \succeq Q.$$

Offensichtlich gilt:

1. $P \succeq Q$ ist äquivalent zu

$$\langle \underline{y}\, \underline{y}^t, P \rangle \geq \langle \underline{y}\, \underline{y}^t, Q \rangle \quad \forall \underline{y} \in \mathbb{R}^N.$$

2. Die Löwner-Halbordnung ist antisymmetrisch, d.h. aus $P \succeq Q$ und $Q \succeq P$ folgt $P = Q$.

3. $\mathcal{Q}_0 = \{Q \in \mathcal{Q}; Q \succeq 0\}$.

Die folgenden zwei Lemmata sind in Stenger, Gabler und Schmidt [27] zu finden.

Lemma 2.1 *Für $Q \in \mathcal{Q}_0$ wächst $\langle Q, P\rangle$ monoton in $P \in \mathcal{Q}$ bezüglich $\succeq$.*

Beweis: Mit $P_2 \succeq P_1$ gilt wegen (2.1)

$$\langle Q, P_2\rangle - \langle Q, P_1\rangle = \langle Q, \underbrace{P_2 - P_1}_{\in \mathcal{Q}_0}\rangle \geq 0.$$

$\diamond$

Lemma 2.2 *Sei $P \succeq Q \succeq 0$. Dann ist*

$$\langle P, P\rangle \geq \langle Q, Q\rangle,$$

d.h. die durch $\langle\, .\, , .\, \rangle$ induzierte Norm ist monoton auf $\mathcal{Q}_0$ bzgl. $\succeq$.

Beweis: Wegen (2.1) gilt

$$\langle P, P\rangle - \langle Q, Q\rangle = \langle P - Q, P\rangle + \langle P - Q, Q\rangle \geq 0.$$

$\diamond$

2.2 Abgeschlossenheit von unten

Dieser technische Abschnitt dient als Vorbereitung für die Diskussion der Zulässigkeit von Strategien des Statistikers. Speziell benötigen wir den Begriff des unteren Grenzpunktes einer Menge $\mathcal{S} \subseteq \mathcal{Q}$. Im nächsten Abschnitt wird sich zeigen, dass die unteren Grenzpunkte von $\mathcal{A}_X$ die dominanten Strategien des Statistikers sind (vgl. Satz 2.16). Für ein $P \in \mathcal{Q}$ definieren wir

$$\mathcal{Q}_P := \{Q \in \mathcal{Q}; P \succeq Q\}$$

und bezeichnen den Abschluss einer beliebigen Teilmenge $\mathcal{S}$ von $\mathcal{Q}$ mit $\overline{\mathcal{S}}$.

Definition 2.3 *$P \in \mathcal{Q}$ heißt* unterer Grenzpunkt *einer Menge $\mathcal{S} \subseteq \mathcal{Q}$, falls*

$$\mathcal{Q}_P \cap \overline{\mathcal{S}} = \{P\}.$$

Die Menge aller unteren Grenzpunkte von $\mathcal{S}$ bezeichnen wir mit $\Lambda(\mathcal{S})$.

Definition 2.4 *Eine Menge $\mathcal{S} \subseteq \mathcal{Q}$ heißt* abgeschlossen von unten*, falls*

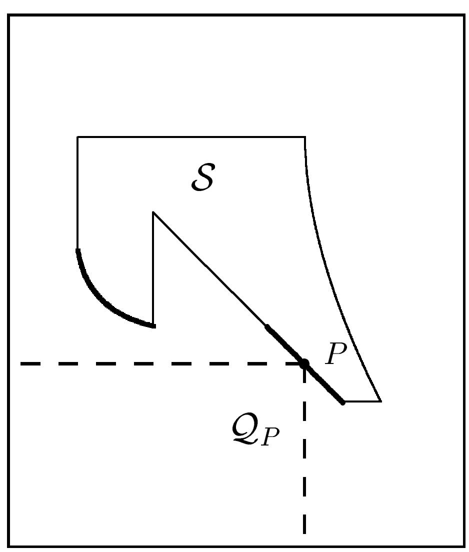
$$\Lambda(\mathcal{S}) \subseteq \mathcal{S}.$$

Bemerkung 2.5 (Interpretation von Skizzen) *Um eine anschauliche Vorstellung der abstrakten Sachverhalte zu bekommen, arbeiten wir mit Skizzen. Dabei fassen wir eine symmetrische Matrix als einen Vektor im $\mathbb{R}^2$ auf und interpretieren wegen*

$$\langle P, Q \rangle = \sum_{i,j} p_{ij} q_{ij} \quad \text{für } P, Q \in \mathcal{Q}$$

den Hilbertraum $(\mathcal{Q}, \langle\,.\,,\,.\,\rangle)$ (bzw. $(\mathcal{Q}_X, \langle\,.\,,\,.\,\rangle)$ mit $\mathcal{Q}_X = \{Q \in \mathcal{Q}; QX = 0\}$) als den zweidimensionalen Euklidischen Raum $(\mathbb{R}^2, (\,.\,,\,.\,)_2)$. Obwohl aus $Q \succeq P$ im Allgemeinen nicht $q_{ij} \geq p_{ij}$ für alle $i, j = 1, \ldots, N$ folgt, interpretieren wir $Q \succeq P$ in den Skizzen als "Q ist komponentenweise größer oder gleich P". In diesem Sinn entspricht die Menge der positiv semidefiniten Matrizen aus $\mathcal{Q}$ (bzw. $\mathcal{Q}_X$) dem ersten Quadrant der Euklidischen Ebene. Unsere Ergebnisse rechtfertigen diese Darstellung.

Skizze 2.6 (Abgeschlossen von unten) *Sei $\mathcal{S}$ eine beliebige Teilmenge von $\mathcal{Q}$. Im Sinne von Bemerkung 2.5 entspricht P einem unteren Grenzpunkt von $\mathcal{S}$ und die Menge der unteren Grenzpunkte $\Lambda(\mathcal{S})$ von $\mathcal{S}$ dem fetten Rand. Gehören diese Punkte zu $\mathcal{S}$, so ist $\mathcal{S}$ abgeschlossen von unten.*



Lemma 2.7 *Sei $\mathcal{S} \subseteq \mathcal{Q}_0$. Dann ist $\Lambda(\mathcal{S}) \neq \emptyset$.*

Beweis: Sei $|\,.\,|_{tr}$ die durch $\langle\,.\,,\,.\,\rangle$ induzierte Norm,

$$b_0 := \inf_{S\in\mathcal{S}} |S|_{tr}$$

und (S_n) eine Folge in $\mathcal{S}$ mit

$$\lim_{n\to\infty} |S_n|_{tr} = b_0.$$

Nach Bolzano-Weierstrass existiert eine konvergente Teilfolge $(S_n)_{n\in\mathbb{N}'}$ mit

$$\lim_{n\in\mathbb{N}'} S_n = S_0 \in \overline{\mathcal{S}}.$$

Offensichtlich ist

$$|S_0|_{tr} = b_0$$

und

$$S_0 \in \mathcal{Q}_{S_0} \cap \overline{\mathcal{S}}.$$

Sei $Q \in \mathcal{Q}_{S_0} \cap \overline{\mathcal{S}}$. Aufgrund der Abgeschlossenheit von $\mathcal{Q}_0$ gilt

$$S_0 \succeq Q \succeq 0.$$

Wegen der Monotonie der Norm auf $\mathcal{Q}_0$ (Lemma 2.2) ist $b_0 = |S_0|_{tr} \geq |Q|_{tr}$, und mit der Definition von b_0 folgt

$$|S_0|_{tr} = |Q|_{tr}.$$

Wegen (2.1) gilt

$$0 \leq \langle S_0 - Q, S_0 - Q\rangle = \langle S_0 - Q, S_0\rangle - \langle S_0 - Q, Q\rangle \leq \langle S_0 - Q, S_0\rangle + \langle S_0 - Q, Q\rangle = 0,$$

d.h. $Q = S_0$. $\diamond$

Lemma 2.8 *Sei $\mathcal{S} \subseteq \mathcal{Q}_0, \neq \emptyset$. Dann gilt*

$$\forall S \in \overline{\mathcal{S}}\ \exists S^* \in \Lambda(\mathcal{S}) :\ \ S \succeq S^*.$$

Beweis: Sei $S \in \overline{\mathcal{S}}$. Es ist

$$\mathcal{S}_1 := \mathcal{Q}_S \cap \overline{\mathcal{S}} \subseteq \mathcal{Q}_0,\ \mathcal{S}_1 \neq \emptyset,$$

und mit Lemma 2.7 folgt $\Lambda(\mathcal{S}_1) \neq \emptyset$. Sei $S^* \in \Lambda(\mathcal{S}_1)$. Dann gilt $S \succeq S^*$ und

$$\{S^*\} = \mathcal{Q}_{S^*} \cap \overline{\mathcal{S}_1} = \mathcal{Q}_{S^*} \cap \overline{\mathcal{Q}_S \cap \overline{\mathcal{S}}} = \mathcal{Q}_{S^*} \cap \overline{\mathcal{S}}$$

(aufgrund der Transitivität von $\succeq$ ist $\mathcal{Q}_{S^*} \subseteq \mathcal{Q}_S$), d.h. $S^* \in \Lambda(\mathcal{S})$. $\diamond$

Satz 2.9 *$\mathcal{S} \subseteq \mathcal{Q}_0$ ist genau dann abgeschlossen von unten, wenn für jede konvergente Folge (S^n) in $\mathcal{S}$ ein $S^* \in \mathcal{S}$ existiert mit*

$$\lim_{n\to\infty} S^n \succeq S^*.$$

Beweis:

"$\Rightarrow$": Sei $\mathcal{S} \neq \emptyset$ abgeschlossen von unten und (S^n) eine Folge in $\mathcal{S}$ mit $S := \lim_{n\to\infty} S^n$ $(\in \overline{\mathcal{S}})$. Nach Lemma 2.8 existiert ein $S^* \in \Lambda(\mathcal{S})$ mit $S \succeq S^*$. Da $\mathcal{S}$ abgeschlossen von unten ist, folgt $S^* \in \mathcal{S}$.

"$\Leftarrow$": Sei $\mathcal{S} \neq \emptyset$ und $S \in \Lambda(\mathcal{S})$ ($\neq \emptyset$ nach Lemma 2.7). Da $S \in \overline{\mathcal{S}}$, existiert eine Folge (S^n) in $\mathcal{S}$ mit $\lim_{n\to\infty} S^n = S$. Nach Voraussetzung existiert ein $S^* \in \mathcal{S}$ mit $S \succeq S^*$; somit gilt

$$S^* \in \mathcal{Q}_S \cap \overline{\mathcal{S}} = \{S\},$$

d.h. $S = S^* \in \mathcal{S}$. $\diamond$

2.3 Der Strategien-Raum des Statistikers

Der *Aktionsraum* des Statistikers ist gegeben durch die Menge aller repräsentativen Schätzvektoren. In der Spieltheorie ist eine (gemischte) Strategie allgemein gegeben durch eine (diskrete) Verteilung auf dem Aktionsraum. Wie in den Abschnitten 1.4 und 1.5 beschrieben, verstehen wir unter einer Strategie des Statistikers eine risikoerzeugende Matrix, also im Sinne der Spieltheorie eine gemischte Strategie, die über einen Auswahlplan definiert ist. Dabei ist die Menge $\mathcal{A}_X$ der risikoerzeugenden Matrizen für $n > 1$ nicht konvex. Sei $\widetilde{\mathcal{A}}_X$ die konvexe Hülle von $\mathcal{A}_X$. Wir werden im Folgenden zeigen, dass die risikoerzeugenden Matrizen eine vollständige Klasse bilden (bzgl. des Spiels $(\mathcal{Y}, \widetilde{\mathcal{A}}_X)$), und zwar für einen beliebigen Parameterraum $\Theta \subseteq \mathbb{R}^N$:

$$\forall \tilde{W} \in \widetilde{\mathcal{A}}_X \, \exists W \in \mathcal{A}_X : \; \langle V, \tilde{W} \rangle \geq \langle V, W \rangle \quad \forall V \in \mathcal{Y},$$

d.h. jede Strategie aus $\widetilde{\mathcal{A}}_X$ wird von einer Strategie aus $\mathcal{A}_X$ dominiert. Als Konsequenz finden wir z.B. eine repräsentative Strategie vom Umfang n, die die RHC-Strategie dominiert (vgl. Gabler und Stenger [7]).

Die folgenden beiden Lemmata sind in Stenger, Gabler und Schmidt [27] zu finden.

Lemma 2.10 *Sei $\tilde{W} \in \widetilde{\mathcal{A}}_X$. Dann existiert ein $W \in \mathcal{A}_X$ mit*

$$\tilde{W} \succeq W.$$

Speziell gilt für einen beliebigen Parameterraum $\Theta \subseteq \mathbb{R}^N$

$$\langle V, \tilde{W} \rangle \geq \langle V, W \rangle \quad \forall V \in \mathcal{Y}.$$

Beweis: Es ist

$$\tilde{W} = \sum_k \lambda_k \sum_s p^k(s)(\underline{a}_s^k - \underline{1})(\underline{a}_s^k - \underline{1})^t$$

mit $\lambda_k \geq 0$, $\sum_k \lambda_k = 1$. Wir definieren

$$p(s) := \sum_k \lambda_k p^k(s)$$

und

$$\underline{a}_s := \sum_k \frac{\lambda_k p^k(s)}{p(s)} \underline{a}_s^k.$$

Dann gilt

$$W := \sum_s p(s)(\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t \in \mathcal{A}_X.$$

Da die Abbildung

$$\underline{y} \mapsto \underline{y}\,\underline{y}^t \quad \forall\, \underline{y} \in \mathbb{R}^N$$

konvex ist bzgl. $\succeq$ (siehe Lemma A.5 im Anhang), folgt

$$\tilde{W} \succeq W.$$

Der zweite Teil des Lemmas ergibt sich aus der Monotonie der Abbildung $\langle V, . \rangle$ bzgl. $\succeq$ für beliebiges $V \in \mathcal{Y} \subseteq \mathcal{Q}_0$ (vgl. Lemma 2.1). $\diamond$

Als nächstes werden wir beweisen, dass $\mathcal{A}_X$ abgeschlossen von unten ist.

Lemma 2.11 *Sei (W^k) eine Folge in $\mathcal{A}_X$ mit $\lim_{k\to\infty} W^k = W$. Dann existiert ein $W^* \in \mathcal{A}_X$ mit*

$$W \succeq W^*.$$

Beweis: Es ist

$$W^k = \sum_s p^k(s)(\underline{a}_s^k - \underline{1})(\underline{a}_s^k - \underline{1})^t \quad \forall\, k \in \mathbb{N}.$$

Offensichtlich existiert eine Teilmenge $\mathbb{N}' \subseteq \mathbb{N}$ mit

$$p(s) := \lim_{k \in \mathbb{N}'} p^k(s) \geq 0 \text{ und } \sum_s p(s) = 1.$$

Wir definieren

$$W^{k*} := \sum_{s \in S(p)} p^k(s)(\underline{a}_s^k - \underline{1})(\underline{a}_s^k - \underline{1})^t \quad \forall\, k \in \mathbb{N}.$$

Dann gilt

$$W^k - W^{k*} \in \mathcal{Q}_0 \quad \forall\, k \in \mathbb{N}.$$

Weiter ist

$$W^k \succeq p^k(s)(\underline{a}_s^k - \underline{1})(\underline{a}_s^k - \underline{1})^t \quad \forall\, k \in \mathbb{N} \forall s \in S(p),$$

also (aufgrund von Lemma 2.2)

$$|W^k|_{tr} \geq p^k(s)|(\underline{a}_s^k - \underline{1})(\underline{a}_s^k - \underline{1})^t|_{tr} \quad \forall\, k \in \mathbb{N} \forall s \in S(p),$$

und es existiert eine Teilmenge $\mathbb{N}'' \subseteq \mathbb{N}'$ mit

$$\underline{a}_s := \lim_{k \in \mathbb{N}''} \underline{a}_s^k \,(\in A_s) \quad \forall\, s \in S(p).$$

Mit

$$W^* := \sum_{s \in S(p)} p(s)(\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t \in \mathcal{A}_X$$

folgt aus der Abgeschlossenheit von $\mathcal{Q}_0$ (vgl. Satz A.6 im Anhang)

$$W - W^* = \lim_{k \in \mathbb{N}''} (W^k - W^{k*}) \in \mathcal{Q}_0.$$

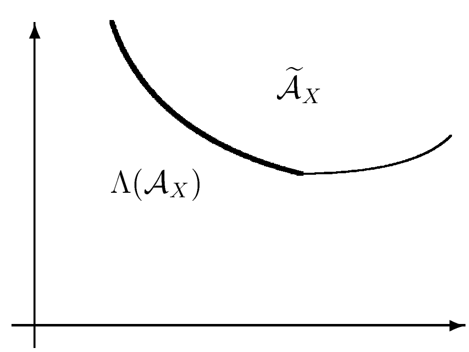
$\diamond$

Korollar 2.12 *$\mathcal{A}_X$ ist abgeschlossen von unten, d.h.*

$$\Lambda(\mathcal{A}_X) \subseteq \mathcal{A}_X.$$

Beweis: Satz 2.9 und Lemma 2.11. $\diamond$

Skizze 2.13 (Die Menge der risikoerzeugenden Matrizen) *Im Sinne von Bemerkung 2.5 stellen wir uns die konvexe Hülle $\widetilde{\mathcal{A}}_X$ des Strategien-Raums $\mathcal{A}_X$ des Statistikers als konvexe Menge im ersten Quadrant vor. Den fetten Rand der konvexen Hülle $\widetilde{\mathcal{A}}_X$ interpretieren wir aufgrund von Lemma 2.10 und Korollar 2.12 als $\Lambda(\mathcal{A}_X)$.*



Wir werden im Folgenden zeigen, dass die zulässigen Strategien in den Spielen $(\mathcal{Y}_U, \mathcal{A}_X)$ und $(\mathcal{Y}_Q, \mathcal{A}_x)$ mit den unteren Grenzpunkten von $\mathcal{A}_X$ übereinstimmen. Eine Strategie $W^* \in \mathcal{A}_X$ ist genau dann zulässig (in dem Spiel $(\mathcal{Y}, \mathcal{A}_X)$), falls für alle $W \in \mathcal{A}_X$ mit

$$\langle \underline{y}\,\underline{y}^t, W\rangle \leq \langle \underline{y}\,\underline{y}^t, W^*\rangle \quad \forall \underline{y} \in \Theta$$

folgt

$$\langle \underline{y}\,\underline{y}^t, W\rangle = \langle \underline{y}\,\underline{y}^t, W^*\rangle \quad \forall \underline{y} \in \Theta.$$

Lemma 2.14 *Sei $\Theta = \Theta_U$ oder $\Theta = \Theta_Q$, $W, W^* \in \mathcal{A}_X$. Folgende Aussagen sind äquivalent:*

(i)

$$W \succeq W^*.$$

(ii)

$$\langle \underline{y}\,\underline{y}^t, W\rangle \geq \langle \underline{y}\,\underline{y}^t, W^*\rangle \quad \forall \underline{y} \in \Theta.$$

Beweis: Dass die erste Aussage die zweite impliziert, ergibt sich aus der Monotonie der Abbildung $\langle Q, . \rangle$ bzgl. $\succeq$ für beliebiges $Q \in \mathcal{Q}_0$ (vgl. Lemma 2.1).

Zum Beweis der Rückrichtung unterscheiden wir zwei Fälle:

1) Sei $\Theta = \Theta_U$ und $\underline{z} \in \mathbb{R}^N \backslash \mathrm{Bild}(X)$. Da $\mathrm{Kern}(U) = \mathrm{Bild}(X)$ ist $c := \langle \underline{z}\,\underline{z}^t, U\rangle > 0$, also $\underline{y} := \frac{1}{\sqrt{c}}\underline{z} \in \Theta_U$ und somit

$$\langle \underline{z}\,\underline{z}^t, W\rangle = c\langle \underline{y}\,\underline{y}^t, W\rangle \geq c\langle \underline{y}\,\underline{y}^t, W^*\rangle = \langle \underline{z}\,\underline{z}^t, W^*\rangle.$$

2) Sei $\Theta = \Theta_Q$ und $\underline{z} \in \mathbb{R}^N$. Mit $\underline{x}_i := x_i \underline{e}_i \in \Theta_Q$ für $i = 1, \ldots, N$ ist $(\underline{x}_1, \ldots, \underline{x}_N)$ eine Orthogonal-Basis des $\mathbb{R}^N$. Es existieren $c_1, \ldots, c_N \in \mathbb{R}$ mit $\underline{z} = \sum_i c_i \underline{x}_i$, und es gilt

$$\langle \underline{z}\,\underline{z}^t, W\rangle = \sum_i c_i^2 \langle \underline{x}_i \underline{x}_i^t, W\rangle \geq \sum_i c_i^2 \langle \underline{x}_i \underline{x}_i^t, W^*\rangle = \langle \underline{z}\,\underline{z}^t, W^*\rangle.$$

◇

Korollar 2.15 *Eine Strategie $W^* \in \mathcal{A}_X$ ist genau dann zulässig in dem Spiel $(\mathcal{Y}_U, \mathcal{A}_X)$ bzw. $(\mathcal{Y}_Q, \mathcal{A}_X)$, wenn aus $W^* \succeq W$ für ein $W \in \mathcal{A}_X$ folgt:*

$$W^* = W.$$

Beweis: Lemma 2.14 und die Antisymmetrie der Löwner-Halbordnung. ◇

Satz 2.16 *In dem Spiel* $(\mathcal{Y}, \mathcal{A}_X)$*, wobei* $\mathcal{Y} = \mathcal{Y}_U$ *oder* $\mathcal{Y} = \mathcal{Y}_Q$*, ist die Menge der zulässigen Strategien gegeben durch* $\Lambda(\mathcal{A}_X)$*.*

Beweis: Sei $W^* \in \Lambda(\mathcal{A}_X)$ und $W \in \mathcal{A}_X$ mit

$$\langle \underline{y}\, \underline{y}^t, W\rangle \leq \langle \underline{y}\, \underline{y}^t, W^*\rangle \quad \forall \underline{y} \in \Theta,$$

also wegen Lemma 2.14

$$W \preceq W^*.$$

Aufgrund der Definition von $\Lambda(\mathcal{A}_X)$ ist

$$W = W^*.$$

Sei andererseits $W^* \in \mathcal{A}_X$ eine zulässige Strategie. Dann existiert wegen Lemma 2.8 ein $W \in \Lambda(\mathcal{A}_X)$ mit $W^* \succeq W$. Da $\mathcal{A}_X$ abgeschlossen von unten ist (Korollar 2.12), ist $W \in \mathcal{A}_X$, und mit Korollar 2.15 folgt

$$W^* = W \in \Lambda(\mathcal{A}_X).$$

$\diamond$

Korollar 2.17 *In dem Spiel* $(\mathcal{Y}, \mathcal{A}_X)$*, wobei* $\mathcal{Y} = \mathcal{Y}_U$ *oder* $\mathcal{Y} = \mathcal{Y}_Q$*, existiert für eine beliebige Strategie* $W \in \mathcal{A}_X$ *eine zulässige Strategie* $W^* \in \mathcal{A}_X$ *mit*

$$W \succeq W^*.$$

Beweis: Korollar 2.12 und Lemma 2.8. $\diamond$

Als Konsequenz finden wir – die Existenz einer Minimax-Strategie vorausgesetzt – immer eine zulässige Minimax-Strategie.

2.4 Der Strategien-Raum der Natur

Wir betrachten den Parameterraum Θ_U. Offensichtlich spielt die Natur nur Vektoren $\underline{y} \in \Theta_U$ mit $\underline{y}^t U \underline{y} = 1$, wir können also ohne Einschränkung

$$\Theta_U := \{\underline{y} \in \mathbb{R}^N; \underline{y}^t U \underline{y} = 1\}$$

setzen. Der Strategien-Raum der Natur ist von vergleichsweise einfacher Struktur, und stellt sich in $\mathcal{Q}_0$ wie folgt dar:

Satz 2.18 *Die konvexe Hülle* $\mathcal{Y}_U$ *von* $\{\underline{y}\,\underline{y}^t; \underline{y} \in \Theta_U\}$ *ist gegeben durch*

$$\{V \in \mathcal{Q}_0; \langle V, U\rangle = 1\}.$$

Beweis: Offensichtlich ist $\mathcal{Y}_U \subseteq \{V \in \mathcal{Q}_0; \langle V, U\rangle = 1\}$. Es verbleibt zu zeigen, dass $\{V \in \mathcal{Q}_0; \langle V, U\rangle = 1\}$ in der konvexen Hülle enthalten ist: Sei $V \in \{V \in \mathcal{Q}_0; \langle V, U\rangle = 1\}$. Es existiert eine Orthonormal-Basis von Eigenvektoren $(\underline{z}^1, \dots, \underline{z}^N)$ bezüglich der Eigenwerte $\lambda_1 \geq \dots \geq \lambda_N \geq 0$ von V. Weiter existiert ein $M \in \{1, \dots, N\}$ mit $\lambda_1 \geq \dots \geq \lambda_M > 0 = \lambda_{M+1} = \dots = \lambda_N$, und es gilt

$$V = \sum_{j=1}^{M} \lambda_j \underline{z}^j (\underline{z}^j)^t.$$

Da $\langle V, U\rangle = 1$, existiert ein $J \subseteq \{1, \dots, M\}$, $J \neq \emptyset$ mit

$$\langle \underline{z}^j (\underline{z}^j)^t, U\rangle > 0 \quad \forall j \in J$$

und

$$\langle \underline{z}^j (\underline{z}^j)^t, U\rangle = 0 \quad \forall j \in J^c := \{1, \dots, M\} \backslash J$$

Für $j \in J$ definieren wir

$$\underline{y}^j := \frac{1}{\sqrt{\langle \underline{z}^j (\underline{z}^j)^t, U\rangle}} \underline{z}^j \in \Theta_U$$

und

$$\mu_j := \lambda_j \langle \underline{z}^j (\underline{z}^j)^t, U\rangle > 0.$$

Es ist

$$\sum_{j \in J} \mu_j = \langle V, U\rangle = 1. \tag{2.2}$$

Falls $J = \{1, \dots, M\}$, ist $V = \sum_{j=1}^{M} \mu_j \underline{y}^j (\underline{y}^j)^t \in \mathcal{Y}_U$.

Sei $J \subsetneq \{1, \dots, M\}$ und $j^* \in J$. Definieren wir für $j \in J^c$

$$\underline{y}^{j1} := \underline{y}^{j^*} + \sqrt{\frac{|J^c|\lambda_j}{\mu_{j^*}}} \underline{z}^j \in \Theta_U, \ \underline{y}^{j2} := \underline{y}^{j^*} - \sqrt{\frac{|J^c|\lambda_j}{\mu_{j^*}}} \underline{z}^j \in \Theta_U,$$

so ist

$$\frac{\mu_{j^*}}{2|J^c|} (\underline{y}^{j1} (\underline{y}^{j1})^t + \underline{y}^{j2} (\underline{y}^{j2})^t) = \frac{\mu_{j^*}}{|J^c|} \underline{y}^{j^*} \underline{y}^{j^*} + \lambda_j \underline{z}^j (\underline{z}^j)^t,$$

und es gilt

$$\sum_{j\in J\setminus\{j^*\}} \mu_j \underline{y}^j(\underline{y}^j)^t + \sum_{j\in J^c} \frac{\mu_{j^*}}{2|J^c|}(\underline{y}^{j1}(\underline{y}^{j1})^t + \underline{y}^{j2}(\underline{y}^{j2})^t) =$$

$$= \sum_{j\in J\setminus\{j^*\}} \lambda_j \underline{z}^j(\underline{z}^j)^t + \sum_{j\in J^c}(\frac{\mu_{j^*}}{|J^c|}\underline{y}^{j^*}\underline{y}^{j^*} + \lambda_j \underline{z}^j(\underline{z}^j)^t) = \sum_{j=1}^{M} \lambda_j \underline{z}^j(\underline{z}^j)^t = V;$$

wegen (2.2) ist $V \in \mathcal{Y}_U$. ◇

Aufgrund der Repräsentativität der Schätzvektoren gilt für eine risikoerzeugende Matrix W offensichtlich

$$WX = 0.$$

In unserem Spiel ist somit für ein $\underline{y} \in \Theta$ nur der orthogonale Anteil bzgl. Bild(X) relevant, und wir können ohne Einschränkung die Projektion von Θ in das orthogonale Komplement von Bild(X) als Parameterraum betrachten. Die Projektion von Θ_U ist wegen $UX = 0$ gegeben durch

$$\Theta_{UX} := \{\underline{y} \in \Theta_U; X^t\underline{y} = \underline{0}\}.$$

Satz 2.19 *Die konvexe Hülle von* $\{\underline{y}\,\underline{y}^t; \underline{y} \in \Theta_{UX}\}$ *ist gegeben durch*

$$\mathcal{Y}_{UX} := \{V \in \mathcal{Q}_0; VX = 0, \langle V, U\rangle = 1\}.$$

Beweis: Offensichtlich ist die konvexe Hülle von $\{\underline{y}\,\underline{y}^t; \underline{y} \in \Theta_{UX}\}$ eine Teilmenge von $\mathcal{Y}_{UX}$. Es verbleibt zu zeigen, dass $\mathcal{Y}_{UX}$ in der konvexen Hülle enthalten ist: Sei $V \in \mathcal{Y}_{UX} \subseteq \mathcal{Y}_U$. Wegen Satz 2.18 ist

$$V = \sum_j \mu_j \underline{y}^j(\underline{y}^j)^t$$

für gewisse $\underline{y}^j \in \Theta_U$ und $\mu_j > 0$ mit $\sum_j \mu_j = 1$. Da

$$0 = X^tVX = \sum_j \mu_j X^t\underline{y}^j(\underline{y}^j)^tX = \sum_j \mu_j X^t\underline{y}^j(X^t\underline{y}^j)^t,$$

ist $X^t\underline{y}^j = 0$ für $j = 1, \ldots, M$, also

$$\underline{y}^j \in \Theta_{UX} \quad \forall j = 1, \ldots, M.$$

◇

Falls X einem Regressionsvektor $\underline{x} \in \mathbb{R}^N$ entspricht, ist eine weitere Darstellung des Parameterraums nützlich: Definieren wir

$$\Theta_{U1} := \{\underline{y} \in \Theta_U; (\underline{y}, \underline{1})_2 = 0\},$$

so ist mit $\underline{y} \in \Theta_U$

$$\underline{y} - y\underline{x} \in \Theta_{U1},$$

und es gilt aufgrund der Repräsentativität

$$\langle (\underline{y} - y\underline{x})(\underline{y} - y\underline{x})^t, W\rangle = \langle \underline{y}\,\underline{y}^t, W\rangle \quad \forall W \in \mathcal{A}_X.$$

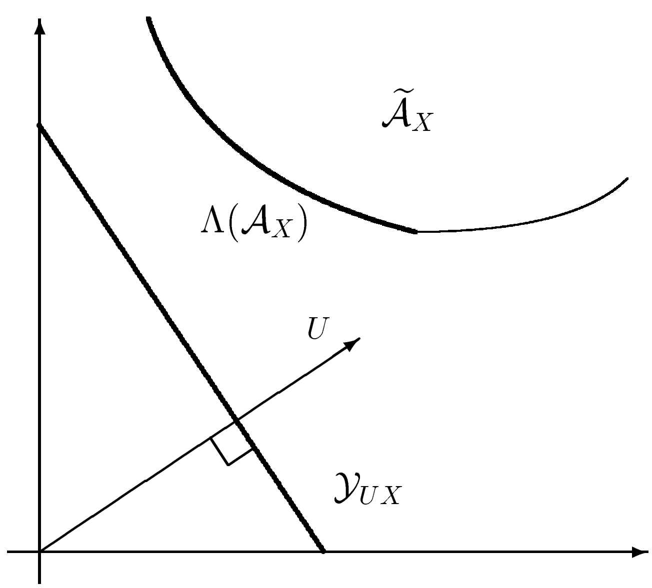
Wir können also ohne Einschränkung den Parameterraum Θ_{U1} betrachten.

Satz 2.20 *Die konvexe Hülle von* $\{\underline{y}\,\underline{y}^t; \underline{y} \in \Theta_{U1}\}$ *ist gegeben durch*

$$\mathcal{Y}_{U1} := \{V \in \mathcal{Q}_0; V\underline{1} = \underline{0}, \langle V, U\rangle = 1\}.$$

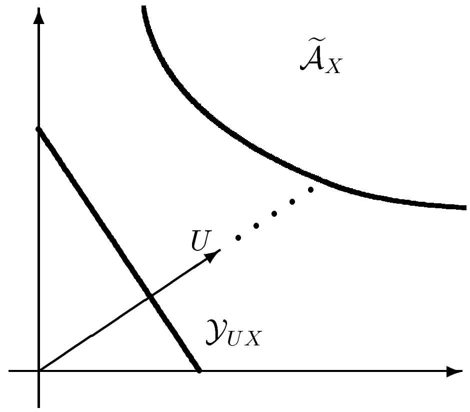
Beweis: Der Beweis verläuft analog zum Beweis von Satz 2.19 mit $X = \underline{1}$. ◇

Skizze 2.21 (Die Strategien-Räume der Spieler) *Aufgrund von Satz 2.19 entspricht der Strategien-Raum* $\mathcal{Y}_{UX}$ *der Natur im Sinne von Bemerkung 2.5 einer zu* U *orthogonalen Geraden im ersten Quadrant. Die konvexe Hülle des Strategien-Raums* $\mathcal{A}_X$ *des Statistikers stellen wir uns als konvexe Menge im ersten Quadrant vor, wobei wir den fetten Rand von* $\widetilde{\mathcal{A}}_X$ *als* $\Lambda(\mathcal{A}_X)$ *interpretieren (vgl. Skizze 2.13).*

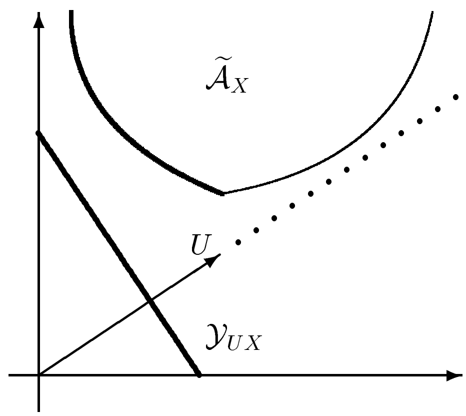


Interessant ist die Lage von U *bzgl.* $\widetilde{\mathcal{A}}_X$ *bzw.* $\Lambda(\mathcal{A}_X)$*. In obiger Skizze schneidet die durch* cU, $c \in \mathbb{R}$, *definierte Gerade die Menge* $\widetilde{\mathcal{A}}_X$, *aber nicht den Rand* $\Lambda(\mathcal{A}_X)$*. Es sind auch folgende Situationen denkbar:*

- *Die Gerade* cU*,* $c \in \mathbb{R}$*, schneidet den Rand* $\Lambda(\mathcal{A}_X)$*.*



- *Die Gerade* cU*,* $c \in \mathbb{R}$*, schneidet* $\widetilde{\mathcal{A}}_X$ *nicht.*



Wir werden im Verlauf dieser Arbeit sehen, dass alle drei Möglichkeiten eintreten können. Welche Bedeutung sie für die Lösung des Spiels haben, diskutieren wir an entsprechender Stelle.

Der Parameterraum Θ_Q wird in Abschnitt 7.1 behandelt.

Kapitel 3

Das Ellipsoid als Parameterraum

Sei X eine Regressionsmatrix mit Rang $K < N$ und U eine symmetrische positiv semidefinite $N \times N$-Matrix mit Rang $N - K$ und $UX = 0$. In diesem Kapitel werden wir den Parameterraum Θ_U betrachten. Die allgemeine Existenz von Minimax-Strategien bzw. die Lösbarkeit des linearen Spiels $(\mathcal{Y}_U, \mathcal{A}_X)$ beweisen wir in Abschnitt 3.1 (vgl. Stenger, Gabler und Schmidt [27]). In den Abschnitten 3.2 bis 3.4 widmen wir uns den so genannten Indifferenzmatrizen, die aussichtsreiche Kandidaten für Minimax bzw. Maximin-Strategien darstellen. Konkrete Berechnungen von Minimax-Strategien für Stichproben vom Umfang Eins finden wir in Abschnitt 3.5.

3.1 Der Existenzsatz

Satz 3.1 (Existenzsatz 1) *Das lineare Spiel $(\mathcal{Y}_U, \mathcal{A}_X)$ ist lösbar. Speziell existiert eine zulässige Minimax-Strategie.*

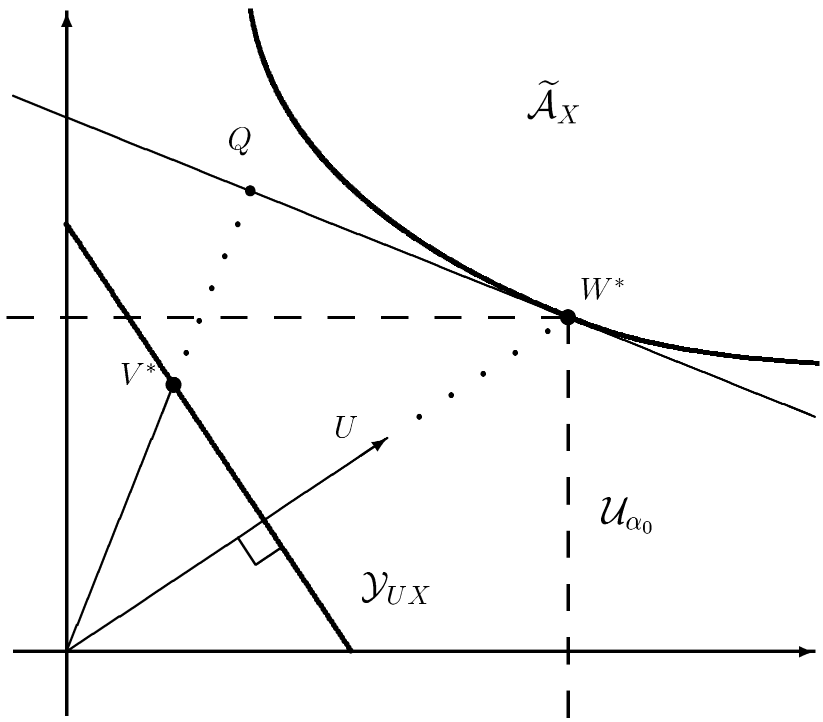
Skizze 3.2 (Beweis-Skizze) *Unter den Vorbehalten von Bemerkung 2.5 geben wir eine Beweis-Skizze von Satz 3.1.*

1. *Konstruiere den Kegel $\mathcal{U}_{\alpha_0} = \{Q \in \mathcal{Q}_X; \alpha_0 U \succeq Q\}$, so dass int$(\mathcal{U}_{\alpha_0}) \cap \widetilde{\mathcal{A}}_X = \emptyset$.*

2. *Nach dem Satz der trennenden Hyperebene existiert eine Matrix $Q \in \mathcal{Q}_X$, die eine Tangential-Hyperebene von $\widetilde{\mathcal{A}}_X$ definiert.*

3. *Aus Konstruktionsgründen ist Q positiv semidefinit, und $V^* := \frac{Q}{\langle Q,U \rangle}$ ist eine Maximin-Strategie.*

4. *Eine Minimax-Strategie ist gegeben durch* $W^* \in \mathcal{U}_{\alpha_0} \cap \mathcal{A}_X$.



Beweis von Satz 3.1: Wir betrachten o.B.d.A. das Spiel $(\mathcal{Y}_{UX}, \mathcal{A}_X)$. Sei

$$\mathcal{Q}_X = \{Q \in \mathcal{Q}; QX = 0\}.$$

Mit

$$\mathcal{U}_\alpha := \{Q \in \mathcal{Q}_X; \alpha U \succeq Q\}$$

für $\alpha \in \mathbb{R}$ ist wegen Lemma A.3 (vgl. Anhang)

$$\{\alpha; \mathcal{U}_\alpha \cap \widetilde{\mathcal{A}}_X \neq \emptyset\} \neq \emptyset.$$

Da $\mathcal{U}_0 \cap \widetilde{\mathcal{A}}_X = \emptyset$, folgt

$$\alpha_0 := \inf\{\alpha; \mathcal{U}_\alpha \cap \widetilde{\mathcal{A}}_X \neq \emptyset\} \geq 0.$$

Weiter ist

$$\mathrm{int}(\mathcal{U}_{\alpha_0}) = \bigcup_{\beta < \alpha_0} \mathcal{U}_\beta$$

(Lemma A.9 im Anhang) konvex und $\mathrm{int}(\mathcal{U}_{\alpha_0}) \cap \widetilde{\mathcal{A}}_X = \emptyset$. Nach dem Satz der trennenden Hyperebene (siehe z.B. Luenberger [15], S. 133) bzgl. des Hilbertraums

$(\mathcal{Q}_X, \langle\,.\,,\,.\,\rangle)$ existiert ein $Q \in \mathcal{Q}_X \backslash \{0\}$ und ein $k \in \mathbb{R}$ mit

$$\begin{aligned} \langle Q, R\rangle &\leq k \quad \forall R \in \mathcal{U}_{\alpha_0}, && (3.1)\\ \langle Q, \tilde{W}\rangle &\geq k \quad \forall \tilde{W} \in \widetilde{\mathcal{A}}_X, && (3.2)\\ k &= \sup_{R\in\mathcal{U}_{\alpha_0}} \langle Q, R\rangle. \end{aligned}$$

Angenommen, $Q \notin \mathcal{Q}_{0X} := \{Q \in \mathcal{Q}_0; QX = 0\}$. Dann existiert ein Eigenvektor $\underline{z} \in \mathbb{R}^N$ von Q mit $X^t\underline{z} = \underline{0}$ und $\underline{z}^t Q \underline{z} < 0$, also

$$\langle Q, \underline{z}\,\underline{z}^t\rangle < 0.$$

Da für alle $n \in \mathbb{N}$

$$-n\underline{z}\,\underline{z}^t \in \mathcal{U}_{\alpha_0}$$

und

$$\lim_{n\to\infty} \langle Q, -n\underline{z}\,\underline{z}^t\rangle = \infty,$$

steht obige Annahme im Widerspruch zur Trennungseigenschaft (3.1), und es folgt

$$Q \in \mathcal{Q}_{0X} \backslash \{0\}.$$

Aufgrund von Korollar A.2 ist

$$\langle Q, U\rangle > 0,$$

also

$$V^* := \frac{Q}{\langle Q, U\rangle} \in \mathcal{Y}_{UX}.$$

Mit der Monotonie von $\langle Q,\,.\,\rangle$ (vgl. Proposition 2.1) erhalten wir

$$k = \sup_{R\in\mathcal{U}_{\alpha_0}} \langle Q, R\rangle = \alpha_0 \langle Q, U\rangle$$

und somit

$$\alpha_0 = \sup_{R\in\mathcal{U}_{\alpha_0}} \frac{\langle Q, R\rangle}{\langle Q, U\rangle} = \sup_{R\in\mathcal{U}_{\alpha_0}} \langle V^*, R\rangle.$$

Aus (3.2) folgt

$$\langle V^*, W\rangle \geq \alpha_0 \quad \forall W \in \mathcal{A}_X. \tag{3.3}$$

Sei (α_k) eine streng monoton fallende Folge mit $\lim_{k\to\infty}\alpha_k = \alpha_0$ und $\tilde{W}^k \in \mathcal{U}_{\alpha_k} \cap \widetilde{\mathcal{A}}_X$. Wegen Lemma 2.10 existieren $W^k \in \mathcal{A}_X$ mit

$$\tilde{W}^k \succeq W^k \quad \forall k \in \mathbb{N}. \tag{3.4}$$

Da

$$\alpha_1 U \succeq W^k \succeq 0 \quad \forall k \in \mathbb{N},$$

ist die Folge (W^k) aufgrund von Lemma 2.2 beschränkt. Nach Bolzano-Weierstrass existiert eine konvergente Teilfolge mit

$$W := \lim_{k\in\mathbb{N}'} W^k.$$

Da $\mathcal{Q}_{0X}$ abgeschlossen ist (vgl. Korollar A.7 im Anhang), gilt mit (3.4)

$$\alpha_0 U - W = \lim_{k\in\mathbb{N}'}(\alpha_k U - W^k) \in \mathcal{Q}_{0X},$$

also

$$\alpha_0 U \succeq W.$$

Wegen Lemma 2.11 existiert ein $W^* \in \mathcal{A}_X$ mit

$$W \succeq W^*$$

und somit

$$\alpha_0 U \succeq W^*.$$

Es folgt

$$\langle V, W^*\rangle \leq \alpha_0 \quad \forall V \in \mathcal{Y}_{UX}$$

aufgrund der Monotonie von $\langle V, .\rangle$ bzgl. $\succeq$ (vgl. Proposition 2.1). Insgesamt erhalten wir mit (3.3)

$$\langle V, W^*\rangle \leq \langle V^*, W^*\rangle \leq \langle V^*, W\rangle \quad \forall V \in \mathcal{Y}_{UX} \forall W \in \mathcal{A}_X,$$

wobei

$$\langle V^*, W^*\rangle = \alpha_0.$$

Die Maximin-Strategie ist gegeben durch V^*, W^* ist Minimax-Strategie und $\langle V^*, W^*\rangle$ der Wert des Spiels. Dass eine zulässige Minimax-Strategie existiert, ergibt sich aus Korollar 2.17. ◇

Korollar 3.3 *Sei D die Menge aller linearen Strategien vom Umfang n. Dann existiert eine zulässige repräsentative Minimax-Strategie bzgl. (Θ_U, D).*

3.2 Indifferenzmatrizen

Sei $\mu \in \mathbb{R}$ und $V^* \in \mathcal{Y}_U$ mit

$$\min_{\underline{a}_s \in A_s} \langle V^*, (\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t \rangle = \mu \quad \forall s \in S(n). \tag{3.5}$$

Spielt die Natur die Matrix V^*, so ist der Verlust unabhängig vom Auswahlplan des Statistikers, solange er nur über die bezüglich V^* optimalen Schätzvektoren $\{\underline{a}_s(V^*) \in \operatorname{argmin}_{\underline{a}_s \in A_s} \langle V^*, (\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t \rangle ; s \in S(n)\}$ mischt. Offensichtlich gilt

$$\langle V^*, W \rangle \geq \mu \quad \forall W \in \mathcal{A}_X.$$

Besitzt der Statistiker wiederum eine Strategie $W^* \in \mathcal{A}_X$ mit

$$\langle V, W^* \rangle = \mu \quad \forall V \in \mathcal{Y}_U, \tag{3.6}$$

so folgt

$$\langle V, W^* \rangle \leq \mu \leq \langle V^*, W \rangle \quad \forall V \in \mathcal{Y}_U \, \forall W \in \mathcal{A}_X,$$

d.h. V^* ist Maximin-Strategie, W^* ist Minimax-Strategie, und der Wert des Spiels ist gegeben durch $\mu = \langle V^*, W^* \rangle$. Offensichtlich gilt

$$W^* = \sum_s p^*(s)(\underline{a}_s(V^*) - \underline{1})(\underline{a}_s(V^*) - \underline{1})^t$$

für einen Auswahlplan p^* und $\underline{a}_s(V^*) \in \operatorname{argmin}_{\underline{a}_s \in A_s} \langle V^*, (\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t \rangle$.

Eine Matrix $V^* \in \mathcal{Y}_U$ mit (3.5) für ein $\mu \in \mathbb{R}$ nennen wir *Indifferenzmatrix* bzw. *Indifferenz-Strategie der Natur*. Falls eine Indifferenz-Strategie existiert, ist sie nicht zwangsläufig eine Maximin-Strategie. Dafür finden wir in Abschnitt 4.6 des folgenden Kapitels Beispiele.

Lemma 3.4 *Eine Strategie $V^* \in \mathcal{Y}_U$ der Natur ist genau dann eine Indifferenzmatrix, wenn für alle Stichproben $s \in S(n)$ gilt*

$$\exists \underline{a}_s^* \in A_s : (\underline{a}_s^* - \underline{1})(\underline{a}_s^* - \underline{1})^t \in \mathcal{BA}(V^*).$$

Korollar 3.5 *Sei $(p^*, \{\underline{a}_s^*\})$ eine Minimax-Strategie mit*

$$S(p^*) = S(n)$$

und V^* eine Maximin-Strategie. Dann ist V^* eine Indifferenzmatrix.

Beweis: Da $p^*(s) > 0$ für alle $s \in S(n)$, gilt (vgl. (1.1))

$$(\underline{a}_s^* - \underline{1})(\underline{a}_s^* - \underline{1})^t \in \mathcal{BA}(V^*) \quad \forall s \in S(n),$$

d.h. V^* ist eine Indifferenzmatrix (vgl. Lemma 3.4). $\diamond$

Eine Matrix $W^* \in \mathcal{A}_X$ mit (3.6) für ein $\mu \in \mathbb{R}$ nennen wir *Indifferenzmatrix* bzw. *Indifferenz-Strategie des Statistikers.*

Lemma 3.6 *Eine Strategie $W^* \in \mathcal{A}_X$ des Statistikers ist genau dann eine Indifferenzmatrix, wenn ein $\mu \in \mathbb{R}$ existiert mit*

$$W^* = \mu U.$$

Beweis: Sei $W^* = \mu U$ für ein $\mu \in \mathbb{R}$. Dann ist

$$\langle V, W^* \rangle = \mu \langle V, U \rangle = \mu \quad \forall V \in \mathcal{Y}_U.$$

Sei andererseits $W^* \in \mathcal{A}_X$ eine Indifferenzmatrix, d.h. es existiert ein $\mu \in \mathbb{R}$ mit

$$\langle V, W^* \rangle = \mu = \langle V, \mu U \rangle \quad \forall V \in \mathcal{Y}_U.$$

Da $\mathcal{Y}_U$ ein Erzeugendensystem von $\mathcal{Q}$ ist, folgt

$$W^* = \mu U.$$

$\diamond$

Indifferenzmatrizen des Statistikers sind also immer gegeben durch ein Vielfaches von U. Wir werden in Abschnitt 3.4 sehen, dass der Statistiker nicht immer ein Vielfaches von U spielen kann.

Satz 3.7 *Sei $V^* \in \mathcal{Y}_U$ eine Indifferenzmatrix der Natur und p^* ein Auswahlplan mit*

$$W^* := \sum_s p^*(s)(\underline{a}_s(V^*) - \underline{1})(\underline{a}_s(V^*) - \underline{1})^t = \lambda U$$

für

$$\underline{a}_s(V^*) \in argmin_{\underline{a}_s \in A_s} \langle V^*, (\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t \rangle.$$

Dann gilt: V^* ist Maximin-Strategie, W^* bzw. $(p^*, \{\underline{a}_s^*(V^*)\})$ ist Minimax-Strategie, und der Wert des Spiels ist gegeben durch

$$\lambda = \langle V^*, (\underline{a}_s^*(V^*) - \underline{1})(\underline{a}_s^*(V^*) - \underline{1})^t \rangle \quad \textit{für } s \in S(n).$$

Insbesondere ist λ eindeutig festgelegt.

Beweis: Aufgrund von Lemma 3.6 ist W^* eine Indifferenzmatrix des Statistikers. Da V^* eine Indifferenzmatrix der Natur ist, existiert ein $\mu \in \mathbb{R}$ mit

$$\langle V^*, (\underline{a}_s^*(V^*) - \underline{1})(\underline{a}_s^*(V^*) - \underline{1})^t \rangle = \mu \quad \forall s \in S(n),$$

und es folgt

$$\lambda = \langle V^*, W^* \rangle = \sum_s p^*(s) \langle V^*, (\underline{a}_s^*(V^*) - \underline{1})(\underline{a}_s^*(V^*) - \underline{1})^t \rangle = \mu.$$

$\diamond$

Offensichtlich erhalten wir mit Hilfe einer Indifferenzmatrix V^* der Natur eine untere Schranke für den Spielwert λ:

$$\lambda \geq \langle V^*, (\underline{a}_s(V^*) - \underline{1})(\underline{a}_s(V^*) - \underline{1})^t \rangle \tag{3.7}$$

für ein $s \in S(n)$. Durch eine Indifferenzmatrix W^* des Statistikers ist eine obere Schranke gegeben:

$$\lambda \leq \langle V, W^* \rangle$$

für ein $V \in \mathcal{Y}_U$.

Beispiel 3.8 *Sei $\underline{x}$ ein Regressionsvektor und $U = diag^{-1}(\underline{x}) - \underline{1}\,\underline{1}^t$. Aufgrund von Lemma 2.10 existiert eine risikoerzeugende Matrix W, die die RHC-Strategie dominiert, und es gilt*

$$\langle V, W \rangle \leq \frac{1}{n} \frac{N-n}{N-1} \quad \forall V \in \mathcal{Y}_U.$$

Eine obere Schranke für den Spielwert ist also gegeben durch $\frac{1}{n}\frac{N-n}{N-1}$.

3.3 Allgemeine Indifferenzmatrizen

Sei X eine Regressionsmatrix. Wir nennen eine Matrix $Q \in \mathcal{Q}_0$ eine *allgemeine Indifferenzmatrix* (bzgl. X), falls

$$\min_{\underline{a}_s \in A_s} (\underline{a}_s - \underline{1})^t Q (\underline{a}_s - \underline{1}) = 1 \quad \forall s \in S(n).$$

Die folgende Bemerkung zeigt, dass die Existenz von Indifferenz-Strategien der Natur unabhängig von der Matrix U ist.

Bemerkung 3.9 *Sei $U \in \mathcal{Q}_0$ mit Kern(U) = Bild(X) und $Q \in \mathcal{Q}_0$ eine allgemeine Indifferenzmatrix. Es ist $\mathcal{A}_X \subseteq$ Kern(X^t) = Bild(U) (vgl. Lemma A.4), also Bild(U) $\nsubseteq$ Kern(Q), und mit Lemma A.1 folgt $\langle Q, U\rangle > 0$. Aufgrund von Satz 2.18 ist*

$$V := \frac{Q}{\langle Q, U\rangle} \in \mathcal{Y}_U$$

eine Indifferenzmatrix der Natur bzgl. des Spiels $(\mathcal{Y}_U, \mathcal{A}_X)$, und für den Spielwert λ gilt (vgl. (3.7))

$$\lambda \geq \frac{1}{\langle Q, U\rangle}.$$

Beispiel 3.10 *Sei $\underline{x} = \frac{1}{N}\underline{1} \in \mathbb{R}^N$ und $U \in \mathcal{Q}_0$ mit Kern(U) = span[$\underline{1}$]. Wir betrachten das Spiel $(\mathcal{Y}_U, \mathcal{A}_x)$. Wie man leicht sieht, ist die Matrix*

$$Q := \frac{n}{N(N-n)} I$$

eine allgemeine Indifferenzmatrix, also

$$V^* := \frac{Q}{\langle Q, U\rangle} = \frac{1}{\sum_i u_{ii}} I \in \mathcal{Y}_U \tag{3.8}$$

eine Indifferenzmatrix der Natur, und für den Wert λ des Spiels gilt

$$\lambda \geq \frac{1}{\langle Q, U\rangle} = \frac{N(N-n)}{n} \frac{1}{\sum_i u_{ii}} =: \mu.$$

Mit

$$a^*_{si} := \frac{N}{n} \quad \forall i \in s$$

für $s \in S(n)$ ist

$$\underline{a}^*_s = argmin_{\underline{a}_s \in A_s} (\underline{a}_s - \underline{1})^t V^* (\underline{a}_s - \underline{1}).$$

Falls ein Auswahlplan p^* existiert mit

$$\sum_s p^*(s)(\underline{a}^*_s - \underline{1})(\underline{a}^*_s - \underline{1})^t = \mu U, \tag{3.9}$$

ist $(p^*, \{\underline{a}^*_s\})$ eine Minimax-Strategie, V^* Maximin-Strategie und der Wert des Spiels ist μ (vgl. Satz 3.7).

Für einen beliebigen Auswahlplan p auf $S(n)$ gilt

$$(\sum_s p(s)(\underline{a}^*_s - \underline{1})(\underline{a}^*_s - \underline{1})^t)_{ij} = \frac{N^2}{n^2}\pi_{ij}(p) - \frac{N}{n}(\pi_{ii}(p) + \pi_{jj}(p)) + 1 \tag{3.10}$$

für alle $i, j = 1, \dots, N$. Falls $N \neq 2n$, erfüllt ein Auswahlplan p genau dann (3.9), wenn seine Inklusionswahrscheinlichkeiten $\pi_{ij}(p)$ gegeben sind durch

$$\varphi_{ij}(p) := \frac{n^2}{N^2}[\mu u_{ij} + \frac{n\mu(u_{ii}+u_{jj}) - N}{N-2n}]$$

für $i, j = 1, \dots, N$. Liegt die Matrix U in einer Umgebung von $I - \frac{1}{N}\underline{1}\,\underline{1}^t$, so ist durch

$$p^*(s) := \frac{1}{\binom{N-4}{n-2}} \sum_{\substack{i,j \in s \\ i<j}} \varphi_{ij} - \frac{n-2}{\binom{N-3}{n-1}} \sum_{i \in s} \varphi_{ii} + \frac{\binom{n-1}{2}}{\binom{N-2}{n}} \quad \forall s \in S(n)$$

ein Auswahlplan definiert mit $\pi_{ij}(p^*) = \varphi_{ij}$ für alle $i, j = 1, \dots, N$ (vgl. Chaudhuri [4]), und $(p^*, \{\underline{a}_s^*\})$ ist eine Minimax-Strategie. (Vgl. Stenger und Gabler [25].)

Sei $Q \in \mathcal{Q}_0$ eine allgemeine Indifferenzmatrix,

$$\underline{a}_s^* \in \mathrm{argmin}_{\underline{a}_s \in A_s} (\underline{a}_s - \underline{1})^t Q (\underline{a}_s - \underline{1})$$

für $s \in S(n)$, p^* ein beliebiger Auswahlplan auf $S(n)$ und

$$\tilde{U} := \sum_s p^*(s)(\underline{a}_s^* - \underline{1})(\underline{a}_s^* - \underline{1})^t.$$

Dabei gilt i.A. nicht $\mathrm{Kern}(\tilde{U}) = \mathrm{Bild}(X)$, sondern nur

$$\mathrm{Kern}(\tilde{U}) \supseteq \mathrm{Bild}(X).$$

Trotzdem ist mit $\mathcal{Y}_{\tilde{U}} = \{V \in \mathcal{Q}_0; \langle V, \tilde{U} \rangle = 1\}$ durch $(\mathcal{Y}_{\tilde{U}}, \mathcal{A}_X)$ in jedem Fall ein lineares Spiel definiert, das eine besonders einfache Lösung besitzt: Es ist $Q \in \mathcal{Y}_{\tilde{U}}$, und für $V \in \mathcal{Y}_{\tilde{U}}$ gilt

$$\begin{aligned} \sum_s p^*(s)(\underline{a}_s^* - \underline{1})^t V (\underline{a}_s^* - \underline{1}) &= \langle V, \tilde{U} \rangle = \\ &= 1 = \langle Q, \tilde{U} \rangle = \sum_s p^*(s)(\underline{a}_s^* - \underline{1})^t Q (\underline{a}_s^* - \underline{1}) \\ &= \sum_s p(s)(\underline{a}_s^* - \underline{1})^t Q (\underline{a}_s^* - \underline{1}) \leq \sum_s p(s)(\underline{a}_s - \underline{1})^t Q (\underline{a}_s - \underline{1}) \end{aligned}$$

für alle repräsentativen Strategien $(p, \{\underline{a}_s\})$, d.h. Q ist eine Maximin-Strategie, $(p^*, \{\underline{a}_s^*\})$ bzw. $W^* := \tilde{U}$ ist Minimax-Strategie, und der Wert des Spiels ist 1.

Beispiel 3.11 (Lahiri-Midzuno-Sen's Auswahlplan) *Sei $\underline{\alpha} \in \mathbb{R}^N$ mit $\alpha_i \geq 0$ ein Wahrscheinlichkeitsvektor. Lahiri-Midzuno-Sen's Auswahlplan p_{LMS} ist wie* *folgt durch $\underline{\alpha}$ definiert: In der ersten Runde wird die i-te Einheit mit Wahrscheinlichkeit α_i ausgewählt; danach werden $n-1$ Einheiten aus den verbleibenden $N-1$ Einheiten ohne Zurücklegen gezogen.*

Die Inklusionswahrscheinlichkeiten sind für $N > 2$ gegeben durch

$$\pi_{ii}(p_{LMS}) = \frac{N-n}{N-1}\alpha_i + \frac{n-1}{N-1} \quad \text{für } i = 1, \dots, N \tag{3.11}$$

und

$$\pi_{ij}(p_{LMS}) = \frac{n-1}{N-1}(\frac{N-n}{N-2}(\alpha_i + \alpha_j) + \frac{n-2}{N-2}) \tag{3.12}$$

für $i, j = 1, \dots, N$ mit $i \neq j$. Definieren wir

$$a^*_{si} := \frac{N}{n} \quad \forall i \in s$$

für $s \in S(n)$ und

$$\widetilde{U} := \sum_s p_{LMS}(s)(\underline{a}^*_s - \underline{1})(\underline{a}^*_s - \underline{1})^t,$$

so ist aufgrund von (3.10)

$$\tilde{u}_{ij} = \frac{N^2}{n^2}\pi_{ij}(p_{LMS}) - \frac{N}{n}(\pi_{ii}(p_{LMS}) + \pi_{jj}(p_{LMS})) + 1,$$

also

$$\tilde{u}_{ii} = \frac{N-n}{n^2(N-1)}[N(N-2n)\alpha_i + n(N+1) - N] \quad \text{für } i = 1, \dots, N$$

und

$$\tilde{u}_{ij} = -\frac{N-n}{n^2(N-1)}\frac{N(N-2n)(\alpha_i + \alpha_j) + Nn - 2(N-n)}{N-2} \quad \text{für } i \neq j.$$

Mit

$$\beta_i := \frac{(N-2n)\alpha_i + n - 1}{(N-2)n} \quad \forall i = 1, \dots, N$$

ist

$$U_\beta := diag(\underline{\beta}) + \frac{1}{N^2}\underline{1}\,\underline{1}^t - \frac{1}{N}(\underline{\beta}\underline{1}^t + \underline{1}\underline{\beta}^t) = \frac{n(N-1)}{N^2(N-n)}\widetilde{U},$$

und es gilt (da $\sum_i \beta_i = \sum_i \alpha_i = 1$)

$$\underline{y}^t U_\beta \underline{y} = \sum_i \beta_i(y_i - \overline{y})^2 \quad \forall \underline{y} \in \mathbb{R}^N,$$

also

$$\Theta_{U_\beta} = \{\underline{y} \in \mathbb{R}^N; \sum_i \beta_i(y_i - \overline{y})^2 = 1\}.$$

Wir betrachten das Spiel $(\mathcal{Y}_{U_\beta}, \mathcal{A}_x)$ *für* $\underline{x} = \frac{1}{N}\underline{1} \in \mathbb{R}^N$. *Eine Maximin-Strategie ist gegeben durch (vgl. 3.8)*

$$V^* := \frac{1}{tr(U_\beta)} I = \frac{N}{N-1} I,$$

$(p_{LMS}, \{\underline{a}_s^*\})$ *ist Minimax-Strategie, und der Wert des Spiels beträgt* $\frac{N^2(N-n)}{n(N-1)}$ *(vgl. Stenger und Gabler [25]).*

Für $N = 2n$ *ist* $\underline{\beta} = \frac{1}{N}\underline{1}$, *und zwar unabhängig von dem Wahrscheinlichkeitsvektor* $\underline{\alpha} \in \mathbb{R}^N$. *In diesem Fall ist*

$$\Theta_{U_\beta} = \{\underline{y} \in \mathbb{R}^N; \frac{1}{N} \sum_i (y_i - \overline{y})^2 = 1\},$$

und jede durch einen beliebigen Wahrscheinlichkeitsvektor aus $\mathbb{R}^N$ *definierte Strategie* $(p_{LMS}, \{\underline{a}_s^*\})$ *ist eine Minimax-Strategie mit risikoerzeugender Matrix*

$$W^* = \frac{1}{N-1}(NI - \underline{1}\,\underline{1}^t).$$

Zusammenfassung Sei X eine Regressionsmatrix und $Q \in \mathcal{Q}_0$ eine allgemeine Indifferenzmatrix. Dann ist für eine beliebige Matrix $U \in \mathcal{Q}_0$ mit $\mathrm{Kern}(U) = \mathrm{Bild}(X)$

$$\frac{Q}{\langle Q, U\rangle} \in \mathcal{Y}_U$$

eine Indifferenzmatrix der Natur bzgl. des Spiels $(\mathcal{Y}_U, \mathcal{A}_X)$ (vgl. Bemerkung 3.9).

Weiter sei

$$\underline{a}_s^* \in \mathrm{argmin}_{\underline{a}_s \in A_s} (\underline{a}_s - \underline{1})^t Q (\underline{a}_s - \underline{1}).$$

Dann existiert zu einem beliebigen Auswahlplan p (vom Umfang n) eine Matrix $\widetilde{U}$, so dass $(p, \{\underline{a}_s^*\})$ eine Minimax-Strategie bzgl. des Spiels $(\mathcal{Y}_{\widetilde{U}}, \mathcal{A}_X)$ ist. Speziell kann

$$\widetilde{U} = \sum_s p(s)(\underline{a}_s^* - \underline{1})(\underline{a}_s^* - \underline{1})^t$$

gewählt werden.

3.4 Indifferenzmatrizen des Statistikers

Wir beschränken uns in diesem Abschnitt auf die Betrachtung eines Regressionsvektors $\underline{x}$ und suchen eine Bedingung für die Existenz von Indifferenz-Strategien des Statistikers.

Der Vektorraum $\mathcal{Q}$ aller symmetrischen $N \times N$-Matrizen besitzt die Dimension $\frac{N(N+1)}{2}$, und

$$\mathcal{Q}_1 := \{Q \in \mathcal{Q}; Q\underline{1} = \underline{0}\}$$

ist ein $\frac{N(N-1)}{2}$-dimensionaler Untervektorraum von $\mathcal{Q}$. Definieren wir

$$Q_{lm} := (\underline{e}_l - \underline{e}_m)(\underline{e}_l - \underline{e}_m)^t \quad \text{für } 1 \leq l < m \leq N,$$

wobei $\underline{e}_l \in \mathbb{R}^N$ den l-ten Einheitsvektor bezeichnet, so ist

$$Q_{lm} \in \mathcal{Q}_1 \cap \mathcal{Q}_0 \quad \forall 1 \leq l < m \leq N.$$

Offensichtlich ist die Familie $(Q_{lm}; l < m)$ linear unabhängig mit

$$|\{Q_{lm}; l < m\}| = \frac{N(N-1)}{2},$$

d.h. $(Q_{lm}; l < m)$ ist eine Basis von $\mathcal{Q}_1$. Da $\text{Kern}(U) = \text{span}[\underline{x}]$, gilt

$$V_{lm} := \frac{Q_{lm}}{\langle U, Q_{lm} \rangle} \in \mathcal{Y}_{U1},$$

und $(V_{lm}; l < m)$ ist ebenfalls eine Basis von $\mathcal{Q}_1$. Für $V \in \mathcal{Y}_{U1}$ folgt

$$\exists \gamma_{lm} \in \mathbb{R} : \; V = \sum_{l<m} \gamma_{lm} V_{lm}$$

und

$$\sum_{l<m} \gamma_{lm} = \langle V, U \rangle = 1.$$

Sei W eine risikoerzeugende Matrix mit

$$\langle V_{lm}, W \rangle = \mu \quad \forall l < m$$

für ein $\mu \in \mathbb{R}$. Dann ist

$$\langle V, W \rangle = \sum_{l<m} \gamma_{lm} \langle V_{lm}, W \rangle = \mu.$$

Wir erhalten folgende Aussage:

Lemma 3.12 *Eine Strategie $W \in \mathcal{A}_x$ des Statistikers ist genau dann eine Indifferenzmatrix, falls gilt*

$$\exists \mu \in \mathbb{R} : \langle V_{lm}, W \rangle = \mu \quad \forall l < m.$$

In diesem Fall ist $W = \mu U$.

3.5 Der Fall Eins aus N

Sei $n = 1$, $N \in \mathbb{N}$ und $\underline{x} \in \mathbb{R}^N$ ein Regressionsvektor. Aufgrund der Repräsentativität ist $\mathcal{A}_x$ gegeben durch die konvexe Hülle von

$$\{(\frac{1}{x_i}\underline{e}_i - \underline{1})(\frac{1}{x_i}\underline{e}_i - \underline{1})^t; i = 1, \ldots, N\}.$$

Die repräsentativen Schätzvektoren sind festgelegt durch

$$a_{\{i\}} := \frac{1}{x_i}\underline{e}_i \quad \text{für } i = 1, \ldots, N.$$

Das Minimax-Problem lässt sich für $U := \operatorname{diag}^{-1}(\underline{x}) - \underline{1}\,\underline{1}^t$ vollständig lösen. Dabei sind die beiden Fälle $\max\{x_1, \ldots, x_N\} \leq \frac{1}{2}$ (siehe z.B Gabler [6], Satz 7 auf S. 101, oder Stenger [22], Satz auf S. 260) bzw. $\max\{x_1, \ldots, x_N\} > \frac{1}{2}$ getrennt zu betrachten.

Satz 3.13 (Die HH-Strategie) *Falls* $\max\{x_1, \ldots, x_N\} \leq \frac{1}{2}$, *ist unter den obigen Voraussetzungen eine Minimax-Strategie des linearen Spiels* $(\mathcal{Y}_U, \mathcal{A}_x)$ *gegeben durch die Hansen-Hurwitz-Strategie*

$$p^*(\{i\}) = x_i, \quad \underline{a}^*_{\{i\}} = \frac{1}{x_i}\underline{e}_i, \quad \textit{für } i = 1, \ldots, N,$$

und der Wert des Spiels ist 1.

Beweis: Sei $\max\{x_1, \ldots, x_N\} < \frac{1}{2}$. Wir definieren $\underline{q} \in \mathbb{R}^N$ durch

$$q_i := \frac{x_i^2}{1 - 2x_i} \quad \text{für } i = 1, \ldots, N$$

und

$$Q := \operatorname{diag}(\underline{q}) \in \mathcal{Q}_0.$$

Dann ist

$$\begin{aligned}
&(\underline{a}^*_{\{i\}} - \underline{1})^t Q (\underline{a}^*_{\{i\}} - \underline{1}) = \\
&= \sum_{j \neq i} q_j + q_i \frac{(1 - x_i)^2}{x_i^2} = \sum_{j \neq i} \frac{x_j^2}{1 - 2x_j} + \frac{x_i^2}{1 - 2x_i}(\frac{1 - 2x_i}{x_i^2} + 1) \\
&= \sum_{j=1}^{N} \frac{x_j^2}{1 - 2x_j} + 1 = \sum_{j=1}^{N} \frac{x_j^2}{1 - 2x_j} + \sum_{j=1}^{N} x_j = \sum_{j=1}^{N} \frac{x_j(1 - x_j)}{1 - 2x_j} \quad \forall i = 1, \ldots, N.
\end{aligned}$$

Weiter gilt

$$\langle Q, U\rangle = \sum_{j=1}^{N} \frac{x_j(1-x_j)}{1-2x_j},$$

also ist

$$V^* := \frac{1}{\sum_{j=1}^{N} \frac{x_j(1-x_j)}{1-2x_j}} Q \in \mathcal{Y}_U \tag{3.13}$$

eine (allgemeine) Indifferenzmatrix der Natur. Andererseits ist (vgl. Bsp. 1.4)

$$\sum_i p^*(\{i\})(\underline{a}^*_{\{i\}} - \underline{1})(\underline{a}^*_{\{i\}} - \underline{1})^t = \operatorname{diag}^{-1}(\underline{x}) - \underline{1}\,\underline{1}^t = U,$$

und mit Satz 3.7 folgt die Behauptung.

Falls $x_{i_0} = \max\{x_1, \ldots, x_N\} = \frac{1}{2}$, ist

$$V^* := (\underline{e}_{i_0} - \underline{x})(\underline{e}_{i_0} - \underline{x})^t \in \mathcal{Y}_U$$

eine Indifferenzmatrix der Natur, und wir folgern wie oben. $\diamond$

Bemerkung 3.14 *Sei* $\max\{x_1, \ldots, x_N\} \leq \frac{1}{2}$ *und* V^* *eine Maximin-Strategie. Da* $p^*(\{i\}) > 0$ *für* $i = 1, \ldots, N$, *ist* V^* *ist eine Indifferenzmatrix der Natur (siehe Korollar 3.5).*

Satz 3.15 (Die bewusste Auswahl) *Sei* $x_{i_0} = \max\{x_1, \ldots, x_N\} \geq \frac{1}{2}$ *für ein* $i_0 \in \{1, \ldots, N\}$,

$$\underline{y}^* = \sqrt{\frac{x_{i_0}}{1-x_{i_0}}}(-x_1, \ldots, 1-x_{i_0}, \ldots, -x_N)^t$$

und

$$p^*(\{i_0\}) = 1, \quad \underline{a}^*_{\{i_0\}} = \frac{1}{x_{i_0}} \underline{e}_{i_0}.$$

Dann ist unter den obigen Voraussetzungen $(\underline{y}^*(\underline{y}^*)^t, (p^*, \underline{a}^*_{\{i_0\}}))$ *ein Gleichgewicht des linearen Spiels* $(\mathcal{Y}_U, \mathcal{A}_x)$, *und der Wert des Spiels ist gegeben durch*

$$\frac{1-x_{i_0}}{x_{i_0}}.$$

Beweis: Sei o.B.d.A. $x_1 = \max\{x_1, \ldots, x_N\} \geq \frac{1}{2}$. Es ist

$$\max_{\underline{y} \in \Theta_U} \mathrm{MSE}(\underline{y}, (p^*, \underline{a}^*_{\{1\}})) = \max_{\underline{y} \in \Theta_{U1}} \mathrm{MSE}(\underline{y}, (p^*, \underline{a}^*_{\{1\}})) = \max_{\underline{y} \in \Theta_{U1}} (\frac{y_1}{x_1})^2 = \frac{1-x_1}{x_1}$$

und wird angenommen in

$$\underline{y}^* = \pm\sqrt{\frac{x_1}{1-x_1}}(1-x_1, -x_2, \ldots, -x_N)^t,$$

d.h.

$$\underline{y}^*(\underline{y}^*)^t \in \mathcal{BA}((\underline{a}^*_{\{1\}} - \underline{1})(\underline{a}^*_{\{1\}} - \underline{1})^t).$$

Für eine beliebige repräsentative Strategie $(p, \{\frac{1}{x_i}\underline{e}_i\})$ des Statistikers gilt

$$\text{MSE}(\underline{y}^*, (p, \{\frac{1}{x_i}\underline{e}_i\})) = \frac{x_1}{1-x_1} + \frac{1-2x_1}{x_1(1-x_1)}p(\{1\}).$$

Da nach Voraussetzung $1 - 2x_1 \leq 0$, wird das Minimum bzgl. p angenommen in $p(\{1\}) = 1$, und somit gilt

$$(\underline{a}^*_{\{1\}} - \underline{1})(\underline{a}^*_{\{1\}} - \underline{1})^t \in \mathcal{BA}(\underline{y}^*(\underline{y}^*)^t).$$

◇

Aus Satz 3.15 ergeben sich zwei interessante Folgerungen:

(i) Im Fall $\max\{x_1, \ldots, x_N\} > \frac{1}{2}$ spielt der Statistiker kein Vielfaches der Matrix U. Speziell ist die Minimax-Strategie keine Indifferenzmatrix (vgl. Skizze 2.21: Die Gerade cU, $c \in \mathbb{R}$, schneidet $\widetilde{\mathcal{A}}_X$, aber nicht den Rand $\Lambda(\mathcal{A}_X)$).

(ii) Falls $x_{i_0} = \max\{x_1, \ldots, x_N\} = \frac{1}{2}$, ist sowohl die HH-Strategie als auch die reine Strategie gegeben durch $p(\{i_0\}) = 1$, $\underline{a}_{\{i_0\}} = \frac{1}{x_{i_0}}\underline{e}_{i_0}$ eine Minimax-Strategie. Sei W_1^* die risikoerzeugende Matrix der HH-Strategie und W_2^* die risikoerzeugende Matrix der reinen Strategie. Für $N > 2$ ist

$$W_1^* \neq W_2^*,$$

und es gilt

$$\langle V, W_1^* \rangle = 1 \geq \langle V, W_2^* \rangle \quad \forall V \in \mathcal{Y}_U.$$

Mit $\tilde{\underline{y}} := \sqrt{\frac{x_2x_3}{x_2+x_3}}(0, 1, -1, 0, \ldots, 0)^t \in \Theta_{U1}$ ist

$$\langle \tilde{\underline{y}}\,\tilde{\underline{y}}^t, W_2^* \rangle = 0,$$

die HH-Strategie wird also strikt dominiert von der (zulässigen) reinen Strategie. Weiter ist die Maximin-Strategie gegeben durch die reine Strategie $\underline{y}^*(\underline{y}^*)^t$ mit

$$\underline{y}^* := \underline{e}_{i_0} - \underline{x}$$

eine Indifferenz-Strategie.

Im Folgenden sei U eine beliebige symmetrische positiv semidefinite $N \times N$-Matrix mit Rang $N - 1$ und $U\underline{x} = \underline{0}$. Wie zuvor mischt der Statistiker aufgrund der Repräsentativität über die Schätzvektoren

$$\underline{a}_{\{i\}} = \frac{1}{x_i}\underline{e}_i \quad \text{für } i = 1, \ldots, N.$$

Versucht er ein Vielfaches von U zu spielen (genau dann stellt er die Natur indifferent), sucht er einen Auswahlplan p mit

$$\sum_i p(i)(\underline{a}_{\{i\}} - \underline{1})(\underline{a}_{\{i\}} - \underline{1})^t = \lambda U$$

für ein $\lambda \in \mathbb{R}$, wobei $p(i) := p(\{i\})$ für $i = 1, \ldots, N$. Da $U\underline{x} = \underline{0}$, sind die Diagonalelemente von U eindeutig festgelegt durch u_{ij} für $i < j$, und wir erhalten das lineare Gleichungssystem

$$p(i)\frac{x_i - 1}{x_i} + p(j)\frac{x_j - 1}{x_j} + \sum_{k \neq i,j} p(k) = \lambda u_{ij} \quad \text{für } i < j,$$

also wegen $\sum_i p(i) = 1$:

$$1 - \frac{p(i)}{x_i} - \frac{p(j)}{x_j} = \lambda u_{ij} \quad \text{für } i < j. \tag{3.14}$$

Das sind $\frac{N(N-1)}{2}$ Gleichungen bei N Unbekannten $p(1), \ldots, p(N-1), \lambda$, wobei $p(N) = 1 - \sum_{i<N} p(i)$.

Das folgende Beispiel zeigt, dass der Statistiker im Allgemeinen kein Vielfaches der Matrix U spielen kann (vgl. Skizze 2.21: Die Gerade cU, $c \in \mathbb{R}$, schneidet $\widetilde{\mathcal{A}}_X$ nicht).

Beispiel 3.16 *Sei $N = 4$, $\underline{x} = \frac{1}{4}\underline{1} \in \mathbb{R}^4$ und*

$$U = \begin{bmatrix} 4 & -1 & -1 & -2 \\ -1 & 4 & -2 & -1 \\ -1 & -2 & 4 & -1 \\ -2 & -1 & -1 & 4 \end{bmatrix}.$$

Die Matrix U ist positiv semidefinit mit Kern(U) = span$[\underline{x}]$. Das Gleichungssystem (3.14) ist äquivalent zu

$$1 - 4(p(i) + p(j)) = \lambda u_{ij} \quad i < j.$$

Speziell folgt wegen $\sum_i p(i) = 1$

$$-2 = \lambda(u_{12} + u_{34}) = -2\lambda$$

und

$$-2 = \lambda(u_{14} + u_{23}) = -4\lambda,$$

d.h. es existiert keine Lösung. Der Statistiker kann kein Vielfaches von U *spielen, besitzt also keine Indifferenzmatrix.*

Sei $\max\{x_1, \ldots, x_N\} < \frac{1}{2}$ und V^* die in (3.13) definierte (allgemeine) Indifferenzmatrix. Aufgrund von Bemerkung 3.9 ist

$$V_U^* := \frac{V^*}{\langle V^*, U \rangle} \in \mathcal{Y}_U$$

eine Indifferenzmatrix der Natur, und eine untere Schranke für den Spielwert ist gegeben durch

$$\frac{1}{\langle V^*, U \rangle} = \frac{\sum_i \frac{x_i(1-x_i)}{1-2x_i}}{\sum_i \frac{x_i^2}{1-2x_i} u_{ii}}.$$

Falls ein Auswahlplan p^* existiert, so dass der Statistiker ein Vielfaches von U spielen kann, gilt

$$p^*(i)(\frac{1-x_i}{x_i})^2 + \sum_{j \neq i} p^*(j) = \lambda u_{ii} \quad \forall i = 1, \ldots, N,$$

also

$$p^*(i) = \frac{x_i^2}{1-2x_i}(\lambda u_{ii} - 1) \quad \forall i = 1, \ldots, N.$$

Eine Minimax-Strategie ist gegeben durch $(p^*, \{\frac{1}{x_i}\underline{e}_i\})$, V_U^* ist Maximin-Strategie, und der Wert des Spiels ist $\lambda = \frac{1}{\langle V^*, U \rangle}$.

Kapitel 4

Der Fall Zwei aus Drei für das Ellipsoid

In diesem Kapitel berechnen wir für $N = 3$ und $n = 2$ bei vorgegebenem Regressionsvektor $\underline{x} \in \mathbb{R}^3$ Minimax-Strategien des Spiels $(\mathcal{Y}_U, \mathcal{A}_x)$, wobei U eine positiv semidefinite 3×3-Matrix mit $U\underline{x} = \underline{0}$ bezeichnet. Aussichtsreiche Kandidaten für Maximin- bzw. Minimax-Strategien sind die in Abschnitt 3.2 eingeführten Indifferenzmatrizen. Es wird sich zeigen, dass die Minimax-Strategie durch ein Vielfaches von U gegeben ist (Satz 4.4), d.h. sie ist eine Indifferenz-Strategie. In Abschnitt 4.2 werden wir sehen, dass eine Indifferenzmatrix der Natur genau dann existiert, wenn alle Komponenten des Regressionsvektors kleiner $\frac{1}{2}$ sind. Wann eine Indifferenz-Matrix eine Maximin-Strategie ist und wie sich in diesen Fällen eine Minimax-Strategie berechnet, klären wir in den Abschnitten 4.3 bis 4.5. In Abschnitt 4.6 lösen wir das Spiel, falls zwei Komponenten des Regressionsvektors gleich sind (Zwei-Schichten-Fall).

4.1 Zusammenhang zwischen Indifferenzmatrizen und Ellipsen

Eine Strategie des Statistikers setzt sich zusammen aus einem Auswahlplan auf $S(2) = \{\{1,2\}, \{1,3\}, \{2,3\}\}$ und drei repräsentativen Schätzvektoren. Um den Statistiker indifferent bzgl. der Stichproben zu stellen, sucht die Natur eine Matrix $V \in \mathcal{Y}_U$ mit

$$\min_{\underline{a}_s \in A_s} (\underline{a}_s - \underline{1})^t V (\underline{a}_s - \underline{1}) = \lambda \quad \forall s \in S(2) \tag{4.1}$$

für ein $\lambda \in \mathbb{R}$. Uns interessiert, unter welchen Umständen (in Abhängigkeit von dem Regressionsvektor $\underline{x}$) eine Indifferenzmatrix $V \in \mathcal{Y}_U$ existiert. Dazu definieren wir

$$\Omega := \{\underline{a} - \underline{1}; \underline{a} \text{ ist repräsentativer Schätzvektor }\}.$$

Ω lässt sich wie folgt geometrisch interpretieren: Sei E_x der zu $\underline{x}$ orthogonale Unterraum, also

$$E_x := \{\underline{z} \in \mathbb{R}^3; (\underline{z}, \underline{x})_2 = 0\}.$$

Weiter sei $E_i := \{\underline{z} \in \mathbb{R}^3; z_i = -1\}$ und

$$g_i := E_i \cap E_x$$

für $i \in \{1, 2, 3\}$. Es gilt

$$g_i = \{\underline{a}_s - \underline{1}; \underline{a}_s \in A_s, s = \{1,2,3\} \setminus \{i\}\} \quad \forall i = 1, 2, 3$$

und somit

$$\Omega = g_1 \cup g_2 \cup g_3.$$

Wir suchen eine positiv semidefinite Matrix Q mit $Q\underline{x} = \underline{0}$, so dass der Zylinder

$$Z := \{\underline{z}; \underline{z}^t Q \underline{z} = 1\}$$

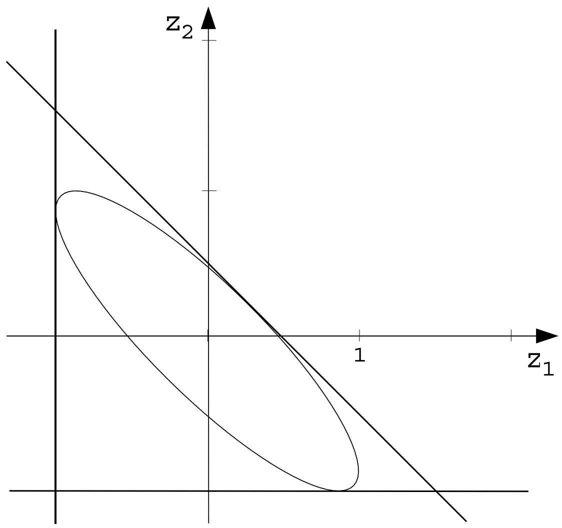
die drei Geraden g_1, g_2, g_3 berührt. Existiert eine solche Matrix Q, so ist $\langle Q, U \rangle > 0$ (siehe Korollar A.2 im Anhang), und aufgrund von Satz 2.19 ist

$$V := \frac{Q}{\langle Q, U \rangle} \in \mathcal{Y}_{Ux}$$

eine Indifferenzmatrix der Natur, denn aus der Berühreigenschaft von Q folgt offensichtlich (4.1) mit $\lambda := \frac{1}{\langle Q, U \rangle}$.

Beschäftigen wir uns mit der Frage, wann ein solcher Tangential-Zylinder existiert: Die drei Geraden g_1, g_2, g_3 beschreiben ein Dreieck in der Ebenen E_x, in dessen Innerem der Null-Punkt liegt. Dieses Dreieck projizieren wir in die z_1-z_2-Ebene. Ein Tangential-Zylinder existiert genau dann, wenn in der z_1-z_2-Ebene eine Ellipse um den Nullpunkt existiert, die tangential ist bzgl. der Geraden $z_1 = -1$, $z_2 = -1$ und $z_2 = -\frac{x_1}{x_2} z_1 + \frac{x_3}{x_2}$.

Beispiel 4.1 *Die folgende Abbildung veranschaulicht obigen Sachverhalt für den Regressionsvektor* $\underline{x} = (\frac{2}{5}, \frac{2}{5}, \frac{1}{5})^t$.



4.2 Existenz von Indifferenzmatrizen

Gesucht ist eine positiv definite Matrix

$$M_E = \begin{bmatrix} a & c \\ c & b \end{bmatrix},$$

so dass die dadurch beschriebene Ellipse

$$E := \{\underline{z} \in \mathbb{R}^2; \underline{z}^t M_E \underline{z} = 1\}$$

die obigen drei Geraden berührt. Sei

$$\Delta := ab - c^2 = \det(M_E) > 0.$$

Eine einfache Rechnung zeigt: E berührt $z_1 = -1$ genau dann, wenn

$$b = \Delta.$$

Der Berührpunkt ist gegeben durch

$$(-1, \frac{c}{\Delta})^t. \tag{4.2}$$

E berührt $z_2 = -1$ genau dann, wenn

$$a = \Delta.$$

Der Berührpunkt ist gegeben durch

$$(\frac{c}{\Delta}, -1)^t. \tag{4.3}$$

Somit ist die gesuchte Matrix von der Gestalt

$$M_E = \begin{bmatrix} \Delta & c \\ c & \Delta \end{bmatrix},$$

und es gilt

$$\Delta^2 - c^2 = \Delta > 0. \tag{4.4}$$

Sei $\beta \in \mathbb{R} \setminus \{0\}$. Offensichtlich berührt E die Gerade $z_2 = \alpha z_1 + \beta$ genau dann, wenn die Gleichung

$$(\Delta\alpha^2 + 2c\alpha + \Delta)z_1^2 + 2\beta(\Delta\alpha + c)z_1 + \Delta\beta^2 = 1$$

nur eine Lösung bzgl. z_1 besitzt. Da $\Delta^2 - c^2 = \Delta > 0$, ist $\Delta\alpha^2 + 2c\alpha + \Delta > 0$ für alle α, und E berührt die Gerade $z_2 = \alpha z_1 + \beta$ genau dann, wenn

$$\Delta\alpha^2 + 2c\alpha + \Delta = \beta^2\Delta.$$

Der Berührpunkt ist gegeben durch

$$(-\frac{\Delta\alpha + c}{\beta\Delta}, \frac{c\alpha + \Delta}{\beta\Delta})^t.$$

Daraus ergibt sich: E berührt $z_2 = -\frac{x_1}{x_2}z_1 + \frac{x_3}{x_2}$ genau dann, wenn

$$\Delta(\frac{x_1}{x_2})^2 - 2c\frac{x_1}{x_2} + \Delta = (\frac{x_3}{x_2})^2\Delta. \tag{4.5}$$

Der Berührpunkt ist gegeben durch

$$(\frac{x_1}{x_3} - \frac{x_2}{x_3}\frac{c}{\Delta}, \frac{x_2}{x_3} - \frac{x_1}{x_3}\frac{c}{\Delta})^t. \tag{4.6}$$

Aus Gleichung (4.5) folgt:

$$c = \Delta K, \tag{4.7}$$

wobei

$$K := \frac{x_1^2 + x_2^2 - x_3^2}{2x_1x_2} = \frac{1 - 2x_3}{2x_1x_2} - 1.$$

Weiter muss wegen (4.4) gelten:

$$0 < \Delta = \Delta^2 - \Delta^2 K^2 = \Delta^2(1 - K^2).$$

Die Tangential-Ellipse existiert somit genau dann, wenn $1 - K^2 > 0$, also falls

$$0 < \frac{1}{2} - x_3 < 2x_1x_2. \tag{4.8}$$

Satz 4.2 *Sei* $\underline{x} \in \mathbb{R}^3$ *ein Regressionsvektor. Eine Indifferenzmatrix der Natur existiert genau dann, wenn*

$$x_0 < \frac{1}{2}, \tag{4.9}$$

wobei $x_0 := \max\{x_1, x_2, x_3\}$.

Beweis: Eine kleine Rechnung zeigt, dass (4.8) äquivalent ist zu (4.9). $\diamond$

Falls $x_0 < \frac{1}{2}$, ist

$$\Delta = \frac{1}{1 - K^2}, \; c = \frac{K}{1 - K^2},$$

und die Matrix M_E ist gegeben durch

$$M_E = \frac{1}{1 - K^2} \begin{bmatrix} 1 & K \\ K & 1 \end{bmatrix}.$$

4.3 Berechnung der optimalen Schätzvektoren

Sei $\underline{x}$ ein Regressionsvektor, der die Bedingung (4.9) erfüllt. Die Natur besitzt also eine Indifferenz-Strategie V, und mit

$$\underline{a}_s^* := \mathrm{argmin}_{\underline{a}_s \in \mathcal{A}_s} \langle V, (\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t \rangle \quad \text{für } s \in S(2)$$

ist jede Mischung über $\{(\underline{a}_s^* - \underline{1})(\underline{a}_s^* - \underline{1})^t; s \in S(2)\}$ eine beste Antwort auf V. Unter welchen Umständen die Indifferenzmatrix V eine Maximin-Strategie ist, werden wir in Abschnitt 4.5 sehen. Vorab berechnen wir die optimalen Schätzvektoren $\underline{a}_s^*$ für $s \in S(2)$.

Berechnung von $\underline{a}_{\{1,2\}}^*$: Die Projektion von $\underline{a}_{\{1,2\}}^* - \underline{1}$ in die z_1-z_2-Ebene ist gegeben durch (4.6), also (mit (4.7))

$$\begin{aligned} \underline{a}_{\{1,2\}}^* - \underline{1} &= (\frac{x_1}{x_3} - \frac{x_2}{x_3}K, \frac{x_2}{x_3} - \frac{x_1}{x_3}K, -1)^t \\ &= (\frac{x_1^2 - x_2^2 + x_3^2}{2x_1x_3}, \frac{-x_1^2 + x_2^2 + x_3^2}{2x_2x_3}, -1)^t, \end{aligned}$$

und somit gilt (wegen $\sum_i x_i = 1$)

$$\underline{a}^*_{\{1,2\}} = (\frac{1-2x_2}{2x_1x_3}, \frac{1-2x_1}{2x_2x_3}, 0)^t.$$

Berechnung von $\underline{a}^*_{\{1,3\}}$: Die Projektion von $\underline{a}^*_{\{1,3\}} - \underline{1}$ ist gegeben durch (4.3), d.h. durch $(K,-1)^t$. Da $\underline{a}^*_{\{1,3\}} - \underline{1} \in g_2$, gilt

$$(\underline{a}^*_{\{1,3\}} - \underline{1})_3 = \frac{x_2 - x_1K}{x_3} = \frac{-x_1^2 + x_2^2 + x_3^2}{2x_2x_3}$$

und somit

$$\underline{a}^*_{\{1,3\}} = (\frac{1-2x_3}{2x_1x_2}, 0, \frac{1-2x_1}{2x_2x_3})^t.$$

Analog ergibt sich mit (4.2)

$$\underline{a}^*_{\{2,3\}} = (0, \frac{1-2x_3}{2x_1x_2}, \frac{1-2x_2}{2x_1x_3})^t.$$

Für spätere Anwendungen ist folgende Darstellung der Berührpunkte nützlich: Sei

$$\xi_i := \frac{\sum x_j^2 - 2x_i^2}{2x_1x_2x_3} x_i \quad \text{für } i = 1,2,3. \tag{4.10}$$

Definieren wir für $i \notin s$, $s \in S(2)$

$$\xi(i) := \underline{a}^*_s - \underline{1},$$

so gilt

$$\xi(1) = (-1, \xi_3, \xi_2)^t,\ \xi(2) = (\xi_3, -1, \xi_1)^t,\ \xi(3) = (\xi_2, \xi_1, -1)^t. \tag{4.11}$$

4.4 Berechnung der Indifferenzmatrix

Wir werden in diesem Abschnitt die Indifferenzmatrix der Natur als Element in $\mathcal{Y}_{U1}$ berechnen.

Sei $x_0 < \frac{1}{2}$ und

$$Q := \begin{bmatrix} q_{11} & q(3) & q(2) \\ q(3) & q_{22} & q(1) \\ q(2) & q(1) & q_{33} \end{bmatrix} \tag{4.12}$$

mit

$$Q\underline{1} = \underline{0} \tag{4.13}$$

und

$$(\underline{a}_s^* - \underline{1})^t Q (\underline{a}_s^* - \underline{1}) = (\underline{a}_s^*)^t Q \underline{a}_s^* = 1 \quad \forall s \in S(2). \tag{4.14}$$

Daraus ergibt sich (vgl. Gabler und Stenger [8])

$$q_{ii} = 4 \frac{x_i(1-2x_i)\prod_j x_j}{\prod_j (1-2x_j)}$$

und

$$q(i) = -2 \frac{\prod_j x_j}{1-2x_i}$$

für $i = 1, 2, 3$, d.h.

$$Q = c \begin{bmatrix} 2x_1(1-2x_1) & -(1-2x_1)(1-2x_2) & -(1-2x_1)(1-2x_3) \\ -(1-2x_1)(1-2x_2) & 2x_2(1-2x_2) & -(1-2x_2)(1-2x_3) \\ -(1-2x_1)(1-2x_3) & -(1-2x_2)(1-2x_3) & 2x_3(1-2x_3) \end{bmatrix}$$

mit

$$c = 2 \frac{\prod_j x_j}{\prod_j (1-2x_j)}.$$

Da die Matrix Q durch (4.12), (4.13) und (4.14) eindeutig festgelegt ist, ist sie eine allgemeine Indifferenzmatrix; somit ist

$$V := \frac{Q}{\langle Q, U \rangle}$$

aufgrund von Satz 2.20 eine Indifferenzmatrix der Natur. (Da $\dim(\mathrm{Kern}(U)) = \dim(\mathrm{Kern}(V)) = 1$, ist $\langle Q, U \rangle > 0$.) Die Natur kann sich den Wert $\frac{1}{\langle Q,U \rangle}$ sichern, falls sie die Indifferenzmatrix V spielt (vgl. Bemerkung 3.9). Für den Spielwert λ gilt also

$$\lambda \geq \frac{1}{\langle Q, U \rangle}.$$

Beispiel 4.3 *Sei $U = diag^{-1}(\underline{x}) - \underline{1}\,\underline{1}^t$. Dann ist*

$$\langle Q, U \rangle = 2c,$$

und für den Wert λ des Spiels gilt

$$\lambda \geq \frac{1}{4} \frac{\prod_j (1-2x_j)}{\prod_j x_j}.$$

Falls $x_1 = x_2 = k$ und $x_3 = 1 - 2k$ für $k \in]\frac{1}{4}; \frac{1}{2}[$, ist

$$V = \frac{1-2k}{2} \begin{bmatrix} 2k & 2k-1 & 1-4k \\ 2k-1 & 2k & 1-4k \\ 1-4k & 1-4k & 8k-2 \end{bmatrix} \tag{4.15}$$

eine Indifferenz-Strategie der Natur, und eine untere Schranke für den Spielwert ist gegeben durch

$$\frac{1}{4}\frac{(1-2k)(4k-1)}{k^2}.$$

In Gabler [6], Bsp. 8 auf S. 97, wurde diese Schranke auf alternative Weise hergeleitet.

4.5 Konstruktion eines Auswahlplans

Im Fall $N = 3$, $n = 2$ lässt sich die Minimax-Strategie wie folgt charakterisieren (vgl. Skizze 2.21: Die Gerade cU, $c \in \mathbb{R}$, schneidet den Rand $\Lambda(\mathcal{A}_X)$).

Satz 4.4 *Sei $\underline{x} \in \mathbb{R}^3$ ein Regressionsvektor, W^* Minimax-Strategie des Spiels $(\mathcal{Y}_U, \mathcal{A}_x)$ und λ der Wert des Spiels. Dann ist*

$$W^* = \lambda U.$$

Beweis: Angenommen, W^* ist kein Vielfaches von U. Dann besitzt das Maximierungsproblem

$$\max_{\underline{y} \in \Theta_{Ux}} \underline{y}^t W^* \underline{y}$$

eine bis auf das Vorzeichen eindeutige Lösung $\underline{y}^* \in \Theta_{Ux}$ mit $(\underline{y}^*)^t W^* \underline{y}^* > 0$: Da $U \in \mathcal{Q}_0$, existiert eine Orthonormal-Basis $(\underline{z}_1, \underline{z}_2, \underline{z}_3)$ von Eigenvektoren bzgl. der Eigenwerte $0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3$ von U. Wegen $\mathrm{Kern}(U) = \mathrm{span}[\underline{x}]$ ist $\underline{z}_1 = \frac{1}{|\underline{x}|_2}\underline{x}$ und $\lambda_1 = 0 < \lambda_2 \leq \lambda_3$, wobei $|\,.\,|_2$ die Euklidische Norm bezeichnet. Sei T die 3×3-Matrix mit den Zeilenvektoren $\underline{z}_1, \underline{z}_2, \underline{z}_3$. Dann ist $T^t T = T T^t = I$, und mit $D := \mathrm{diag}((0, \lambda_2, \lambda_3)^t)$ gilt

$$U = T^t D T,$$

also

$$D = T U T^t.$$

Sei $W := T W^* T^t \in \mathcal{Q}_0$. Es ist $W \underline{e}_1 = \frac{1}{|\underline{x}|_2} T W^* \underline{x} = \underline{0}$, also

$$W = \begin{bmatrix} 0 & 0 \\ 0 & \tilde{W} \end{bmatrix},$$

wobei $\tilde{W}$ die 2×2-Matrix mit den Zeilenvektoren $(w_{22}, w_{23})^t$ und $(w_{23}, w_{33})^t$ bezeichnet. Offensichtlich ist $\tilde{W}$ positiv semidefinit. Mit $\tilde{D} := \mathrm{diag}((\lambda_1, \lambda_2)^t)$ gilt

$$\max_{\underline{y} \in \Theta_{Ux}} \underline{y}^t W^* \underline{y} = \max_{\underline{v}^t \tilde{D} \underline{v} = 1} \underline{v}^t \tilde{W} \underline{v}.$$

Da $cU \neq W^*$ und somit $c\tilde{D} \neq \tilde{W}$ für alle $c \in \mathbb{R}$, besitzt die rechte Seite der obigen Gleichung eine bis auf das Vorzeichen eindeutige Lösung $\underline{v}^* \in \mathbb{R}^2$; offensichtlich ist $\underline{y}^* := T^t(0, v_1^*, v_2^*)^t$ die bis auf das Vorzeichen eindeutige Lösung des ursprünglichen Maximierungsproblems.

Die Maximin-Strategie V^* ist also gegeben durch

$$V^* = \underline{y}^* (\underline{y}^*)^t.$$

Andererseits existiert offensichtlich ein $s \in S(2)$, $\underline{\tilde{a}}_s \in A_s$ mit

$$(\underline{y}^*, \underline{\tilde{a}}_s - \underline{1})_2 = 0,$$

was im Widerspruch zur Minimax-Eigenschaft von W^* steht. $\diamond$

Ein Auswahlplan p ist im Fall $N = 3$, $n = 2$ festgelegt durch die Wahrscheinlichkeiten p_i, $i = 1, 2, 3$, mit der die i-te Einheit nicht in die Auswahl gelangt. Sei $x_0 < \frac{1}{2}$, es existiert also eine Indifferenzmatrix der Natur. Aufgrund von Satz 4.4 suchen wir einen Wahrscheinlichkeitsvektor $\underline{p} = (p_1, p_2, p_3)^t$ mit

$$\sum p_i \xi(i) \xi(i)^t = \lambda U \tag{4.16}$$

für ein $\lambda \in \mathbb{R}$, wobei $\xi(i) = \underline{a}_s^* - \underline{1}$ für $i \notin s$ (vgl. (4.11)). Erfüllt ein Wahrscheinlichkeitsvektor $\underline{p}$ die Gleichung (4.16) für ein $\lambda \in \mathbb{R}$, so ist aufgrund von Satz 3.7 $(p, \{\underline{a}_s^*\})$ eine Minimax-Strategie, $V^* := \frac{Q}{\langle Q, U \rangle}$ eine Maximin-Strategie (wobei Q die allgemeine Indifferenzmatrix aus dem letzten Abschnitt bezeichnet), und der Wert des Spiels ist $\lambda = \frac{1}{\langle Q, U \rangle}$.

Sei

$$U = \begin{bmatrix} u_{11} & u(3) & u(2) \\ u(3) & u_{22} & u(1) \\ u(2) & u(1) & u_{33} \end{bmatrix}.$$

Bei Vorgabe von $u(1), u(2), u(3)$ sind die Diagonalelemente von U eindeutig festgelegt, da $U\underline{x} = \underline{0}$. Aus (4.16) ergibt sich folgendes Gleichungssystem:

$$\begin{aligned} -\xi_3 p_1 - \xi_3 p_2 + \xi_1 \xi_2 p_3 &= \lambda u(3), \\ -\xi_2 p_1 + \xi_1 \xi_3 p_2 - \xi_2 p_3 &= \lambda u(2), \\ \xi_2 \xi_3 p_1 - \xi_1 p_2 - \xi_1 p_3 &= \lambda u(1) \end{aligned} \tag{4.17}$$

unter der Nebenbedingung

$$\sum_i p_i = 1.$$

Somit ist

$$p_i \frac{\xi_1 \xi_2 \xi_3}{\xi_i} - \xi_i (1 - p_i) = \lambda u(i) \quad \text{für } i = 1, 2, 3,$$

also

$$p_i = \frac{\xi_i}{\xi_1 \xi_2 \xi_3 + \xi_i^2} (\xi_i + \lambda u(i)) \quad \text{für } i = 1, 2, 3, \tag{4.18}$$

wobei aufgrund der Nebenbedingung

$$\lambda = \frac{1 - \sum_i \frac{\xi_i^2}{\xi_1 \xi_2 \xi_3 + \xi_i^2}}{\sum_i \frac{u(i)\xi_i}{\xi_1 \xi_2 \xi_3 + \xi_i^2}}.$$

Falls $p_i \geq 0$ für $i = 1, 2, 3$, haben wir wie oben beschrieben ein Gleichgewicht des Spiels gefunden. Allerdings ist dies nicht immer der Fall, d.h. wir können auf diese Art und Weise i.a. keine Minimax-Strategie konstruieren (siehe Abschnitt 4.6).

Lemma 4.5 *Sei $U = diag^{-1}(\underline{x}) - \underline{1}\,\underline{1}^t$. Eine obere Schranke λ_M für den Wert des Spiels ist gegeben durch*

$$\lambda_M = \frac{x_1 x_2 x_3}{x_1 x_2^2 + x_2 x_3^2 + x_3 x_1^2}.$$

Beweis: Wir definieren

$$\underline{a}_{\{1,2\}} := (1, 1 + \frac{x_3}{x_2}, 0)^t, \ \underline{a}_{\{1,3\}} := (1 + \frac{x_2}{x_1}, 0, 1)^t, \ \underline{a}_{\{2,3\}} := (0, 1, 1 + \frac{x_1}{x_3})^t.$$

Mit

$$A_i := (\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t \quad \text{für } i \notin s, \ i = 1, 2, 3$$

bildet $(A_i; i = 1, 2, 3)$ eine Basis des Vektorraums der symmetrischen 3×3-Matrizen mit Eigenvektor $\underline{x}$ bzgl. des Eigenwerts 0. Es existieren also $\lambda_i \in \mathbb{R}$ mit

$$U = \lambda_1 A_1 + \lambda_2 A_2 + \lambda_3 A_3.$$

Aufgrund der Gestalt von A_i und U erhalten wir

$$\lambda_1 = \frac{x_3}{x_1}, \lambda_2 = \frac{x_1}{x_2}, \lambda_3 = \frac{x_2}{x_3}.$$

Mit

$$\lambda_M := \frac{1}{\sum_i \lambda_i} = \frac{x_1x_2x_3}{x_1x_2^2 + x_2x_3^2 + x_3x_1^2},$$
$$p_i := \lambda_i\lambda_M \quad \text{für } i = 1, 2, 3$$

kann der Statistiker die Matrix

$$W_M := \sum_i p_iA_i = \lambda_M U$$

spielen, und es gilt

$$\langle V, W_M\rangle = \lambda_M\langle V, U\rangle \leq \lambda_M \quad \forall V \in \mathcal{Y}_{Ux}.$$

$\diamond$

Bemerkung 4.6 *Für eine beliebige Matrix U gilt wie oben*

$$U = \sum_i \lambda_iA_i$$

mit

$$\lambda_1 = -\frac{u_{23}x_3}{x_1}, \quad \lambda_2 = -\frac{u_{13}x_1}{x_2}, \quad \lambda_3 = -\frac{u_{12}x_2}{x_3}.$$

Falls $u_{ij} < 0$ für alle $i \neq j$, erhalten wir eine obere Schranke für den Wert des Spiels entsprechend aus

$$\lambda_M := \frac{1}{\sum_i \lambda_i}.$$

Bemerkung 4.7 *Definieren wir*

$$\underline{a}_{\{1,2\}} := (1 + \frac{x_3}{x_1}, 1, 0)^t,\ \underline{a}_{\{1,3\}} := (1, 0, 1 + \frac{x_2}{x_3})^t,\ \underline{a}_{\{2,3\}} := (0, 1 + \frac{x_1}{x_2}, 1)^t$$

und

$$A_i := (\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t \quad \text{für } i \notin s,\ i = 1, 2, 3,$$

so gilt wie im Beweis des obigen Lemmas

$$U = \sum_i \lambda_iA_i,$$

wobei

$$\lambda_1 = \frac{x_2}{x_1},\ \lambda_2 = \frac{x_3}{x_2},\ \lambda_3 = \frac{x_1}{x_3}.$$

Somit ist

$$\frac{x_1x_2x_3}{x_1x_3^2 + x_2x_1^2 + x_3x_2^2}$$

ebenfalls eine obere Schranke für den Wert des Spiels. Für allgemeines U erhalten wir

$$\lambda_1 = -\frac{u_{23}x_2}{x_3}, \ \lambda_2 = -\frac{u_{13}x_3}{x_2}, \ \lambda_3 = -\frac{u_{12}x_1}{x_3},$$

und falls $\lambda_i > 0$ für $i = 1, 2, 3$, ist

$$\frac{1}{\sum_i \lambda_i}$$

eine obere Schranke für den Wert des Spiels.

Alternativ können wir zur Bestimmung des Auswahlplans wie folgt vorgehen (vgl. Lemma 3.12): Wir definieren

$$Q_1 := \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ Q_2 := \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \ Q_3 := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

und

$$V_i := \frac{Q_i}{\langle Q_i, U \rangle} \in \mathcal{Y}_{U1} \quad \forall i = 1, 2, 3.$$

Offensichtlich ist $\mathcal{Y}_{U1} \subseteq \text{span}[V_i; i = 1, 2, 3]$, für ein $V \in \mathcal{Y}_{U1}$ existieren also $\gamma_i \in \mathbb{R}$ mit $\sum_i \gamma_i = 1$ und

$$V = \sum_i \gamma_i V_i.$$

Sei $V \in \mathcal{Y}_{U1}$ die in Abschnitt 4.4 berechnete Indifferenzmatrix der Natur und $\underline{a}_s^*$ die in Abschnitt 4.3 berechneten Schätzvektoren. Wir definieren

$$A_i^* := \underline{a}^*_{\{1,2,3\}\setminus\{i\}} (\underline{a}^*_{\{1,2,3\}\setminus\{i\}})^t \quad \forall i = 1, 2, 3$$

und

$$\lambda := \langle V, A_i^* \rangle.$$

Falls das Gleichungssystem

$$\sum_{i=1}^{3} p_i \langle V_j, A_i^* \rangle = \lambda \quad \text{für } j = 1, 2, 3$$

lösbar ist mit $p_i^* \geq 0$ für $i = 1, 2, 3$ ($\sum_i p_i = 1$ ist für jede Lösung erfüllt), ist V Maximin-Strategie, $(p^*, \{\underline{a}_s^*\})$ Minimax-Strategie und λ der Wert des Spiels.

4.6 Der Zwei-Schichten-Fall

Zu einer gegebenen Indifferenzmatrix der Natur können die aus dem obigen Gleichungssystem resultierenden p_i negativ sein. Um diese Problematik besser zu verstehen, untersuchen wir den Fall

$$x_1 = x_2 := k,\ x_3 = 1 - 2k =: h$$

für $k \in]\frac{1}{4}; \frac{1}{2}[$. (Genau dann existiert eine Indifferenzmatrix der Natur.) Die Matrix U sei gegeben durch

$$U := \operatorname{diag}^{-1}(\underline{x}) - \underline{1}\,\underline{1}^t = \begin{bmatrix} \frac{1-k}{k} & -1 & -1 \\ -1 & \frac{1-k}{k} & -1 \\ -1 & -1 & \frac{2k}{h} \end{bmatrix}.$$

Es ist (vgl. (4.10))

$$\xi_1 = \xi_2 = \frac{1}{2k} - 1,\ \xi_3 = \frac{4k-1}{2k^2} - 1,$$

und für die aus (4.18) resultierenden Wahrscheinlichkeiten gilt

$$\begin{aligned} p_1 = p_2 &= -\frac{k(2k-1)}{4k-1} = \frac{kh}{2k-h}, \\ p_3 &= \frac{4k^2+2k-1}{4k-1} = \frac{4k^2-h}{2k-h}. \end{aligned}$$

Offensichtlich ist $p_1 = p_2 > 0$ für alle $k \in]\frac{1}{4}; \frac{1}{2}[$ und p_3 genau dann gleich Null, wenn $k = \frac{\pm\sqrt{5}-1}{4}$; somit ist das Minimax-Problem für $k \in [\frac{\sqrt{5}-1}{4}; \frac{1}{2}[$ gelöst. (Vgl. Gabler [6], Bsp. 8 auf S. 97. Hier wird die Minimax-Strategie mittels Eigenwert-Berechnungen hergeleitet.) Die Maximin-Strategie ist gegeben durch (vgl. (4.15))

$$V^* = \frac{1-2k}{2} \begin{bmatrix} 2k & 2k-1 & 1-4k \\ 2k-1 & 2k & 1-4k \\ 1-4k & 1-4k & 8k-2 \end{bmatrix}.$$

Beispiel 4.8 *Falls $k = \frac{3}{8}$, ist die Maximin-Strategie gegeben durch*

$$V^* = \frac{1}{32} \begin{bmatrix} 3 & -1 & -2 \\ -1 & 3 & -2 \\ -2 & -2 & 4 \end{bmatrix} \in \mathcal{Y}_{U1}.$$

Allerdings ist sie keine Maximin-Strategie im Fall 1 aus N, denn

$$\frac{1}{x_i^2}\underline{e}_i V^* \underline{e}_i = \begin{cases} \frac{2}{3} & : \quad i = 1,2 \\ 2 & : \quad i = 3 \end{cases} ,$$

d.h. V^* ist keine Indifferenzmatrix der Natur und somit auch keine Maximin-Strategie (vgl. Bemerkung 3.14).

Wie soll der Statistiker vorgehen, wenn $k \in]\frac{1}{4}; \frac{\sqrt{5}-1}{4}[$ bzw. $k \in]0; \frac{\sqrt{5}-1}{4}[$? Aufgrund obiger Beobachtungen (p_3 ist negativ für $k \in]\frac{1}{4}; \frac{\sqrt{5}-1}{4}[$) lässt sich vermuten, dass die Minimax-Strategie die Stichprobe $s = \{1,2\}$ nicht auswählt. Unter Berücksichtigung der Repräsentativität und der symmetrischen Bedingungen definieren wir

$$\underline{a}_{\{1,3\}}(\alpha) := (\frac{1-h\alpha}{k}, 0, \alpha)^t, \ \underline{a}_{\{2,3\}}(\alpha) := (0, \frac{1-h\alpha}{k}, \alpha)^t$$

für $\alpha \in \mathbb{R}$ und $p_1^* := p_2^* := \frac{1}{2}$. Wegen Satz 4.4 ist jede Minimax-Strategie eine Indifferenz-Strategie. Um die Natur indifferent zu stellen, suchen wir ein $\alpha \in \mathbb{R}$ und ein $\lambda \in \mathbb{R}$, so dass

$$\frac{1}{2}\langle V_i, \underline{a}_{\{1,3\}}(\alpha)(\underline{a}_{\{1,3\}}(\alpha))^t + \underline{a}_{\{2,3\}}(\alpha)(\underline{a}_{\{2,3\}}(\alpha))^t\rangle = \lambda \quad \forall i = 1,2,3,$$

wobei die V_i definiert sind wie im Abschnitt zuvor, also

$$V_1 = \frac{k}{2}Q_1, \ V_2 = \frac{kh}{k+h}Q_2, \ V_3 = \frac{kh}{k+h}Q_3.$$

Somit ist obiges Gleichungssystem äquivalent zu:

$$\begin{aligned} \lambda &= \frac{1}{2}\frac{(1-h\alpha)^2}{k}, \\ \lambda &= \frac{1}{2}\frac{h(1-2(1-k)\alpha + (2k^2+h)\alpha^2)}{k(1-k)}. \end{aligned}$$

Durch Gleichsetzen erhält man

$$h\alpha^2 = 1,$$

also

$$\alpha = \frac{1}{\sqrt{h}} = \frac{1}{\sqrt{1-2k}}$$

und somit

$$\frac{1-h\alpha}{k} = \frac{1-\sqrt{h}}{k} = \frac{2}{1+\sqrt{h}}.$$

Für den potentiellen Spielwert λ ergibt sich

$$\lambda := \frac{1}{2}\frac{(\sqrt{h}-1)^2}{k} = \frac{1-\sqrt{h}}{1+\sqrt{h}}.$$

(Die Wahl $\alpha = -\frac{1}{\sqrt{h}}$ führt zu einem größeren potentiellen Spielwert.)

Sei

$$\begin{aligned}\underline{a}^*_{\{1,3\}} &:= \underline{a}_{\{1,3\}}(\frac{1}{\sqrt{h}}) = (\frac{1-\sqrt{h}}{k}, 0, \frac{\sqrt{h}}{h})^t,\\ \underline{a}^*_{\{2,3\}} &:= \underline{a}_{\{2,3\}}(\frac{1}{\sqrt{h}}) = (0, \frac{1-\sqrt{h}}{k}, \frac{\sqrt{h}}{h})^t.\end{aligned}$$

Falls eine Matrix $V \in \mathcal{Y}_U$ existiert, die den Statistiker indifferent bzgl. der Stichproben $\{1,3\}$ und $\{2,3\}$ stellt mit

$$\underline{a}^*_s = \mathrm{argmin}_{\underline{a}_s \in A_s} \langle V, (\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t \rangle$$

für $s = \{1,3\}$,$\{2,3\}$, und falls

$$\langle V, (\underline{a} - \underline{1})(\underline{a} - \underline{1})^t \rangle \geq \lambda \quad \forall \underline{a} \in A_{\{1,2\}},$$

ist V eine Maximin-Strategie, $W^* := \lambda U$ bzw.

$$(p^*, \{\underline{a}^*_s; s = \{1,3\}, \{2,3\}\})$$

Minimax-Strategie, und der Wert des Spiels ist gegeben durch λ.

Betrachten wir das projizierte Problem, so suchen wir eine Ellipse $E := \{\underline{z} \in \mathbb{R}^2; \underline{z}^t M_E \underline{z} = 1\}$ in der z_1-z_2-Ebene , die die Geraden $z_1 = -1$ und $z_2 = -1$ in den Punkten $(-1, \frac{1-k-\sqrt{1-2k}}{k})$ bzw. $(\frac{1-k-\sqrt{1-2k}}{k}, -1)$ berührt. Aufgrund der Tangential-Bedingung gilt (vgl. Abschnitt 4.2)

$$M_E = \begin{bmatrix} \Delta & c \\ c & \Delta \end{bmatrix},$$

wobei

$$\Delta = \det(M_E) = \frac{1}{2} + \sqrt{\frac{1}{4} + c^2}.$$

Aus $(\frac{1-k-\sqrt{1-2k}}{k}, -1) M_E (\frac{1-k-\sqrt{1-2k}}{k}, -1)^t = 1$ ergibt sich

$$c = \frac{(\sqrt{1-2k}+k-1)k}{(\sqrt{1-2k}+2k-1)(\sqrt{1-2k}-1)}.$$

Die Projektion der repräsentativen Strategien bzgl. der Stichprobe $\{1,2\}$ ist gegeben durch die Gerade $g: z_2 = -z_1 + \frac{1-2k}{k}$; es verbleibt zu zeigen, dass g die Ellipse nicht schneidet. Aufgrund der symmetrischen Bedingungen verläuft eine Achse der Ellipse entlang der Winkelhalbierenden. Weiterhin ist die Gerade g orthogonal zur Winkelhalbierenden, und der Schnittpunkt ist gegeben durch $(\frac{1}{2k}-1, \frac{1}{2k}-1)$. Es zeigt sich, dass der Schnittpunkt (z_E, z_E) der Ellipse mit der Winkelhalbierenden aus dem ersten Quadranten gegeben ist durch

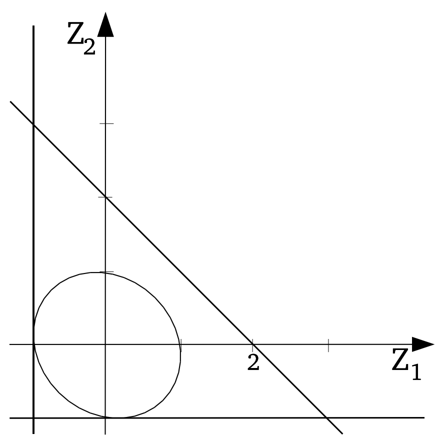
$$z_E = \frac{1}{\sqrt{2(\Delta + c)}},$$

und es gilt

$$z_E < \frac{1}{2k} - 1 \quad \forall k \in]0; \frac{\sqrt{5}-1}{4}[.$$

Somit ist W^* für $0 < k < \frac{\sqrt{5}-1}{4} \approx 0.31$ eine Minimax-Strategie.

Beispiel 4.9 *Die folgende Abbildung veranschaulicht obigen Sachverhalt für den Regressionsvektor* $\underline{x} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{2})^t$.



Zusammenfassung Obwohl für $k \in]\frac{1}{4}; \frac{\sqrt{5}-1}{4}[$ eine Indifferenzmatrix der Natur existiert, ist diese keine Maximin-Strategie. Wir diskutieren im Folgenden die Eigenschaften der Minimax-Strategie.

Falls $k \in]0; \frac{\sqrt{5}-1}{4}]$, ist $p^*(\{1,2\}) = 0$. Die dritte Einheit gelangt immer in die Auswahl, und wegen

$$\alpha - \frac{1-h\alpha}{k} = \frac{(1-\sqrt{h})^2}{2\sqrt{h}k} = \frac{1}{\sqrt{h}}\frac{1-\sqrt{h}}{1+\sqrt{h}} > 0$$

wird der ihr zugeordnete y-Wert stärker gewichtet.

Falls $k \in]\frac{\sqrt{5}-1}{4}; \frac{1}{2}[$, ist $p^*(s) > 0$ für alle Stichproben s vom Umfang zwei:

$$p_1^* = p_2^* = \frac{kh}{2k-h}, \; p_3^* = \frac{4k^2-h}{2k-h}.$$

Dabei sind die optimalen Schätzvektoren gegeben durch

$$\underline{a}^*_{\{1,2\}} = (\frac{1}{2k}, \frac{1}{2k}, 0)^t, \; \underline{a}^*_{\{1,3\}} = (\frac{2k-h}{2k^2}, 0, \frac{1}{2k})^t, \; \underline{a}^*_{\{2,3\}} = (0, \frac{2k-h}{2k^2}, \frac{1}{2k})^t.$$

Sei $h > k$, also $k \in]\frac{\sqrt{5}-1}{4}; \frac{1}{3}[$. Bezeichnen wir mit π_i^* die Inklusionswahrscheinlichkeiten erster Ordnung des Auswahlplans p^*, so ist

$$\pi_3^* - \pi_j^* = \frac{2k+1}{4k-1}(h-k) > 0 \quad \text{für } j = 1, 2.$$

Da

$$\frac{1}{2k} - \frac{2k-h}{2k^2} = \frac{h-k}{2k^2} > 0$$

wird der y-Wert der dritten Einheit stärker gewichtet. Dies ist insofern bemerkenswert, da der Horvitz-Thompson-Schätzer und der HH-Schätzer Merkmalswerte von Einheiten mit kleinen Auswahlwahrscheinlichkeiten stärker gewichten.

Kapitel 5

Allgemeiner Ansatz für Stichproben vom Umfang Zwei

Sei $n = 2$ und $N \geq 2$. In Abschnitt 5.1 geben wir eine hinreichende Bedingung für die Existenz von Indifferenzmatrizen der Natur. In Abschnitt 5.2 bestimmen wir eine Lösung des Spiels $(\mathcal{Y}_U, \mathcal{A}_x)$, falls der Parameterraum gegeben ist durch den HH-Raum und

$$\underline{x} = (k, \ldots, k, h)^t \in \mathbb{R}^N$$

für entsprechend großes h.

5.1 Existenz von Indifferenzmatrizen

Für $i, j \in \{1, \ldots, N\}$, $i < j$ definieren wir die Geraden g_{ij} durch

$$g_{ij} : \nu \frac{1}{x_i} \underline{e}_i + (1 - \nu) \frac{1}{x_j} \underline{e}_j \quad \text{für } \nu \in \mathbb{R}.$$

Die Vereinigung dieser Geraden ist die Menge aller repräsentativen Schätzvektoren. Gesucht ist eine symmetrische positiv semidefinite Matrix Q mit:

1. Die durch $\{\underline{z} \in \mathbb{R}^N; \underline{z}^t Q \underline{z} = 1\}$ definierte Mannigfaltigkeit ist tangential an g_{ij} für alle $i < j$.

2. $Q\underline{1} = \underline{0}$.

Weiter fordern wir, dass Q *nicht entartet* ist, was bedeutet

3. $\dim(\text{Kern}(Q)) = 1$.

Diese Bedingung ist nicht zwingend, wie wir im Fall Eins aus N mit $\max_i x_i = \frac{1}{2}$ gesehen haben und was auch Beispiel 6.8 zeigen wird. Andererseits existiert in einer Umgebung von $\underline{x} = \frac{1}{N}\underline{1}$ eine nicht entartete Indifferenzmatrix, falls der Parameterraum durch den HH-Raum gegeben ist (vgl. Gabler und Stenger [9]). Falls eine solche Matrix Q existiert, ist $V := \frac{Q}{\langle U,Q\rangle} \in \mathcal{Y}_{U1}$ eine Indifferenzmatrix (vgl. Bemerkung 3.9).

Die Gleichung

$$(\nu\frac{1}{x_i}\underline{e}_i - (1-\nu)\frac{1}{x_j}\underline{e}_j)^t Q(\nu\frac{1}{x_i}\underline{e}_i - (1-\nu)\frac{1}{x_j}\underline{e}_j) = 1$$

ist äquivalent zu

$$(\frac{q_{ii}}{x_i^2} + \frac{q_{jj}}{x_j^2} - 2\frac{q_{ij}}{x_ix_j})\nu^2 - 2(\frac{q_{jj}}{x_j^2} - \frac{q_{ij}}{x_ix_j})\nu + \frac{q_{jj}}{x_j^2} - 1 = 0.$$

Aufgrund von Bedingung 2 und 3 ist

$$\frac{q_{ii}}{x_i^2} + \frac{q_{jj}}{x_j^2} - 2\frac{q_{ij}}{x_ix_j} = (\frac{1}{x_i}\underline{e}_i - \frac{1}{x_j}\underline{e}_j)^t Q(\frac{1}{x_i}\underline{e}_i - \frac{1}{x_j}\underline{e}_j) > 0$$

Somit ist Bedingung 1 äquivalent zu

$$(q_{ij} - x_ix_j)^2 = (q_{ii} - x_i^2)(q_{jj} - x_j^2) \quad \forall i \neq j, \tag{5.1}$$

also

$$q_{ij} - x_ix_j = \sigma_{ij}\sqrt{(q_{ii} - x_i^2)(q_{jj} - x_j^2)} \quad \forall i \neq j, \tag{5.2}$$

wobei $\sigma_{ij} = \sigma_{ji} \in \{-1, 1\}$. Weiter ist $q_{ii} - x_i^2 > 0$ für alle i oder $q_{ii} - x_i^2 < 0$ für alle i, d.h.

$$\sigma(q_{ii} - x_i^2) > 0 \quad \forall i = 1, \ldots, N,$$

wobei $\sigma \in \{-1, 1\}$. Der Berührpunkt berechnet sich mittels

$$\nu = \nu_{ij} = \frac{x_i^2 q_{jj} - x_ix_jq_{ij}}{x_i^2q_{jj} + x_j^2q_{ii} - 2x_ix_jq_{ij}}. \tag{5.3}$$

Die zweite Bedingung ist genau dann erfüllt, wenn (da x=1)

$$q_{ii} - x_i^2 + x_i + \sum_{j\neq i}(q_{ij} - x_ix_j) = 0 \quad \forall i = 1, \ldots, N. \tag{5.4}$$

Einsetzen von (5.2) in (5.4) ergibt

$$q_{ii} - x_i^2 + x_i + \sum_{j\neq i}\sigma_{ij}\sqrt{(q_{ii} - x_i^2)(q_{jj} - x_j^2)} = 0 \quad \forall i = 1, \ldots, N,$$

und wir erhalten folgendes Gleichungssystem:

$$x_i + \sqrt{\sigma(q_{ii} - x_i^2)} \sum_{j=1}^{N} \sigma_{ij} \sqrt{\sigma(q_{jj} - x_j^2)} = 0 \quad \forall i = 1, \ldots, N, \tag{5.5}$$

wobei $\sigma_{ii} = \sigma$ für alle i und $\sigma_{ij} = \sigma_{ji} \in \{-1, 1\}$. Definieren wir

$$z_i := \sqrt{\sigma(q_{ii} - x_i^2)} \quad \text{für } i = 1, \ldots, N, \tag{5.6}$$

so ergibt sich aus (5.5) das Gleichungssystem

$$\sum_{j=1}^{N} \sigma_{ij} z_j = -\frac{x_i}{z_i} \quad \forall i = 1, \ldots, N.$$

Also existiert zu einem gegebenen Regressionsvektor $\underline{x}$ eine nicht entartete Indifferenzmatrix der Natur höchstens dann, wenn Vorzeichen $\sigma, \sigma_{ij} \in \{-1, 1\}$ für $i < j$ existieren, so dass das Gleichungssystem gegeben durch

$$\begin{bmatrix} \sigma & \sigma_{12} & \ldots & \sigma_{1N} \\ \sigma_{12} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \sigma_{1N} & \ldots & \ldots & \sigma \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ \vdots \\ z_N \end{bmatrix} = - \begin{bmatrix} \frac{x_1}{z_1} \\ \vdots \\ \vdots \\ \frac{x_N}{z_N} \end{bmatrix}$$

lösbar ist mit $z_i > 0$. Dann ist wegen (5.6)

$$q_{ii} = \sigma z_i^2 + x_i^2$$

für alle $i = 1, \ldots, N$ und wegen (5.2)

$$q_{ij} = \sigma_{ij} z_i z_j + x_i x_j$$

für alle $i \neq j$. Falls die dadurch definierte Matrix Q positiv semidefinit ist, ist Q eine allgemeine Indifferenzmatrix.

Der Berührpunkt der Mannigfaltigkeit mit der Geraden g_{ij} ist gegeben durch $\nu_{ij} \frac{1}{x_i} \underline{e}_i + (1 - \nu_{ij}) \frac{1}{x_j} \underline{e}_j$, wobei

$$\begin{aligned} \nu_{ij} &= \frac{x_i^2(q_{jj} - x_j^2) - x_i x_j (q_{ij} - x_i x_j)}{x_j^2(q_{ii} - x_i^2) + x_i^2(q_{jj} - x_j^2) - 2x_i x_j (q_{ij} - x_i x_j)} \\ &= \frac{\sigma x_i^2 z_j^2 - \sigma_{ij} x_i x_j z_i z_j}{\sigma x_j^2 z_i^2 + \sigma x_i^2 z_j^2 - 2\sigma_{ij} x_i x_j z_i z_j} = \frac{\frac{x_i}{z_i}}{\frac{x_i}{z_i} - \frac{\sigma_{ij}}{\sigma} \frac{x_j}{z_j}}. \end{aligned} \tag{5.7}$$

Falls $\underline{x}$ in einer Umgebung von $\frac{1}{N}\underline{1}$ liegt und der Parameterraum gegeben ist durch den HH-Raum, sind die Komponenten der optimalen Schätzvektoren positiv. (Dass zulässige Strategien mit negativen Gewichten existieren, zeigt uns der zweite Teil von Beispiel 5.2.) Fordern wir also

$$\nu_{ij} \in]0;1[\quad \forall i \neq j,$$

so gilt (wegen $\frac{x_i}{z_i} > 0$ für alle $i = 1, \ldots, N$)

$$\sigma_{ij} = -\sigma \quad \forall i \neq j,$$

also

$$\nu_{ij} = \frac{z_j x_i}{z_j x_i + z_i x_j}, \tag{5.8}$$

und es gibt nur zwei relevante Gleichungssysteme:

$$(-1)^m \begin{bmatrix} 1 & -1 & \ldots & -1 \\ -1 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ -1 & \ldots & \ldots & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ \vdots \\ z_N \end{bmatrix} = - \begin{bmatrix} \frac{x_1}{z_1} \\ \vdots \\ \vdots \\ \frac{x_N}{z_N} \end{bmatrix}$$

für $m \in \{0, 1\}$. Eine kleine Überlegung zeigt, dass das Gleichungssystem für $m = 1$ nicht lösbar ist mit $z_i > 0$ für $i = 1, \ldots, N$. Somit ist nur der Fall $m = 0$ von Interesse:

$$z_i - \sum_{j \neq i} z_j = -\frac{x_i}{z_i} \quad \forall i = 1, \ldots, N.$$

Mit $z := \sum_j z_j$ ergibt sich aus obigem Gleichungssystem

$$z_i^2 - \frac{z}{2} z_i + \frac{x_i}{2} = 0 \quad \forall i = 1, \ldots, N,$$

also

$$z_i = \frac{z + \epsilon_i \sqrt{z^2 - 8x_i}}{4} \quad \forall i = 1, \ldots, N,$$

wobei $\epsilon_i \in \{-1, 1\}$. Durch Summation beider Seiten über i erhalten wir folgendes Fixpunktproblem (vgl. Gabler und Stenger [9]):

$$z = \frac{Nz + \sum_i \epsilon_i \sqrt{z^2 - 8x_i}}{4}$$

bzw.

$$(N-4)z = -\sum_i \epsilon_i \sqrt{z^2 - 8x_i}. \tag{5.9}$$

Existieren Vorzeichen ϵ_i derart, dass obiges Fixpunktproblem lösbar ist mit $z > 0$, so sind die daraus resultierenden $z_i > 0$, da $x_i > 0$. Wir erhalten folgenden Satz:

Satz 5.1 *Sei $N \geq n = 2$, $\underline{x} \in \mathbb{R}^N$ ein Regressionsvektor und $z > 0$ eine Lösung von*

$$(N-4)z = -\sum_i \epsilon_i \sqrt{z^2 - 8x_i}$$

für gewisse Vorzeichen $\varepsilon_i \in \{-1, 1\}$, $i = 1, \ldots, N$. Falls die durch

$$q_{ii} := x_i^2 + z_i^2 \quad \text{für } i = 1, \ldots, N, \tag{5.10}$$

$$q_{ij} := x_i x_j - z_i z_j \quad \text{für } i \neq j \tag{5.11}$$

definierte Matrix Q positiv semidefinit ist, wobei

$$z_i := \frac{z + \epsilon_i \sqrt{z^2 - 8x_i}}{4} \quad \text{für } i = 1, \ldots, N, \tag{5.12}$$

gilt

$$V = \frac{Q}{\langle U, Q \rangle} \in \mathcal{Y}_U \tag{5.13}$$

ist eine Indifferenzmatrix der Natur.

Das folgende Beispiel zeigt, dass in manchen Spielen $(\mathcal{Y}_U, \mathcal{A}_x)$ Indifferenzmatrizen der Natur mit unterschiedlichen Risiken existieren und Schätzvektoren mit negativen Koeffizienten zulässig sein können.

Beispiel 5.2 *Sei $N = 4$, $n = 2$ und $\underline{x} = \frac{1}{4}\underline{1}$. Dann ist das Fixpunktproblem (5.9) gegeben durch*

$$\sum_{i=1}^{4} \epsilon_i \sqrt{z^2 - 2} = 0.$$

Eine spezielle Lösung ist $z = 2$, $\epsilon_1 = \epsilon_3 = 1$ und $\epsilon_2 = \epsilon_4 = -1$. Für die daraus resultierenden z_i ergibt sich mit (5.12):

$$z_1 = z_3 = \frac{2 + \sqrt{2}}{4}, \; z_2 = z_4 = \frac{2 - \sqrt{2}}{4}.$$

Nach (5.10) und (5.11) ist die Matrix Q gegeben durch

$$Q := \frac{1}{16}\begin{bmatrix} a^2 & -2 & -a^2 & -2 \\ -2 & b^2 & -2 & -b^2 \\ -a^2 & -2 & a^2 & -2 \\ -2 & -b^2 & -2 & b^2 \end{bmatrix} + \frac{1}{16}\underline{1}\,\underline{1}^t,$$

wobei

$$a := 2+\sqrt{2},\ \ b := 2-\sqrt{2}.$$

Die Berührpunkte der Geraden g_{ij} mit der durch Q definierten Mannigfaltigkeit $\{\underline{z};\underline{z}^tQ\underline{z}=1\}$ stimmen mit den bzgl. (5.8) berechneten überein; wir erhalten

$$\begin{aligned} \underline{a}_{\{1,2\}} &= (b,a,0,0)^t,\ \underline{a}_{\{1,3\}} = (2,0,2,0)^t,\ \underline{a}_{\{1,4\}} = (b,0,0,a)^t, \\ \underline{a}_{\{2,3\}} &= (0,a,b,0)^t,\ \underline{a}_{\{2,4\}} = (0,2,0,2)^t,\ \underline{a}_{\{3,4\}} = (0,0,b,a)^t. \end{aligned}$$

Tatsächlich ist Q eine Indifferenzmatrix mit den gewünschten Eigenschaften; allerdings ist $V = \frac{Q}{\langle U,Q\rangle}$ i.a. keine Maximin-Strategie, was z.B. die Wahl $U := diag^{-1}(\underline{x}) - \underline{1}\,\underline{1}^t$ zeigt. Nach Beispiel 3 ist die Maximin-Strategie gegeben durch die Indifferenzmatrix $\frac{1}{12}I$, und der Spielwert ist $\frac{1}{3}$; hingegen gilt

$$\langle U,Q\rangle = 7,$$

also

$$\min_{W\in\mathcal{A}_x}\langle V,W\rangle = \frac{1}{7}.$$

Mit $U = diag(\underline{x})^{-1} - \underline{1}\,\underline{1}^t$ ist

$$\Theta_{U1} = \{\underline{y}\in\mathbb{R}^4; \sum_i y_i = 0, \sum_i y_i^2 = \frac{1}{4}\},$$

und es gilt

$$\underline{y}_1 := \frac{1}{2\sqrt{206}}(1,5,6,-12)^t \in \Theta_{U1},\ \underline{y}_2 := \frac{1}{2\sqrt{206}}(1,5,-6,12)^t \in \Theta_{U1}.$$

Daraus folgt

$$V := \frac{1}{2}(\underline{y}_1\underline{y}_1^t + \underline{y}_2\underline{y}_2^t) = \frac{1}{824}\begin{bmatrix} 1 & 5 & -3 & -3 \\ 5 & 25 & -15 & -15 \\ -3 & -15 & 90 & -72 \\ -3 & -15 & -72 & 90 \end{bmatrix} \in \mathcal{Y}_{U1}.$$

Mit $\underline{a}^*_{\{1,2\}} := (5,-1,0,0)^t$ *ist* $\underline{a}^*_{\{1,2\}}$ *repräsentativ und*

$$(\underline{a}^*_{\{1,2\}})^t V \underline{a}^*_{\{1,2\}} = 0.$$

Für alle Stichproben $s \neq \{1,2\}$ *vom Umfang zwei gilt*

$$\underline{a}_s^t V \underline{a}_s > 0 \quad \forall \underline{a}_s \in A_s,$$

d.h. die reine Strategie $\underline{a}^*_{\{1,2\}} = (5,-1,0,0)^t$ *ist zulässig.*

Wählen wir in (5.8) die Vorzeichen $\varepsilon_i := -1$ für $i = 1, \ldots, N$, erhält man folgende Aussage (vgl. Gabler u. Stenger [9], Lemma 3):

Lemma 5.3 (Lösungsbereich des Fixpunktproblems) *Sei* $\varepsilon_i := -1$ *für* $i = 1, \ldots, N$ *und* $x_0 := max\{x_1, \ldots, x_N\}$. *Das Fixpunktproblem (5.8) ist genau dann lösbar mit* $z > 0$, *wenn*

$$N - 4 \geq \sum_{i=1}^{N} \sqrt{1 - \frac{x_i}{x_0}}.$$

Speziell ist die Lösung eindeutig.

Beweis: Sei $z > 2\sqrt{2x_0}$ und

$$f(z) := (N-4)z, \;\; g(z) := \sum_i \sqrt{z^2 - 8x_i}.$$

Es gilt

$$g'(z) = \sum_i \frac{z}{\sqrt{z^2 - 8x_i}} > \sum_i \frac{z}{\sqrt{z^2}} > N > N - 4 = f'(z),$$

d.h. die Funktion $g - f$ ist streng monoton wachsend auf $]2\sqrt{2x_0}, \infty[$ und besitzt somit genau dann eine (eindeutige) Nullstelle, wenn $f(2\sqrt{2x_0}) \geq g(2\sqrt{2x_0})$, also

$$N - 4 \geq \sum_{i=1}^{N} \sqrt{1 - \frac{x_i}{x_0}}.$$

◇

Beispiel 5.4 *Sei* $N > 4$,

$$\underline{x} = (k, \ldots, k, h, h)^t \in \mathbb{R}^N$$

und $h = \frac{1-(N-2)k}{2} \in]0; \frac{1}{4}[$ *bzw.* $k \in]\frac{1}{2(N-2)}; \frac{1}{N-2}[$. *Setzen wir*

$$\epsilon_i = -1 \quad \forall i = 1, \ldots, N,$$

so ist das Fixpunktproblem gegeben durch

$$(N-4)z = (N-2)\sqrt{z^2-8k} + 2\sqrt{z^2-8h}$$

und wird gelöst durch

$$z = \frac{N(N-2)k-2}{\sqrt{(N-2)(N-4)[2(N-2)k-1]}}.$$

Somit gilt (vgl. (5.12))

$$z_i = \begin{cases} \sqrt{\frac{2(N-2)k-1}{(N-2)(N-4)}} = \sqrt{\frac{1-4h}{(N-2)(N-4)}} & : \quad i = 1,\ldots,N-2 \\ \frac{1}{2}\frac{\sqrt{N-4}[1-(N-2)k]}{\sqrt{N-2}\sqrt{2(N-2)k-1}} = \sqrt{\frac{N-4}{(N-2)(1-4h)}}h & : \quad i = N-1, N \end{cases}$$

und (vgl. (5.7))

$$\nu_{ij} \quad = \quad \frac{(N-4)k}{(N-4)k+2(N-2)k-1} = \frac{(N-4)k}{(N-4)k+1-4h}$$

für alle $i = 1,\ldots,N-2, j = N-1, N$*. Die daraus resultierenden Schätzvektoren sind gegeben durch*

$$\underline{a}^*_{\{i,j\}} = \begin{cases} \frac{1}{2k}(\underline{e}_i + \underline{e}_j) & : i,j \in \{1,\ldots,N-2\} \\ \frac{1}{2h}(\underline{e}_i + \underline{e}_j) & : i = N-1, j = N \\ \\ \frac{(N-4)k\frac{\underline{e}_i}{k} + (1-4h)\frac{\underline{e}_j}{h}}{(N-4)k+1-4h} & : i \in \{1,\ldots,N-2\}, j \in \{N-1,N\} \end{cases}.$$

Im Allgemeinen ist die durch z_i *via (5.10), (5.11) und (5.13) definierte Matrix* V_z *keine Maximin-Strategie. Falls der Parameterraum durch den HH-Raum gegeben ist und* $\underline{x}$ *in der Nähe von* $\frac{1}{N}\underline{1}$ *liegt, ist* V_z *eine Maximin-Strategie und die Vektoren* $\underline{a}^*_{\{i,j\}}$ *definieren den Schätzer der Minimax-Strategie. (Vgl. Gabler u. Stenger [9], Bsp. 2.1) Wir werden diese Umgebung bestimmen und die Minimax-Strategie berechnen: Sei* $U := diag(\underline{x})^{-1} - \underline{1}\,\underline{1}^t$*, also*

$$U = \begin{bmatrix} \frac{1-k}{k} & -1 & \ldots & \ldots & \ldots & -1 \\ -1 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \frac{1-k}{k} & \ddots & \vdots \\ \vdots & & & \ddots & \frac{1-h}{h} & -1 \\ -1 & \ldots & \ldots & \ldots & -1 & \frac{1-h}{h} \end{bmatrix}.$$

Aus Symmetriegründen gilt für den Auswahlplan p^* der Minimax-Strategie (vgl. Lemma 1.5)

$$\begin{aligned} p^*(\{i,j\}) &= p_{kk} \quad \forall i < j \le N-2, \\ p^*(\{i,j\}) &= p_{kh} \quad \forall i \le N-2 < j, \\ p^*(\{N-1,N\}) &= p_{hh} \end{aligned}$$

für gewisse $p_{kk}, p_{kh}, p_{hh} \ge 0$ mit

$$\binom{N-2}{2} p_{kk} + p_{hh} + 2(N-2)p_{kh} = 1. \tag{5.14}$$

Wir definieren wie in Abschnitt 3.3

$$Q_{lm} := (\underline{e}_l - \underline{e}_m)(\underline{e}_l - \underline{e}_m)^t \quad \forall 1 \le l < m \le N.$$

Um die Natur indifferent zu stellen, muss gelten (vgl. Lemma 3.12)

$$\sum_s p^*(s)(\underline{a}_s^*)^t Q_{lm} \underline{a}_s^* = \lambda \langle U, Q_{lm} \rangle \quad \forall l < m$$

für ein $\lambda \in \mathbb{R}$. Wegen

$$\begin{aligned} \langle U, Q_{lm} \rangle &= \frac{2}{k} \quad \forall l < m \le N-2, \\ \langle U, Q_{lm} \rangle &= \frac{k+h}{kh} \quad \forall l \le N-2 < m, \\ \langle U, Q_{N-1N} \rangle &= \frac{2}{h}, \end{aligned}$$

erhalten wir die drei Gleichungen:

$$p_{kk}\left(\frac{1}{2k}\right)^2 2(N-4) + p_{kh} 4 \frac{(N-4)^2}{[(N-4)k + (1-4h)]^2} = \frac{2}{k}\lambda, \tag{5.15}$$

$$\begin{aligned} p_{kk}\left(\frac{1}{2k}\right)^2 2(N-3) + p_{kh}[(N-4)k + (1-4h)]^{-2} \Big[2(N-4)^2 + \\ + \frac{(1-4h)^2(N-2)}{h^2} - 2\frac{(N-4)(1-4h)}{h}\Big] + p_{hh}\left(\frac{1}{2h}\right)^2 = \frac{k+h}{kh}\lambda \end{aligned} \tag{5.16}$$

und

$$p_{kh} 2(N-2)\frac{(1-4h)^2}{h^2}[(N-4)k + (1-4h)]^{-2} = \frac{2}{h}\lambda. \tag{5.17}$$

Durch sukzessives Lösen der Gleichungen (5.14), (5.15), (5.16) und (5.17) ergibt sich

$$\begin{aligned}
p_{kh} &= 4\lambda h \frac{(1-k-3h)^2}{(N-2)(1-4h)^2}, \\
p_{kk} &= 4\lambda k \frac{N-2-(10N^2-48N+64)kh}{(N-2)(N-4)(1-4h)^2}, \\
p_{hh} &= 4\lambda h^2 \frac{(3N-8)Nk-(N-2)-(N-2)(N^2+8N-32)k^2}{k(N-2)(N-4)(1-4h)^2}, \\
\lambda &= \frac{1}{2}\frac{(N-2)(N-4)(1-4h)k}{(N-2)(1-k-h)-(5N^2-20N+16)kh}.
\end{aligned}$$

Diese Darstellungen leiten sich nicht trivial durch Einsetzen ab. Es verbleibt zu klären, für welche Werte von k durch p_{kk}, p_{hh}, p_{kh} ein Auswahlplan definiert ist: Sei

$$k_1 := \frac{1}{2(N-2)} + \frac{1}{2(N-2)}\frac{N-4}{\sqrt{5N^2-24N+32}}$$

und

$$k_2 := \frac{1}{2}\frac{N(3N-8)+(N-4)\sqrt{5N^2-24N+32}}{(N-2)(N^2+8N-32)}.$$

Für $4 \le N \le 8$ ist der relevante Bereich gegeben durch $[k_1; k_2]$ (falls $k \in]\frac{1}{2(N-2)}; k_1[$, ist p_{kk} negativ; falls $k \in]k_2; \frac{1}{N-2}[$, ist p_{hh} negativ) und für $N > 8$ durch $[k_1; \frac{1}{N-2}[$. In diesen Fällen ist die durch z_i via (5.10), (5.11) und (5.13) definierte Matrix V eine Maximin-Strategie, $(p^*, \{\underline{a}_s^*\})$ Minimax-Strategie, und der Wert des Spiels ist gegeben durch λ.

5.2 Der einfache Zwei-Schichten-Fall

Wie wir in Abschnitt 4.6 gesehen haben, wird man unter gewissen Umständen Einheiten bewusst auswählen. In Hinblick auf dieses Ergebnis untersuchen wir den einfachen Zwei-Schichten-Fall: Sei

$$\underline{x} := (k, \ldots, k, h)^t \in \mathbb{R}^N$$

und $h = 1-(N-1)k > k > 0$. Für $1 \le l < m \le N$ sei (vgl. Abschnitt 3.3)

$$Q_{lm} = (\underline{e}_l - \underline{e}_m)(\underline{e}_l - \underline{e}_m)^t.$$

Für eine beliebige Matrix $B \in \mathcal{Q}$ ist

$$\langle B, Q_{lm} \rangle = b_{ll} + b_{mm} - 2b_{lm}. \tag{5.18}$$

Aufgrund der symmetrischen Bedingungen definieren wir

$$\underline{a}_{\{i,N\}} := \beta \underline{e}_i + \alpha \underline{e}_N,$$

und

$$A_i := \underline{a}_{\{i,N\}} \underline{a}^t_{\{i,N\}}$$

für $i = 1, \ldots, N-1$, wobei wegen der Repräsentativität gilt

$$\beta = \frac{1 - h\alpha}{k}. \tag{5.19}$$

Die Spur von $A_i Q_{lm}$ ist gegeben durch

$$\langle A_i, Q_{lm} \rangle = \begin{cases} 0 & : \; i \notin \{l, m\}, m \neq N \\ \beta^2 & : \; i \in \{l, m\}, m \neq N \\ \alpha^2 & : \; i \neq l, m = N \\ (\alpha - \beta)^2 & : \; i = l, m = N \end{cases},$$

und es folgt

$$\sum_{i=1}^{N-1} \langle A_i, Q_{lm} \rangle = \begin{cases} 2\beta^2 & : \; m < N \\ (N-1)\alpha^2 + \beta^2 - 2\alpha\beta & : \; m = N \end{cases}.$$

Unter der Annahme, dass die N-te Einheit immer in die Auswahl gelangt, definieren wir

$$p^*(s) := \begin{cases} \frac{1}{N-1} & : \; N \in s \\ 0 & : \; \text{sonst} \end{cases}$$

und erhalten das Gleichungssystem

$$\sum_s p^*(s) \underline{a}_s^t Q_{lm} \underline{a}_s = \lambda \langle U, Q_{lm} \rangle \quad \forall l < m$$

bzw.

$$\frac{1}{N-1} \sum_{i=1}^{N-1} \frac{\langle A_i, Q_{lm} \rangle}{\langle U, Q_{lm} \rangle} = \lambda \quad \forall l < m.$$

Wählen wir speziell

$$U := \operatorname{diag}(\underline{x})^{-1} - \underline{1}\,\underline{1}^t,$$

so ist

$$\langle U, Q_{lm} \rangle = \begin{cases} \frac{2}{k} & : \; m < N \\ \frac{k+h}{kh} & : \; m = N \end{cases},$$

und obiges Gleichungssystem reduziert sich auf die folgenden zwei Gleichungen:

$$\lambda = \frac{k}{N-1}\beta^2, \tag{5.20}$$

$$\lambda = \frac{kh}{k+h}(\alpha^2 + \frac{1}{N-1}(\beta^2 - 2\alpha\beta)). \tag{5.21}$$

Durch Gleichsetzen ergibt sich

$$\alpha = \frac{\sqrt{h}}{h}, -\frac{\sqrt{h}}{h},$$

also

$$\underline{a}^*_{\{i,N\}} = \frac{1-\sqrt{h}}{k}\underline{e}_i + \frac{\sqrt{h}}{h}\underline{e}_N \quad \text{für } i = 1, \ldots, N-1,$$

und der potentielle Spielwert ist gegeben durch

$$\lambda = \frac{(1-\sqrt{h})^2}{(N-1)k}.$$

Dabei ist λ in jedem Fall eine obere Schranke für den Spielwert. (Die Wahl $\alpha = -\frac{\sqrt{h}}{h}$ führt zu dem potentiellen Spielwert $\lambda_- = \frac{(1+\sqrt{h})^2}{(N-1)k} > \lambda$, kann also ausgeschlossen werden.) Damit $(p^*, \{\underline{a}^*_s\})$ eine Minimax-Strategie ist, muss eine positiv semidefinite Matrix Q existieren mit:

1. Die Mannigfaltigkeit $\{\underline{z}, \underline{z}^t Q \underline{z} = 1\}$ berührt die Geraden g_{iN} in

   $$(0, \ldots, 0, \frac{1-\sqrt{h}}{k}, 0, \ldots, 0, \frac{\sqrt{h}}{h})^t \in \mathbb{R}^N$$

   für $i = 1, \ldots, N-1$.

2. $\{\underline{z}, \underline{z}^t Q \underline{z} = 1\}$ schneidet nicht die Geraden g_{ij} für $i < j < N$.

3. $Q\underline{1} = \underline{0}$.

Aufgrund der symmetrischen Bedingungen können wir annehmen, dass Q von der Gestalt

$$Q = \begin{bmatrix} q_d & b & \ldots & b & c \\ b & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & b & \vdots \\ b & \ldots & b & q_d & c \\ c & \ldots & \ldots & c & q_{NN} \end{bmatrix} \tag{5.22}$$

ist, wobei aufgrund von Bedingung 3 gilt

$$q_{NN} = -(N-1)c \tag{5.23}$$

und

$$b = -\frac{q_d + c}{N-2}. \tag{5.24}$$

Die Matrix Q besitzt die Eigenvektoren $\underline{e}_1 - \underline{e}_i$, $i = 2, \ldots, N-1$, zum Eigenwert $q_d - b = \frac{(N-1)q_d + c}{N-2}$, $(1, \ldots, 1, -(N-1))^t$ zum Eigenwert $-Nc$ und $\underline{1}$ zum Eigenwert 0. Damit $Q \in \mathcal{Q}_0$, muss also gelten

$$0 \geq c \geq -(N-1)q_d.$$

Weiter ist (da $Q \neq 0$)

$$\frac{q_d}{k^2} + \frac{q_{NN}}{h^2} - 2\frac{c}{kh} = \frac{q_d}{k^2} - \frac{1+h}{h^2 k}c > 0,$$

und die Tangential-Bedingung aus 1. ist äquivalent zu (vgl. (5.1))

$$(c - kh)^2 = -(q_d - k^2)((N-1)c - h^2). \tag{5.25}$$

Die Berührpunkt-Bedingung aus 1. ist äquivalent zu (vgl. (5.3))

$$\sqrt{h} = \frac{h^2 q_d - khc}{h^2 q_d - (1+h)kc}. \tag{5.26}$$

Die Lösung der Gleichungen (5.25) und (5.26) ist gegeben durch

$$c = -(1 + \sqrt{h})\sqrt{h}\frac{h}{N-1} \tag{5.27}$$

und

$$q_d = \frac{(1 + h\sqrt{h})}{1 - \sqrt{h}} \frac{k}{N-1}. \tag{5.28}$$

Speziell folgt

$$(N-1)q_d + c = \frac{1 + \sqrt{h}}{N-1} > 0,$$

und somit ist die durch (5.22), (5.23), (5.24), (5.27) und (5.28) definierte Matrix Q positiv semidefinit.

Überlegen wir uns, für welche Werte von h die zweite Bedingung erfüllt ist: Sei $i < j < N$. Es ist

$$\begin{aligned}(\nu\frac{1}{k}\underline{e}_i + (1-\nu)\frac{1}{k}\underline{e}_j)^t Q(\nu\frac{1}{k}\underline{e}_i + (1-\nu)\frac{1}{k}\underline{e}_j) = \\ = \frac{1}{k^2}[2(q_d - b)\nu^2 - 2(q_d-b)\nu + q_d].\end{aligned}$$

Da $q_d - b = \frac{1+\sqrt{h}}{(N-1)(N-2)} > 0$, ist der kleinste Vektor auf g_{ij} bezüglich der durch Q definierten quadratischen Form gegeben durch $\nu = \nu_{ij} = \frac{1}{2}$. Somit ist Q eine Maximin-Strategie (und $(p^*, \{\underline{a}_s^*\})$ eine Minimax-Strategie), wenn

$$1 \leq \frac{1}{2k^2}(q_d + b) = \frac{1}{2}\frac{N-3+2(N-2)h\sqrt{h}}{(N-2)(1-h)(1-\sqrt{h})},$$

was für alle

$$1 - \frac{\sqrt{(3N-4)(N-2)}-1}{2(N-2)} < h < 1$$

(bzw. $k \in]0; \frac{1}{2}\frac{\sqrt{(3N-4)(N-2)}-1}{(N-1)(N-2)}[$) erfüllt ist. Für großes N folgt näherungsweise

$$0.134 \approx \frac{2-\sqrt{3}}{2} < h < 1.$$

Zusammenfassend erhalten wir folgenden Satz:

Satz 5.5 *Sei* $\underline{x} := (k, \ldots, k, h) \in \mathbb{R}^N$ *ein Regressionsvektor,* $U := diag^{-1}(\underline{x}) - \underline{1}\,\underline{1}^t$. *Wir betrachten das Spiel* $(\mathcal{Y}_U, \mathcal{A}_x)$. *Für* $k \in]0; \frac{1}{2}\frac{\sqrt{(3N-4)(N-2)}-1}{(N-1)(N-2)}[$ *ist* $(p^*, \{\underline{a}_s^*\})$ *definiert durch*

$$p^*(s) := \begin{cases} \frac{1}{N-1} & : \quad N \in s \\ 0 & : \quad sonst \end{cases},$$

$$\underline{a}^*_{\{i,N\}} := \frac{1-\sqrt{h}}{k}\underline{e}_i + \frac{\sqrt{h}}{h}\underline{e}_N \quad \text{für } i = 1, \ldots, N-1$$

eine Minimax-Strategie, eine Maximin-Strategie ist gegeben durch (5.22), (5.23), (5.24), (5.27) und (5.28), und $\frac{(1-\sqrt{h})^2}{(N-1)k}$ *ist der Wert des Spiels.*

Kapitel 6

Schichtungsstrategien als Minimax-Strategien

In den zwei Abschnitten dieses Kapitels diskutieren wir, unter welchen Bedingungen Schichtungsstrategien Minimax-Strategien sind. Dies ist nicht automatisch der Fall, wenn sich die Gesamtheit mittels eines Regressionsvektors in Schichten zerlegen lässt, wie wir in Abschnitt 5.2 gesehen haben (vgl. Satz 5.5).

6.1 Die Regressionsmatrix als Blockmatrix

Sei $H \in \mathbb{N}$, $N(h) \in \mathbb{N}$ mit $N(h) \geq 2$ für $h = 1, \dots, H$, so dass $\sum_{h=1}^{H} N(h) = N$. Definieren wir

$$\mathcal{U}(h) := \{\sum_{k=1}^{h-1} N(k) + 1, \dots, \sum_{k=1}^{h} N(k)\},$$

so ist $\{\mathcal{U}(h); h = 1, \dots, H\}$ eine Partition von $\mathcal{U}$ mit $N(h) = |\mathcal{U}(h)|$. Mit

$$\underline{1}_{N(h)} := \underline{1} \in \mathbb{R}^{N(h)},\ \underline{0}_{N(h)} := \underline{0} \in \mathbb{R}^{N(h)}$$

ist durch

$$X := \begin{bmatrix} \underline{1}_{N(1)} & \underline{0}_{N(1)} & \cdots & \underline{0}_{N(1)} \\ \underline{0}_{N(2)} & \underline{1}_{N(2)} & & \underline{0}_{N(2)} \\ \vdots & & \ddots & \vdots \\ \underline{0}_{N(H)} & \cdots & \cdots & \underline{1}_{N(H)} \end{bmatrix}$$

offensichtlich eine Regressionsmatrix gegeben. Bei dieser Wahl von X finden wir nicht zu jeder Stichprobe $s \in S(n)$ repräsentative Schätzvektoren: Als notwendige und hinreichende Bedingung für $A_s \neq \emptyset$ ergibt sich offensichtlich

$$|s \cap \mathcal{U}(h)| \geq 1 \quad \forall h = 1, \ldots, H,$$

d.h. die Stichprobe muss aus jeder Schicht mindestens eine Einheit enthalten. Somit muss der Stichprobenumfang n größer oder gleich der Anzahl H der Schichten sein, falls wir repräsentativ schätzen wollen.

Sei

$$U(h) := \frac{1}{N(h)-1} I_{N(h)} - \frac{1}{N(h)(N(h)-1)} \underline{1}_{N(h)} \underline{1}^t_{N(h)} \quad \text{für } h = 1, \ldots, H,$$

wobei $I_{N(h)} \in M(N(h) \times N(h))$ die Einheitsmatrix bezeichnet. Für einen Vektor $\underline{\alpha} \in \mathbb{R}^H$ mit $\alpha_h > 0$ und $\sum_{h=1}^H \alpha_h = 1$ definieren wir die Blockmatrix U_α durch

$$U_\alpha := \begin{bmatrix} \alpha_1 U(1) & 0 & \ldots & 0 \\ 0 & \alpha_2 U(2) & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & \alpha_H U(H) \end{bmatrix}.$$

Betrachten wir das Superpopulationsmodell

$$\underline{Y} = X\underline{\beta} + \underline{\epsilon}$$

mit $\underline{\beta} \in \mathbb{R}^H$, $\mathrm{E}(\underline{\epsilon}) = \underline{0}$ und $\mathrm{var}(\underline{\epsilon}) = \Omega_\alpha$ für

$$\Omega_\alpha := \begin{bmatrix} \frac{N(1)-1}{\alpha_1} I_{N(1)} & 0 & \ldots & 0 \\ 0 & \frac{N(2)-1}{\alpha_2} I_{N(2)} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & \frac{N(H)-1}{\alpha_2} I_{N(H)} \end{bmatrix},$$

so gilt

$$U_\alpha = \Omega_\alpha^{-1} - \Omega_\alpha^{-1} X (X^t \Omega_\alpha^{-1} X)^{-1} X^t \Omega_\alpha^{-1}.$$

Mit

$$\overline{y}(h) := \frac{1}{N(h)} \sum_{i \in \mathcal{U}(h)} y_i,$$

$$s_{yy}(h) := \frac{1}{N(h)-1} \sum_{i \in \mathcal{U}(h)} (y_i - \overline{y}(h))^2$$

für $\underline{y} \in \mathbb{R}^N$ gilt

$$\underline{y}^t U_\alpha \underline{y} = \sum_{h=1}^H \alpha_h s_{yy}(h).$$

Wir erhalten

$$\Theta_\alpha := \{\underline{y} \in \mathbb{R}^N; \sum_{h=1}^{H} \alpha_h s_{yy}(h) = 1\}$$

als Parameterraum der Natur.

Aufgrund der symmetrischen Bedingungen des Parameterraums Θ_α in den Schichten $\mathcal{U}(1), \ldots, \mathcal{U}(H)$ und der Repräsentativität ist die Minimax-Strategie eine Schichtungsstrategie bzgl. einer (zufälligen) Aufteilung: Sei $\underline{n}(.)$ ein H-dimensionaler Zufallsvektor mit Realisationen $(n(1), \ldots, n(H))$ in $\mathbb{N}^H$, so dass $\sum_h n(h) = n$ und $n(h) \leq N(h)$ für $h = 1, \ldots, H$. In $\mathcal{U}(h)$ werden $n(h)$ Einheiten ohne Zurücklegen gezogen, und man benutzt wegen der Repräsentativität

$$\sum_h N(h) \overline{y}_s(h)$$

als Schätzung für y, wobei $\overline{y}_s(h)$ für $h = 1, \ldots, H$ die aus $\underline{n}(.)$ resultierenden Stichprobenmittel bezeichnen. Der erwartete quadratische Verlust ist gegeben durch

$$\sum_h s_{yy}(h) N(h)^2 \mathrm{E}(\frac{1}{n(h)} - \frac{1}{N(h)}).$$

Offensichtlich ist die Maximin-Strategie von der Gestalt (vgl. Beispiel 1.7)

$$V^* = \begin{bmatrix} s_{yy}(1) I_{N(1)} & 0 & \ldots & 0 \\ 0 & s_{yy}(2) I_{N(2)} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & s_{yy}(H) I_{N(H)} \end{bmatrix}$$

für gewisse $s_{yy}(h) > 0$, $h = 1, \ldots, H$, mit

$$\sum_{h=1}^{H} \alpha_h s_{yy}(h) = \langle V^*, U_\alpha \rangle = 1.$$

Wir können uns also vorstellen, dass die Natur einen Varianzvektor $\underline{s}_{yy} \in \widehat{\Theta}_\alpha$ wählt, wobei

$$\widehat{\Theta}_\alpha := \{\underline{s}_{yy} = (s_{yy}(1), \ldots, s_{yy}(H))^t \in \mathbb{R}_+^H; \sum_h \alpha_h s_{yy}(h) = 1\}.$$

Für einen festen Varianzvektor $\underline{s}_{yy} \in \widehat{\Theta}_\alpha$ wird das Minimum von

$$\underline{n} \mapsto \sum_h N(h)^2 (\frac{1}{n(h)} - \frac{1}{N(h)}) s_{yy}(h)$$

angenommen in $\underline{n}^*$ definiert durch

$$n^*(h) := n \frac{N(h)\sqrt{s_{yy}(h)}}{\sum_k N(k)\sqrt{s_{yy}(k)}} \quad \text{für } h = 1, \ldots, H, \tag{6.1}$$

d.h. $\underline{n}^*$ entspricht der so genannten varianzoptimalen Aufteilung (vgl. z.B. Stenger [22], (6) auf S. 120). Dabei vernachlässigen wir, dass $n^*(h)$ i.a. keine natürliche Zahl ist.

Für eine (verallgemeinerte) Aufteilung $\underline{n} = (n(1), \ldots, n(H))^t \in \mathbb{R}_+^H$ definieren wir folgende (konvexe) Transformation t:

$$\begin{aligned} t : \mathbb{R}_+^H &\to \mathbb{R}_+^H, \\ n(h) &\mapsto N(h)^2\left(\frac{1}{n(h)} - \frac{1}{N(h)}\right). \end{aligned}$$

Wir erhalten ein lineares Spiel: Die Natur wählt einen Varianzvektor $\underline{s}_{yy} \in \widehat{\Theta}_\alpha$ und der Statistiker eine transformierte Aufteilung $t(\underline{n}) \in t(\mathcal{N})$, wobei $\mathcal{N}$ die Menge aller Aufteilungen von n bezeichnet. Der Verlust ist gegeben durch

$$\text{MSE}(\underline{s}_{yy}, \underline{n}) = (\underline{s}_{yy}, t(\underline{n}))_2.$$

Spielt der Statistiker eine Aufteilung $\underline{n}$ mit $t(\underline{n}) = \lambda\underline{\alpha}$ für ein $\lambda > 0$, also

$$n(h) = \frac{N(h)^2}{\lambda\alpha_h + N(h)} \quad \text{für } h = 1, \ldots, H,$$

so gilt

$$\text{MSE}(\underline{s}_{yy}, \underline{n}) = \lambda(\underline{s}_{yy}, \underline{\alpha})_2 = \lambda \quad \forall \underline{s}_{yy} \in \widehat{\Theta}_\alpha,$$

d.h. der Statistiker stellt die Natur indifferent. Dabei ist λ eindeutig festgelegt durch

$$\underline{1}^t t^{-1}(\lambda\underline{\alpha}) = \sum_{h=1}^{H} \frac{N(h)^2}{\lambda\alpha_h + N(h)} = n. \tag{6.2}$$

(Offensichtlich ist die Funktion $\lambda \mapsto \sum_{h=1}^{H} \frac{N^2(h)}{\lambda\alpha_h + N(h)}$ streng monoton fallend, nimmt in $\lambda = 0$ den Wert N an und konvergiert für λ gegen ∞ gegen 0.)

Zur Bestimmung der Maximin-Strategie konstruieren wir die Tangentialebene an den Strategienraum des Statistikers in $\lambda\underline{\alpha}$: Sei

$$\phi \ : \ \{(n(1), \ldots, n(H-1)); n(h) > 0, \sum_{h=1}^{H-1} n(h) \leq n\} \to \mathbb{R}^H,$$

$$(n(1), \ldots, n(H-1)) \mapsto (n(1), \ldots, n(H-1), n - \sum_{h=1}^{H-1} n(h)).$$

Dann ist

$$\tilde{t} := t \circ \phi$$

eine Parameterdarstellung des Strategienraums des Statistikers, und die Tangentialebene in $\lambda\underline{\alpha}$ wird aufgespannt durch die Vektoren

$$\frac{\partial \tilde{t}}{\partial n(h)}(n^*(1), \ldots, n^*(H-1)) = (\frac{N(H)}{n^*(H)})^2 \underline{e}_H - (\frac{N(h)}{n^*(h)})^2 \underline{e}_h,$$

$h = 1, \ldots, H-1$, wobei

$$\underline{n}^* = t^{-1}(\lambda\underline{\alpha}).$$

Ein Orthogonalenvektor $\underline{z}$ an die Tangentialebene ist gegeben durch

$$z_h := (\frac{n^*(h)}{N(h)n^*(H)})^2 \quad \text{für } h = 1, \ldots, H.$$

Offensichtlich existiert ein $c > 0$ mit

$$\underline{s}^*_{yy} := c\underline{z} \in \widehat{\Theta}_\alpha;$$

somit ist $\underline{s}^*_{yy}$ wegen (6.1) Maximin-Strategie, $\underline{n}^*$ Minimax-Strategie und λ der Wert des Spiels.

Satz 6.1 (Die Minimax-Aufteilung) *Sei $N \geq n \geq H \in \mathbb{N}$, $\underline{\alpha} \in \mathbb{R}^H$ ein Wahrscheinlichkeitsvektor mit $\alpha_h > 0$ für $h = 1, \ldots, H$ und $\lambda > 0$ mit*

$$\sum_{h=1}^{H} \frac{N(h)^2}{\lambda\alpha_h + N(h)} = n.$$

Dann ist die Schichtungsstrategie bzgl. der Aufteilung $\underline{n}^*$ definiert durch

$$n^*(h) := \frac{N(h)^2}{\lambda\alpha_h + N(h)} \quad \textit{für } h = 1, \ldots, H$$

(unter Vernachlässigung der Ganzzahligkeit) eine Minimax-Strategie bzgl. des Parameterraums $\Theta_\alpha = \{\underline{y} \in \mathbb{R}^N; \sum_{h=1}^H \alpha_h s_{yy}(h) = 1\}$ und $\mathcal{A}_X$. Der Wert des Spiels ist λ.

Im Allgemeinen wird sich die so berechnete Minimax-Strategie $\underline{n}^*$ nicht aus ganzzahligen Komponenten zusammensetzen. Um ganzzahlige Aufteilungen zu erhalten, wird man zu zufälligen Aufteilungen übergehen oder entsprechende Rundungsverfahren anwenden. Bei den folgenden Beispielen vernachlässigen wir das Ganzzahligkeitsproblem.

Beispiel 6.2 (Die proportionale Aufteilung) *Sei $\underline{\alpha}$ gegeben durch*

$$\alpha_h := \frac{N(h)}{N} \quad \text{für } h = 1, \ldots, H.$$

Dann ist

$$\lambda = \frac{N(N-n)}{n}$$

und somit

$$n^*(h) = n\frac{N(h)}{N} \quad \text{für } h = 1, \ldots, H,$$

d.h. die Minimax-Strategie entspricht der so genannten proportionalen Aufteilung, und die Maximin-Strategie ist gegeben durch

$$\underline{s}^*_{yy} = \underline{1} \in \mathbb{R}^H.$$

Beispiel 6.3 *Sei $H = 2$ und $\underline{\alpha} \in \mathbb{R}^2_+$ definiert durch*

$$\alpha_h := \frac{N(h)^2}{N(1)^2 + N(2)^2} \quad \text{für } h = 1, 2.$$

Aus (6.2) ergibt sich für den Spielwert

$$\begin{aligned} \lambda &= \frac{(\sqrt{n^2N^2 + 4N(1)N(2)[N(1)N(2) - n^2]} + 2N(1)N(2) - nN)}{2nN(1)N(2)} \times \\ &\quad \times (N(1)^2 + N(2)^2), \end{aligned}$$

und die Minimax-Strategie ist gegeben durch die Aufteilung

$$n^*(1) = \frac{2nN(1)N(2)}{\sqrt{n^2N^2 + 4N(1)N(2)[N(1)N(2) - n^2]} + 2N(1)N(2) + n(N(2) - N(1))},$$

$$n^*(2) = \frac{2nN(1)N(2)}{\sqrt{n^2N^2 + 4N(1)N(2)[N(1)N(2) - n^2]} + 2N(1)N(2) + n(N(1) - N(2))}.$$

Beispiel 6.4 *Sei $H = 2$, $\gamma \in]0;1[$ und $n \leq \min\{\frac{N(1)}{\gamma}, \frac{N(2)}{1-\gamma}\}$. Dann ist $\underline{\alpha} \in \mathbb{R}^2$ definiert durch*

$$\begin{aligned} \alpha_1 &:= \frac{(1-\gamma)N(1)(N(1) - n\gamma)}{(1-\gamma)N(1)^2 + \gamma N(2)^2 - \gamma(1-\gamma)nN}, \\ \alpha_2 &:= 1 - \alpha_1 = \frac{\gamma N(2)(N(2) - n(1-\gamma))}{(1-\gamma)N(1)^2 + \gamma N(2)^2 - \gamma(1-\gamma)nN} \end{aligned}$$

ein Wahrscheinlichkeitsvektor. Aus (6.2) ergibt sich für den Spielwert

$$\lambda = \frac{(1-\gamma)N(1)^2 + \gamma N(2)^2 - \gamma(1-\gamma)nN}{\gamma(1-\gamma)nN^2},$$

und die Minimax-Strategie ist gegeben durch die Aufteilung

$$n(1)^* = n\gamma,\ n(2)^* = n(1-\gamma).$$

Oftmals ist der Stichprobenumfang im Verhältnis zur Größe der einzelnen Schichten sehr klein. Unter diesen Umständen liegt die Aufteilung $\underline{\tilde{n}}$ definiert durch

$$\tilde{n}(h) := n\frac{\frac{N(h)^2}{\alpha_h}}{\sum_{k=1}^{H}\frac{N(k)^2}{\alpha_k}} \quad \text{für } h = 1, \ldots, H \tag{6.3}$$

in der Nähe der Minimax-Aufteilung:

Satz 6.5 (Abschätzung der Minimax-Aufteilung) *Sei*

$$\frac{N(h)}{n} \geq \gamma \quad \forall h = 1, \ldots, H$$

für ein $\gamma \in \mathbb{N}$. Dann gilt unter den Voraussetzungen von Satz 6.1 für die Minimax-Aufteilung $\underline{n}^*$

$$\tilde{n}(h) - \frac{\tilde{n}(h)}{\gamma+1} < n^*(h) < \tilde{n}(h) + \frac{n-\tilde{n}(h)}{\gamma+1} \quad \forall h = 1, \ldots, H,$$

wobei $\tilde{n}(h) = n\frac{\frac{N(h)^2}{\alpha_h}}{\sum_{k=1}^{H}\frac{N(k)^2}{\alpha_k}}$ für $h = 1, \ldots, H$, und für den Spielwert λ

$$\lambda < \frac{1}{n}\sum_{h=1}^{H}(\frac{N(h)}{N})^2\frac{1}{\alpha_h}.$$

Beweis: Sei $\tau_\alpha : [0;\infty[\to \mathbb{R}$ definiert durch

$$\tau_\alpha(z) := \underline{1}^t t^{-1}(z\underline{\alpha}) = \sum_{h=1}^{H}\frac{N(h)^2}{z\alpha_h N^2 + N(h)}.$$

Mit

$$\tilde{\lambda} := \frac{1}{n}\sum_{h=1}^{H}(\frac{N(h)}{N})^2\frac{1}{\alpha_h} = \frac{\sum_{h=1}^{H}\frac{N(h)^2}{\alpha_h}}{N^2 n} \tag{6.4}$$

ist

$$\begin{aligned}\tau_\alpha(\tilde{\lambda}) &= \sum_{h=1}^{H} \frac{N(h)^2}{\frac{\alpha_h}{n}\sum_{k=1}^{H}\frac{N(k)^2}{\alpha_k} + N(h)} \\ &< n\sum_{h=1}^{H} \frac{N(h)^2}{\alpha_h \sum_{k=1}^{H}\frac{N(k)^2}{\alpha_k}} = n = \tau_\alpha(\lambda).\end{aligned}$$

Aufgrund der Monotonie von τ_α gilt

$$\lambda < \tilde{\lambda}, \tag{6.5}$$

d.h. $\tilde{\lambda}$ ist eine obere Schranke für den Spielwert λ. Da nach Voraussetzung

$$\frac{N(h)}{\gamma} \geq n \quad \forall h = 1, \ldots, H$$

für ein $\gamma \in \mathbb{N}$, ist

$$\begin{aligned}\sum_{k=1}^{H}\frac{N(k)^2}{\alpha_k} + n\frac{N(h)}{\alpha_h} &= \sum_{k\neq h}\frac{N(k)^2}{\alpha_k} + (N(h)+n)\frac{N(h)}{\alpha_h} \\ &\leq \sum_{k\neq h}\frac{N(k)^2}{\alpha_k} + \frac{\gamma+1}{\gamma}\frac{N(h)}{\alpha_h} \\ &< \frac{\gamma+1}{\gamma}\sum_{k=1}^{H}\frac{N(k)^2}{\alpha_k},\end{aligned} \tag{6.6}$$

und wir erhalten mit (6.5) eine untere Schranke für die Minimax-Aufteilung $\underline{n}^*$:

$$\begin{aligned} n^*(h) &= \frac{N(h)^2}{\lambda\alpha_h N^2 + N(h)} > \frac{N(h)^2}{\tilde{\lambda}\alpha_h N^2 + N(h)} = n\frac{N(h)^2}{\alpha_h(\sum_{k=1}^{H}\frac{N(k)^2}{\alpha_k} + n\frac{N(h)}{\alpha_h})} \\ &> \frac{\gamma}{\gamma+1}n\frac{N(h)^2}{\alpha_h\sum_{k=1}^{H}\frac{N(k)^2}{\alpha_k}} = \frac{\gamma}{\gamma+1}\tilde{n}(h)\end{aligned}$$

für $h = 1, \ldots, H$. Da $\sum_h n^*(h) = \sum_h \tilde{n}(h) = n$, folgt

$$\tilde{n}(h) = n - \sum_{k\neq h}\tilde{n}(k) > n - \frac{\gamma+1}{\gamma}\sum_{k\neq h} n^*(k) = \frac{\gamma+1}{\gamma}n^*(h) - \frac{n}{\gamma}$$

und somit

$$n^*(h) - \frac{n}{\gamma+1} < \frac{\gamma}{\gamma+1}\tilde{n}(h) < n^*(h),$$

also

$$n^*(h) - \frac{n-n^*(h)}{\gamma} < \tilde{n}(h) < n^*(h) + \frac{n^*(h)}{\gamma} \quad \forall h = 1, \ldots, H$$

bzw.

$$\tilde{n}(h) - \frac{\tilde{n}(h)}{\gamma+1} < n^*(h) < \tilde{n}(h) + \frac{n-\tilde{n}(h)}{\gamma+1} \quad \forall h = 1,\ldots,H.$$

$\diamond$

Weiter erhalten wir eine untere Schranke für den Spielwert: Wegen (6.6) ist

$$\begin{aligned} \tau_\alpha(\tilde{\lambda}) &= n\sum_{h=1}^{H} \frac{N(h)^2}{\alpha_h(\sum_{k=1}^{H}\frac{N(k)^2}{\alpha_k} + n\frac{N(h)}{\alpha_h})} \\ &> \frac{\gamma}{\gamma+1} n \sum_{h=1}^{H} \frac{N(h)^2}{\alpha_h \sum_{k=1}^{H}\frac{N(k)^2}{\alpha_k}} = \frac{\gamma}{\gamma+1} n, \end{aligned}$$

also

$$\tau_\alpha(\lambda) > \tau_\alpha(\tilde{\lambda}) > \frac{\gamma}{\gamma+1}\tau_\alpha(\lambda).$$

Da

$$\tau'_\alpha(z) = \frac{d}{dz}\tau_\alpha(z) = -N^2\sum_{h=1}^{H}\frac{N(h)^2\alpha_h}{(z\alpha_h N^2 + N(h))^2},$$

ist $|\tau'_\alpha|$ streng monoton fallend auf $[0;\infty[$, und es gilt

$$|\tau'_\alpha(\tilde{\lambda})| = N^2\sum_{h=1}^{H}\frac{N(h)^2\alpha_h}{(\frac{\alpha_h}{n}\sum_k\frac{N(k)^2}{\alpha_k} + N(h))^2} \geq \frac{\gamma^2}{(\gamma+1)^2}\frac{N^2 n^2}{\sum_h \frac{N(h)^2}{\alpha_h}}.$$

Mit dem Mittelwertsatz der Differentialrechnung folgt

$$\tilde{\lambda} - \lambda \leq \frac{\tau_\alpha(\lambda) - \tau_\alpha(\tilde{\lambda})}{|\tau'_\alpha(\tilde{\lambda})|} \leq \frac{n}{\gamma+1}\frac{1}{|\tau'_\alpha(\tilde{\lambda})|} = \frac{\gamma+1}{\gamma^2}\frac{\sum_h\frac{N(h)^2}{\alpha_h}}{nN^2},$$

also

$$\lambda < \tilde{\lambda} < \lambda + \frac{\gamma+1}{\gamma^2}\sum_h(\frac{N(h)}{N})^2\frac{1}{n\alpha_h}.$$

Falls der Wahrscheinlichkeitsvektor $\underline{\alpha}$ nicht zu extrem ist, lässt sich der Spielwert wie folgt abschätzen:

Bemerkung 6.6 *Falls $\alpha_h \geq \frac{1}{n}$, ist $\sum_h(\frac{N(h)}{N})^2\frac{1}{n\alpha_h} \leq 1$,und für den Spielwert λ gilt $\lambda < \tilde{\lambda} < \lambda + \frac{\gamma+1}{\gamma^2}$ bzw. $\tilde{\lambda} - \frac{\gamma+1}{\gamma^2} < \lambda < \tilde{\lambda}$, also*

$$\lambda < \frac{1}{n}\sum_{h=1}^{H}(\frac{N(h)}{N})^2\frac{1}{\alpha_h} < \lambda + \frac{\gamma+1}{\gamma^2}$$

bzw.

$$\frac{1}{n}\sum_{h=1}^{H}(\frac{N(h)}{N})^2\frac{1}{\alpha_h} - \frac{\gamma+1}{\gamma^2} < \lambda < \frac{1}{n}\sum_{h=1}^{H}(\frac{N(h)}{N})^2\frac{1}{\alpha_h}.$$

Beispiel 6.7 *Sei $H = 2$ und $\underline{\alpha} \in \mathbb{R}^2_+$ wie in Beispiel 6.3 definiert durch*

$$\alpha_h := \frac{N(h)^2}{N(1)^2 + N(2)^2} \quad \textit{für } h = 1, 2.$$

Wegen $\frac{N(h)^2}{\alpha_h} = N(1)^2 + N(2)^2$ für $h = 1, 2$, ist (vgl. 6.4)

$$\tilde{\lambda} = 2\frac{N(1)^2 + N(2)^2}{nN^2}$$

und (vgl. 6.3)

$$\tilde{n}(h) = \frac{n}{2} \quad \textit{für } h = 1, 2.$$

Sei $\gamma \in \mathbb{N}$ mit $\gamma \leq \frac{N(h)}{n}$, $h = 1, 2$. Dann gilt für die Minimax-Aufteilung $\underline{n}^*$

$$\frac{\gamma}{\gamma + 1}\frac{n}{2} < n^*(h) < \frac{\gamma + 2}{\gamma + 1}\frac{n}{2} \quad \textit{für } h = 1, 2.$$

6.2 Eine Regressionsmatrix vom Rang Zwei

Sei X die Regressionsmatrix mit den Spalten $\underline{1}, \underline{x} \in \mathbb{R}^N$, wobei $\underline{x} \notin \operatorname{span}[\underline{1}]$ mit $x = 1$. Wir gehen aus von dem Superpopulationsmodell

$$\underline{Y} = \beta_1 \underline{1} + \beta_2 \underline{x} + \underline{\epsilon} = X\underline{\beta} + \underline{\epsilon}$$

mit $\mathrm{E}(\underline{\epsilon}) = \underline{0}$ und $\operatorname{var}(\underline{\epsilon}) = \Omega := I$. Für $U := \Omega^{-1} - \Omega^{-1}X(X^t\Omega^{-1}X)^{-1}X^t\Omega^{-1}$ ergibt sich

$$\begin{aligned} U &= I - \frac{1}{N\sum_k x_k^2 - 1}[\sum_k x_k^2 \underline{1}\,\underline{1}^t + N\underline{x}\underline{x}^t - (x_i + x_j)_{i,j}] \\ &= I - \frac{1}{N\underline{x}^t\underline{x} - 1}[\underline{x}^t\underline{x}\underline{1}\,\underline{1}^t + N\underline{x}\underline{x}^t - \underline{x}\underline{1}^t - \underline{1}\underline{x}^t], \end{aligned} \tag{6.7}$$

also

$$\Theta_{UX} = \{\underline{y} \in \mathbb{R}^N; X^t\underline{y} = 0, \sum_i y_i^2 = 1\} = \{\underline{y} \in \mathbb{R}^N; X^t\underline{y} = 0, \sigma_{yy} = \frac{1}{N}\}.$$

Beispiel 6.8 *Sei $N = 4$, $n = 2$ und $\underline{x} := (\frac{1}{4}, \frac{1}{4}, \frac{1}{6}, \frac{1}{3})^t$. Dann ist*

$$U = \frac{1}{4}\begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}.$$

Aufgrund der Repräsentativität sind folgende Schätzvektoren eindeutig festgelegt:

$$\underline{a}^*_{\{1,3\}} = \underline{a}^*_{\{1,4\}} = (4,0,0,0)^t,\ \underline{a}^*_{\{2,3\}} = \underline{a}^*_{\{2,4\}} = (0,4,0,0)^t\ \underline{a}^*_{\{3,4\}} = (0,0,2,2)^t.$$

Für $s = \{1,2\}$ *wählen wir aus Symmetriegründen*

$$\underline{a}^*_{\{1,2\}} = (2,2,0,0)^t.$$

Mit

$$\underline{y}^* := \frac{1}{2}(1,1,-1,-1)^t \in \Theta_{UX}$$

gilt

$$(\underline{a}^*_{\{i,j\}})^t \underline{y}^* (\underline{y}^*)^t \underline{a}^*_{\{i,j\}} = 4 \quad \forall i<j,$$

d.h. die Natur besitzt die Indifferenzmatrix $V^* := \underline{y}^*(\underline{y}^*)^t$. *Definieren wir*

$$W_{ij} := (\underline{a}^*_{\{i,j\}} - \underline{1})(\underline{a}^*_{\{i,j\}} - \underline{1})^t \quad \forall i<j,$$

so ist

$$4\,U = \mu(W_{13}+W_{23}) + (\frac{1}{4}-\mu)(W_{14}+W_{24}) + \nu W_{12} + (\frac{1}{2}-\nu)W_{34} \quad \forall \mu,\nu.$$

Mit $p^*_{(\mu,\nu)} : \mathcal{S} \to [0;1]$ *definiert durch*

$$\begin{aligned}
p^*_{(\mu,\nu)}(\{1,2\}) &:= \nu,\\
p^*_{(\mu,\nu)}(\{3,4\}) &:= \frac{1}{2}-\nu,\\
p^*_{(\mu,\nu)}(\{1,3\}) &:= p^*_{(\mu,\nu)}(\{2,3\}) := \mu,\\
p^*_{(\mu,\nu)}(\{1,4\}) &:= p^*_{(\mu,\nu)}(\{2,4\}) := \frac{1}{4}-\mu
\end{aligned}$$

für $\mu \in [0;\frac{1}{4}]$, $\nu \in [0;\frac{1}{2}]$, *ist* $(p^*_{(\mu,\nu)}, \{\underline{a}^*_s\})$ *Minimax-Strategie,* V^* *Maximin-Strategie, und der Wert des Spiels ist* $\lambda = 4$.

Wählen wir speziell $\mu = 0$, $\nu = \frac{1}{2}$, *so ist die Minimax-Strategie* $(p^*_{(0,\frac{1}{2})}, \{\underline{a}^*_s\})$ *eine Indifferenz-Strategie, die die dritte Einheit nicht berücksichtigt. Tatsächlich wird auch die vierte Einheit nicht berücksichtigt. Diese kann zwar ausgewählt werden, jedoch wird der ihr zugeordnete Merkmalswert mit* 0 *gewichtet. Entsprechendes gilt für die Minimax-Strategie* $(p^*_{(\frac{1}{4},0)}, \{\underline{a}^*_s\})$.

Im Zwei-Schichten-Fall ist die Matrix U (vgl. (6.7)) von besonders einfacher Gestalt: Falls $\underline{x} := (k\underline{1}^t_{N(1)}, h\underline{1}^t_{N(2)})^t$, wobei $N(1)+N(2) = N$ und $N(1)k+N(2)h = 1$, ist U gegeben durch die Blockmatrix

$$U := \begin{bmatrix} U(1) & 0 \\ 0 & U(2) \end{bmatrix}$$

mit

$$U(h) := I_{N(h)} - \frac{1}{N(h)} \underline{1}_{N(h)} \underline{1}_{N(h)}^t.$$

Für Θ_U ergibt sich

$$\Theta_U = \{\underline{y} \in \mathbb{R}^N; N(1)\sigma_{yy}(1) + N(2)\sigma_{yy}(2) = 1\}.$$

Wir untersuchen das zugehörige Minimax-Problem für $n = 2$: Da $\underline{x} \notin \mathrm{span}[\underline{1}]$, ist $k \neq h$, und die repräsentativen Schätzvektoren sind gegeben durch

$$\underline{a}_{\{i,j\}} := N(1)\underline{e}_i + N(2)\underline{e}_j \quad \text{für } i \leq N(1) < j.$$

Wählt der Statistiker (den aus Symmetriegründen offensichtlichen) Auswahlplan p^* definiert durch

$$p^*(\{i,j\}) := \frac{1}{N(1)N(2)} \quad \forall i \leq N(1) < j,$$

so gilt

$$W^* := \sum_s p^*(s)(\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t = \begin{bmatrix} N(1)U(1) & 0 \\ 0 & N(2)U(2) \end{bmatrix}.$$

Für $\underline{y} = (\underline{y}_{N(1)}^t, \underline{y}_{N(2)}^t)^t \in \Theta_{UX}$ ist

$$\sum_i y_{N(1)i} = \sum_i y_{N(2)i} = 0,$$

also

$$\underline{y}^t W^* \underline{y} = N(1)\underline{y}_{N(1)}^t \underline{y}_{N(1)} + N(2)\underline{y}_{N(2)}^t \underline{y}_{N(2)} = N(1)^2\sigma_{yy}(1) + N(2)^2\sigma_{yy}(2).$$

Falls $N(1) \geq N(2)$, ist die Maximin-Strategie gegeben durch

$$V^* := \frac{1}{N(1)-1} \begin{bmatrix} U(1) & 0 \\ 0 & 0 \end{bmatrix},$$

W^* ist Minimax-Strategie, und der Wert des Spiels ist $N(1)$. (Der Fall $N(2) > N(1)$ wird entsprechend behandelt.)

Im Folgenden betrachten wir das Minimax-Problem für beliebiges $n \geq 2$ und nehmen eine kleine Modifikation von U vor: Definieren wir

$$\widetilde{U} := \begin{bmatrix} \frac{N(1)}{N(1)-1}U(1) & 0 \\ 0 & \frac{N(2)}{N(2)-1}U(2) \end{bmatrix},$$

so ist

$$\Theta_{\widetilde{U}} = \{\underline{y} \in \mathbb{R}^N; N(1)s_{yy}(1) + N(2)s_{yy}(2) = 1\}.$$

Aufgrund der symmetrischen Bedingungen und der Repräsentativität ist die Minimax-Strategie von der Gestalt

$$\sum_h N(h)\overline{y}_s(h)$$

bzgl. einer zufälligen Aufteilung $\underline{n}(.)$ mit Realisationen $(n(1), n(2)) \in \mathbb{N}^2$ (wegen $k \neq h$ muss aus jeder Schicht mindestens eine Einheit ausgewählt werden), und der erwartete quadratische Verlust ist gegeben durch

$$E[\sum_h N(h)^2(\frac{1}{n(h)} - \frac{1}{N(h)})s_{yy}(h)].$$

Wählen wir speziell die proportionale Aufteilung (unter Vernachlässigung der Ganzzahligkeit von $n(h)$)

$$n(h) = n\frac{N(h)}{N} \quad \text{für } h = 1, 2,$$

so ist der quadratische Verlust gegeben durch

$$\frac{N-n}{n}\sum_h N(h)s_{yy}(h).$$

Offensichtlich ist die zu der proportionalen Aufteilung gehörige risikoerzeugende Matrix W^* gegeben durch

$$W^* = \frac{N-n}{n}\widetilde{U},$$

d.h. der Statistiker stellt die Natur indifferent.

Die Maximin-Strategie ist aufgrund der symmetrischen Bedingungen von der Gestalt

$$\begin{bmatrix} s_{yy}(1)U(1) & 0 \\ 0 & s_{yy}(2)U(2) \end{bmatrix}$$

für einen Varianzvektor $(s_{yy}(1), s_{yy}(2))^t \in \mathbb{R}^2_+$ mit $N(1)s_{yy}(1) + N(2)s_{yy}(2) = 1$. (Da $\langle U(h), U(h)\rangle = N(h) - 1$ für $h = 1, 2$, ist obige Matrix in $\mathcal{Y}_{\widetilde{U}X}$.) Wählt die Natur

$$s^*_{yy}(h) = \frac{1}{N} \quad \text{für } h = 1, 2,$$

so gilt

$$\begin{aligned}
\frac{n}{N}(N(1), N(2)) &\in \operatorname{argmin}_{n(1)+n(2)=n} \sum_h (N(h))^2 (\frac{1}{n(h)} - \frac{1}{N(h)}) s^*_{yy}(h) \\
&= \operatorname{argmin}_{n(1)+n(2)=n} \sum_h \frac{N(h)}{N} (\frac{N(h)}{n(h)} - 1).
\end{aligned}$$

Somit ist W^* Minimax-Strategie, $V^* = \frac{1}{N}U$ Maximin-Strategie, und der Wert des Spiels beträgt $\frac{N-n}{n}$. (Vgl. Beispiel 6.2.) Es verbleibt zu bemerken, dass das Ergebnis nur von der Größe der Schichten abhängt, also nicht von den Parametern k und h.

Kapitel 7

Der Quader als Parameterraum

Angenommen, wir kennen die Beschäftigtenzahlen von N Betrieben einer Branche und interessieren uns für die Krankmeldungen an einem bestimmten Tag. Dieses Beispiel führt uns offensichtlich zu dem Parameterraum Θ_Q. Es list zu vermuten, dass das Verhältnis zwischen Krankmeldungen und Beschäftigtenzahl von Betrieb zu Betrieb wenig schwankt, was den Einsatz von repräsentativen Schätzern rechtfertigt.

Wie bereits in Kapitel 1 bemerkt, ist die Verwendung von repräsentativen Strategien im Fall $\Theta = \Theta_Q$ nicht zwingend. Vielmehr lassen sich unter gewissen Umständen mit nicht-repräsentativen Strategien kleinere Minimax-Risiken erzeugen (siehe z.B. Gabler [6], Bsp. 1 auf S. 27). Aufgrund von Bemerkung 1.2 in Kapitel 1 werden wir uns trotzdem auf die Menge der repräsentativen Schätzer beschränken.

7.1 Der Existenzsatz

Seien $N, n \in \mathbb{N}$ mit $N \geq n$. Für Vektoren $\underline{s}, \underline{t} \in \mathbb{R}^N$ definieren wir folgende Ordnungsrelation auf $\mathbb{R}^N$:

$$\underline{s} \preceq \underline{t} \text{ falls } s_i \leq t_i \quad \forall i = 1, \ldots, N,$$

bzw.

$$\underline{s} \prec \underline{t} \text{ falls } s_i < t_i \quad \forall i = 1, \ldots, N.$$

Wir benutzen dieselbe Symbolik wie bei der Löwner-Halbordnung auf dem Raum der symmetrischen Matrizen. Da Matrizen mit Großbuchstaben gekennzeichnet sind, sollten keine Missverständnisse entstehen.

Sei $\underline{x} \in \mathbb{R}^N$ ein Regressionsvektor,

$$\Theta_Q = \{\underline{y} \in \mathbb{R}^N; \underline{0} \preceq \underline{y} \preceq \underline{x}\}$$

der Parameterraum der Natur und $\mathcal{Y}_Q$ die konvexe Hülle von $\{\underline{y}\,\underline{y}^t; \underline{y} \in \Theta_Q\}$. Offensichtlich ist der Quader Θ_Q die konvexe Hülle seiner 2^N Ecken, die wir mit $\underline{c}_i$, $i = 1, \ldots, 2^N =: k$, bezeichnen. Aufgrund der Konvexität der MSE-Verlustfunktion können wir die Ecken als die Aktionsmenge der Natur auffassen, d.h. die Natur mischt nur über die Ecken des Quaders Θ_Q, denn für alle $\underline{y} \in \Theta_Q$ existiert ein Wahrscheinlichkeitsvektor $\underline{\mu} \in \mathbb{R}^k$ mit $\underline{y} = \sum_{i=1}^k \mu_i \underline{c}_i$ und

$$(\underline{y}, \underline{z})_2^2 \le \sum_{i=1}^k \mu_i (\underline{c}_i, \underline{z})_2^2 \quad \forall \underline{z} \in \mathbb{R}^N.$$

Wir befinden uns also im klassischen Kontext der S-Spiele (vgl. Blackwell und Girshick [3], S. 47). Eine Strategie der Natur entspricht einer Verteilung μ auf der Menge der Ecken bzw. einem Wahrscheinlichkeitsvektor $\underline{\mu} \in \mathbb{R}^k$, und es gilt

$$\mathrm{MSE}(\mu, \underline{a}_s) = \mathrm{MSE}(\underline{\mu}, \underline{a}_s) = \sum_{i=1}^k \mu_i (\underline{c}_i, \underline{a}_s - \underline{1})_2^2.$$

Dabei nennen wir μ eine *Indifferenz-Strategie* der Natur, falls für ein $\lambda \in \mathbb{R}$ gilt

$$\min_{\underline{a}_s \in A_s} \mathrm{MSE}(\mu, \underline{a}_s) = \lambda \quad \forall s \in S(n).$$

Eine Matrix $W \in \mathcal{A}_x$ bezeichnen wir als *Indifferenzmatrix* bzw. *Indifferenz-Strategie* des Statistikers (bzgl. des Spiels $(\mathcal{Y}_Q, \mathcal{A}_x)$), falls für ein $\lambda \in \mathbb{R}$ gilt

$$\underline{c}_i^t W \underline{c}_i = \lambda \quad \forall i = 1, \ldots, k.$$

Die Existenz von Minimax-Strategien in S-Spielen ist geklärt (vgl. z.B. Berger [1], Satz 15 auf S. 206).

Satz 7.1 (Existenzsatz 2) *Das lineare Spiel $(\mathcal{Y}_Q, \mathcal{A}_x)$ ist lösbar. Speziell existiert eine zulässige Minimax-Strategie.*

Beweis: Sei

$$S := \{\underline{s} \in \mathbb{R}^k; \exists \tilde{W} \in \widetilde{\mathcal{A}}_x : \; s_j = \langle C_j, \tilde{W} \rangle \text{ für } j = 1, \ldots, k\},$$

wobei

$$C_j := \underline{c}_j \underline{c}_j^t \quad j = 1, \ldots, k.$$

Eine kleine Überlegung zeigt, dass die Matrizen C_j ein Erzeugendensystem von $\mathcal{Q}$ bilden. Da sowohl $\tilde{W}$ als auch C_j positiv semidefinit sind, gilt mit Lemma A.1 (vgl. Anhang)

$$\underline{s} \succeq \underline{0} \quad \forall \underline{s} \in S.$$

Sei

$$Q_\alpha := \{\underline{z} \in \mathbb{R}^k; \underline{z} \preceq \alpha \underline{1}\}$$

und

$$\alpha_0 := \inf\{\alpha \in \mathbb{R}; Q_\alpha \cap S \neq \emptyset\} \geq 0.$$

Die Menge $S \subseteq \mathbb{R}^k$ ist offensichtlich konvex; da $\operatorname{int}(Q_{\alpha_0}) \cap S = \emptyset$, existiert nach dem Satz der trennenden Hyperebene ein $\underline{\xi} \in \mathbb{R}^k \setminus \{\underline{0}\}$ und ein $c \in \mathbb{R}$ mit

$$\begin{aligned} (\underline{\xi}, \underline{s})_2 &\geq c \quad \forall \underline{s} \in S, & (7.1)\\ (\underline{\xi}, \underline{z})_2 &\leq c \quad \forall \underline{z} \in Q_{\alpha_0}, & (7.2)\\ c &= \sup_{\underline{z} \in Q_{\alpha_0}} (\underline{\xi}, \underline{z})_2. \end{aligned}$$

Angenommen, es gilt nicht $\underline{\xi} \succeq \underline{0}$. Dann existiert ein $i \in \{1, \ldots, k\}$ mit $\xi_i < 0$. Da für alle $n \in \mathbb{N}$

$$\underline{z}^n := -n\underline{e}_i \in Q_{\alpha_0}$$

und

$$\lim_{n \to \infty} (\underline{\xi}, \underline{z}^n)_2 = \infty,$$

steht die obige Annahme im Widerspruch zur Trennungseigenschaft (7.1). Es gilt also $\underline{\xi} \succeq \underline{0}$, und mit

$$\mu_i := \frac{\xi_i}{\sum_j \xi_j} \quad \text{für } i = 1, \ldots, k$$

ist $\underline{\mu} \in \mathbb{R}^k$ ein Wahrscheinlichkeitsvektor. Da

$$c = \sup_{\underline{z} \in \mathcal{Q}_{\alpha_0}} (\underline{\xi}, \underline{z})_2 = (\underline{\xi}, \alpha_0 \underline{z})_2 = \alpha_0 \sum_i \xi_i,$$

ist

$$\alpha_0 = \frac{c}{\sum_i \xi_i},$$

und wegen (7.2) folgt

$$(\underline{\mu}, \underline{s})_2 \geq \alpha_0 \quad \forall \underline{s} \in S$$

bzw.

$$\langle \sum_i \mu_i C_i, \tilde{W} \rangle \geq \alpha_0 \quad \forall \tilde{W} \in \widetilde{\mathcal{A}}_x.$$

Aufgrund der Definition von α_0 existiert eine Folge $(\underline{s}^n)$ in S mit

$$\lim_{n\to\infty} \max_i \{s_i^n\} = \alpha_0.$$

Weiter existiert eine Folge $(\tilde{W}^n)$ in $\widetilde{\mathcal{A}}_x$ mit

$$s_i^n = \langle C_i, \tilde{W}^n \rangle \quad \forall n \in \mathbb{N} \forall i = 1, \ldots, k.$$

Nach Lemma 2.10 existiert eine Folge (W^n) in $\mathcal{A}_x$ mit

$$\tilde{W}^n \succeq W^n.$$

Da die Folgen $(\langle C_i, W^n \rangle)_n$ für $i = 1, \ldots, k$ beschränkt sind und die Matrizen C_i ein Erzeugendensystem bilden, existiert nach Bolzano-Weierstrass eine Teilfolge $(W^n)_{n\in\mathbb{N}'}$ und ein $W \in \mathcal{Q}$ mit

$$\lim_{n\in\mathbb{N}'} W^n = W.$$

Wegen Lemma 2.11 existiert ein $W^* \in \mathcal{A}_x$ mit

$$W \succeq W^*,$$

für das aufgrund der Definition von α_0 und der Monotonie von $\langle C_i, . \rangle$ (vgl. Proposition 2.1) gilt

$$\max_i \langle C_i, W^* \rangle = \alpha_0$$

und somit

$$\langle \sum_i \nu_i C_i, W^* \rangle \leq \alpha_0$$

für einen beliebigen Wahrscheinlichkeitsvektor $\underline{\nu} \in \mathbb{R}^k$. Wir erhalten insgesamt

$$\langle \sum_i \nu_i C_i, W^* \rangle \leq \langle \sum_i \mu_i C_i, W^* \rangle \leq \langle \sum_i \mu_i C_i, W \rangle \quad \forall \underline{\nu}, \forall W \in \mathcal{A}_x,$$

wobei

$$\langle \sum_i \mu_i C_i, W^* \rangle = \alpha_0.$$

Die Maximin-Strategie ist gegeben durch $\underline{\mu}$, W^* ist Minimax-Strategie, und α_0 der Wert des Spiels. Dass eine zulässige Minimax-Strategie existiert, ergibt sich aus Korollar 2.17. $\diamond$

Im Folgenden bezeichnen wir die Ecken des Quaders mit

$$\underline{x}_K := \sum_{i \in K} x_i \underline{e}_i,$$

wobei $K \subseteq \{1, \ldots, N\}$ und setzen

$$x_K := \sum_{i \in K} x_i.$$

Speziell definieren wir

$$\underline{x}_i := \underline{x}_{\{i\}} = x_i \underline{e}_i \quad \text{für } i = 1, \ldots, N.$$

Sei K^c das Komplement von $K \subseteq \{1, \ldots, N\}$.

Lemma 7.2 *Für jede repräsentative Strategie $(p, \{\underline{a}_s\})$ gilt*

$$MSE(\underline{x}_{K^c}, (p, \{\underline{a}_s\})) = MSE(\underline{x}_K, (p, \{\underline{a}_s\})).$$

Beweis: Da

$$\underline{x}_{K^c} = \underline{x} - \underline{x}_K,$$

ist

$$(\underline{x}_{K^c}, \underline{a} - \underline{1})_2^2 = (\underline{x} - \underline{x}_K, \underline{a} - \underline{1})_2^2 = (\underline{x}_K, \underline{a} - \underline{1})_2^2$$

für alle repräsentativen Schätzvektoren $\underline{a} \in \mathbb{R}^N$. $\diamond$

Falls eine allgemeine Indifferenzmatrix in Diagonalgestalt existiert, erhalten wir wie folgt eine Indifferenz-Strategie der Natur bzgl. $(\mathcal{Y}_Q, \mathcal{A}_x)$:

Lemma 7.3 *Sei $\underline{x} \in \mathbb{R}^N$ ein Regressionsvektor und $Q = diag(\underline{q}) \in \mathcal{Q}_0$ eine allgemeine Indifferenzmatrix bzgl. $\underline{x}$. Dann ist die Verteilung μ definiert durch*

$$\mu(\underline{x}_i) := \frac{q_i}{x_i^2} \frac{1}{\sum_j \frac{q_j}{x_j^2}} \text{ für } i = 1, \ldots, N, \ 0 \text{ sonst}$$

eine Indifferenz-Strategie der Natur bzgl. $(\mathcal{Y}_Q, \mathcal{A}_x)$, und eine untere Schranke für den Spielwert ist gegeben durch

$$\frac{1}{\sum_j \frac{q_j}{x_j^2}}.$$

Beweis: Sei $s \in S(n)$. Dann gilt

$$\begin{aligned}\min_{\underline{a}_s \in A_s} \mathrm{MSE}(\mu, \underline{a}_s) &= \min_{\underline{a}_s \in A_s} (\underline{a}_s - \underline{1})^t \sum_i \mu(\underline{x}_i) \underline{x}_i \underline{x}_i^t (\underline{a}_s - \underline{1}) \\ &= \frac{1}{\sum_j \frac{q_j}{x_j^2}} \min_{\underline{a}_s \in A_s} (\underline{a}_s - \underline{1})^t \mathrm{diag}(\underline{q}) (\underline{a}_s - \underline{1}) = \frac{1}{\sum_j \frac{q_j}{x_j^2}}.\end{aligned}$$

$\diamond$

Ein Anwendung dieser Aussage finden wir im nächsten Abschnitt (siehe Bemerkung 7.6).

7.2 Der Fall Eins aus N

Im Fall Eins aus N sind die repräsentativen Schätzvektoren festgelegt durch

$$\underline{a}_{\{i\}} := \frac{1}{x_i} \underline{e}_i \quad \text{für } i = 1, \ldots, N,$$

und für einen beliebigen Auswahlplan p gilt

$$\begin{aligned}\mathrm{MSE}(\underline{x}_K, (p, \{\underline{a}_{\{i\}}\})) &= \sum_{i=1}^N p_i (\underline{x}_K, \frac{1}{x_i}\underline{e}_i - \underline{1})_2^2 = \sum_{i=1}^N p_i [(\underline{x}_K, \frac{1}{x_i}\underline{e}_i)_2 - x_K]^2 \\ &= (1 - x_K)^2 \sum_{i \in K} p_i + x_K^2 \sum_{i \notin K} p_i \\ &= x_K^2 + (1 - 2x_K) \sum_{i \in K} p_i \quad \forall K \subseteq \{1, \ldots, N\}.\end{aligned}$$

Bemerkung 7.4 *Spielt der Statistiker die HH-Strategie, also $p_{HH}(\{i\}) = x_i$ für $i = 1, \ldots, N$, so gilt für alle $K \subseteq \{1, \ldots, N\}$*

$$MSE(\underline{x}_K, (p_{HH}, \{\frac{1}{x_i}\underline{e}_i\})) = x_K^2 + (1 - 2x_K) x_K = x_K(1 - x_K) \leq \frac{1}{4},$$

d.h. eine obere Schranke für den Spielwert ist gegeben durch $\frac{1}{4}$.

A.J. Scott und T.M.F. Smith bewiesen folgenden Satz (siehe Scott und Smith [19] oder Gabler [6], Satz 1 auf S. 26):

Satz 7.5 (Scott und Smith) *Sei $\underline{x} \in \mathbb{R}^N$ ein Regressionsvektor; es existiere eine Teilmenge $K_0 \subseteq \{1, \ldots, N\}$ mit*

$$x_{K_0} = \frac{1}{2}.$$

Dann ist die HH-Strategie $(p_{HH}, \{\frac{1}{x_i}\underline{e}_i\})$ eine Minimax-Strategie bzgl. des Parameterraums Θ_Q und der Menge der repräsentativen Strategien vom Umfang 1.

Beweis: Sei $x_{K_0} = \frac{1}{2}$. Dann ist

$$(\underline{x}_{K_0}, \frac{1}{x_i}\underline{e}_i - \underline{1})_2^2 = \begin{cases} x_{K_0}^2 + 1 - 2x_{K_0} = \frac{1}{4} & : \quad i \in K_0 \\ x_{K_0}^2 = \frac{1}{4} & : \quad i \notin K_0 \end{cases},$$

d.h. die reine Strategie $\underline{x}_{K_0}$ stellt den Statistiker indifferent, und es gilt

$$\mathrm{MSE}(\underline{x}_{K_0}, (p_{HH}, \{\frac{1}{x_i}\underline{e}_i\})) = \frac{1}{4}.$$

Wegen Bemerkung 7.4 ist $\underline{x}_{K_0}$ Maximin-Strategie, $(p_{HH}, \{\frac{1}{x_i}\underline{e}_i\})$ Minimax-Strategie, und der Wert des Spiels ist $\frac{1}{4}$. $\diamond$

Bemerkung 7.6 *Sei* $\max\{x_1, \ldots, x_N\} < \frac{1}{2}$. *Dann ist die in (3.13) definierte Matrix*

$$\frac{1}{\sum_j \frac{x_j(1-x_j)}{1-2x_j}} \begin{bmatrix} \frac{x_1^2}{1-2x_1} & & & & \\ & \ddots & & 0 & \\ & & \frac{x_i^2}{1-2x_i} & & \\ & 0 & & \ddots & \\ & & & & \frac{x_N^2}{1-2x_N} \end{bmatrix}$$

eine allgemeine Indifferenzmatrix. Aufgrund von Lemma 7.3 ist

$$\mu^*(\underline{x}_i) := \frac{\frac{1}{1-2x_i}}{\sum_j \frac{1}{1-2x_j}} \text{ für } i = 1, \ldots, N, \ 0 \text{ sonst}$$

eine Indifferenz-Strategie der Natur, und mit Bemerkung 7.4 folgt für den Spielwert λ

$$\frac{\sum_j \frac{x_j(1-x_j)}{1-2x_j}}{\sum_j \frac{1}{1-2x_j}} \leq \lambda \leq \frac{1}{4}.$$

7.3 Der einfache Zwei-Schichten-Fall

Bei einer Inventur auf Stichprobenbasis werden in der Praxis Positionen mit hohen vorläufigen Werten total erfasst, während man Positionen mit niedrigen vorläufigen Werten anhand einer Stichprobe abschätzt. Wir untersuchen in diesem Abschnitt für den einfachen Zwei-Schichten-Fall, wann dieses Vorgehen (verbunden mit einer entsprechenden Schätzung) das Minimax-Prinzip erfüllt.

Sei $n = 2$, $N > 2$ und

$$\underline{x} = (1, \ldots, 1, r)^t \in \mathbb{R}^N$$

für ein $r \geq 1$. (Wir verzichten hier auf die Normierung des Regressionsvektors.) Als Parameterraum betrachten wir den Quader $\Theta_Q = \{\underline{y} \in \mathbb{R}^N; \underline{0} \preceq \underline{y} \preceq \underline{x}\}$ und gehen davon aus, dass der optimale Auswahlplan p^* gegeben ist durch

$$p^*(s) = \frac{1}{N-1} \quad \forall s = \{i, N\},\ i = 1, \ldots, N-1.$$

(Diese Annahme ist – wie wir sehen werden – für entsprechend großes r erfüllt.) Wegen der symmetrischen Bedingungen betrachten wir nur Schätzvektoren der Form

$$\underline{a}_s = \beta \underline{e}_i + \alpha \underline{e}_N \quad \forall s = \{i, N\},\ i = 1, \ldots, N-1.$$

Aufgrund der Repräsentativität ist

$$0 = (\underline{a}_s - \underline{1})^t \underline{x} = \beta + \alpha r - (N - 1 + r),$$

also

$$\alpha = \frac{N-1+r-\beta}{r} = 1 + \frac{N-1-\beta}{r}.$$

Sei $K \subseteq \{1, \ldots, N\}$ und $\underline{x}_K = \sum_{j \in K} x_j \underline{e}_j \in \Theta_Q$. Wegen Lemma 7.2 genügt es, Teilmengen K zu betrachten, die N nicht als Element enthalten. Sei also o.B.d.A. $K \subseteq \{1, \ldots, N-1\}$; dann ist

$$\underline{x}_K = \sum_{j \in K} \underline{e}_j,$$

und es gilt

$$\begin{aligned}
\mathrm{MSE}(\underline{x}_K, (p^*, \{\underline{a}_s\})) &= \frac{1}{N-1} \sum_{i=1}^{N-1} (\underline{x}_K, \underline{a}_{\{i,N\}} - \underline{1})_2^2 \\
&= \frac{1}{N-1} \sum_{i=1}^{N-1} [(\underline{x}_K, \underline{a}_{\{i,N\}})_2 - |K|]^2 \\
&= \frac{1}{N-1} \sum_{i=1}^{N-1} [\beta \mathbf{1}_K(i) - |K|]^2 \\
&= \frac{1}{N-1} [|K|[\beta - |K|]^2 + (N - 1 - |K|)|K|^2] \\
&= \frac{1}{N-1} [|K|\beta^2 - 2|K|^2\beta + (N-1)|K|^2],
\end{aligned}$$

wobei

$$\mathbf{1}_K(i) := \begin{cases} 1 & : \quad i \in K \\ 0 & : \quad i \notin K \end{cases}.$$

Für eine fixierte Teilmenge K wird das Minimum offensichtlich angenommen in

$$\tilde{\beta} = |K|, \tag{7.3}$$

und für den resultierenden quadratischen Verlust gilt

$$\mathrm{MSE}(\underline{x}_K, (p^*, \{\underline{\tilde{a}_s}\})) = \frac{N-1-|K|}{N-1}|K|^2.$$

Spielt der Statistiker die Schätzvektoren $\underline{a}_s^*$ gegeben durch

$$\beta^* := \frac{2}{3}(N-1),$$

so ist

$$\begin{aligned}\max_K \mathrm{MSE}(\underline{x}_K, (p^*, \{\underline{a}_s^*\})) &= \max_{|K|=1,\ldots,N-1}[\frac{4}{9}(N-1)|K| - \frac{1}{3}|K|^2] \\ &= \frac{4}{27}(N-1)^2,\end{aligned}$$

und das Maximum wird in den Ecken $\underline{x}_{K^*}$ angenommen mit

$$|K^*| = \beta^* = \frac{2}{3}(N-1). \tag{7.4}$$

(Wir vernachlässigen das Ganzzahligkeitsproblem.) Spielt die Natur eine Gleichverteilung auf den Ecken

$$\mathcal{E} := \{\underline{x}_K; K \subseteq \{1,\ldots,N-1\}, |K| = \frac{2}{3}(N-1)\}, \tag{7.5}$$

so ist der optimale Schätzer (unter der Voraussetzung, dass der Auswahlplan durch p^* festgelegt ist) aufgrund von (7.3) gegeben durch die Schätzvektoren

$$\begin{aligned}\underline{a}_s^* &= \beta^*\underline{e}_i + (1 + \frac{N-1-\beta^*}{r})\underline{e}_N \\ &= \frac{2}{3}(N-1)\underline{e}_i + (1 + \frac{N-1}{3r})\underline{e}_N \quad \forall s = \{i, N\},\ i = 1,\ldots,N-1.\end{aligned}$$

Spielt der Statistiker andererseits die Strategie $(p^*, \{\underline{a}_s^*\})$, so sind alle Ecken aus $\mathcal{E}$ optimale Ecken (d.h. $\underline{x}_K \in \mathrm{argmax}_{\underline{y}\in\Theta_Q}\mathrm{MSE}(\underline{y}, (p^*, \{\underline{a}_s^*\}))$ für alle $K \in \mathcal{E}$). Der maximale Verlust der Strategie $(p^*, \{\underline{a}_s^*\})$ ist gegeben durch

$$\frac{4}{27}(N-1)^2. \tag{7.6}$$

Weiter ist $(p^*, \{\underline{a}_s^*\})$ für entsprechend großes r eine Minimax-Strategie: Sei μ^* die Gleichverteilung auf $\mathcal{E}$, $M := \frac{2}{3}(N-1)$ und $s = \{i, j\}$ mit $i < j < N$.

Aufgrund der symmetrischen Bedingungen und der Repräsentativität ist nur der Schätzvektor

$$\underline{a}_s = \gamma \underline{e}_i + \gamma \underline{e}_j$$

von Interesse, wobei

$$\gamma := \frac{N-1+r}{2}.$$

Es gilt

$$\begin{aligned}
\mathrm{MSE}(\mu^*, \underline{a}_s) &= \sum_{\underline{x}_K} \mu^*(\underline{x}_K)(\underline{x}_K, \underline{a}_s - \underline{1})_2^2 = \sum_{\underline{x}_K} \mu^*(\underline{x}_K)[(\underline{x}_K, \underline{a}_s)_2 - |K|]^2 \\
&= \binom{N-1}{M}^{-1} \sum_{\underline{x}_K \in \mathcal{E}} [(\underline{x}_K, \underline{a}_s)_2 - M]^2 \\
&= \binom{N-1}{M}^{-1} [\binom{N-3}{M} M^2 + 2\binom{N-3}{M-1}(\gamma - M)^2 + \binom{N-3}{M-2}(2\gamma - M)^2] \\
&= \frac{1}{(N-1)(N-2)}[(N-1-M)(N-2-M)M^2 + \\
&\quad +2M(N-1-M)(\gamma - M)^2 + M(M-1)(2\gamma - M)^2].
\end{aligned}$$

Falls

$$r \geq \frac{2\sqrt{6(N-1)(N-2)} - 3(N-3)}{15N-33}(N-1),$$

ist

$$\mathrm{MSE}(\mu^*, \underline{a}_s) \geq \frac{4}{27}(N-1)^2,$$

und $(\mu^*, (p^*, \{\underline{a}_s^*\}))$ ist ein Gleichgewicht des Spiels. Dabei verhält sich die Schranke $\frac{2\sqrt{6(N-1)(N-2)}-3(N-3)}{15N-33}(N-1)$ für große N wie $\frac{2\sqrt{6}-3}{15}N \approx 0.1266\,N$. Zusammenfassend erhalten wir:

Satz 7.7 *Sei $N \geq n = 2$ und $\underline{x} = (1, \ldots, 1, r) \in \mathbb{R}^N$. Wir betrachten das Spiel $(\mathcal{Y}_Q, \mathcal{A}_x)$. Falls $r \geq \frac{2\sqrt{6(N-1)(N-2)}-3(N-3)}{15N-33}(N-1)$, ist $(p^*, \{\underline{a}_s^*\})$ definiert durch*

$$p^*(s) := \begin{cases} \frac{1}{N-1} & : \quad N \in s \\ 0 & : \quad \mathit{sonst} \end{cases},$$

$$\underline{a}^*_{\{i,N\}} := \frac{2}{3}(N-1)\underline{e}_i + (1 + \frac{N-1}{3r})\underline{e}_N \quad \textit{für } i = 1, \ldots, N-1$$

eine Minimax-Strategie, eine Maximin-Strategie ist gegeben durch die Gleichverteilung auf den Ecken $\mathcal{E}$ (vgl. (7.5)), und $\frac{4}{27}(N-1)^2$ ist der Wert des Spiels.

Im Folgenden verallgemeinern wir dieses Ergebnis für beliebige Stichprobenumfänge: Sei $N > n > 2$ und

$$p^*(s) = \binom{N-1}{n-1}^{-1} \quad \forall s = \{\tilde{s}, N\},$$

wobei $\tilde{s} \subseteq \{1, \ldots, N-1\}$, $|\tilde{s}| = n-1$. Aufgrund der symmetrischen Bedingungen betrachten wir nur Schätzvektoren der Gestalt

$$\underline{a}_s = \beta \sum_{i \in \tilde{s}} \underline{e}_i + \alpha \underline{e}_N \quad \forall s = \{\tilde{s}, N\},$$

wobei wegen der Repräsentativität gilt

$$\alpha = \frac{r + N - 1 - (n-1)\beta}{r} = 1 + \frac{N - 1 - (n-1)\beta}{r}.$$

Aus den Momenten der hypergeometrischen Verteilung ergibt sich für eine Ecke $\underline{x}_K \in \Theta_Q$ mit $K \subseteq \{1, \ldots, N-1\}$ (vgl. Gabler [10])

$$\begin{aligned}
&\mathrm{MSE}(\underline{x}_K, (p^*, \{\underline{a}_s\})) = \\
&= \binom{N-1}{n-1}^{-1} \sum_{\substack{\tilde{s} \subseteq \{1,\ldots,N-1\} \\ |\tilde{s}| = n-1}} (\underline{x}_K, \underline{a}_{\{\tilde{s},N\}} - \underline{1})_2^2 = \binom{N-1}{n-1}^{-1} \sum_{\tilde{s}} [\beta |K \cap \tilde{s}| - |K|]^2 \\
&= \binom{N-1}{n-1}^{-1} \sum_{i=0}^{n-1} \binom{|K|}{i} \binom{N-1-|K|}{n-1-i} [\beta i - |K|]^2 \\
&= \frac{n-1}{(N-1)(N-2)} |K|((n-2)|K| + N - n)\beta^2 - 2\frac{n-1}{N-1}|K|^2\beta + |K|^2 \\
&= \frac{(n-1)(n-2)\beta^2 - 2(n-1)(N-2)\beta + (N-1)(N-2)}{(N-1)(N-2)} |K|^2 + \\
&\quad + \frac{(n-1)(N-n)\beta^2}{(N-1)(N-2)} |K|.
\end{aligned}$$

Für eine fixierte Teilmenge K wird das Minimum bzgl. β offensichtlich angenommen in

$$\tilde{\beta}_K = \frac{(N-2)|K|}{(n-2)|K| + N - n},$$

und für den resultierenden quadratischen Verlust gilt

$$\mathrm{MSE}(\underline{x}_K, (p^*, \{\underline{\tilde{a}}_s\})) = |K|^2 \frac{(N-n)(N-1-|K|)}{(N-1)((n-2)|K| + N - n)}.$$

Spielt der Statistiker die Schätzvektoren $\underline{a}_s^*$ gegeben durch

$$\beta^* := \frac{N-2}{n-2} \frac{(N+2)n - 5N + 2 + c(n,N)}{(N-2)(n-1) + c(n,N)},$$

wobei

$$c(n,N) := \sqrt{(N-2)(n-1)((n+7)N-10n+2)},$$

so ist

$$K \in \operatorname{argmax}_{K\subseteq\{1,\ldots,N-1\}} \mathrm{MSE}(\underline{x}_K, (p^*, \{\underline{a}_s^*\}))$$

genau dann, wenn

$$|K| = M_n := \frac{(N+2)n-5N+2+c(n,N)}{4(n-2)}.$$

(Wir vernachlässigen auch hier das Ganzzahligkeitsproblem.) Andererseits gilt für alle Teilmengen $K \subseteq \{1,\ldots,N-1\}$ mit $|K| = M_n$

$$\tilde{\beta}_K = \beta^*.$$

Für hinreichend großes r ist $(p^*, \{\underline{a}_s^*\})$ eine Minimax-Strategie, eine Maximin-Strategie ist gegeben durch die Gleichverteilung μ^* auf der Menge der Ecken

$$\mathcal{E}_n := \{\underline{x}_K; K \subseteq \{1,\ldots,N-1\}, |K| = M_n\},$$

und

$$\begin{aligned}
\lambda &= \mathrm{MSE}(\mu^*, (p^*, \{\underline{a}_s^*\})) \\
&= M_n^2 \frac{(N-n)(N-1-M_n)}{(N-1)((n-2)M_n+N-n)} \\
&= \frac{3Nn(N-1)-3(N+n^2)(N-2)-(N-n)c(n,N)-6n}{16(N-1)(n-2)^3((N-2)(n-1)+c(n,N))} \times \\
&\quad \times((N+2)n-5N+2+c(n,N))^2
\end{aligned}$$

ist der Spielwert:

Sei $s \in S(n)$, $N \notin s$. Aufgrund der symmetrischen Bedingungen und der Repräsentativität ist nur der Schätzvektor

$$\underline{a}_s = \gamma_n \sum_{j\in s} \underline{e}_j$$

von Interesse, wobei

$$\gamma_n := \frac{N-1+r}{n}.$$

Es gilt

$$\begin{aligned}
&\mathrm{MSE}(\mu^*, \underline{a}_s) = \\
&= \binom{N-1}{M_n}^{-1} \sum_{\underline{x}_K \in \mathcal{E}} [(\underline{x}_K, \underline{a}_s)_2 - M_n]^2 \\
&= \binom{N-1}{M_n}^{-1} \sum_{j=0}^{M_n} \binom{n}{j} \binom{N-1-n}{M_n - j} (j\gamma_n - M_n)^2 \\
&= M_n[(M_n(n-1) + N - n - 1)\frac{(N-1+r)^2}{(N-1)(N-2)n} - 2(1 + \frac{r}{N-1}) + M_n].
\end{aligned}$$

Falls r so gewählt wird, dass

$$M_n[(M_n(n-1) + N - n - 1)\frac{(N-1+r)^2}{(N-1)(N-2)n} - 2(1 + \frac{r}{N-1}) + M_n] \geq \lambda,$$

ist $(\mu^*, (p^*, \{\underline{a}_s^*\}))$ ein Gleichgewicht des Spiels.

Bemerkung 7.8 *Es ist auffällig, dass das Gewicht β^* als auch die Maximin-Strategie nicht von r abhängen – im Gegensatz zum Ellipsen-Fall (vgl. Satz 5.5) – sondern nur von n und N.*

7.4 Starke Repräsentativität im Schichtungsfall

Sei $H \in \mathbb{N}$, $\underline{x} = (x(1), \ldots, x(1), \ldots, x(H), \ldots, x(H))^t \in \mathbb{R}^N$ mit $x(h) > 0$, $\Theta_Q = \{\underline{y} \in \mathbb{R}^N; \underline{0} \preceq \underline{y} \preceq \underline{x}\}$ und $N(h) := |\{x_i; x_i = x(h)\}|$ für $h = 1, \ldots, H$. Wir betrachten das Spiel $(\mathcal{Y}_Q, \mathcal{A}_X)$, wobei die Regressionsmatrix X gegeben ist durch

$$X := \begin{bmatrix} \underline{1}_{N(1)} & \underline{0}_{N(1)} & \cdots & \underline{0}_{N(1)} \\ \underline{0}_{N(2)} & \underline{1}_{N(2)} & & \underline{0}_{N(2)} \\ \vdots & & \ddots & \vdots \\ \underline{0}_{N(H)} & \cdots & \cdots & \underline{1}_{N(H)} \end{bmatrix}.$$

Wir fordern also, dass die Strategien des Statistikers repräsentativ sind bzgl. X. (Offensichtlich ist $\mathcal{A}_X \subseteq \mathcal{A}_x$.) Weiter sei $\frac{N(h)}{2} \in \mathbb{N}$ für $h = 1, \ldots, H$ und

$$\max_{h=1,\ldots,H} x(h)\sqrt{\frac{N(h)}{N(h)-1}} \leq \frac{\sum_k N(k)x(k)\sqrt{\frac{N(k)}{N(k)-1}}}{n}. \tag{7.7}$$

Unter diesen Bedingungen lässt sich die Minimax-Strategie konkret angeben: Aufgrund der symmetrischen Bedingungen und der Repräsentativität der Schätzer bzgl. X ist die Minimax-Strategie von der Gestalt

$$\sum_h N(h)\overline{y}_s(h)$$

bzgl. einer zufälligen Aufteilung $\underline{n}(.)$ (vgl. Abschnitt 6.1), und der erwartete quadratische Verlust ist gegeben durch

$$\mathrm{E}[\sum_h N(h)^2(\frac{1}{n(h)} - \frac{1}{N(h)})s_{yy}(h)] = \sum_h s_{yy}(h)N(h)^2\mathrm{E}[\frac{1}{n(h)} - \frac{1}{N(h)}].$$

Mit

$$C := \{\underline{y} \in \Theta_Q; |\{i; y_i = 0\}| = \frac{N}{2}, |\{i; y_i = x(h)\}| = \frac{N(h)}{2} \text{ für } h = 1, \ldots, H\}$$

ist

$$C = \operatorname{argmax}_{\underline{y}\in\Theta_Q} \mathrm{E}[\sum_h N(h)^2(\frac{1}{n(h)} - \frac{1}{N(h)})s_{yy}(h)]$$

(falls die Zufallsvariablen $n(h) \neq N(h)$ für $h = 1, \ldots, H$), und es gilt

$$s_{yy}(h) = \frac{1}{4}\frac{N(h)}{N(h)-1}x(h)^2 \quad \forall \underline{y} \in C$$

für $h = 1, \ldots, H$. Der erwartete quadratische Verlust bzgl. einer Ecke $\underline{y} \in C$ und einer Aufteilung $\underline{n}$ ist gegeben durch

$$\frac{1}{4}\sum_h \frac{N(h)^3}{N(h)-1}(\frac{1}{n(h)} - \frac{1}{N(h)})x(h)^2,$$

und das Minimum wird angenommen in $\underline{n}^*$ definiert durch

$$n^*(h) := n\frac{N(h)x(h)\sqrt{\frac{N(h)}{N(h)-1}}}{\sum_k N(k)x(k)\sqrt{\frac{N(k)}{N(k)-1}}} \quad \text{für } h = 1, \ldots, H.$$

Aufgrund von (7.7) ist $n^*(h) \leq N(h)$ für $h = 1, \ldots, H$. Eine Maximin-Strategie ist gegeben durch die Gleichverteilung auf C, und die Schichtungsstrategie bzgl. der Aufteilung $\underline{n}^*$ ist Minimax-Strategie.

Im Spiel $(\mathcal{Y}_Q, \mathcal{A}_x)$ ist die Minimax-Strategie im Allgemeinen keine Schichtungsstrategie. Andererseits zeigte Stenger [23], dass die Schichtungsstrategie bzgl. der Aufteilung $\underline{\tilde{n}}$ definiert durch

$$\tilde{n}(h) := n\frac{N(h)x(h)}{\sum_k N(k)x(k)} \quad \text{für } h = 1, \ldots, H$$

in einem asymptotischen Sinn minimax ist.

Kapitel 8

Der Fall Zwei aus Drei für den Quader

Sei $N = 3$ und $n = 2$. Wie in Kapitel 4 für das Ellipsoid wollen wir in diesem Kapitel für den Quader bei vorgegebenem Regressionsvektor $\underline{x} \in \mathbb{R}^3$ Minimax-Strategien berechnen. In Abschnitt 8.4 stellen wir einen Zusammenhang zwischen dem Quader und dem Ellipsoid her.

Der Quader $\Theta_Q = \{\underline{y} \in \mathbb{R}^3; \underline{0} \preceq \underline{y} \preceq \underline{x}\}$ besitzt die 8 Ecken

$$\underline{0},\ \underline{x},\ \underline{x}_1,\ \underline{x}_2,\ \underline{x}_3,\ (x_1, x_2, 0)^t,\ (x_1, 0, x_3)^t,\ (0, x_2, x_3)^t$$

mit $\underline{x}_i = x_i \underline{e}_i$ für $i = 1, 2, 3$. Aufgrund von Lemma 7.2 ist es eine dominante Strategie der Natur, nur über die 3 Ecken $\underline{x}_1, \underline{x}_2, \underline{x}_3$ zu mischen. Eine Strategie der Natur entspricht somit einem Wahrscheinlichkeitsvektor $\underline{\mu} \in \mathbb{R}^3$.

8.1 Indifferenz-Strategien der Natur

Um den Statistiker indifferent bzgl. der Stichproben $s \in S(2)$ zu stellen, sucht die Natur einen Wahrscheinlichkeitsvektor $\underline{\mu} \in \mathbb{R}^3$, so dass

$$\min_{\underline{a} \in A_s} \sum_i \mu_i (\underline{x}_i, \underline{a} - \underline{1})_2^2 = \lambda \quad \forall s \in S(2)$$

für ein $\lambda > 0$. Eine Indifferenz-Strategie $\underline{\mu}$ der Natur existiert also genau dann, wenn das Gleichungssystem

$$\sum_i \mu_i = 1,$$

$$\min_{\underline{a}\in A_s} \sum_i \mu_i(\underline{x}_i, \underline{a}-\underline{1})_2^2 = \lambda \quad \forall s \in S(2)$$

lösbar ist mit $\mu_i \geq 0$ für ein $\lambda > 0$.

Bemerkung 8.1 *Sei $a > 0$ und $b, c \in \mathbb{R}$. Mit $f(\alpha) := a\alpha^2 - 2b\alpha + c$ ist*

$$\min_\alpha f(\alpha) = f(\frac{b}{a}) = \frac{ac-b^2}{a}.$$

Betrachten wir die Stichprobe $s = \{1,2\}$. Sei $\underline{a} \in A_{\{1,2\}}$ und

$$(\beta, \alpha, -1)^t = \underline{a} - \underline{1}.$$

Aufgrund der Repräsentativität ist

$$\beta = \frac{x_3 - x_2\alpha}{x_1},$$

also

$$\begin{aligned}\sum_i \mu_i(\underline{x}_i, (\beta,\alpha,-1)^t)_2^2 &= \mu_1 x_1^2\beta^2 + \mu_2 x_2^2\alpha^2 + \mu_3 x_3^2 \\ &= \mu_1(x_3 - x_2\alpha)_2^2 + \mu_2 x_2^2\alpha^2 + \mu_3 x_3^2 \\ &= (\mu_1+\mu_2)x_2^2\alpha^2 - 2\mu_1 x_2 x_3\alpha + (\mu_1+\mu_3)x_3^2,\end{aligned}$$

und mit Bemerkung 8.1 folgt

$$\begin{aligned}\min_{\underline{a}\in A_{\{1,2\}}} \sum_i \mu_i(\underline{x}_i, \underline{a}-\underline{1})_2^2 &= \frac{(\mu_1+\mu_2)(\mu_1+\mu_3)x_2^2x_3^2 - \mu_1^2x_2^2x_3^2}{(\mu_1+\mu_2)x_2^2} \\ &= \frac{\mu_1\mu_2+\mu_1\mu_3+\mu_2\mu_3}{1-\mu_3}x_3^2, \end{aligned} \tag{8.1}$$

wobei das Minimum angenommen wird in

$$\alpha^*_{\{1,2\}} := \frac{\mu_1 x_3}{(1-\mu_3)x_2}.$$

Für $s = \{1,3\}$ ergibt sich mit

$$(\beta, -1, \alpha) = \underline{a} - \underline{1}$$

für $\underline{a} \in A_{\{1,3\}}$ entsprechend:

$$\beta = \frac{x_2 - x_3\alpha}{x_1},$$

$$\min_{\underline{a}\in A_{\{1,3\}}} \sum_i \mu_i(\underline{x}_i, \underline{a}-\underline{1})_2^2 = \frac{\mu_1\mu_2+\mu_1\mu_3+\mu_2\mu_3}{1-\mu_2}x_2^2,$$

wobei das Minimum angenommen wird in

$$\alpha^*_{\{1,3\}} := \frac{\mu_1 x_2}{(1-\mu_2)x_3}. \tag{8.2}$$

Ebenso erhalten wir für $s = \{2,3\}$ mit

$$(-1, \beta, \alpha) = \underline{a} - \underline{1}$$

für $\underline{a} \in A_{\{2,3\}}$:

$$\beta = \frac{x_1 - x_3\alpha}{x_2},$$

$$\min_{\underline{a}\in A_{\{2,3\}}} \sum_i \mu_i(\underline{x}_i, \underline{a} - \underline{1})_2^2 = \frac{\mu_1\mu_2 + \mu_1\mu_3 + \mu_2\mu_3}{1-\mu_1} x_1^2,$$

wobei das Minimum angenommen wird in

$$\alpha^*_{\{2,3\}} := \frac{\mu_2 x_1}{(1-\mu_1)x_3}. \tag{8.3}$$

Mit

$$K := \mu_1\mu_2 + \mu_1\mu_3 + \mu_2\mu_3$$

erhalten wir das Gleichungssystem

$$\sum_i \mu_i = 1,$$

$$\frac{K x_i^2}{1-\mu_i} = \lambda \quad \text{für } i = 1,2,3.$$

Speziell ist

$$K \sum_i x_i^2 = \lambda \sum_i (1-\mu_i) = 2\lambda,$$

also

$$\lambda = \frac{K \sum_i x_i^2}{2}$$

und somit

$$\mu_i = \frac{\sum_j x_j^2 - 2x_i^2}{\sum_j x_j^2} \quad \text{für } i = 1,2,3.$$

Da $x_i \neq 0$, ist $\mu_i < 1$ für $i = 1,2,3$, und wir erhalten folgenden Satz:

Satz 8.2 *Sei $\underline{x} \in \mathbb{R}^3$ ein Regressionsvektor. Eine Indifferenz-Strategie der Natur existiert genau dann, wenn*

$$\sum_j x_j^2 - 2x_0^2 \geq 0, \tag{8.4}$$

wobei $x_0 := \max\{x_1, x_2, x_3\}$.

Bemerkung 8.3 *Die Existenz-Bedingung für Indifferenz-Strategien der Natur ist für den Quader schärfer ist als für das Ellipsoid (vgl. Satz 4.2). Dies zeigt auch folgende Überlegung: Sei (8.4) erfüllt, $\underline{\mu}^*$ eine Indifferenz-Strategie der Natur bzgl. $(\mathcal{Y}_Q, \mathcal{A}_x)$ und $\underline{d} \in \mathbb{R}^3$ definiert durch*

$$d_i := \mu_i^* x_i^2 \geq 0.$$

Dann ist

$$V_U^* := \frac{1}{\sum_i \mu_i^* x_i^2 u_{ii}} diag(d) \in \mathcal{Y}_U$$

eine Indifferenzmatrix der Natur bzgl. $(\mathcal{Y}_U, \mathcal{A}_x)$ für eine beliebige Matrix $U \in \mathcal{Q}_0$ mit $Kern(U) = span[\underline{x}]$. Allgemein gilt im Fall $N = 3$

$$\mathcal{Y}_Q = \{diag(\underline{d}) \in \mathcal{Q}_0; \sum_{i=1}^{3} \frac{d_i}{x_i^2} = 1\}. \tag{8.5}$$

Betrachten wir den Regressionsvektor $\underline{x} = (k, k, h)^t$, so ist (8.4) genau dann erfüllt, wenn $1 - \frac{\sqrt{2}}{2} \leq k < \frac{1}{2}$, wobei $0.29 < 1 - \frac{\sqrt{2}}{2}$. Im Ellipsen-Fall existiert eine Indifferenzmatrix der Natur genau dann, wenn $\frac{1}{4} < k < \frac{1}{2}$ (vgl. Satz 4.2).

Unter der Bedingung (8.4) ist

$$\lambda := (\mu_1\mu_2 + \mu_1\mu_3 + \mu_2\mu_3)\frac{\sum_i x_i^2}{2} = \frac{2(x_1^2x_2^2 + x_1^2x_3^2 + x_2^2x_3^2) - \sum_i x_i^4}{2\sum_i x_i^2} \tag{8.6}$$

eine untere Schranke für den Spielwert, die optimalen Schätzvektoren des Statistikers sind gegeben durch

$$\underline{a}^*_{\{1,2\}} := (\xi_2, \xi_1, -1)^t + \underline{1},\ \underline{a}^*_{\{1,3\}} := (\xi_3, -1, \xi_1)^t + \underline{1},\ \underline{a}^*_{\{2,3\}} := (-1, \xi_3, \xi_2)^t + \underline{1},$$

wobei

$$\xi_i := \frac{\sum_j x_j^2 - 2x_i^2}{2\prod_j x_j} x_i \quad \text{für } i = 1, 2, 3,$$

und sie stimmen mit den optimalen Schätzvektoren im Ellipsen-Fall überein (vgl. (4.10)).

8.2 Konstruktion eines Auswahlplans

Wir gehen davon aus, dass (8.4) für den Regressionsvektor $\underline{x}$ erfüllt ist; es existiert also eine Indifferenz-Strategie $\underline{\mu}$ der Natur. Falls ein Auswahlplan p existiert mit

$$\sum_s p(s)(\underline{x}_i, \underline{a}_s^* - \underline{1})_2^2 = \lambda \quad \forall i = 1, 2, 3,$$

ist $\underline{\mu}$ eine Maximin-Strategie, $(p, \{\underline{a}_s^*\})$ eine Minimax-Strategie, und der Wert des Spiels ist gegeben durch (8.6) (vgl. Satz 3.7 für das Ellipsoid). Dabei existiert ein solcher Auswahlplan genau dann, wenn das Gleichungssystem

$$\begin{aligned} p(1) + \xi_3^2 p(2) + \xi_2^2 p(3) &= \frac{\lambda}{x_1^2}, \\ \xi_3^2 p(1) + p(2) + \xi_1^2 p(3) &= \frac{\lambda}{x_2^2}, \\ \xi_2^2 p(1) + \xi_1^2 p(2) + p(3) &= \frac{\lambda}{x_3^2} \end{aligned} \tag{8.7}$$

lösbar ist mit $p(i) \geq 0$ für alle $i = 1, 2, 3$, und der Auswahlplan p ist gegeben durch

$$p(\{1,2,3\} \setminus \{i\}) := p(i) \quad \text{für } i = 1, 2, 3.$$

(Für jede Lösung von (8.7) gilt $\sum_i p(i) = 1$.)

Beispiel 8.4 *Sei $\underline{x} = \frac{1}{3}\underline{1}$. Dann ist*

$$\xi_i = \frac{1}{2}, \ \mu_i = \frac{1}{3} \quad \textit{für } i = 1, 2, 3$$

und $\lambda = \frac{1}{18}$. Das lineare Gleichungssystem (8.7) ist gegeben durch

$$\frac{1}{4} \sum_{j \neq i} p(j) + p(i) = \frac{1}{2} \quad \textit{für } i = 1, 2, 3$$

und wird offensichtlich gelöst durch

$$p(i) = \frac{1}{3} \quad \forall i = 1, 2, 3.$$

Die Gewichte a_{si}^* für $i \in s$ der Schätzvektoren $\underline{a}_s^*$ sind alle gleich $\frac{3}{2}$. Eine Maximin-Strategie ist also gegeben durch $\underline{\mu} = \frac{1}{3}\underline{1}$, die Minimax-Strategie ist die Standardstrategie, und der Wert des Spiels ist $\lambda = \frac{1}{18}$.

Definieren wir den affin-linearen Schätzer t^* durch

$$t^*(s, \underline{y}) := \sum_{i \in s} y_i + \frac{1}{6} \quad \forall s \in S(2), \forall \underline{y} \in \Theta_Q,$$

so gilt

$$\max_{\underline{y} \in \Theta_Q} MSE(\underline{y}, (p_0, t^*)) = \frac{1}{36},$$

wobei p_0 die einfache Zufallsauswahl bezeichnet. Tatsächlich ist (p_0, t^*) eine Minimax-Strategie bzgl. der Strategien-Menge

$$\{(p, t); t \textit{ ist affin-linearer Schätzer}, \sum_s |s| p(s) = 2\}.$$

Die allgemeinere Aussage findet man in Gabler [6], Satz 5 auf Seite 41.

8.3 Der Zwei-Schichten-Fall

Sei $x_1 = x_2 := k$, $x_3 = 1 - 2k =: h$. Nach (8.4) existiert eine Indifferenz-Strategie der Natur genau dann, wenn

$$1 - \frac{\sqrt{2}}{2} \leq k < \frac{1}{2}.$$

Wir gehen davon aus, dass k der obigen Bedingung genügt. Die optimalen Schätzvektoren bzgl. der Indifferenz-Strategie berechnen sich mittels (vgl. Abschnitt 8.1)

$$\xi_1 = \xi_2 = \frac{1}{2k} - 1,\ \xi_3 = \frac{4k-1}{2k^2} - 1,$$

und der potentielle Spielwert ist gegeben durch (vgl. (8.6))

$$\lambda = \frac{1}{2}\frac{h^2(4k^2 - h^2)}{2k^2 + h^2} = \frac{1}{2}\frac{(1-2k)^2(4k-1)}{6k^2 - 4k + 1}.$$

Das lineare Gleichungssystem (8.7) wird gelöst durch

$$p(1) = p(2) = 3\frac{k^2h^2}{(4k^2 - h^2)(2k^2 + h^2)} = 3\frac{k^2(1-2k)^2}{(4k-1)(6k^2 - 4k + 1)}$$

und

$$p(3) = \frac{8k^4 - 4k^2h^2 - h^4}{(4k^2 - h^2)(2k^2 + h^2)} = \frac{-24k^4 + 48k^3 - 28k^2 + 8k - 1}{(4k-1)(6k^2 - 4k + 1)}.$$

Offensichtlich ist $p(1) = p(2) > 0$ für $1 - \frac{\sqrt{2}}{2} \leq k < \frac{1}{2}$. Allerdings ist $p(3)$ für $k < \frac{1}{2} + \frac{1}{6}\sqrt{3} - \frac{1}{6}\sqrt{3 + 3\sqrt{3}}$ negativ. (Es gilt $1 - \frac{\sqrt{2}}{2} < 0.2929 < 0.3115 < \frac{1}{2} + \frac{1}{6}\sqrt{3} - \frac{1}{6}\sqrt{3 + 3\sqrt{3}}$.) Also besteht auch im Quader-Fall das Problem, dass zu vorgegebener Indifferenz-Strategie der Natur im Allgemeinen kein Auswahlplan existiert, der das Gleichungssystem (8.7) löst. Wie soll der Statistiker vorgehen, wenn $p(3)$ negativ ist? Aufgrund der Repräsentativität und der symmetrischen Bedingungen definieren wir

$$p(1)^* := p(2)^* := \frac{1}{2}$$

und

$$\xi(1,\alpha) := (-1, \frac{k - h\alpha}{k}, \alpha)^t,\ \xi(2,\alpha) := (\frac{k - h\alpha}{k}, -1, \alpha)^t.$$

Um die Natur indifferent zu stellen, suchen wir ein $\alpha \in \mathbb{R}$ und ein $\lambda \in \mathbb{R}_+$ mit

$$\frac{1}{2}\sum_{i=1}^{2}(\underline{x}_j, \xi(i,\alpha))_2^2 = \lambda \quad \text{für } j = 1, 2, 3.$$

Wir erhalten das Gleichungssystem

$$\begin{aligned}2\lambda &= 2k^2 - 2kh\alpha + h^2\alpha^2,\\ 2\lambda &= 2h^2\alpha^2.\end{aligned}$$

Durch Gleichsetzen ergibt sich

$$\alpha = -\frac{k}{h}(1 \pm \sqrt{3});$$

wir wählen

$$\alpha = \frac{k}{h}(\sqrt{3} - 1) =: \alpha^*$$

und erhalten

$$\lambda = k^2(\sqrt{3} - 1)^2.$$

(Die Wahl $\alpha = -\frac{k}{h}(\sqrt{3} + 1)$ führt zu $\lambda_- = k^2(\sqrt{3} + 1)^2 > \lambda$.)

Es verbleibt zu zeigen, dass eine Strategie $\underline{\mu}$ der Natur existiert mit

$$\sum_i \mu_i(\underline{x}_i, \xi(j, \alpha^*))_2^2 = \min_{\underline{a} \in A_j} \sum_i \mu_i(\underline{x}_i, \underline{a} - \underline{1})_2^2 = \lambda \quad \text{für } j = 1, 2 \tag{8.8}$$

und

$$\min_{\underline{a} \in A_3} \sum_i \mu_i(\underline{x}_i, \underline{a} - \underline{1})_2^2 \geq \lambda, \tag{8.9}$$

wobei (vgl. Definition 1.3)

$$A_j := A_{\{1,2,3\}\setminus\{j\}} \quad \text{für } j = 1, 2, 3.$$

Aufgrund der symmetrischen Bedingungen setzen wir

$$\mu_1 = \mu_2, \ \ \mu_3 = 1 - 2\mu_1.$$

Wegen (8.2), (8.3) und (8.8) muss gelten

$$\alpha^* = \frac{\mu_1 k}{(1 - \mu_1)h},$$

also

$$\mu_1 = \mu_2 = 1 - \frac{1}{\sqrt{3}}, \ \ \mu_3 = \frac{2}{\sqrt{3}} - 1.$$

(Die Bedingung $\sum_i \mu_i(\xi(j, \alpha^*), \underline{x}_i)_2^2 = \lambda$ für $j = 1, 2$ ist trivialerweise erfüllt.) Weiter gilt (vgl. (8.1))

$$\min_{\underline{a} \in A_3} \sum_i \mu_i(\underline{a} - \underline{1}, \underline{x}_i)_2^2 = \frac{\mu_1^2 + 2\mu_1\mu_3}{1 - \mu_3} h^2 = \frac{\sqrt{3} - 1}{2} h^2.$$

Somit ist (8.9) genau dann erfüllt, wenn

$$k \in]0; \frac{1}{2} + \frac{1}{6}\sqrt{3} - \frac{1}{6}\sqrt{3 + 3\sqrt{3}}].$$

Zusammenfassung Im Fall $k \in [\frac{1}{2} + \frac{1}{6}\sqrt{3} - \frac{1}{6}\sqrt{3 + 3\sqrt{3}}, \frac{1}{2}[$ ist die Minimax-Strategie festgelegt durch

$$p^*(1) = p^*(2) = 3\frac{k^2(1-2k)^2}{(4k-1)(6k^2-4k+1)},$$

$$\underline{a}^*_{\{1,2\}} = (\frac{1}{2k}, \frac{1}{2k}, 0)^t,\ \underline{a}^*_{\{1,3\}} = (\frac{4k-1}{2k^2}, 0, \frac{1}{2k})^t,\ \underline{a}^*_{\{2,3\}} = (0, \frac{4k-1}{2k^2}, \frac{1}{2k})^t.$$

Obwohl für $k \in [1 - \frac{\sqrt{2}}{2}, \frac{1}{2} + \frac{1}{6}\sqrt{3} - \frac{1}{6}\sqrt{3 + 3\sqrt{3}}[$ eine Indifferenz-Strategie der Natur existiert, ist diese keine Maximin-Strategie.

Für $k \in]0, \frac{1}{2} + \frac{1}{6}\sqrt{3} - \frac{1}{6}\sqrt{3 + 3\sqrt{3}}]$ ist die Minimax-Strategie gegeben durch

$$p^*(1) = p^*(2) = \frac{1}{2},$$

$$\underline{a}^*_{\{1,3\}} = (3 - \sqrt{3}, 0, \frac{1 - (3 - \sqrt{3})k}{h})^t,\ \underline{a}^*_{\{2,3\}} = (0, 3 - \sqrt{3}, \frac{1 - (3 - \sqrt{3})k}{h})^t.$$

Die Maximin-Strategie ist gegeben durch

$$\underline{\mu}^* = \frac{1}{\sqrt{3}}(\sqrt{3} - 1, \sqrt{3} - 1, 2 - \sqrt{3})^t,$$

und das Spiel hat den Wert

$$\lambda = k^2(\sqrt{3} - 1)^2.$$

8.4 Der äußere Zylinder

Sei $\underline{x} \in \mathbb{R}^3$ ein Regressionsvektor und $U \in \mathcal{Q}$ mit

$$U\underline{x} = \underline{0} \tag{8.10}$$

und

$$\underline{x}_i^t U \underline{x}_i = 1 \quad \forall i = 1, 2, 3. \tag{8.11}$$

Durch (8.10) und (8.11) ist die Matrix U eindeutig festgelegt, und es gilt

$$U = \begin{bmatrix} \frac{1}{x_1^2} & -\frac{1}{2x_1x_2} & -\frac{1}{2x_1x_3} \\ -\frac{1}{2x_1x_2} & \frac{1}{x_2^2} & -\frac{1}{2x_2x_3} \\ -\frac{1}{2x_1x_3} & -\frac{1}{2x_2x_3} & \frac{1}{x_3^2} \end{bmatrix}.$$

Offensichtlich ist U positiv semidefinit, $\text{Kern}(U) = \text{span}[\underline{x}]$, und Θ_U ist der kleinste Zylinder entlang der Geraden $L_x := \{\beta\underline{x}; \beta \in \mathbb{R}\}$, der Θ_Q enthält. Wir werden im Folgenden die Spiele $(\mathcal{Y}_U, \mathcal{A}_x)$ und $(\mathcal{Y}_Q, \mathcal{A}_x)$ vergleichen. Dazu bezeichnen wir mit

$$(V^*, W_U^*, \lambda_U)$$

eine Lösung des Spiels $(\mathcal{Y}_U, \mathcal{A}_x)$ und mit

$$(\underline{\mu}^*, W_Q^*, \lambda_Q)$$

eine Lösung des Spiels $(\mathcal{Y}_Q, \mathcal{A}_x)$.

Aufgrund von Satz 4.4 ist $W_U^* = \lambda_U U$. Da $\underline{x}_i \in \Theta_U$ für $i = 1, 2, 3$, ist W_U^* eine Indifferenzmatrix des Statistikers bezüglich des Spiels $(\mathcal{Y}_Q, \mathcal{A}_x)$, und es gilt

$$\lambda_Q \leq \lambda_U.$$

Diese Ungleichung folgt auch aus $\mathcal{Y}_Q \subsetneq \mathcal{Y}_U$ (vgl. (8.5)).

Bemerkung 8.5 *Da $u_{ij} = -\frac{1}{x_i x_j} < 0$ für $i \neq j$, gilt aufgrund von Bemerkung 4.6*

$$\lambda_Q \leq \lambda_U \leq 2x_1x_2x_3.$$

Ist andererseits $W_Q^* = \lambda_Q U$, so ist W_Q^* eine Minimax-Strategie des Spiels $(\mathcal{Y}_U, \mathcal{A}_x)$, und es gilt

$$\lambda_Q = \lambda_U.$$

Als eine Konsequenz aus (8.10) und (8.11) ist jede Indifferenzmatrix des Statistikers bzgl. $\mathcal{Y}_Q$ ein Vielfaches von U. Somit ist das Gleichungssystem (8.7) aus Abschnitt 8.2 äquivalent zu

$$p(i) = \frac{\xi_i}{\xi_1\xi_2\xi_3 + \xi_i^2}(\xi_i + \lambda u(i)) \quad \text{für } i = 1, 2, 3,$$

wobei

$$u(i) = -\frac{x_i}{2x_1x_2x_3} \quad \text{für } i = 1, 2, 3$$

und

$$\lambda = \frac{1 - \sum_i \frac{\xi_i^2}{\xi_1\xi_2\xi_3 + \xi_i^2}}{\sum_i \frac{u(i)\xi_i}{\xi_1\xi_2\xi_3 + \xi_i^2}}$$

(vgl. Abschnitt 4.5, (4.18)). Im Fall der Lösbarkeit mit $p(i) \geq 0$ ist die dadurch definierte Strategie eine Minimax-Strategie beider Spiele (unter der Voraussetzung, dass $\sum_j x_j^2 - 2x_0^2 \geq 0$). Betrachten wir den Zwei-Schichten-Fall mit den üblichen Bezeichnungen, so folgt aus den Berechnungen des letzten Abschnitts für $k \in]0, \frac{1}{2} + \frac{1}{6}\sqrt{3} - \frac{1}{6}\sqrt{3+3\sqrt{3}}]$:

$$W^* = k^2(\sqrt{3}-1)^2 U$$

ist Minimax-Strategie des Spiels $(\mathcal{Y}_Q, \mathcal{A}_x)$ und somit von beiden Spielen.

Kapitel 9

Zusammenfassung

Dieses Kapitel fasst die wichtigsten Ergebnisse dieser Arbeit zusammen.

Gegeben sei eine Gesamtheit bestehend aus den Einheiten $1, \ldots, N$, wobei jeder Einheit ein uns nicht bekannter Wert $y_i \in \mathbb{R}$, $i = 1, \ldots, N$ eines interessierenden Merkmales zugeordnet ist. Wir interessieren uns für die Merkmalsumme

$$y := \sum_i y_i.$$

Um y zu schätzen wählen wir eine Stichprobe

$$s = \{i_1, \ldots, i_n\} \subseteq \{1, \ldots, N\}$$

vom Umfang n zufällig aus und verwenden eine lineare Funktion

$$\sum_{i \in s} a_{si} y_i$$

der beobachteten Merkmalswerte y_i, $i \in s$, als Schätzwert. Mit

$$\underline{y} := (y_1, \ldots, y_N)^t \in \mathbb{R}^N$$

und

$$\underline{a}_s := (a_{s1}, \ldots, a_{sN})^t \in \mathbb{R}^N,$$

wobei $a_{si} := 0 \quad$ für $i \notin s$, gilt

$$\sum_{i \in s} a_{si} y_i = \underline{a}_s^t \underline{y} = (\underline{a}_s, \underline{y})_2.$$

Wie wir auswählen und wie wir schätzen wird stark von unseren a-priori-Informationen abhängen. Oftmals kennt man die Werte eines oder mehrerer Hilfsmerkmale, die die Menge der a-priori möglichen Parameter eingrenzen.

Der Quader Angenommen, wir kennen die Beschäftigtenzahlen von N Betrieben einer Branche und interessieren uns für die Krankmeldungen an einem bestimmten Tag. Bezeichnen wir mit x_i die Beschäftigtenzahl des i-ten Betriebs und mit y_i die Krankmeldungen im i-ten Betrieb, dann gilt für den Parameter $\underline{y} \in \mathbb{R}^N$

$$0 \leq y_i \leq x_i \quad \forall i = 1, \ldots, N.$$

Den resultierenden Parameterraum bezeichnen wir mit Θ_Q, also

$$\Theta_Q := \{\underline{y} \in \mathbb{R}^N; 0 \leq y_i \leq x_i \quad \forall i = 1, \ldots, N\}.$$

Das Ellipsoid Angenommen, man interessiert sich für den Jahresumsatz y von N Betrieben einer Branche und kennt deren Mitarbeiterzahlen. Sei x_i die Anzahl der Mitarbeiter des i-ten Betriebs und $\underline{x} = (x_1, \ldots, x_N)^t \in \mathbb{R}^N$. Es lässt sich vermuten, dass für den Parameter $\underline{y}$ in etwa gilt

$$\underline{y} = \beta \underline{x}$$

für ein (unbekanntes) $\beta \in \mathbb{R}$. Allgemeiner gehen wir von der folgenden Situation aus: Sei X eine $N \times K$-Matrix mit vollem Spaltenrang (eine Regressionsmatrix). Wir nehmen an, dass der Parameter $\underline{y}$ in der Nähe der Unterraums

$$\text{Bild}(X) = \{X\underline{\beta}; \underline{\beta} \in \mathbb{R}^K\}$$

liegt. Der Begriff der Nähe wird modelliert durch eine symmetrische positiv semidefinite Matrix U vom Rang $N - K$ mit $UX = 0$: $\underline{y}$ liegt in der Nähe von Bild(X), falls $\underline{y}^t U \underline{y}$ klein ist. Den resultierenden Parameterraum bezeichnen wir mit Θ_U, also

$$\Theta_U := \{\underline{y} \in \mathbb{R}^N; \underline{y}^t U \underline{y} \leq 1\}.$$

Ein Beispiel Sei $\underline{x} \in \mathbb{R}^N$ ein Regressionsvektor und $U = \text{diag}^{-1}(\underline{x}) - \underline{1}\,\underline{1}^t$. Dann ist

$$\Theta_U = \{\underline{y} \in \mathbb{R}^N; \sum_{i=1}^{N} x_i (\frac{y_i}{x_i} - y)^2 \leq 1\}.$$

In diesem Fall bezeichnet man Θ_U als HH-Raum.

Die Stichprobe s wird mittels eines Auswahlplans p ermittelt, und eine Strategie des Statistikers ist gegeben durch einen Auswahlplan p und einer Menge

$\{\underline{a}_s; p(s) > 0\}$ von Schätzvektoren. Das Risiko einer Strategie $(p, \{\underline{a}_s\})$ bzgl. eines Parameters $\underline{y} \in \mathbb{R}^N$ ist gegeben durch den quadratischen Verlust (MSE):

$$\begin{aligned} \mathrm{MSE}(\underline{y}, (p, \{\underline{a}_s\})) &= \sum_s p(s)(\underline{y}^t \underline{a}_s - y)^2 = \sum_s p(s)(\underline{y}, \underline{a}_s - \underline{1})_2^2 \\ &= \sum_s p(s)(\underline{a}_s - \underline{1})^t \underline{y}\, \underline{y}^t (\underline{a}_s - \underline{1}) \\ &= \mathrm{tr}[\underline{y}\, \underline{y}^t \sum_s p(s)(\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t], \end{aligned}$$

wobei tr den Spuroperator bezeichnet. Offensichtlich sind

$$\underline{y}\, \underline{y}^t, \ \sum_s p(s)(\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t$$

symmetrische positiv semidefinite $N \times N$-Matrizen. Dabei ist

$$\sup_{\underline{y} \in \Theta_U} \mathrm{MSE}(\underline{y}, (p, \{\underline{a}_s\}))$$

genau dann endlich, wenn die Schätzvektoren repräsentativ sind, d.h.

$$X^t(\underline{a}_s - \underline{1}) = \underline{0} \quad \forall s : p(s) > 0.$$

In diesem Fall nennen wir $\sum_s p(s)(\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t$ eine risikoerzeugende Matrix, und wir bezeichnen mit

$$\mathcal{A}_X$$

die Menge aller risikoerzeugenden Matrizen.

Bemerkung Die Schätzung mittels einer repräsentativen Strategie ist für alle Parameter aus $\mathrm{Bild}(X)$ exakt, d.h. das Risiko ist Null. Unter der Annahme, dass der Parameter in der Nähe des Unterraums $\mathrm{Bild}(X)$ liegt, ist die Verwendung von repräsentativen Strategien naheliegend.

Für einen Parameterraum Θ sei $\mathcal{Y}$ die konvexe Hülle aller $\underline{y}\,\underline{y}^t$ mit $\underline{y} \in \Theta$. Falls $\Theta = \Theta_U$ bzw. $\Theta = \Theta_Q$, schreiben wir $\mathcal{Y}_U$ bzw. $\mathcal{Y}_Q$ für den Strategienraum der Natur. Bezeichnen wir mit $\mathcal{Q}$ den Vektorraum aller symmetrischen $N \times N$-Matrizen und mit $\mathcal{Q}_0 \subseteq \mathcal{Q}$ die Menge aller positiv semidefiniten Matrizen, so gilt

$$\mathcal{Y}, \ \mathcal{A}_X \subseteq \mathcal{Q}_0 \subseteq \mathcal{Q}.$$

Durch

$$\langle Q_1, Q_2 \rangle := \mathrm{tr}(Q_1 Q_2) \quad \forall Q_1, Q_2 \in \mathcal{Q}$$

ist ein Skalarprodukt auf $\mathcal{Q}$ definiert, und $(\mathcal{Q}, \langle\, .\, ,\, .\, \rangle)$ ist ein Hilbertraum. Das Auswertungsproblem lässt sich als lineares Spiel auffassen: $\mathcal{Y}$ ist der Strategienraum der Natur und $\mathcal{A}_X$ der Strategienraum des Statistikers. Die Natur wählt eine Matrix $V \in \mathcal{Y}$ und der Statistiker eine Matrix $W \in \mathcal{A}_X$. Der Verlust des Statistikers (bzw. der Gewinn der Natur) ist gegeben durch $\mathrm{tr}(VW) = \langle V, W \rangle$. Wenn eine Maximin-Strategie V^* und eine Minimax-Strategie W^* existiert, heißt das Spiel $(\mathcal{Y}, \mathcal{A}_X)$ lösbar, und wir bezeichnen mit

$$\langle V^*, W^* \rangle$$

den Wert des Spiels. Das Tupel (V^*, W^*) nennen wir ein Gleichgewicht.

Nach Satz 3.1 bzw. Satz 7.1 besitzt das Spiel $(\mathcal{Y}_U, \mathcal{A}_X)$ bzw. $(\mathcal{Y}_Q, \mathcal{A}_x)$ ein Gleichgewicht (V^*, W^*); speziell existiert in beiden Fällen eine zulässige Minimax-Strategie. Dabei haben wir in Beispiel 1.8, Beispiel 3.11 und Beispiel 6.8 gesehen, dass sich die Minimax-Strategie i. A. nicht eindeutig in einen Auswahlplan p^* und einen Schätzer t^* zerlegen lässt. Dabei zeigt Beispiel 6.8, dass man p^* sogar so wählen kann, dass gewisse Einheiten nicht berücksichtigt werden.

Das Indifferenz-Prinzip Im Allgemeinen existiert keine Strategie $V^* \in \mathcal{Y}_U$ mit

$$\langle V^*, W \rangle = \lambda \quad \forall W \in \mathcal{A}_X$$

für ein $\lambda \in \mathbb{R}$. Allerdings existiert unter gewissen Umständen (in Abhängigkeit von X und n) eine Matrix $V^* \in \mathcal{Y}_U$ mit

$$\min_{\underline{a}_s \in A_s} \langle V^*, (\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t \rangle = \lambda \quad \forall s \in S(n).$$

In diesem Fall nennen wir V^* eine Indifferenz-Strategie der Natur. Seien

$$\underline{a}_s^* \in \mathrm{argmin}_{\underline{a}_s \in A_s} \langle V^*, (\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t \rangle \quad \forall s \in S(n).$$

Falls ein Auswahlplan p^* existiert mit

$$W^* := \sum_s p^*(s)(\underline{a}_s^* - \underline{1})(\underline{a}_s^* - \underline{1})^t = \mu U,$$

ist $\mu = \lambda$ der Wert des Spiels und (V^*, W^*) ein Gleichgewicht (vgl. Satz 3.7).

Im Verlauf dieser Arbeit hat sich herausgestellt, dass das Indifferenz-Prinzip für die Lösung des Spiels $(\mathcal{Y}_U, \mathcal{A}_X)$ sehr hilfreich ist. Es stellen sich folgende Fragen:

(i) Unter welchen Umständen existieren Indifferenz-Strategien?

(ii) Können reine Strategien Indifferenz-Strategien sein?

(iii) Sind Minimax- bzw. Maximin-Strategien immer Indifferenz-Strategien?

(iv) Existieren Indifferenz-Strategien mit unterschiedlichem Risiko?

ad (i): In Abschnitt 3.1 haben wir gesehen, dass die Existenz von Indifferenz-Strategien der Natur unabhängig ist von der Matrix U. (Dass die Existenz vom Stichprobenumfang n abhängt, zeigt uns Beispiel 4.8). Dies gilt nicht für den Statistiker: Eine risikoerzeugende Matrix ist genau dann eine Indifferenz-Strategie des Statistikers, wenn sie ein Vielfaches von U ist (vgl. Lemma 3.12). Beispiel 3.16 zeigt, dass der Statistiker je nach Wahl von U kein Vielfaches von U spielen kann, er besitzt also im Allgemeinen keine Indifferenz-Strategie (vgl. Skizze 2.21: Die Gerade cU, $c \in \mathbb{R}$, schneidet $\widetilde{\mathcal{A}}_X$ nicht).

ad (ii): Für eine reine Strategie $(\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t$, $\underline{a}_s \in A_s$, des Statistikers gilt

$$\dim(\mathrm{Kern}((\underline{a}_s - \underline{1})(\underline{a}_s - \underline{1})^t)) = N - 1.$$

Falls $\dim(\mathrm{Kern}(U)) = \dim(\mathrm{Bild}(X)) = K < N - 1$, ist keine reine Strategie des Statistikers eine Indifferenz-Strategie.

Eine reine Strategie $\underline{y}^*(\underline{y}^*)^t$, $\underline{y}^* \in \Theta_U$, der Natur kann eine Indifferenz-Strategie sein, wie wir in Abschnitt 3.5 (der HH-Raum mit $\max\{x_1, \ldots, x_N\} = \frac{1}{2}$) und Beispiel 6.8 gesehen haben.

ad (iii): Dies gilt weder für den Statistiker noch für die Natur: Im Fall Eins aus N mit $\max\{x_1, \ldots, x_N\} > \frac{1}{2}$ ist für den HH-Raum die Minimax-Strategie W^* gegeben durch die bewusste Auswahl der Einheit mit dem größten x-Wert, obwohl die HH-Strategie eine Indifferenz-Strategie ist, und es gilt $W^* \neq \lambda U$ (vgl. Satz 3.15 und Skizze 2.21: Die Gerade cU, $c \in \mathbb{R}$, schneidet $\widetilde{\mathcal{A}}_X$, aber nicht den Rand $\Lambda(\mathcal{A}_X)$). Im Fall Zwei aus Drei ist jede Minimax-Strategie eine Indifferenz-Strategie des Statistikers, also ein Vielfaches von U (vgl. Satz 4.4 und Skizze 2.21: Die Gerade cU, $c \in \mathbb{R}$, schneidet den Rand $\Lambda(\mathcal{A}_X)$). Falls $\underline{x} = (k, k, h)$ und der Parameterraum durch den HH-Raum gegeben ist, ist die Maximin-Strategie für $k \in]\frac{1}{4}; \frac{\sqrt{5}-1}{4}[$ keine Indifferenz-Strategie der Natur, obwohl eine solche existiert (siehe Zusammenfassung in Abschnitt 4.6).

ad (iv): Ja, wie uns Beispiel 5.2 für die Natur zeigt, und was für den Statistiker z.B. aus dem Beweis von Lemma 4.5 folgt.

Für den HH-Raum ist das Spiel $(\mathcal{Y}_U, \mathcal{A}_X)$ für folgende Fälle gelöst:

- $n = 1$ und N, $\underline{x} \in \mathbb{R}^N$ beliebig: Satz 3.13 und Satz 3.15.

- $n = 2$, $N = 3$ und $\underline{x} \in \mathbb{R}^3$ in einer Umgebung von $\frac{1}{3}\underline{1} \in \mathbb{R}^3$: Kapitel 4.

- $n = 2$ und $N = 3$, falls zwei Komponenten des Regressionsvektors $\underline{x} \in \mathbb{R}^3$ gleich sind: Zusammenfassung in Abschnitt 4.6.

- $n = 2$, N beliebig und $\underline{x} = (k, \ldots, k, h, h) \in \mathbb{R}^N$ für große k-Werte: Beispiel 5.4.

- $n = 2$, N beliebig und $\underline{x} = (k, \ldots, k, h) \in \mathbb{R}^N$ für kleine k-Werte: Satz 5.5.

Falls die Regressionsmatrix eine Blockmatrix ist, haben wir das Spiel für den Parameterraum

$$\Theta_\alpha = \{\underline{y} \in \mathbb{R}^{\mathbb{N}}; \sum_{h=1}^{H} \alpha_h s_{yy}(h) = 1\}$$

untersucht, wobei $\underline{\alpha} \in \mathbb{R}^H$ einen Wahrscheinlichkeitsvektor mit positiven Komponenten bezeichnet. In diesem Fall ist die Minimax-Strategie eine Schichtungsstrategie (vgl. Satz 6.1), und die optimale Aufteilung lässt sich mittels des Wahrscheinlichkeitsvektors $\underline{\alpha}$ und der Größe $N(h)$, $h = 1, \ldots, H$, der H Schichten abschätzen (vgl. Satz 6.5).

Für den Quader Θ_Q wurden folgende Fälle behandelt:

- $n = 1$, N beliebig und $x_{K_0} = \frac{1}{2}$ für ein $K_0 \subseteq \{1, \ldots, N\}$: Satz 7.5 (Scott und Smith).

- $n = 2$, $N = 3$ und $\underline{x} \in \mathbb{R}^3$ in einer Umgebung von $\frac{1}{3}\underline{1} \in \mathbb{R}^3$: Kapitel 8.

- $n = 2$ und $N = 3$, falls zwei Komponenten des Regressionsvektors $\underline{x} \in \mathbb{R}^3$ gleich sind: Zusammenfassung in Abschnitt 8.3.

- n, N beliebig und $\underline{x} = (1, \ldots, 1, r) \in \mathbb{R}^N$ für große r-Werte: Abschnitt 7.3.

Unter der Voraussetzung, dass der Regressionsvektor die Gesamtheit in $H < n$ Schichten zerlegt, und wir nur "stark repräsentativ" schätzen, wurde in Abschnitt 7.4 gezeigt, dass die Minimax-Strategie eine Schichtungsstrategie ist und die optimale Aufteilung berechnet.

Anhang A

Die Lemmata A.1, A.3, A.5 und Korollar A.2 sind in ähnlicher Form bei Stenger, Gabler und Schmidt [27] zu finden.

Lemma A.1 *Seien $P, Q \in \mathcal{Q}_0$. Dann gilt*

$$tr(PQ) \geq 0,$$

und folgende Aussagen sind äquivalent:

$$(i)\ tr(PQ) = 0. \quad (ii)\ Bild(Q) \subseteq Kern(P) \quad (iii)\ Bild(P) \subseteq Kern(Q).$$

Beweis: Da Q positiv semidefinit ist, existiert eine Orthogonal-Basis von Eigenvektoren $(\underline{z}^1, \ldots, \underline{z}^N)$ bezüglich der Eigenwerte $0 \leq \lambda_1 \leq \cdots \leq \lambda_N$ mit

$$Q = \sum_{i=1}^{N} \lambda_i \underline{z}^i (\underline{z}^i)^t.$$

Es gilt

$$\mathrm{tr}(PQ) = \sum_{i=1}^{N} \lambda_i \mathrm{tr}(P\underline{z}^i(\underline{z}^i)^t) = \sum_{i=1}^{N} \lambda_i \underbrace{(\underline{z}^i)^t P \underline{z}^i}_{\geq 0} \geq 0,$$

und der erste Teil des Lemmas ist bewiesen.

Sei $\dim(\mathrm{Kern(Q)}) = M$, also $Q = \sum_{i=M+1}^{N} \lambda_i \underline{z}^i (\underline{z}^i)^t$ und

$$\mathrm{tr}(PQ) = \sum_{i=M+1}^{N} \underbrace{\lambda_i}_{>0} \underbrace{(\underline{z}^i)^t P \underline{z}^i}_{\geq 0}.$$

Somit ist

$$\begin{aligned} & \mathrm{tr}(PQ) = 0 \\ \Leftrightarrow\quad & (\underline{z}^i)^t P \underline{z}^i = 0 \quad \text{für } i = M+1, \ldots, N \\ \Leftrightarrow\quad & \mathrm{Bild}(Q) \subseteq \mathrm{Kern}(P), \end{aligned}$$

da $\mathrm{Bild}(Q) = \mathrm{span}[\underline{z}^{M+1}, \ldots, \underline{z}^N]$, und wir erhalten (i) $\Leftrightarrow$ (ii). Aus Symmetriegründen gilt (i) $\Leftrightarrow$ (iii). $\diamond$

Korollar A.2 *Seien $U, Q \in \mathcal{Q}_0 \backslash \{0\}$ mit Kern$(U) \subseteq$ Kern(Q). Dann ist*

$$tr(UQ) > 0.$$

Beweis: Da $Q \in \mathcal{Q}_0 \backslash \{0\}$, existiert ein Eigenvektor $\underline{z} \notin \mathrm{Kern}(Q)$ von Q mit $\underline{z} \in \mathrm{Bild}(Q)$. Wegen $\mathrm{Kern}(U) \subseteq \mathrm{Kern}(Q)$ gilt $\underline{z} \notin \mathrm{Kern}(U)$, also $\mathrm{Bild}(Q) \not\subseteq \mathrm{Kern}(U)$. Mit Lemma A.1 folgt die Behauptung. $\diamond$

Lemma A.3 *Seien $U, Q \in \mathcal{Q}_0$ und Kern$(U) \subseteq$ Kern(Q). Dann existiert ein $\alpha > 0$ mit*

$$\alpha U \succeq Q.$$

Beweis: Falls $Q = 0$, ist die Aussage offensichtlich richtig. Sei $Q \in \mathcal{Q}_0 \backslash \{0\}$,

$$\lambda_m = \min\{\underline{y}^t U \underline{y}; \underline{y} \in \mathrm{Bild}(U), |\underline{y}|_2 = 1\} > 0$$

der kleinste von Null verschiedene Eigenwert von U,

$$\alpha_M := \max_{|\underline{y}|_2 = 1} \underline{y}^t Q \underline{y} > 0$$

und

$$\alpha := \frac{\alpha_M}{\lambda_m} > 0.$$

Dann ist

$$\underline{y}^t \alpha U \underline{y} = \alpha \underline{y}^t U \underline{y} \geq \alpha \lambda_m = \alpha_M \geq \underline{y}^t Q \underline{y}$$

für alle $\underline{y} \in \mathrm{Bild}(U)$ mit $|\underline{y}|_2 = 1$, also wegen $\mathrm{Kern}(U) \subseteq \mathrm{Kern}(Q)$

$$\underline{y}^t \alpha U \underline{y} \geq \underline{y}^t Q \underline{y} \quad \forall \underline{y} \in \mathbb{R}^N.$$

$\diamond$

Lemma A.4 *Sei X eine $N \times K$-Matrix, wobei $1 \leq K < N$, und $U \in \mathcal{Q}_0$ mit Kern(U) = Bild(X). Dann gilt*

$$\mathit{Kern}(X^t) = \mathit{Bild}(U).$$

Beweis:

"$\supseteq$": Sei $\underline{u} \in \text{Bild}(U)$, also $\underline{u} = U\underline{z}$ für ein $\underline{z} \in \mathbb{R}^N$. Dann ist

$$X^t\underline{u} = X^tU\underline{z} = \underline{0},$$

also $\underline{u} \in \text{Kern}(X^t)$.

"$\subseteq$": Sei andererseits $\underline{u} \in \text{Kern}(X^t)$, also $X^t\underline{u} = \underline{0}$. Da $\text{Kern}(U) = \text{Bild}(X)$ und U positiv semidefinit ist, existieren Eigenvektoren $\underline{u}^1, \ldots, \underline{u}^M$ von U bzgl. der Eigenwerte $0 < \lambda_1 \leq, \ldots, \leq \lambda_M$ mit

$$\underline{u} = \sum_{i=1}^{M} a_i\underline{u}^i$$

für gewisse $a_i \in \mathbb{R}$, $i = 1, \ldots, M$. Sei

$$\underline{z} := \sum_{i=1}^{M} \frac{a_i}{\lambda_i}\underline{u}^i.$$

Dann ist

$$U\underline{z} = \sum_{i=1}^{M} \frac{a_i}{\lambda_i}U\underline{u}^i = \sum_{i=1}^{M} a_i\underline{u}^i = \underline{u},$$

also $\underline{u} \in \text{Bild}(U)$. $\diamond$

Lemma A.5 *Die Abbildung $\mathbb{R}^N \to \mathcal{Q}_0$, $\underline{y} \mapsto \underline{y}\,\underline{y}^t$ ist konvex bzgl. der Löwner-Halbordnung $\succeq$, d.h.*

$$\lambda\underline{y}\,\underline{y}^t + (1-\lambda)\underline{z}\,\underline{z}^t \succeq [\lambda\underline{y} + (1-\lambda)\underline{z}][\lambda\underline{y} + (1-\lambda)\underline{z}]^t$$

für alle $\lambda \in [0;1]$ und $\underline{y}, \underline{z} \in \mathbb{R}^N$.

Beweis: Sei $\lambda \in [0;1]$ und $\underline{y}, \underline{z} \in \mathbb{R}^N$. Dann gilt

$$\begin{aligned}
&\lambda\underline{y}\,\underline{y}^t + (1-\lambda)\underline{z}\,\underline{z}^t - [\lambda\underline{y} + (1-\lambda)\underline{z}][\lambda\underline{y} + (1-\lambda)\underline{z}]^t \\
&\quad = \lambda\underline{y}\,\underline{y}^t + (1-\lambda)\underline{z}\,\underline{z}^t - \lambda^2\underline{y}\,\underline{y}^t - \lambda(1-\lambda)\underline{z}\,\underline{y}^t - \lambda(1-\lambda)\underline{y}\,\underline{z}^t - (1-\lambda)^2\underline{z}\,\underline{z}^t \\
&\quad = \lambda(1-\lambda)(\underline{y} - \underline{z})(\underline{y} - \underline{z})^t \in \mathcal{Q}_0.
\end{aligned}$$

$\diamond$

Satz A.6 *$\mathcal{Q}_0$ ist abgeschlossen.*

Beweis: Wir zeigen, dass das Komplement $\mathcal{Q}_0^c$ von $\mathcal{Q}_0$ offen ist. Sei $A \notin \mathcal{Q}_0$. Dann existiert ein $\underline{z} \in \mathbb{R}^N$ mit

$$\alpha := \frac{\underline{z}^t A \underline{z}}{\underline{z}^t \underline{z}} < 0.$$

Sei $B \in \mathcal{Q}$ mit $|B|_{tr} < -\frac{\alpha}{2}$. Aufgrund der Schwarzschen Ungleichung gilt

$$\underline{z}^t B \underline{z} = \langle B, \underline{z}\,\underline{z}^t \rangle \leq |B|_{tr} \underline{z}^t \underline{z},$$

also

$$\frac{\underline{z}^t B \underline{z}}{\underline{z}^t \underline{z}} \leq -\frac{\alpha}{2}.$$

Daraus folgt

$$\frac{\underline{z}^t (A+B) \underline{z}}{\underline{z}^t \underline{z}} = \frac{\underline{z}^t A \underline{z}}{\underline{z}^t \underline{z}} + \frac{\underline{z}^t B \underline{z}}{\underline{z}^t \underline{z}} \leq \alpha - \frac{\alpha}{2} = \frac{\alpha}{2} < 0,$$

d.h.

$$\mathcal{U}_{-\frac{\alpha}{2}}(A) = \{A + B; |B|_{tr} < -\frac{\alpha}{2}\} \subseteq \mathcal{Q}_0^c.$$

$\diamond$

Sei X eine Regressionsmatrix (vom Rang K), $\mathcal{Q}_X := \{Q \in \mathcal{Q}; X^t Q = 0\}$, $\mathcal{Q}_{0X} := \{Q \in \mathcal{Q}_0; X^t Q = 0\}$ und $U \in \mathcal{Q}_{0X}$ mit $\dim(\text{Kern}(U))) = K$.

Korollar A.7 *$\mathcal{Q}_{0X}$ ist abgeschlossen (in $\mathcal{Q}_X$).*

Beweis: Der Beweis ergibt sich aus dem Beweis des vorhergehenden Satzes. $\diamond$

Lemma A.8 *Sei $c > 0$. Dann existiert ein $\epsilon > 0$ mit*

$$cU \succeq Q \quad \forall Q \in \mathcal{B}(\epsilon),$$

wobei $\mathcal{B}(\epsilon) := \{Q' \in \mathcal{Q}_X; |Q'|_{tr} \leq \epsilon\}$.

Beweis: Sei

$$\lambda_m = \min\{\underline{y}^t U \underline{y}; \underline{y} \in \text{Bild}(U), |\underline{y}|_2 = 1\} > 0$$

der kleinste von Null verschiedene Eigenwert von U und $\epsilon := c\lambda_m$. Dann gilt für ein $Q \in \mathcal{B}(\epsilon)$ aufgrund der Schwarzschen Ungleichung

$$\underline{y}^t cU \underline{y} \geq c\lambda_m \geq |Q|_{tr} |\underline{y}\,\underline{y}^t|_{tr} \geq \langle Q, \underline{y}\,\underline{y}^t \rangle = \underline{y}^t Q \underline{y}$$

für alle $\underline{y} \in \text{Bild}(U)$ mit $|\underline{y}|_2 = 1$ und somit

$$\underline{y}^t cU \underline{y} \geq \underline{y}^t Q \underline{y} \quad \forall \underline{y} \in \mathbb{R}^N.$$

$\diamond$

Lemma A.9 *Mit* $\mathcal{U}_\alpha := \{Q \in \mathcal{Q}_X ; \alpha U \succeq Q\}$ *für* $\alpha \in \mathbb{R}$ *ist*

$$int(\mathcal{U}_\alpha) = \bigcup_{\beta<\alpha} \mathcal{U}_\beta.$$

Beweis:

"$\subseteq$": Sei $R \in \mathrm{int}(\mathcal{U}_\alpha)$. Nach Definition des Inneren existiert ein $\epsilon > 0$ mit

$$\alpha U \succeq Q \quad \forall Q \in \mathcal{B}(R, \epsilon),$$

wobei $\mathcal{B}(R,\epsilon) = \{Q \in \mathcal{Q}_X ; |R - Q|_{tr} \leq \epsilon\}$. Mit $\epsilon' := \frac{\epsilon}{|U|_{tr}}$ ist

$$|R - (R + \epsilon' U)|_{tr} = \epsilon,$$

also $R + \epsilon' U \in \mathcal{B}(R, \epsilon)$ und

$$(\alpha - \epsilon')U - R = \alpha U - (R + \epsilon' U) \succeq 0,$$

d.h.

$$R \in \mathcal{U}_\beta$$

für $\beta := \alpha - \epsilon'$.

"$\supseteq$": Sei $\beta < \alpha$ und $R \in \mathcal{U}_\beta$. Es ist

$$\alpha U - (\alpha - \beta)U - R = \beta U - R \succeq 0$$

und somit

$$\alpha U - R \succeq (\alpha - \beta)U.$$

Aufgrund von Lemma A.8 existiert ein $\epsilon > 0$ mit

$$(\alpha - \beta)U \succeq Q \quad \forall Q \in \{Q' \in \mathcal{Q}_X ; |Q'|_{tr} \leq \epsilon\},$$

und es folgt

$$\mathcal{B}(R, \epsilon) \subseteq \mathcal{U}_\alpha.$$

$\diamond$

Symbolverzeichnis

$\underline{0}$	Der Nullvektor.
$\underline{1}$	Der Einsvektor.
$]a;b[$	Das offene Intervall von a bis b.
$[a;b]$	Das abgeschlossene Intervall von a bis b.
$\langle P, Q\rangle$	Die Spur des Produkts zweier symmetrischer Matrizen P, Q: $\mathrm{tr}(PQ)$.
$\lvert\,.\,\rvert_{tr}$	Die durch $\langle\,.\,,\,.\,\rangle$ induzierte Norm.
A	Die Menge der repräsentativen Schätzvektoren.
A_s	Die Menge der repräsentativen Schätzvektoren bzgl. der Stichprobe s.
$\mathcal{A}_x$, $\mathcal{A}_X$	Die Menge der risikoerzeugenden Matrizen.
$\widetilde{\mathcal{A}}_X$	Die konvexe Hülle von $\mathcal{A}_X$.
argmax	Die Menge der maximierenden Elemente.
argmin	Die Menge der minimierenden Elemente.
Bild	Das Bild einer Matrix: Der durch die Spaltenvektoren aufgespannte Vektorraum.
$\mathcal{BA}$	Die Menge der besten Antworten.
$\mathrm{diag}(\underline{z})$	Die Diagonalmatrix mit den Diagonalelementen $z_1, \ldots, z_N$.
$\underline{e}_i$	Der i-te Einheitsvektor.
HH-Raum	Der Parameterraum $\{\underline{y} \in \mathbb{R}^N; \sum_{i=1}^N x_i(\frac{y_i}{x_i} - y)^2 \leq 1\}$.
I	Die Einheitsmatrix.
Kern	Der Kern einer Matrix.
$\Lambda(\mathcal{S})$	Die Menge der unteren Grenzpunkte einer Menge $\mathcal{S} \subseteq \mathcal{Q}$.

MSE	Der erwartete quadratische Verlust.
n	Der Stichprobenumfang.
$\underline{n}$	Eine Aufteilung.
$\underline{n}(\,.\,)$	Eine zufällige Aufteilung.
$n(h)$	Der Stichprobenumfang aus der h-ten Schicht.
N	Die Größe der Gesamtheit.
$N(h)$	Die Größe der h-ten Schicht.
$\mathbb{N}$	Die Menge der natürlichen Zahlen.
p	Ein Auswahlplan, i.A. vom Umfang n.
$\pi_i(p)$, $\pi_{ii}(p)$	Die Inklusionswahrscheinlichkeit erster Ordnung von p.
$\pi_{ij}(p)$	Die Inklusionswahrscheinlichkeit zweiter Ordnung von p.
$\mathcal{Q}$	Der Vektorraum der symmetrischen Matrizen.
$\mathcal{Q}_0$	Die Menge der positiv semidefiniten Matrizen.
$\mathcal{Q}_X$	Der Vektorraum der symmetrischen Matrizen Q mit $\mathrm{Bild}(X) \subseteq \mathrm{Kern}(Q)$.
$\mathbb{R}$	Die Menge der reellen Zahlen.
$\mathbb{R}^N$	Die N-dimensionale Euklidische Raum.
$\mathbb{R}^H_+$	Die Menge der H-dimensionalen Vektoren mit positiven Komponenten.
$S(n)$	Die Menge der Stichproben vom Umfang n: $\{s \in 2^{\mathcal{U}}; \lvert s \rvert = n\}$.
$S(p)$	Der Träger von p: $\{s \in 2^{\mathcal{U}}; p(s) > 0\}$.
s_{yy}	Die korrigierte Varianz von $\underline{y}$: $\frac{1}{N-1}\sum_i (y_i - \overline{y})^2$.
$s_{yy}(h)$	Die korrigierte Varianz der h-ten Schicht: $\frac{1}{N-1}\sum_i (y_i(h) - \overline{y}(h))^2$.
$\mathrm{span}[\underline{z}^1, \ldots, \underline{z}^M]$	Der durch die Vektoren $\underline{z}^1, \ldots, \underline{z}^M$ aufgespannte Vektorraum.
Θ	Der Parameterraum.
Θ_Q	Der Quader als Parameterraum: $\{\underline{y} \in \mathbb{R}^N; 0 \leq \underline{y}_i \leq \underline{x}_i,\ i = 1, \ldots, N\}$.
Θ_U	Das Ellipsoid als Parameterraum: $\{\underline{y} \in \mathbb{R}^N; \underline{y}^t U \underline{y}{=}1\}$.
Θ_{U1}	Der Parameterraum $\{\underline{y} \in \Theta_U; y = 0\}$.

Θ_{UX}	Der Parameterraum $\{\underline{y} \in \Theta_U; X^t y = \underline{0}\}$.
tr	Der Spuroperator: $\mathrm{tr}(M) = \sum_i m_{ii}$.
U	Eine positiv semidefinite Matrix mit Kern(U)=Bild(X).
x	Die Summe der x-Werte: $\sum_{i=1}^N x_i$.
$\underline{x}$	Ein Regressionsvektor: $N \times 1$-Matrix mit $x_i > 0$ und $\sum_{i=1}^N x_i = 1$.
x_K	Die Summe der x_K-Werte: $\sum_{i \in K} x_i$.
$\underline{x}_K$	Eine Ecke des Quaders: $\underline{x}_K = \sum_{i \in K} x_i \underline{e}_i$.
X	Eine Regressionsmatrix: $N \times K$-Matrix mit dim(Bild(X))=K.
y	Die Summe der y-Werte: $\sum_{i=1}^N y_i$.
$\overline{y}$	Der Mittelwert von $\underline{y}$: $\frac{y}{N}$.
$\underline{y}$	Der Parameter $\underline{y} \in \mathbb{R}^N$.
$\overline{y}(h)$	Der Mittelwert der h-ten Schicht: $\frac{1}{N(h)} \sum_i y(h)_i$.
$\mathcal{Y}$	Die konvexe Hülle von $\{\underline{y}\underline{y}^t; \underline{y} \in \Theta\}$.
$\mathcal{Y}_U$	Die konvexe Hülle von $\{\underline{y}\underline{y}^t; \underline{y} \in \Theta_U\}$: $\{V \in \mathcal{Q}_0; \langle V, U \rangle = 1\}$.
$\mathcal{Y}_{U1}$	Die konvexe Hülle von $\{\underline{y}\underline{y}^t; \underline{y} \in \Theta_{U1}\}$: $\{V \in \mathcal{Q}_0; V\underline{1} = \underline{0}, \langle V, U \rangle = 1\}$.
$\mathcal{Y}_{UX}$	Die konvexe Hülle von $\{\underline{y}\underline{y}^t; \underline{y} \in \Theta_{UX}\}$: $\{V \in \mathcal{Q}_0; VX = 0, \langle V, U \rangle = 1\}$.

Literaturverzeichnis

[1] Berger, J.O.: Statistical Decision Theory: Foundations, Concepts, and Methods. Springer (1980).

[2] Bickel, P.J. und E.L. Lehmann: A minimax property of the sample mean in finite populations. Annals of Statistics 9, 1119-1122 (1981).

[3] Blackwell, D. und M.A. Girshick: Theory of Games and Statitical Decisions. Wiley, New York (1954).

[4] Chaudhuri, A.: Some sampling schemes to use Horvitz-Thompson estimator in estimating a finite population total. Calcutta Statistical Association Bulletin 20, 37-66 (1971).

[5] Cheng, C.S. und K.C. Li: A minimax approach to sample surveys. Annals of Statistics 11, 552-563 (1983).

[6] Gabler, S.: Minimax Solutions in Sampling from Finite Populations. Lecture Notes in Statistics 64, Springer: New York (1990).

[7] Gabler, S. und H. Stenger: Improving the RHC-strategie. Statistical Papers 36, 327-336 (1995).

[8] Gabler, S. und H. Stenger: Minimax strategies in survey sampling. Discussion Paper Nr. 572-99, Mannheim (1999).

[9] Gabler, S. und H. Stenger: Minimax strategies in survey sampling. Journal of Statistical Planning and Inference 90, 305-321 (2000).

[10] Gabler, S.: Warum die größten Elemente auswählen. Unveröffentlichtes Paper (2002).

[11] Godambe, V.P.: A unified theory of sampling from finite populations. Journal of the Royal Statistical Society, Ser. B, 17, 269-278 (1955).

[12] Hajek, J.: Optimum strategy and other problems in probability sampling. Cas. Pest. Mat. 84, 387-473 (1959).

[13] Hansen, M.H. und W.N. Hurwitz: On the theory of sampling from a finite population. Annals of Mathematical Statistics, 14, 333-362 (1943).

[14] Horwitz, D.G. und D.J. Thompson: A generalization of sampling without replacement from a finite universe. Journal of the American Statistical Association, 47, 663-685 (1952).

[15] Luenberger, D.G.: Optimization by Vektor Space Methods. Wiley: New York (1969).

[16] Marshall, A.W. und I. Olkin: Inequalities: Theory of Majorization and Its Applications. Mathematics in Science and Engineering, Volume 143. Academic Press, Inc.

[17] Osborne, M.J. und A. Rubinstein: A Course in Game Theory. MIT Press, Cambridge, Massachusetts (1994).

[18] Rao, J.N.K., Hartley, H.O. und W.G. Cochran: On a simple procedure of unequal probability sampling without replacement. Journal of the Royal Statistical Society, Ser. B, 24, 482-491 (1962).

[19] Scott, A.J. und T.M.F. Smith: Minimax designs for sample surveys. Biometrica, 62, 353-361 (1975).

[20] Stenger, H.: A minimax approach to randomization and estimation in survey sampling. Annals of Mathematical Statistics, 7, 395-399 (1979).

[21] Stenger, H.: Minimax strategies and asymptotical minimax strategies in survey sampling. Discussion Paper Nr. 219-82, Mannheim (1982).

[22] Stenger, H.: Stichproben. Physica Verlag Heidelberg Wien (1986).

[23] Stenger, H.: Asymptotic expansion of the minimax value in survey sampling. Metrika 35, 77-92 (1988).

[24] Stenger, H.: Asymptotic minimaxity of the ratio strategy. Biometrika 77, 389-395 (1990).

[25] Stenger, H. und S. Gabler: A minimax property of Lahiri-Midzuno-Sen's sampling scheme. Metrika 43, 213-220 (1996).

[26] Stenger, H.: Begründung der Zufallsauswahl im Rahmen des Prognoseansatzes der Stichprobentheorie. Wirtschafts- und Sozialstatistik heute, v.d. Lippe et al.. Wissenschaft & Praxis, Dr. Brauner GmbH: Ludwigsburg/Berlin, 52-62 (1997).

[27] Stenger, H., Gabler, S. und J. Schmidt: Survey sampling: A linear game. Erscheint in: Statistics & Decisions (2002).

Lebenslauf

Persönliche Daten

Name:	Jochen Schmidt
Geboren:	am 25.02.1967 in Mannheim
Familienstand:	verheiratet mit Christine Schmidt, geb. Maierhofer
Kinder:	Philipp Schmidt, geb. am 24.10.1999

Werdegang

1973-1977:	Besuch der Feudenheimer Grundschule
1977-1986:	Besuch des Feudenheim-Gymnasiums mit Abschluss Abitur
1986-1988:	Zivildienst
1988-1995:	Studium der Mathematik (Nebenfach Informatik) an der Universität Mannheim mit Abschluss Diplom-Mathematiker
Seit 1996:	Wissenschaftlicher Mitarbeiter am Lehrstuhl Statistik I, Fakultät für VWL und Statistik, Universität Mannheim