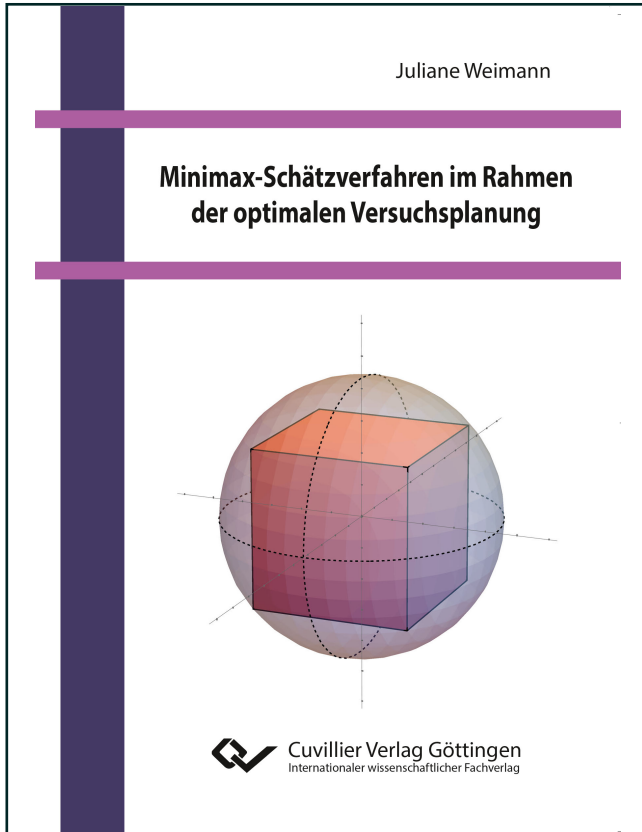




Juliane Weimann (Autor)

# Minimax-Schätzverfahren im Rahmen der optimalen Versuchsplanung



<https://cuvillier.de/de/shop/publications/6794>

Copyright:

Cuvillier Verlag, Inhaberin Annette Jentsch-Cuvillier, Nonnenstieg 8, 37075 Göttingen,  
Germany

Telefon: +49 (0)551 54724-0, E-Mail: [info@cuvillier.de](mailto:info@cuvillier.de), Website: <https://cuvillier.de>



# 1 Einleitung

Ein wichtiges Modell zur Darstellung ökonomischer Zusammenhänge ist das lineare Regressionsmodell

$$y = X\beta + u$$

mit  $y$  als  $(n \times 1)$ -Vektor der Beobachtungen der endogenen Größe,  $X$  als beobachtbare  $(n \times k)$ -Regressormatrix,  $\beta$  als unbekanntem  $(k \times 1)$ -Parametervektor und  $u$  als  $(n \times 1)$ -Vektor der unbeobachtbaren Störgrößen. Einen zentralen Punkt der Schätztheorie für dieses Modell bildet das Gauß-Markov-Theorem. Es basiert im einfachsten Fall auf den folgenden Modellannahmen:

Die Erwartungswerte  $E(u)$  und  $E(uu')$  existieren, und es gilt:

- (i)  $E(u) = 0_n$ .
- (ii)  $E(uu') = \sigma^2 I_n$ .
- (iii) Die Regressormatrix  $X$  ist nichtstochastisch und hat vollen Spaltenrang.
- (iv) Es gibt keine Vorinformationen über  $\beta$  und  $\sigma^2$ .

Unter diesen sogenannten Standardannahmen (i) bis (iv) ist der Kleinste-Quadrate-Schätzer

$$\hat{\beta} = (X'X)^{-1}X'y$$

der beste lineare unverzerrte Schätzer für  $\beta$ . Er besitzt also in der Klasse aller linearen unverzerrten Schätzer die „kleinste“ Varianz-Kovarianzmatrix. Auch unter Berücksichtigung korrekter linearer Vorinformationen  $R\beta = r$  ist der restringierte KQ-Schätzer

$$b_r = \hat{\beta} - (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}(R\hat{\beta} - r)$$

in der Klasse aller linearen unverzerrten Schätzer effizient. Zur Einbeziehung von Vorinformationen unterschiedlichster Gestalt lassen sich Minimax-Schätzer verwenden. Vorinformationen in Form von Ellipsoiden werden beispielsweise von Bunke (1975), Hoffmann (1979) und Läuter (1975) thematisiert. Drygas (1991),



Stahlecker (1987) und Stahlecker und Trenkler (1991) untersuchen auch linear-affine Vorinformationen beziehungsweise die Kombination von Ellipsoid- und linearen Restriktionen. Minimax-Schätzer unter Berücksichtigung von unscharfen Vorinformationen erläutern Arnold und Stahlecker (1997a). Klintworth (2004) betrachtet partielle Parameterkenntnisse. Wie Vorinformationen bei Testverfahren berücksichtigt werden können, beschreibt Becker (1998).

Zur Beurteilung der Güte eines Schätzers dienen in den genannten Arbeiten die Varianz beziehungsweise der mittlere quadratische Fehler. Arnold und Stahlecker (2000) messen die Qualität eines linear-affinen Schätzers mittels des relativen quadratischen Fehlers (RSE). Dieser Ansatz wurde in einigen anderen Arbeiten erweitert. Arnold und Stahlecker (2010b,c) liefern zum Beispiel beste Minimax-Schätzer, einerseits für eine Klasse unscharfer Vorinformationen und andererseits für den optimalen Schätzer bezüglich des relativen quadratischen Fehlers.

Die Güte einer Schätzung wird jedoch nicht nur von den Eigenschaften der Schätzfunktion bestimmt, sondern auch von der Wahl der Messpunkte. Verfahrensweisen zur Auswahl eines Experiments aus einer Menge von möglichen Versuchen nach einem gegebenen Optimalitätskriterium bietet die Theorie der optimalen Versuchsplanung. Diese beruht auf Methoden und Resultaten der Algebra, Entscheidungstheorie und der nichtlinearen Optimierung.

Aufgrund von finanziellen und zeitlichen Restriktionen ist in der Praxis oft nur eine begrenzte Anzahl von Messungen realisierbar. Die optimale Versuchsplanung erlaubt es, den Anteil der Daten mit wesentlichem Informationsgehalt zu erhöhen und steigert somit die Effizienz von Experimenten.

Unter dem Begriff „Versuchsplanung“ werden in der Literatur unterschiedliche Ansätze verstanden (vgl. Melas 2006):

- Als einer der ersten Ansätze wird die Arbeit von Fisher (1935) erwähnt. Bei landwirtschaftlichen Experimenten erforschte er die Einflüsse unterschiedlicher Faktoren auf den Ertrag mithilfe von kombinatorischen Verfahren. In diesem Kontext wurde die Varianzanalyse entwickelt.
- Die Methode von Box und Wilson (1951) basiert auf fraktionellen faktoriellen Versuchsplänen. Faktorielle Versuchspläne bezeichnen Designs, in denen mehrere Faktoren mit verschiedenen Faktorausprägungen untersucht werden. Bei fraktionellen faktoriellen Designs wird dagegen nur ein Teil der möglichen Kombinationen von Faktorstufen überprüft (siehe auch Montgomery 2009).

- Im dritten Ansatz, der speziell von Kiefer (1959) geprägt wurde, wird ein Versuchsplan als diskretes Wahrscheinlichkeitsmaß betrachtet. Die Gewichte bestimmen, wie viele Versuche unter den jeweiligen experimentellen Bedingungen getätigt werden sollten. Bei einer festen Anzahl von Experimenten soll die Genauigkeit von statistischen Aussagen verbessert werden.

Ziel dieser Arbeit ist es, einige Methoden beziehungsweise Ergebnisse der optimalen Versuchsplanung auf Minimax-Schätzer anzuwenden. Hierzu werden Algorithmen zur Ermittlung exakter sowie diskreter Versuchspläne betrachtet, welche diese Schätzer bezüglich des relativen quadratischen Fehlers weiter optimieren.

Die Dissertation gliedert sich in sieben Kapitel. Nach der Einleitung beleuchtet das zweite Kapitel zwei Typen von Vorinformationsmengen, Ellipsoide und Fuzzy-Mengen, sowie deren Eigenschaften näher. Danach liegt das Augenmerk auf dem Minimax-Schätzverfahren und dem Ansatz des relativen quadratischen Fehlers.

Der Theorie der optimalen Versuchsplanung widmet sich das dritte Kapitel. Dabei werden die Informationsmatrix sowie deren Eigenschaften näher beschrieben und verschiedene Optimalitätskriterien thematisiert. Zwischen D- und G-optimalen Designs schaffen Kiefer und Wolfowitz (1960) im Äquivalenztheorem eine Verbindung. Zudem sollen explizite Lösungen für einige Spezialfälle entwickelt werden. Mit der Erläuterung von Algorithmen zur Berechnung von optimalen Versuchsplänen und deren Implementierung in Computerprogramme schließt der Abschnitt.

Im vierten Kapitel werden Vorinformationen über den unbekannt Parameter in Form von Fuzzy-Mengen unterstellt und mittels des Extensionsprinzips von Zadeh Fuzzymengen für die Schätzfehler hergeleitet. Hieraus resultiert das Konzept der gleichmäßig besten  $\Gamma$ -kompatiblen linear-affinen Schätzfunktionen. Nach einigen Zwischenergebnissen zeigt sich, dass der beste lineare Schätzer bezüglich des RSE-Ansatzes nicht nur den gleichmäßig besten Schätzer in der Klasse aller linear-affinen Schätzfunktionen widerspiegelt, sondern sich dieses Resultat auf eine weit größere Klasse verallgemeinern lässt.

Diesen Ansatz erweitert das fünfte Kapitel, indem weder Stochastik noch Vorinformationen über die Parameter vorausgesetzt werden.

Gegenstand des sechsten Kapitels ist die Implementierung des relativen quadratischen Fehlers als Gütekriterium bei der optimalen Versuchsplanung. Dafür sollen zunächst die neuen Optimalitätskriterien definiert und zwei Spezialfälle betrachtet werden. Danach gilt es, den im Statistikprogramm R implementierten Algorithmus zur Berechnung von exakten RSE-optimalen Versuchsplänen zu plausibilisie-



ren und eine Möglichkeit herauszustellen, mit MATLAB diskrete RSE-optimale Designs zu berechnen. Es folgen Beispiele, in welchen die exakten und diskreten RSE-optimalen Versuchspläne bestimmt und mit den klassischen Kriterien aus dem dritten Kapitel verglichen werden.

Eine Zusammenfassung mit einem Ausblick auf mögliche weitere Forschungsschwerpunkte rundet im siebten Kapitel den Hauptteil der Dissertation ab. Im Anhang werden einige mathematische Hintergründe, die in der Arbeit gebraucht werden, aufgezeigt und die Programmcodes der implementierten Algorithmen dokumentiert.



## 2 Grundlagen

### 2.1 Vorinformationen

Im klassischen linearen Regressionsmodell gehen wir davon aus, dass außer den Beobachtungen  $(y, X)$  keine Informationen über den Parametervektor  $\beta$  und die Varianz der Störgrößen  $\sigma^2$  existieren. Diese Bedingung wollen wir nun durch die Annahme, dass Vorinformation verfügbar ist, ersetzen.

Die a-priori Kenntnisse über  $\beta$  können beispielsweise in Gestalt von linearen Parameterrestriktionen, Ellipsoiden und Fuzzy-Mengen vorliegen. Stahlecker (1987) zeigt verschiedene Typen von Vorinformation über den Parametervektor  $\beta$  auf. Eine Auswahl soll hier vorgestellt werden.

#### 2.1.1 Ellipsoide

Vorinformationen können beispielsweise durch Ellipsoide beschrieben werden.

**Definition 2.1:** Eine Menge  $E_{T,c} \subseteq \mathbb{R}^r$  ist ein Ellipsoid, wenn ein Vektor  $c \in \mathbb{R}^r$  und eine positiv definite<sup>1</sup> Matrix  $T \in \mathbb{R}^{r \times r}$  existieren, sodass

$$E_{T,c} = \{x \in \mathbb{R}^r \mid (x - c)' T (x - c) \leq 1\}$$

*gilt. Der Vektor  $c$  entspricht dem Mittelpunkt des Ellipsoids.*

Zwischen den geometrischen Eigenschaften des Ellipsoids  $E_{T,c}$  und den algebraischen Eigenschaften der Matrix  $T^{-1}$  mit den Eigenwerten  $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_r$  gibt es folgende Verbindungen, welche unter anderem in Grötschel, Lovász und Schrijver (1993, S. 64 ff.) dargelegt werden:

- Durchmesser von  $E_{T,c} \hat{=} \text{Länge der längsten Achse} \hat{=} 2\sqrt{\lambda_r}$
- längste Achse von  $E_{T,c}$  verläuft in Richtung des Eigenvektors zu  $\lambda_r$

---

<sup>1</sup>Anhang A.5 liefert eine Definition von positiv semidefiniten und positiv definiten Matrizen.  
Dieses Werk ist copyrightgeschützt und darf in keiner Form vervielfältigt werden noch an Dritte weitergegeben werden.  
Es gilt nur für den persönlichen Gebrauch.



- Weite von  $E_{T,c} \hat{=}$  Länge der kürzesten Achse  $\hat{=}$   $2\sqrt{\lambda_1}$
- kürzeste Achse von  $E_{T,c}$  verläuft in Richtung des Eigenvektors zu  $\lambda_1$

Das Volumen eines Ellipsoids ist gegeben durch

$$\text{vol}(E_{T,c}) = \mathfrak{B} \frac{1}{\sqrt{\det T}},$$

dabei bezeichnet  $\mathfrak{B}$  das Volumen der Einheitskugel  $\{x \in \mathbb{R}^r \mid \|x\| \leq 1\}$  im  $\mathbb{R}^r$ .

Ziel soll es sein, für eine gegebene nichtleere, konvexe und kompakte Menge  $K \subseteq \mathbb{R}^r$  mit  $\dim(K) = r$  ein Ellipsoid  $E_{T,c}$  so zu bestimmen, dass  $K \subseteq E_{T,c}$  und das Volumen von  $E_{T,c}$  minimal ist. Somit lautet die zu minimierende Zielfunktion:

$$(\det T)^{-\frac{1}{2}},$$

vgl. Grötschel, Lovász und Schrijver (1993, S. 69) Da die Logarithmusfunktion streng monoton wächst, ergibt sich das äquivalente Optimierungsproblem<sup>2</sup>

$$\begin{aligned} \min_{T,c} \quad & -\log(\det T) & (2.1) \\ \text{s.t.} \quad & (x-c)'T(x-c) \leq 1 \quad \forall x \in K \\ & T \text{ p. d.} \end{aligned}$$

Nun betrachten wir einige Spezialfälle dieses Problems:

- Sei die Menge  $K$  in Form von Intervallen

$$K = \{x \in \mathbb{R}^r \mid a_i \leq x_i \leq b_i, a_i \neq b_i, i = 1, \dots, r\}$$

bestimmt. Dann zeigen Toutenburg und Rao (1997, S. 73 f.), dass das volumenminimale Ellipsoid gegeben ist durch

$$c^* = \frac{1}{2} \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_r + b_r \end{pmatrix} \quad \text{und} \quad T^* = \frac{4}{r} \begin{pmatrix} \frac{1}{(b_1 - a_1)^2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{(b_r - a_r)^2} \end{pmatrix}.$$

<sup>2</sup>Sei  $x^*$  eine Optimallösung eines Optimierungsproblems. Dann ist  $x^*$  auch eine optimale Lösung des äquivalenten Optimierungsproblems. Die optimalen Zielfunktionswerte müssen dagegen nicht übereinstimmen.

**Beispiel 2.1:** Gegeben sei die Menge

$$K_1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid 1 \leq x_1 \leq 2, 1 \leq x_2 \leq 3 \right\}.$$

Das volumenminimale Ellipsoid lässt sich angeben durch

$$c^* = \begin{pmatrix} \frac{3}{2} \\ 2 \end{pmatrix} \text{ und } T^* = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

- Stahlecker und Schmidt (1989) betrachten die Menge

$$K = \left\{ x \in \mathbb{R}^r \mid x = \sum_{j=1}^m \mu_j \pi_j, \mu_j \geq 0, \sum_{j=1}^m \mu_j = 1 \right\},$$

dabei beschreiben  $\pi_1, \dots, \pi_m$  die Eckpunkte eines konvexen und kompakten Polyeders mit  $\dim(K) = r$  und  $\mu_j$  Gewichte. Als approximative Lösungen des volumenminimalen Ellipsoids  $E_{T^*, c^*}$  ergeben sich:

$$c^* = \frac{1}{m} \sum_{j=1}^m \pi_j$$

$$T^* = \frac{1}{\max_{j=1, \dots, m} (b'_j (BB')^{-1} b_j)} (BB')^{-1}$$

mit  $b_j = \pi_j - c^*$  und  $B = (b_1 \dots b_m) \in \mathbb{R}^{r \times m}$ ,  $\text{Rang}(B) = r$ .

Dabei bewirkt der Faktor  $\frac{1}{\max_{j=1, \dots, m} (b'_j (BB')^{-1} b_j)}$ , dass mindestens ein Eckpunkt des Polyeders auf dem Rand des Ellipsoids liegt. Für

$$\frac{m}{r} b'_j (BB')^{-1} b_j = 1 \quad \forall j = 1, \dots, m$$

ist diese Näherung identisch zur optimalen Lösung.

**Beispiel 2.2:** Gegeben seien die Eckpunkte

$$\pi_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \pi_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ und } \pi_3 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

eines Polyeders  $K_2$ . Dann folgen

$$c^* = \begin{pmatrix} \frac{4}{3} \\ \frac{5}{2} \end{pmatrix},$$

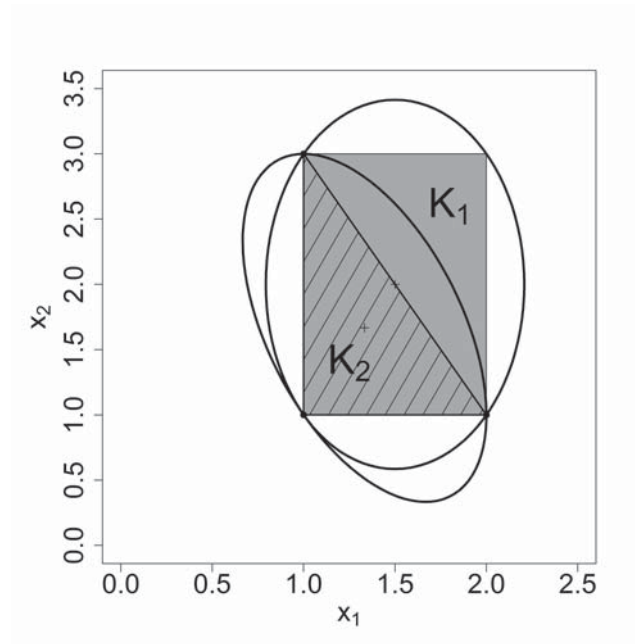


$$(BB')^{-1} = \begin{pmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

und

$$T^* = \begin{pmatrix} 3 & \frac{3}{4} \\ \frac{3}{4} & \frac{3}{4} \end{pmatrix}.$$

Die Mengen  $K_1$  und  $K_2$  sowie deren volumenminimale Ellipsen sind in der Abbildung 2.1 veranschaulicht.



**Abbildung 2.1:** Volumenminimale Ellipsen der Mengen  $K_1$  und  $K_2$

- Sei  $K$  eine endliche Menge von  $d$  Punkten  $p_i \in \mathbb{R}^r$ ,  $K = \{p_1, \dots, p_d\} \subseteq \mathbb{R}^r$ ,  $V = (p_1 \dots p_d)$ . Die affine Hülle von  $p_1, \dots, p_d$ ,

$$\text{aff}(K) = \left\{ z \mid z = \sum_{i=1}^d \vartheta_i p_i, \sum_{i=1}^d \vartheta_i = 1, p_i \in K \right\},$$

sei  $\mathbb{R}^r$ . Dann ist das Optimierungsproblem (2.1) äquivalent zu

$$\begin{aligned} \min_{\check{T}} \quad & -\log \det \check{T} \\ \text{s.t.} \quad & \left( (p_i)' \ 1 \right) \check{T} \begin{pmatrix} p_i \\ 1 \end{pmatrix} \leq 1, \quad i = 1, \dots, d \\ & \check{T} \in \mathbb{R}^{(r+1) \times (r+1)}, \check{T} \text{ ist positiv definit.} \end{aligned}$$

Für das duale Problem folgt

$$\begin{aligned} \max_u \quad & \log \left[ \det \sum_{i=1}^d u_i \begin{pmatrix} p_i \\ 1 \end{pmatrix} (p_i' \ 1) \right] \\ \text{s.t.} \quad & \sum_{i=1}^d u_i = 1 \\ & u \geq 0_d. \end{aligned}$$

Sei  $u^*$  eine Optimallösung des dualen Problems. Dann ergibt sich das volumenminimale Ellipsoid  $E_{T^*,c^*}$  nach Kumar und Yildirim (2005) aus

$$T^* = \frac{1}{n} (V \operatorname{diag}(u^*) V' - V u^* (V u^*)')^{-1}, \quad c^* = V u^*.$$

Sind die Punkte  $p_1, \dots, p_d$  affin unabhängig, nutzen Kumar und Yildirim (2005) den Algorithmus von Khachiyan (1996) zur Berechnung eines optimalen Vektors  $u^*$ . Dieser ist im MATLAB-Programm<sup>3</sup> von Moshtagh (2005) implementiert.

Khachiyan (1979) hat gezeigt, dass ein Algorithmus zur Lösung eines linearen Optimierungsproblems mit polynomialer Laufzeit existiert. Er liefert als Ergebnis aber auch das volumenminimale Ellipsoid  $E_{T,c}$  mit

$$\begin{aligned} K &= \{p_1, \dots, p_d\} \subseteq E_{T,c} \text{ und} \\ \operatorname{vol}(E_{T,c}) &\leq (1 + \epsilon) \operatorname{vol}(\operatorname{VME}(K)), \quad \epsilon > 0 \end{aligned}$$

dabei ist  $\operatorname{VME}(K)$  das volumenminimale Ellipsoid, welches die Punkte der Menge  $K$  einschließt.

Der Algorithmus von Khachiyan (1996) testet, ob ein gegebenes Polyeder

$$\mathcal{V} = \operatorname{conv}(V) = \left\{ \sum_i \vartheta_i p_i \mid \vartheta_i \geq 0, \sum_i \vartheta_i = 1 \right\}$$

leer ist. Dafür wird zuerst ein Ellipsoid  $E_1$  erzeugt, welches das Polyeder  $\mathcal{V}$  komplett enthält. Liegt der Mittelpunkt des Ellipsoids  $E_1$  im Polyeder, dann ist dieser nichtleer. Sonst wird ein neues Ellipsoid  $E_2$  erzeugt, welches  $\mathcal{V}$  enthält, aber ein kleineres Volumen besitzt als  $E_1$ . Nach polynomial vielen Iterationen haben wir entweder einen Punkt gefunden, den das Polyeder enthält, oder wir gehen davon aus, dass  $\mathcal{V}$  leer ist.

<sup>3</sup>Für eine Beschreibung des Programms siehe <http://www.mathworks.com/matlabcentral/fileexchange/9542>.



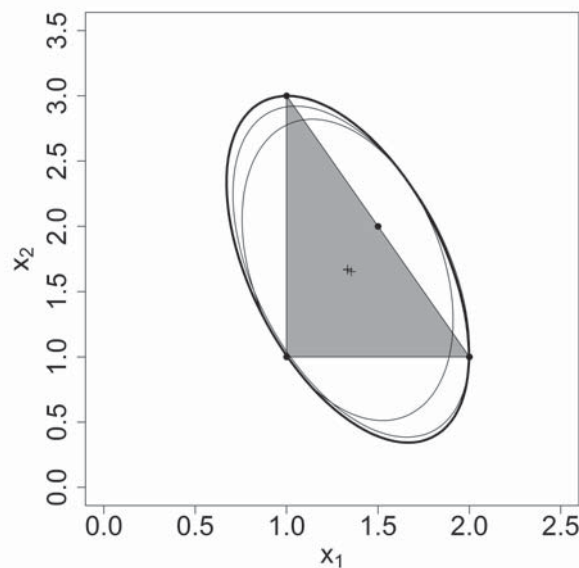
**Beispiel 2.3:** Betrachten wir das Polyeder mit den Eckpunkten aus dem Beispiel 2.2, dann erhalten wir nach einem Durchlauf des Khachiyan-Algorithmus folgende Ergebnisse:

$$u^* = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}, c^* = \begin{pmatrix} 4 \\ 3 \\ 3 \end{pmatrix} \text{ und } T^* = \begin{pmatrix} 3 & \frac{3}{4} \\ \frac{3}{4} & 3 \end{pmatrix}.$$

Fügen wir nun den Punkt  $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$  zu der Menge  $K_2$  hinzu, dann benötigt der Algorithmus von Khachiyan für  $\epsilon = 0,01$  insgesamt 192 Durchläufe. Als approximative Optimallösungen ergeben sich:

$$u^* = \begin{pmatrix} 0,330 \\ 0,329 \\ 0,331 \\ 0,010 \end{pmatrix}, c^* = \begin{pmatrix} 1,334 \\ 1,672 \end{pmatrix} \text{ und } T^* = \begin{pmatrix} 3,022 & 0,752 \\ 0,752 & 0,754 \end{pmatrix}.$$

In der Abbildung 2.2 sind die volumenminimalen Ellipsen für den ersten, 17. und 192. Durchlauf dargestellt.



**Abbildung 2.2:** Volumenminimale Ellipsen für die Durchläufe 1, 17 und 192

Die Schritte des Khachiyan Algorithmus sind in der Abbildung 2.3 dargestellt.

Wir halten uns dabei an die Darstellung von Todd und Yildirim (2007).

Dieses Werk ist copyrightgeschützt und darf in keiner Form vervielfältigt werden noch an Dritte weitergegeben werden.

Es gilt nur für den persönlichen Gebrauch.