

TA ed.

FORSCHUNG - AUSBILDUNG - WEITERBILDUNG

BERICHT Nr. 24

EINFACHE VERFAHREN ZUR PLANUNG UND
AUSWERTUNG VON NAVIGATIONSVERSUCHSFAHRTEN

MICHAEL SCHEUTZOW

UNIVERSITÄT KAISERSLAUTERN
FACHBEREICH MATHEMATIK
ERWIN-SCHRÖDINGER-STRASSE
D - 6750 KAISERSLAUTERN

JULI 1985

EINFACHE VERFAHREN ZUR PLANUNG UND
AUSWERTUNG VON NAVIGATIONSVERSUCHSFAHRTEN

1. Problemstellung, allgemeines Modell
2. Kostenfunktionen, allgemeine Vorgehensweise
3. Einfache Fehlerabschätzungen mit (verallgemeinerter) Tschebychevungleichung
4. Anwendungen des Zentralen Grenzwertsatzes zur Fehlerabschätzung
5. Verallgemeinerungen
6. Zusammenfassung
7. Nachteile bzw. Einschränkungen des vorgeschlagenen Verfahrens und Ausblick auf genauere Modelle, insbesondere im Hinblick auf Parameteroptimierung bei unendlich vielen Parameterwerten

1. Problemstellung, allgemeines Modell

Bei einem von VDO entwickelten Navigationsgerät stehen unterschiedliche Verfahren zur Durchführung der Navigation zur Verfügung. Das Problem besteht darin, daß die von dem jeweiligen Verfahren erzielten Navigationsendpunkte von dem gewünschten Ziel mehr oder weniger stark abweichen. Diese Abweichungen hängen nicht nur von Anfangs- und Endpunkt der Testfahrt und vom Navigationsverfahren ab, sondern ändern sich auch bei der Wiederholung solcher Testfahrten.

Ordnet man nun jedem Navigationsverfahren einen Parameter (vektor) zu und bezeichnet man die Menge aller betrachteten Parameter (werte) mit M (M kann zunächst endlich oder unendlich sein), so besteht das Ziel darin, den Parameter $p^* \in M$ aufzufinden, für den das zugehörige Verfahren im Mittel am besten abschneidet. Hierzu ist es notwendig, zunächst ein Gütekriterium festzulegen.

Wir gehen von folgendem allgemeinen Modell aus:

Sei S die Menge aller Fahrtstrecken (definiert durch Anfangspunkt, Endpunkt und Streckenverlauf), auf denen potentiell das Navigationsgerät später eingesetzt wird und P ein Wahrscheinlichkeitsmaß auf S , welches für jede Teilmenge \tilde{S} von Strecken (z.B. kann \tilde{S} alle Strecken mit Länge höchstens 20 km umfassen) den relativen Anteil aller Fahrten angibt, die auf Strecken aus \tilde{S} entfallen (es ist nicht notwendig S und P tatsächlich exakt anzugeben).

Wird das Navigationsgerät auf einer Strecke $s \in S$ eingesetzt, so wird bei Fahrtende eine gewisse (zum Teil zufällige) Abweichung $(A_s(p), B_s(p))$ auftreten, wobei $A_s(p)$ und $B_s(p)$ die Abweichungen in Ost- bzw. Nordrichtung angeben, wenn das Verfahren mit Parameter (vektor) p angewendet wird.

2. Kostenfunktion, allgemeine Vorgehensweise

Zur Beurteilung der Güte der verschiedenen Verfahren ist es sinnvoll, jeder Fahrt einen "Kosten"betrag $K(A_s(p), B_s(p), s)$ zuzuordnen, der angibt, wie gravierend die Abweichung $(A_s(p), B_s(p))$ ist. $K(A_s(p), B_s(p), s)$ kann etwa als Maß für die Unzufriedenheit des Kunden über das Ergebnis der Navigation angesehen werden. Man wird nun ein Verfahren p dem Verfahren q vorziehen, wenn die mittleren (erwarteten) Kosten beim Verfahren p geringer sind als

bei q , d.h.

$$E K(A_S(p), B_S(p), s) < E K(A_S(q), B_S(q), s)$$

wobei der Erwartungswert E bezüglich des Wahrscheinlichkeitsmaßes P gebildet wird. Es wird also die Funktion K zunächst über jede einzelne Strecke s gemittelt und dann über alle Strecken $s \in S$ gemäß der Häufigkeit ihres Befahrens. Wir werden weiter unten auf geeignete Wahlen von Kostenfunktionen näher eingehen.

Das Problem besteht nun darin, den Parameter $p^* \in M$ zu finden, für den $a(p) := E K(A_S(p), B_S(p), s)$ minimal ist. Leider kennt man aber die Funktion $a(p)$ nicht. Man wird daher so vorgehen (Alternativen werden in Abschnitt 5 vorgestellt), daß man sich zufällig N Strecken auswählt und zwar gemäß dem Wahrscheinlichkeitsmaß P (d.h. typische Strecken sollten auch bei Testfahrten entsprechend häufig vertreten sein) und unabhängig voneinander (man sollte z.B. nicht sehr ähnliche Strecken oft durchfahren). Wir werden im folgenden immer annehmen, daß diese beiden Voraussetzungen exakt erfüllt sind, auch wenn dies in der Praxis kaum verifizierbar ist. Wir möchten aber an dieser Stelle darauf hinweisen, daß alle mathematischen Aussagen und Fehlerabschätzungen nur so exakt sind wie die Erfüllung der Voraussetzungen. Werden zum Beispiel Testfahrten nur im Rhein-Main-Gebiet, nur mit einem Wagentyp, nur bei Tag o.ä. durchgeführt, so haben auch alle Aussagen nur unter diesen Bedingungen Gültigkeit.

Es liegt nun nahe, das Verfahren \tilde{p}_N auszuwählen, das sich bei den N Testfahrten als das beste herausgestellt hat, d.h. für das

$$a_N(p) = \frac{1}{N} \sum_{i=1}^N X_i(p)$$

minimal ist. Hier bezeichnet $X_i(p) := K(A_{S_i}(p), B_{S_i}(p), s_i)$ die Kosten bei der i -ten Fahrt und Parameter p . Ist M eine endliche (nicht zu große) Parametermenge, so kann man $X_i(p)$ für jedes $i = 1, \dots, n$ und jedes $p \in M$ berechnen und damit auch \tilde{p}_N bestimmen. Im unendlichen Fall ist dies dagegen nicht möglich, und man wird sich mit Näherungslösungen für \tilde{p}_N zufriedengeben müssen. Da wir hier nur einfache Modelle behandeln, werden wir im weiteren annehmen, daß M endlich und nicht zu groß ist und somit \tilde{p}_N aus den Daten der Versuchsfahrten ausgerechnet werden kann.

Wir wollen nun noch einige Hinweise zur Wahl einer Kostenfunktion K geben. Mögliche Wahlen sind

$$K_1(x,y,s) = \sqrt{x^2+y^2} \quad , \quad \text{d.h. absolute Abweichung,}$$

$$K_2(x,y,s) = \frac{\sqrt{x^2+y^2}}{\|s\|} \quad , \quad \text{d.h. relative Abweichung (\|s\| \text{ sei die Streckenlänge)}$$

oder auch Kombinationen von beiden, z.B.

$$K_3(x,y,s) = \lambda K_1(x,y,s) + (1-\lambda)K_2(x,y,s) \quad \text{mit } 0 \leq \lambda \leq 1.$$

Man kann K auch nicht nur von $A_s(p), B_s(p)$ und s , sondern auch von Größen vor dem Ende der Fahrt abhängen lassen, z.B. von der absoluten oder relativen Abweichung der Anzeige während der gesamten Fahrt. Mathematisch macht dies keinerlei Unterschied, nur die numerische Berechnung wird aufwendiger.

Statt absoluter oder relativer Abweichung ist es möglicherweise sinnvoller, eine geeignete nichtlineare Funktion \bar{K} derselben als Kostenfunktion zu wählen, d.h.

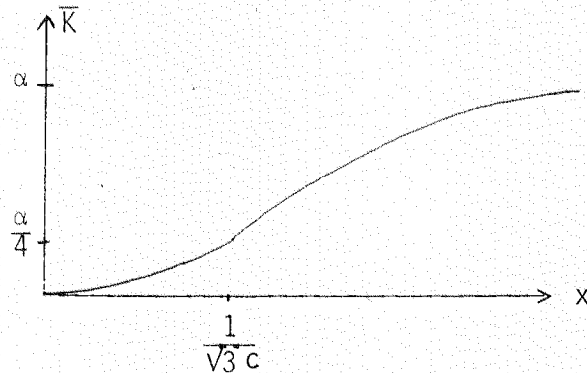
$$\bar{K}_1(x,y,s) = \bar{K}(\sqrt{x^2+y^2})$$

oder analog für K_2 oder K_3 . Man wird nämlich bei sehr großen Abweichungen annehmen können, daß es kaum eine Rolle spielt, wie groß die Abweichung genau ist, z.B. wird man für $\bar{K}(50 \text{ km})$ weniger als das Zehnfache von $\bar{K}(5 \text{ km})$ wählen, d.h. \bar{K} sollte für große Werte relativ flach verlaufen. Sicherlich ist es sinnvoll $\bar{K}(0) = 0$ zu wählen und anzunehmen, daß \bar{K} monoton wachsend ist. Will man erreichen, daß bei gleichem Mittelwert der absoluten oder relativen Abweichung ein solches Verfahren als besser angesehen wird, das geringere Fluktuationen aufweist, so wird man \bar{K} im Bereich kleiner (realistischer) Werte konvex wählen, d.h.

$$\frac{d^2 \bar{K}}{dx^2}(x) > 0.$$

Beispiel 1:

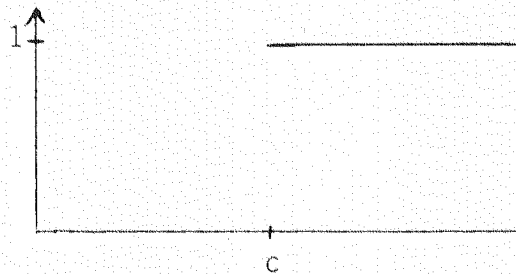
$$\bar{K}(x) = \alpha \cdot \frac{(cx)^2}{1+(cx)^2}$$



Die Parameter c und α können noch frei gewählt werden. Die Wahl von α spielt keine Rolle, da α nur multiplikativ eingeht. \bar{K} ist konvex im Intervall $0 \leq x \leq \frac{1}{\sqrt{3}c}$. Daher sollte c so gewählt werden, daß etwa 75-99% aller Abweichungen (bei guten Parametern p) im Intervall $[0, \frac{1}{\sqrt{3}c}]$ liegen.

Beispiel 2:

$$\bar{K}(x) = \begin{cases} 0 & x < c \\ 1 & x \geq c \end{cases}$$



Mit dieser Kostenfunktion wird das Verfahren am besten abschneiden, für das die Wahrscheinlichkeit einer Abweichung $\geq c$ am kleinsten ist. Es wird aber nicht unterschieden, wie groß die Abweichungen $\geq c$ oder die Abweichungen $< c$ sind.

Sowohl aus den genannten (praktischen) Gründen als auch aus mathematischen Gründen, die im folgenden Abschnitt erläutert werden, empfehlen wir, die Funktion K bzw. \bar{K} als beschränkt anzunehmen.

Sollte man bei der Wahl von K oder \bar{K} unsicher sein, so kann man auch verschiedene Wahlen treffen und die Ergebnisse vergleichen.

3. Einfache Fehlerabschätzungen mit (verallgemeinerter) Tschebychevungleichung

Führt man N Testfahrten durch und entscheidet sich dann für den Parameter \tilde{p}_N , für den $a_N(p)$ minimal ist (vgl. letzter Abschnitt), so stellt sich die Frage, wie groß die Wahrscheinlichkeit ist, daß man sich für ein falsches (d.h. nicht für das beste) Verfahren entschieden hat. Man wird dann N so groß wählen, daß diese Wahrscheinlichkeit unterhalb einer vorgegebenen Schranke (z.B. 5 %) liegt. Nun wird allerdings in dem Fall, daß mehrere Parameter fast genausogut sind wie der beste, die Zahl N extrem hoch sein müssen, um mit großer Sicherheit den besten aus den Daten zu erkennen. Zwar kann man zeigen, daß im Fall einer endlichen Parametermenge M , das (zufällige) \tilde{p}_N für $N \rightarrow \infty$ gegen den besten Wert p^* konvergiert und damit die Irrtumswahrscheinlichkeit gegen Null geht, aber die Konvergenzgeschwindigkeit hängt entscheidend davon ab, wie nahe andere Werte $a(p)$ bei $a(p^*)$ liegen. Da man dies aber *a-priori* nicht weiß, lassen sich hier prinzipiell keine zuverlässigen Irrtumswahrscheinlichkeiten angeben.

Man muß daher die Fragestellung ändern, um Irrtumswahrscheinlichkeiten angeben zu können. Die "richtige" Frage lautet: Gegeben sei ein $\varepsilon > 0$; wie groß ist die Wahrscheinlichkeit $p_{N,\varepsilon}$, daß bei N Testfahrten $a(\tilde{p}_N) > a(p^*) + \varepsilon$ ist, d.h. daß der beste Parameter beim Test \tilde{p}_N im Durchschnitt mehr als ε Kosten mehr verursacht als das tatsächlich beste Verfahren p^* ? Dieses ε sollte vor den Versuchen festgelegt werden. Es beschreibt die Größe der Trennschärfe, die man erreichen will. Selbstverständlich wird die notwendige Zahl N von Versuchen bei vorgegebenem $p_{N,\varepsilon}$ (z.B. 5 %) umso größer sein, je kleiner die Trennschärfe ε gewählt wird.

Wir werden nun einfache Abschätzungen für $p_{N,\varepsilon}$ herleiten. In Abschnitt 4 werden wir dann mit Hilfe von detaillierten Modellen bessere Abschätzungen für $p_{N,\varepsilon}$ erhalten.

Wir bezeichnen mit

$$\bar{X}_N(p) = \frac{1}{N} \sum_{i=1}^N X_i(p)$$

die mittleren Kosten bei N Versuchsfahrten und Parameter p und mit M_ε die (zunächst unbekannte) Menge aller $q \in M$, für die $a(q) > a(p^*) + \varepsilon$ ist. Dann ist

$$\begin{aligned}
 p_{N,\epsilon} &= P\{\bar{X}_N(p^*) > \bar{X}_N(q) \quad \text{für ein } q \in M_\epsilon\} \\
 &\leq \sum_{q \in M_\epsilon} P\{\bar{X}_N(p^*) - \bar{X}_N(q) - (a(p^*) - a(q)) > -(a(p^*) - a(q))\} \\
 &\leq \sum_{q \in M_\epsilon} P\{|\bar{X}_N(p^*) - \bar{X}_N(q) - (a(p^*) - a(q))| > a(q) - a(p^*)\} \\
 &\leq \sum_{q \in M_\epsilon} \frac{E(\bar{X}_N(p^*) - \bar{X}_N(q) - (a(p^*) - a(q)))^2}{(a(q) - a(p^*))^2}
 \end{aligned}$$

Die letzte Ungleichung ist die sogenannte Tschebychevungleichung, die besagt, daß

$$P\{|Y| > \beta\} \leq \frac{EY^2}{\beta^2}$$

ist und für beliebige Zufallsvariablen Y gilt. Sie ist allerdings in den meisten Fällen sehr grob.

Nun ist

$$\begin{aligned}
 &E(\bar{X}_N(p^*) - a(p^*) - (\bar{X}_N(q) - a(q)))^2 \\
 &= \frac{1}{N^2} E\left(\sum_{i=1}^N (X_i(p^*) - a(p^*)) - (X_i(q) - a(q))\right)^2 \\
 &= \frac{1}{N} E(X_1(p^*) - a(p^*) - (X_1(q) - a(q)))^2 \\
 &= \frac{1}{N} \text{var}(X_1(p^*) - X_1(q)) ,
 \end{aligned}$$

wobei die vorletzte Gleichheit gilt, weil $X_i(p) - X_i(q)$ und $X_j(p) - X_j(q)$ für $i \neq j$ nach Voraussetzung unabhängig sind und dieselbe Wahrscheinlichkeitsverteilung haben. Somit gilt

$$(1) \quad p_{N,\epsilon} \leq \frac{1}{N\epsilon^2} \sum_{q \in M_\epsilon} \text{var}(X_1(p^*) - X_1(q)) .$$

Leider kennt man nun aber weder p^* noch $\text{var}(X_1(p^*) - X_1(q))$. Das erste Problem kann man dadurch lösen, daß man rechts statt p^* p schreibt und (vorsichtshalber) das Maximum über alle p bildet und über alle $q \neq p^*$ summiert, d.h. mit dem schlimmsten Fall rechnet:

$$(2) \quad p_{N,\epsilon} \leq \frac{1}{N\epsilon^2} \max_{p \in M} \sum_{\substack{q \in M \\ q \neq p}} \text{var}(X_1(p) - X_1(q))$$

$\text{var}(X_1(p) - X_1(q))$ kann man nun entweder aus den Daten schätzen (wie, wird im folgenden Abschnitt dargelegt) oder grob nach oben abschätzen für den Fall,

daß die Kostenfunktion durch eine Größe α nach oben beschränkt ist.

Da $0 \leq X_1(p) \leq \alpha$ für alle $p \in M$ gilt, folgt

$$\text{var}(X_1(p) - X_1(q)) = E(X_1(p) - X_1(q))^2 - (a(p) - a(q))^2 \leq \alpha^2 - (a(p) - a(q))^2$$

und daher

$$(3) \quad P_{N,\varepsilon} \leq \frac{1}{N\varepsilon} 2^{(n-1)(\alpha^2 - \varepsilon^2)},$$

wobei $n = |M|$ die Anzahl der betrachteten Parameter ist. Diese Abschätzung ist sehr grob, gilt aber dafür immer wenn die beiden Generalvoraussetzungen Unabhängigkeit und identische Verteilung für alle Versuchsfahrten erfüllt ist und wenn die Kostenfunktion beschränkt ist.

Beispiel 3:

Wenn z.B. $\alpha = 1$ ist, $\varepsilon = \frac{1}{100}$ und $n = 20$, so ist

$$P_{N,\varepsilon} \leq \frac{189981}{N}.$$

Man braucht also $20 \cdot 189981 \approx 3,8$ Millionen Versuche, um eine Fehlerwahrscheinlichkeit von höchstens 5 % zu garantieren.

Dieses (wohl nicht völlig unrealistische) Beispiel zeigt, daß die obige Abschätzung absolut unbefriedigend ist. Glücklicherweise gibt es aber im Fall beschränkter Kostenfunktionen wesentlich bessere Schranken als die in (3) angegebene. Die sogenannte verallgemeinerte Tschebychevungleichung (siehe z.B. A. Renyi: Wahrscheinlichkeitsrechnung, Berlin, 1973) liefert in diesem Fall

$$(4) \quad P_{N,\varepsilon} \leq \sum_{q \in M_\varepsilon} 2 \exp\left(-\frac{N(m(q))^2}{2D_q^2\left(1 + \frac{(m(q)(m(q)+\alpha))^2}{2D_q^2}\right)}\right),$$

wobei $m(q) = a(q) - a(p^*)$

und $D_q^2 = \text{var}(X_1(q) - X_1(p^*))$ ist.

Obige Ungleichung gilt unter der Voraussetzung, daß $X_1(p)$ für alle p durch α nach oben und Null nach unten beschränkt ist und daß $m(q) \cdot (m(q) + \alpha) \leq D_q^2$ für alle $q \in M_\varepsilon$ gilt. Da $m(q)$, D_q und M_ε nicht a priori bekannt sind,

empfehlte sich folgende Abschätzung:

$$(5) \quad P_{N,\epsilon} \leq 2(n-1) \exp\left[-N \frac{2\left(\frac{\alpha}{\epsilon} - 1\right)}{\left(\frac{\alpha}{\epsilon} + 1\right)\left(2\frac{\alpha}{\epsilon} - 1\right)^2}\right],$$

die man aus (4) herleiten kann und die auch ohne die Voraussetzung $m(q)(m(q)+\alpha) \leq D_q^2$ gilt (allerdings muß $\frac{\alpha}{\epsilon} \geq 5$ sein, was aber in der Praxis ohnehin sinnvoll ist). Vergleicht man (5) mit (3), so sieht man, daß die rechte Seite von (5) mit $N \rightarrow \infty$ (unter denselben Voraussetzungen wie (3)) wesentlich schneller gegen Null geht.

Beispiel 3a:

Mit denselben Daten wie vorher liefert (5)

$$(6) \quad P_{N,\epsilon} \leq 38 \cdot \exp\left(-\frac{N}{20200,5}\right).$$

Die erforderliche Anzahl N von Versuchen, um $P_{N,\epsilon} \leq 0,05$ zu garantieren, beträgt daher $N = 20200,5 \cdot \log \frac{38}{0,05} \approx 134000$, was natürlich immer noch unbefriedigend ist und hauptsächlich dadurch zu erklären ist, daß die Abschätzung der Varianzen D_q nach oben durch $\alpha^2 - \epsilon^2$ in der Regel sehr grob ist (in Extremfällen allerdings angenommen werden kann). Man beachte, daß (4) und (5) im Gegensatz zu (1) und (2) nur im Fall einer beschränkten Kostenfunktion gültig sind.

4. Anwendungen des Zentralen Grenzwertsatzes zur Fehlerabschätzung

Aus der Wahrscheinlichkeitstheorie ist bekannt, daß für eine Folge Y_1, Y_2, \dots von unabhängigen identisch verteilten Zufallsgrößen mit endlicher Varianz $\sigma^2 = E(Y_1 - EY_1)^2$ der Zentrale Grenzwertsatz gilt, d.h.

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N (Y_i - EY_1)$$

konvergiert für $N \rightarrow \infty$ in Verteilung gegen eine Normalverteilung mit Mittelwert Null und Varianz σ^2 . Dieser Satz gilt analog auch für Zufallsvariable mit Werten im \mathbb{R}^n und ist daher in unserem Fall anwendbar. Es gilt

$$\left(\sqrt{N} (\bar{X}_N(p_1) - a(p_1)), \dots, \sqrt{N} (\bar{X}_N(p_n) - a(p_n))\right)$$

ist asymptotisch (für $N \rightarrow \infty$) normalverteilt mit Mittelwert Null und

Kovarianzmatrix

$$\gamma(p_i, p_j) = E(X_1(p_i) - a(p_i))(X_1(p_j) - a(p_j)).$$

Wir werden im weiteren annehmen, daß N so groß ist, daß die Näherung durch den Zentralen Grenzwertsatz (ZGS) gerechtfertigt ist. Leider gibt es keine allgemeingültige Regel, wie groß N dazu gewählt werden muß. Dies hängt von der (unbekannten) Verteilung von $(X_1(p_1), \dots, X_1(p_n))$ ab. In der Praxis sollte man auf jeden Fall N groß gegenüber n wählen. In der Regel ist die Näherung für beschränkte oder nur langsam wachsende Kostenfunktionen besser als für linear oder schneller wachsende. Wie im letzten Abschnitt wollen wir die Irrtumswahrscheinlichkeit $P_{N,\epsilon}$ berechnen bzw. nach oben abschätzen. Wir gehen davon aus, daß die Normalapproximation nicht nur eine Approximation, sondern exakt ist, d.h.

$$(Z_1, \dots, Z_n) = (\sqrt{N}(\bar{X}_N(p_1) - a(p_1)), \dots, \sqrt{N}(\bar{X}_N(p_n) - a(p_n)))$$

ein Gaußvektor mit Erwartungswert Null und Kovarianzmatrix γ (wie oben) ist. p^* sei (wie vorher) der optimale Parameter, sein Index sei ℓ ($1 \leq \ell \leq n$) und $M_\epsilon = \{q \in M: a(q) > a(p^*) + \epsilon\}$. Dann gilt

$$\begin{aligned} P_{N,\epsilon} &= P\{\bar{X}_N(p^*) > \bar{X}_N(q) \text{ für ein } q \in M_\epsilon\} \\ &= P\{\frac{1}{\sqrt{N}} Z_\ell + a(p^*) > \frac{1}{\sqrt{N}} Z_j + a(p_j) \text{ für ein } p_j \in M_\epsilon\} \\ &= P\{Z_\ell - Z_j > \sqrt{N} (a(p_j) - a(p^*)) \text{ für ein } p_j \in M_\epsilon\} \end{aligned}$$

Da man weder p^* , noch M_ϵ , noch $a(p)$ kennt, empfiehlt sich die Abschätzung

$$(7) \quad P_{N,\epsilon} \leq \max_{1 \leq k \leq n} P\{Z_k - Z_j > \sqrt{N} \epsilon \text{ für ein } j \neq k\}.$$

Dieser ungünstigste Fall tritt dann ein, wenn alle anderen Verfahren gerade um ϵ schlechter sind als p^* und p^* die Eigenschaft hat, daß das obige Maximum angenommen wird.

Die rechte Seite ist nun gegeben durch

$$(8) \quad \max_{1 \leq k \leq n} (1 - \int_{-\infty}^{\epsilon\sqrt{N}} \dots \int_{-\infty}^{\epsilon\sqrt{N}} f_k(t_1, \dots, t_{n-1}) dt_1 \dots dt_{n-1}),$$

wobei

$$f_k(t_1, \dots, t_{n-1}) = \frac{1}{(2\pi)^{\frac{n-1}{2}} \sqrt{\det C_k}} e^{-\frac{1}{2}(t_1, \dots, t_{n-1}) C_k^{-1} \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix}}$$

die Dichte des Gaußvektors $(Z_k - Z_1, \dots, Z_k - Z_{k-1}, Z_k - Z_{k+1}, \dots, Z_k - Z_n)$ mit Kovarianzmatrix $(C_k)_{ij} = E(Z_k - Z_i)(Z_k - Z_j)$ ist.

Das Integral ist nicht analytisch lösbar. Man kann es aber numerisch (wenn entsprechende Programme vorhanden sind, würden wir diese Methode empfehlen) oder durch Simulation lösen, oder das Integral durch leichter berechenbare Größen abschätzen. Wir werden nur die letzten beiden Möglichkeiten weiter diskutieren. In jedem Fall stellt sich aber die Frage, wie man die zunächst unbekannte Kovarianzfunktion $\gamma(p, q)$ oder die Kovarianzmatrizen C_k erhält. Da sie unbekannt ist, kann man entweder für den Fall beschränkter Kostenfunktionen wie im vorigen Abschnitt alle Varianzen nach oben abschätzen und damit eine obere Schranke für $P_{N, \epsilon}$ erhalten, oder man schätzt die Kovarianzmatrix aus den Daten, und zwar durch

$$\tilde{\gamma}(p, q) = \frac{1}{N_0} \sum_{i=1}^{N_0} (X_i(p) - \tilde{a}_{N_0}(p))(X_i(q) - \tilde{a}_{N_0}(q)),$$

wobei

$$\tilde{a}_{N_0}(p) = \frac{1}{N_0} \sum_{i=1}^{N_0} X_i(p)$$

und N_0 die Zahl der in einem Vortest durchgeführten Versuchsfahrten ist. N_0 sollte dabei nicht zu klein gewählt werden ($N_0 \gg n$), um γ einigermaßen genau schätzen zu können. Da die Genauigkeit der Schätzung von a wichtiger ist als die von γ , werden wir $\tilde{a}_{N_0}(p)$ nicht als Schätzung für $a(p)$ verwenden.

Wir wollen nun $P_{N, \epsilon}$ weiter abschätzen. Sei $A_{kj} = \{Z_k - Z_j > \sqrt{n} \epsilon\}$. Dann gilt

$$\begin{aligned} (9) \quad P_{N, \epsilon} &\leq \max_{1 \leq k \leq n} P\left(\bigcup_{\substack{j=1 \\ j \neq k}}^n A_{kj}\right) \\ &= \max_{1 \leq k \leq n} \left(\sum_{\substack{j=1 \\ j \neq k}}^n P(A_{kj}) - \sum_{\substack{j=2 \\ j \neq k}}^n \sum_{\substack{\ell=1 \\ \ell \neq k}}^{j-1} P(A_{kj} \cap A_{k\ell}) + \sum_{j=3}^n \sum_{\substack{\ell=2 \\ \ell \neq k}}^{j-1} \sum_{\substack{m=1 \\ j, \ell, m \neq k}}^{\ell-1} P(A_{kj} \cap A_{k\ell} \cap A_{km}) \right. \\ &\quad \left. - \dots (-1)^n P(A_{k1} \cap \dots \cap A_{kn}) \right) \\ &\leq \max_{1 \leq k \leq n} \sum_{\substack{j=1 \\ j \neq k}}^n P(A_{kj}) \\ &= \max_{1 \leq k \leq n} \sum_{\substack{j=1 \\ j \neq k}} P\left\{ \frac{Z_k - Z_j}{\sigma_{kj}} > \frac{\epsilon \sqrt{n}}{\sigma_{kj}} \right\} \quad \text{mit } \sigma_{kj} = \sqrt{E(Z_k - Z_j)^2} = \sqrt{\gamma_{kk} + \gamma_{jj} - 2\gamma_{jj}} \end{aligned}$$

$$\begin{aligned}
 &= \max_{1 \leq k \leq n} \sum_{\substack{j=1 \\ j \neq k}}^n \int_{\frac{\epsilon \sqrt{N}}{\sigma_{kj}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
 (10) \quad &\leq \frac{2}{\sqrt{2\pi}} \max_{1 \leq k \leq n} \sum_{j=1}^n \frac{e^{-\frac{\epsilon^2 N}{2\sigma_{kj}^2}}}{\sqrt{\frac{\epsilon^2 N}{2\sigma_{kj}^2} + 2} + \frac{\epsilon \sqrt{N}}{\sigma_{kj}}}
 \end{aligned}$$

Die letzte Abschätzung folgt aus der bekannten Ungleichung

$$2 \frac{1}{\sqrt{a^2+4} + a} \leq e^{\frac{a^2}{2}} \int_a^{\infty} e^{-\frac{b^2}{2}} db \leq 2 \frac{1}{\sqrt{a^2+2} + a}, \quad a \geq 0.$$

Der Ausdruck (10) läßt sich problemlos auf einem (Taschen)rechner auswerten. Etwas gröbere, aber schneller berechenbare Abschätzungen erhält man, wenn man die Größen σ_{kj} nach oben durch ihr Maximum \bar{M} oder, noch größer, durch $2\sqrt{\gamma}$ abschätzt, wobei $\gamma = \max_{1 \leq j \leq n} \gamma_{jj}$ ist.

Dann folgt

$$(11) \quad P_{N,\epsilon} \leq \frac{2}{\sqrt{2\pi}} (n-1) e^{-\frac{\epsilon^2 N}{2\bar{M}^2}} \left(\sqrt{\frac{\epsilon^2 N}{\bar{M}^2} + 2} + \frac{\epsilon \sqrt{N}}{\bar{M}} \right)^{-1}.$$

Verwendet man diese Fehlerabschätzung, so hat dies den Vorteil, daß man lediglich das Maximum der Varianzen σ_{kj} , nicht aber alle Kovarianzen γ_{ij} schätzen muß. Ist $\frac{\epsilon \sqrt{N}}{\bar{M}}$ nicht zu klein, dann kann man ohne großen Verlust an Genauigkeit die rechte Seite von (11) nach oben abschätzen durch

$$(12) \quad P_{N,\epsilon} \leq \frac{1}{\sqrt{2\pi}} (n-1) e^{-\frac{\epsilon^2 N}{2\bar{M}^2}} \cdot \frac{\bar{M}}{\epsilon \sqrt{N}}.$$

Wenn die Kostenfunktion durch α beschränkt ist, erhält man $\sqrt{E(X_k - X_j)^2} \leq \alpha$ und damit $\sigma_{kj}^2 \leq \bar{M}^2 \leq \alpha^2 - \epsilon^2$ und aus (12)

$$(13) \quad P_{N,\epsilon} \leq \frac{1}{\sqrt{2\pi}} (n-1) e^{-\frac{\epsilon^2 N}{2(\alpha^2 - \epsilon^2)}} \frac{\sqrt{\alpha^2 - \epsilon^2}}{\epsilon \sqrt{N}}.$$

Vergleicht man diesen Ausdruck mit (5) in Abschnitt 3, so sieht man, daß hier die Konvergenz von $P_{N,\epsilon}$ gegen Null für $N \rightarrow \infty$ schneller ist.

Beispiel 4:

Mit denselben Daten wie in Beispiel 3, d.h. $\alpha = 1$, $\epsilon = \frac{1}{100}$ und $n = 20$ folgt

$$P_{N,\epsilon} \leq \frac{1}{\sqrt{2\pi}} \cdot 19 \cdot e^{-\frac{N}{20000}} \cdot \frac{100}{\sqrt{N}}.$$

In diesem Fall wären etwa $N = 80000$ Testfahrten (statt 134000) notwendig, um 0,05 auf der rechten Seite zu erhalten.

Dieses ebenfalls unbefriedigende Ergebnis legt es nahe, statt (13) die σ_{jk} aus den Daten zu schätzen.

Beispiel 4a:

Ist z.B. $\max_{j \neq k} \sigma_{jk} = \frac{1}{10}$, so erhält man aus (12)

$$P_{N,\epsilon} \leq \frac{1}{\sqrt{2\pi}} \cdot 19 \cdot e^{-\frac{N}{200}} \cdot \frac{10}{\sqrt{N}}.$$

Es genügen also $\frac{1}{100}$ der oberen Anzahl, d.h. $N = 800$ Fahrten. Gibt man sich mit $\epsilon = \frac{1}{20}$ zufrieden, so genügen bereits $N = 32$ Fahrten. Man kann davon ausgehen, daß selbst bei Verwendung der besseren Abschätzung (10) und sogar der numerischen Auswertung des $(n-1)$ -fachen Integrals die wirklichen $P_{N,\epsilon}$ deutlich unter den angegebenen Schranken liegen, da ja immer der ungünstigste Fall angenommen wurde.

Hat man nun N Versuche (einschließlich der ursprünglichen N_0 Fahrten zur Bestimmung der Kovarianzfunktion) durchgeführt, so bietet es sich an, die Schätzung von γ anhand der größeren Anzahl N von Versuchen noch einmal zu überprüfen und $P_{N,\epsilon}$ damit neu zu berechnen. Ergeben sich deutliche Veränderungen, d.h. war die ursprüngliche Schätzung zu schlecht, so sollte man dies berücksichtigen und gegebenenfalls weitere Testfahrten durchführen. Andernfalls entscheide man sich für den Parameter \tilde{p} , der bei den N Testfahrten am besten abschnitt. Will man nun $a(\tilde{p})$, d.h. die Qualität des Verfahrens bestimmen, so ist diese näherungsweise durch die Schätzung

$\frac{1}{N} \sum_{i=1}^N X_i(\tilde{p})$ gegeben. Man sollte aber beachten, daß man bei nicht sehr großem N und nicht sehr kleinem n durch diese Schätzung $a(\tilde{p})$, ja sogar $a(p^*)$ in der Regel unterschätzt. Ist man an genauen Abschätzungen von $a(\tilde{p})$ und evtl. auch der Varianz $E(X_1(\tilde{p}) - a(\tilde{p}))^2$ interessiert, so sollte man danach noch einmal m Fahrten nur mit Parameter \tilde{p} durchführen. Dann ist

$\tilde{a}(\tilde{p}) = \frac{1}{m} \sum_{i=N+1}^{N+m} X_i(\tilde{p})$ eine unverfälschte Schätzung für $a(\tilde{p})$ und

$\frac{1}{m-1} \sum_{i=N+1}^{N+m} (X_i(\tilde{p}) - \tilde{a}(\tilde{p}))^2$ eine Schätzung für die Varianz σ^2 von $X_1(\tilde{p})$.

Man sollte sich dabei $\epsilon > 0$ und $\delta > 0$ vorgeben und m dann so groß wählen, daß

$$P\{|\tilde{a}(\tilde{p}) - a(\tilde{p})| < \epsilon\} \geq 1 - \delta$$

ist. Mit der Normalapproximation heißt dies gerade, daß m so zu wählen ist, daß

$$\frac{\frac{\epsilon\sqrt{m}}{\sigma}}{\frac{\epsilon\sqrt{1}}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \geq 1 - \delta$$

ist. Aus Tabellen kann man dann m berechnen.

Wir möchten nun noch kurz erläutern, wie man das Integral (8), oder was dasselbe ist, die rechte Seite von (9) durch Monte-Carlo-Simulation näherungsweise bestimmen kann. Sei $Q^{(k)}$ die geschätzte Kovarianzmatrix ($i, j \neq k$)

$$Q_{ij}^{(k)} = \frac{1}{N_0} \sum_{\ell=1}^{N_0} (X_\ell(p_k) - X_\ell(p_j) - (a_N(p_k) - a_N(p_j)))(X_\ell(p_k) - X_\ell(p_i) - (a_N(p_k) - a_N(p_i)))$$

(wir werden im folgenden den oberen Index (k) weglassen).

Dann läßt sich die (positiv-semidefinite) Matrix Q in der Form $Q = AA^T$ zerlegen, wobei A eine $(n-1) \times (n-1)$ untere Dreiecksmatrix ist (d.h. $a_{ij} = 0$ für $i < j$).

A läßt sich schnell aus Q berechnen, indem man rekursiv $\frac{n(n-1)}{2}$ Gleichungen für die a_{ij} ($i \geq j$) löst. Mit dieser Matrix A gilt:

Sind V_1, \dots, V_{n-1} unabhängig normalverteilt mit Varianz 1, dann ist der Vektor

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_{n-1} \end{pmatrix} = A \begin{pmatrix} V_1 \\ \vdots \\ V_{n-1} \end{pmatrix}$$

normalverteilt mit Kovarianz Q , denn $EYY^T = E(AV)(AV)^T = AEVV^T A^T = AA^T = Q$,

wobei $V = \begin{pmatrix} V_1 \\ \vdots \\ V_{n-1} \end{pmatrix}$ und $Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_{n-1} \end{pmatrix}$.

Somit ist die Aufgabe, $N(0,Q)$ -verteilte Vektoren auf die Aufgabe $N(0,E)$ -verteilte Vektoren zu simulieren ($E =$ Einheitsmatrix) zurückgeführt. Dies ist z.B. aus unabhängigen und auf $[0,1]$ gleichverteilten Zufallsvariablen U_1 und U_2 (die meisten Rechenanlagen enthalten Unterprogramme zur Erzeugung solcher Zufallszahlen) möglich durch die Transformation

$$\begin{aligned} V_1 &= (-2 \ln U_1)^{1/2} \cos 2\pi U_2 \\ V_2 &= (-2 \ln U_1)^{1/2} \sin 2\pi U_2, \end{aligned}$$

die zwei unabhängige $N(0,1)$ -verteilte Zufallsgrößen liefert.

Man gehe also folgendermaßen vor:

- Berechne die Matrix A wie oben beschrieben,
- simulierte U_1 und U_2 mit dem Rechner,
- berechne V_1 und V_2 nach obiger Transformation,
- simulierte 2 weitere gleichverteilte Zufallsgrößen,
- berechne daraus V_3 und V_4 ,

usw..

Hat man V_1, \dots, V_{n-1} erzeugt, dann berechne man $\begin{pmatrix} Y_1 \\ \vdots \\ Y_{n-1} \end{pmatrix} = A \begin{pmatrix} V_1 \\ \vdots \\ V_n \end{pmatrix}$

- Ist $Y_i \leq \epsilon \sqrt{N}$ für alle $1 \leq i \leq n-1$, notiere $P_1 = 0$, sonst $P_1 = 1$,
- wiederhole die Simulation M -mal und bilde $\bar{P}_k = \frac{1}{M} \sum_{i=1}^M P_i$,
- führe die gesamte Berechnung für alle anderen k durch,
- $\max_k \bar{P}_k$ liefert dann eine Schätzung für $P_{N,\epsilon}$.

Ist z.B. $P_{N,\epsilon} = 0,05$ angestrebt, so sollte man $M \approx 1000$ wählen, für kleineres $P_{N,\epsilon}$ entsprechend höher.

Zum Abschluß des Kapitels möchten wir noch einmal auf das Problem der Schätzung der Kovarianzmatrix eingehen. Da die Schätzung von den Daten abhängig ist, ist sie zufällig und damit auch die nach dem vorgeschlagenen Verfahren bestimmte Anzahl N von Versuchsfahrten.

Um ein Gefühl dafür zu bekommen, wie diese (zufällige) Zahl N variiert, wenn man verschiedene Blöcke von je N_0 Strecken gefahren wäre, kann man verschiedene in den letzten zehn Jahren entwickelte Verfahren verwenden, unter denen das "Bootstrapping" wohl das bekannteste ist (ein guter einführender Artikel ist "Computer-intensive Methods in Statistics" von P. Diaconis und B. Efron im Scientific American, Mai 1983, S. 96-108). Man ziehe dazu zufällig mit Zurücklegen aus den N_0 Zufallsvektoren $(X_i(p_1), \dots, X_i(p_n))$ N_0 Stück (wobei meist einige mehrfach auftreten) und tue so, als ob diese N_0 Strecken gefahren worden wären. Man schätze dafür die Kovarianzfunktion und berechne (z.B.) numerisch das Integral (8) und zwar sei N so gewählt, daß die rechte Seite von (8) z.B. 0,05 ist. Man wiederhole dies mehrmals und vergleiche die Werte für N . Variieren sie sehr stark, dann sollte man N lieber etwas größer als den Durchschnitt dieser Werte wählen und nach dem Versuch mit N Fahrten noch einmal "bootstrappen".

5. Verallgemeinerungen

Wir möchten in diesem Abschnitt kurz darauf eingehen, wie man das vorgeschlagene Verfahren modifizieren sollte, wenn man auch Daten während der Fahrtstrecke in die Berechnungen einbezieht, denn statt des tatsächlichen Endpunktes hätte der Kunde ja auch einen Endpunkt auf der Strecke als Ziel haben können. In diesem Fall sollte man Streckenlängen $s_1 < s_2 < s_3 < \dots < s_\ell$ festlegen derart, daß s_1 und s_ℓ ungefähr die minimale bzw. maximale Streckenlänge angeben, auf denen Kunden das Navigationsgerät benutzen werden und die Dichte der Punkte umso größer ist, je häufiger der entsprechende Bereich von Streckenlängen gefahren wird (gemäß dem Wahrscheinlichkeitsmaß P).

Denkbar wäre z.B. (in km) $s_1 = 2, s_2 = 5, s_3 = 7, s_4 = 8, s_5 = 9, s_6 = 11, s_7 = 14, s_8 = 18$. Nun berechnet man für jede Fahrt statt $X_i(p) = K(A_S(p), B_S(p), s)$

die Größe $X_i(p) = \frac{1}{\ell} \sum_{i=1}^{\ell} K(s_i)$, wobei $K(s_i)$ die Kosten sind für den Fall, daß das Ziel im Abstand von s_i km vom Startpunkt auf der Fahrtstrecke liegt.

Mit dieser Größe $X_j(p)$ kann man nun alle Rechnungen wie vorher durchführen. Der Vorteil liegt darin, daß zu erwarten ist, daß alle Varianzen und Kovarianzen sich durch diese Mittelung verringern und damit eine kleinere Anzahl von Versuchsfahrten genügt, um geringe Irrtumswahrscheinlichkeiten zu erzielen.

Um Probleme bei der Berechnung von $X_j(p)$ zu vermeiden, empfehlen wir bei diesem Ansatz, immer die maximale Streckenlänge s_0 zu fahren. Ist dies nicht möglich (z.B. weil ältere Versuchsfahrten die Bedingung nicht erfüllen), so kann man (wenn die Fahrten nicht zu kurz sind - in diesem Fall sollten sie unberücksichtigt bleiben), X_j dadurch berechnen, daß man nur über die s_j mittelt, die kleiner als die Teststreckenlänge sind. Man sollte sich aber klarmachen, daß dadurch kürzere Strecken stärker gewichtet werden. Dringend abraten möchten wir davor, eine einzige Testfahrt als λ getrennte Fahrten, nämlich vom Start bis s_1 , vom Start bis s_2 usw. anzusehen und dann wie in Abschnitt 4 zu verfahren. Man spart auf diese Weise zwar viele Testfahrten ($\frac{N}{\lambda}$ statt N), beachte aber, daß die Voraussetzung der Unabhängigkeit der verschiedenen Fahrten u.U. grob verletzt ist.

6. Zusammenfassung

Wir fassen noch einmal zusammen:

1. Man überlege sich eine geeignete Kostenfunktion, die ausdrückt, wie gravierend eine bestimmte Abweichung vom Zielpunkt ist.
2. Man wähle n (n endlich und nicht zu groß) "Parameter" (d.h. verschiedene Navigationsverfahren) aus.
3. Man wähle für Voruntersuchungen N_0 Teststrecken aus, und zwar "unabhängig" voneinander und "typische" Strecken (z.B. sollte die relative Häufigkeit der Teststrecken zwischen 10 und 20 km mit der relativen Häufigkeit der Strecken zwischen 10 und 20 km etwa übereinstimmen, auf denen später das Gerät benutzt werden wird; will man das Verfahren in Abschnitt 5 verwenden, dann sollten die Strecken ebenfalls unabhängig voneinander, aber alle mit ungefähr derselben Streckenlänge s_0 gewählt werden). Die Zahl N_0 wähle man nicht zu klein. Für $n=10$ könnte man etwa $N_0=50$ nehmen (N_0 sollte wesentlich größer sein als n).

4. Man schätze die Kovarianzfunktion wie in Abschnitt 4 beschrieben.
5. Durch Abschätzungen, Simulation oder numerisch berechne man für verschiedene N obere Abschätzungen für die Irrtumswahrscheinlichkeit $P_{N,\epsilon}$, wobei ϵ eine vorgegebene Irrtumstoleranz ist (siehe Abschnitt 5).
6. Man wähle ein N , für das $P_{N,\epsilon}$ genügend klein ist.
7. Man führe N Testfahrten durch (dabei kann man die N_0 Vortests mit verwenden).
8. Man vergleiche die ursprünglichen Schätzungen der Kovarianzfunktion mit der neuen, auf den N Fahrten basierenden und führe ein "Bootstrapping" durch. Gegebenenfalls wiederhole man Punkt 5 mit der neuen Schätzung.
9. Man wähle den Parameter \tilde{p} , der bei den N Versuchen am besten abschnitt, d.h. für den $\frac{1}{N} \sum_{i=1}^N X_i(p)$ minimal ist.
10. Um $EX_1(\tilde{p})$ zu bestimmen, führe man noch weitere Versuche durch (siehe Abschnitt 4).

7. Nachteile und Einschränkungen des vorgeschlagenen Verfahrens und Ausblick auf genauere Modelle, insbesondere im Hinblick auf Parameteroptimierung bei unendlichen Parametermengen

Das in Abschnitt 5 erläuterte Verfahren erlaubt es, aus n zur Auswahl stehenden Navigationsverfahren aufgrund von Daten aus Testfahrten einen Parameter auszuwählen, der bis auf ein vorgegebenes Toleranzintervall der Länge ϵ mit vorgegebener Höchstirrtumswahrscheinlichkeit der beste ist. Neben den Grundannahmen, daß die Testfahrtstrecken unabhängig und gemäß derselben Wahrscheinlichkeitsverteilung P ausgewählt wurden (diese Annahmen wird man auch bei komplizierteren Verfahren machen müssen), nahmen wir im vierten Abschnitt auch an, daß die Normalapproximation gerechtfertigt ist und daß die Kovarianzfunktionen mit hinreichender Genauigkeit geschätzt werden können. Beides ist zwar allgemein für hinreichend großes N (oder N_0) richtig, ob aber für realistische Daten und realistisches N dies so ist, können wir nicht mit absoluter Sicherheit sagen. Sicher ist aber, daß die Schätzung umso besser wird, je kleiner die Zahl n ist. Wenn man daher zunächst sehr viele Parameter zur Auswahl hat, sollte man in einer Voruntersuchung erst einmal möglichst viele "aussichtslose" Kandidaten aussortieren und nur auf

die restlichen n (≤ 20) die Rechnungen in Abschnitt 5 anwenden (mit neuen Daten). Natürlich geht man damit das Risiko ein, ein evtl. gutes Verfahren vorzeitig abzulehnen, nur weil es zufällig bei einigen Testfahrten schlecht abgeschnitten hat.

Ist die Zahl n von konkurrierenden Parametern sehr groß (evtl. unendlich), so sind die hier vorgestellten Algorithmen nicht anwendbar. Hier entsteht neben dem Problem, daß bei großem n viele der bisherigen Abschätzungen schlecht werden und einige Näherungen nicht mehr gerechtfertigt erscheinen auch das Problem, daß die Berechnung von $X_j(p)$ zeitaufwendig ist und nicht für alle Parameter durchgeführt werden kann. Hier sind geeignete Suchalgorithmen zu entwickeln (oder in der Literatur zu finden), die näherungsweise globale Minima finden. Ein grobes Muster eines Verfahrens für $n = \infty$ und einer Parametermenge $M \subseteq \mathbb{R}^d$ könnte etwa so aussehen:

- Wähle ein grobes Raster von Parametern in $M \subseteq \mathbb{R}^d$.
- Führe einige Testfahrten durch und berechne $X_j(p)$ für die Rasterpunkte.
- Wähle einige aussichtsreiche Parameter (in der Nähe des Minimums von $X_j(p)$) außerhalb des Rasters aus und berechne $X_j(p)$ dort.
- Führe weitere Testfahrten durch und berechne $X_j(p)$ aus allen bisherigen Fahrten für alle aussichtsreichen Parameter.

Entweder

- wähle nach einigen Iterationen die aussichtsreichsten Parameter aus und wende auf sie die Ergebnisse von Abschnitt 5 an.

oder

- versuche mit Hilfe eines zentralen Grenzwertsatzes den stochastischen Prozeß $Y_N(p) = \frac{1}{\sqrt{N}} \left(\sum_{i=1}^N (X_j(p)) - \frac{1}{N} \sum_{i=1}^N X_j(p) \right)$ zu beschreiben, seine Kovarianzfunktion zu schätzen (dies ist im allgemeinen recht problematisch) und (wiederum unter ungünstigen Umständen) die Wahrscheinlichkeit $P_{N,\epsilon}$ zu berechnen oder abzuschätzen. Da die Abschätzungen in Abschnitt 4 linear in n sind, sind sie hier völlig ungeeignet. Aus der Literatur sind uns keine Verfahren bekannt, die gute näherungsweise Berechnungen von $P_{N,\epsilon}$ erlauben. Evtl. könnte man wiederum $P_{N,\epsilon}$ durch Simulation näherungsweise bestimmen.

Weiterhin könnte man das Datenmaterial dadurch besser ausschachten, daß man den stochastischen Prozeß der Abweichungen in Ost- und Nordrichtung während einer ganzen Testfahrt genauer untersucht. Dies könnte auch Hinweise auf den Einfluß lokaler Störquellen liefern, sowie darauf, welche Parameter unter welchen Bedingungen besonders gut oder schlecht sind. Bei einer solchen Untersuchung könnte sich auch herausstellen, ob evtl. ein dynamischer Wechsel von einem Parameter zu einem anderen während der Fahrt, abhängig von gewissen Bedingungen (z.B. Fahrtgeschwindigkeit), sinnvoll ist.