

Wie schnell arbeitet das Simplexverfahren normalerweise? Oder: das Streben nach (stochastischer) Unabhängigkeit

Karl-Heinz Borgwardt

Angaben zur Veröffentlichung / Publication details:

Borgwardt, Karl-Heinz. 2014. "Wie schnell arbeitet das Simplexverfahren normalerweise? Oder: das Streben nach (stochastischer) Unabhängigkeit." *Mitteilungen der Deutschen Mathematiker-Vereinigung* 22 (2): 80–92. <https://doi.org/10.1515/dmvm-2014-0037>.

Nutzungsbedingungen / Terms of use:

licgercopyright



Wie schnell arbeitet das Simplexverfahren normalerweise?

Oder: Das Streben nach (stochastischer) Unabhängigkeit

Karl Heinz Borgwardt

1972 stand die Welt der mathematischen Optimierung
3 unter Schock, weil für das wichtigste Optimierungsver-
4 fahren, das Simplexverfahren zur Lösung linearer Opti-
5 mierungsprobleme, mit den Klee-Minty-Polytopen Bei-
6 spielprobleme gefunden worden waren, bei denen Vari-
7 anten dieses Verfahrens einen exponentiellen Rechenauf-
8 wand benötigen. Dies stand in krassem Gegensatz zu den
9 bis dahin gemachten – äußerst positiven – Erfahrungen
10 mit diesem Algorithmus und dessen Geschwindigkeit.

11 Infolgedessen setzte eine Forschungsbewegung ein, die
12 eine Welle von Publikationen und von wertvollen Ergeb-
13 nissen hervorbrachte. Es ging darum, analytisch nachzu-
14 weisen, dass die entdeckten Probleme „Ausreißer“ sind
15 und dass das Simplexverfahren normalerweise viel bes-
16 ser ist als es diese Worst-Case-Resultate besagen. Damit
17 begab man sich auf das Feld der probabilistischen Ana-
18 lyse von Algorithmen. Dort werden bei Unterstellung
19 von Verteilungsannahmen für Probleme Erwartungswer-
20 te für den Rechenaufwand analytisch ermittelt oder ab-
21 geschätzt. Dieser Aufsatz beschäftigt sich mit dieser For-
22 schungsrichtung in Bezug auf das Simplexverfahren, geht
23 aber aus von einer grundsätzlichen Erörterung des Kon-
24 zepts der probabilistischen Analyse von Algorithmen.

25 Im Zeitraum von 1975–1998 wurden viele derartige Un-
26 tersuchungen zum Average-Case-Verhalten des Simplex-
27 verfahrens angestellt. Maßgeblich involviert in diese For-
28 schungsrichtung der analytischen Ermittlung von Erwar-
29 tungswerten und durchschnittlichen Rechenzeiten waren
30 unter anderem Stephen Smale (Fields-Medaille 1966), Ri-
31 chard Karp (Fulkerson-Preis 1979), Michael Todd (Geor-
32 ge Dantzig Prize 1988, John von Neumann-Prize 2003),
33 Ilan Adler und Nimrod Megiddo (Lanchester Prize 1988)
34 und der Verfasser dieses Artikels (Lanchester Prize
35 1982). In all diesen Analysen konnte festgestellt werden,
36 dass die ermittelten Erwartungswerte sehr klein waren.

37 Ab Beginn des neuen Jahrhunderts entwickelte sich noch
38 eine andere Sicht auf die Beurteilung. Man wollte nun
39 vermeiden, dass der Aufwand bei schweren Problemen
40 dadurch verschleiert werden könnte, dass viele leichte
41 Probleme diese ausgleichen. Anders gesehen, wollte man
42 wissen, ob (oder ob nicht) sich die schweren Probleme
43 irgendwo häufen. Dazu unterzog man die einzelnen fest-
44 ten Probleme leichten Störungen und mittelte dann in
45 diesem kleinen Streubereich. Man war also gespannt dar-
46 auf zu sehen, ob sich bei dieser Art von Mittelung schon
47 eine Mäßigung des Rechenaufwands erkennen lässt. Die-
48 se Forschungsrichtung läuft unter dem Schlagwort Glät-
49 tungsanalyse (Smoothed Analysis).

50 Es war das Verdienst von Dan Spielman, zusammen mit
51 Shang-Hua Teng als erste dieses Prinzip erfolgreich auf
52 das Simplexverfahren angewendet zu haben und bewie-
53 sen zu haben, dass auch aus dieser neuen Sicht das Nor-
54 malverhalten besser als das Worst-Case-Verhalten ist.
55 Dafür erhielt Dan Spielman 2010 den Rolf Nevanlinna-
56 Preis.

57 In diesem Aufsatz sollen die Vorteile, die Nachteile,
58 aber auch die Missdeutbarkeiten und die manchmal über-
59 zogene Erwartungshaltung bei probabilistischen Analy-
60 sen von Algorithmusgeschwindigkeiten erläutert werden.
61 Dies geschieht konkret am Fall der probabilistischen Ana-
62 lyse des Simplexverfahrens, an deren Entwicklung der
63 Autor maßgeblich beteiligt war. Hier wird gezeigt, wie
64 weit man in den Forschungsbemühungen kommen kann-
65 te und welche systematischen Barrieren mathematischer
66 Art einer weitergehenden Durchleuchtung im Wege ste-
67 hen. Es wird verdeutlicht, dass der Analyseerfolg weit-
68 gehend davon abhängt, ob die verschiedenen ineinander-
69 greifenden Einflussfaktoren auf den algorithmischen Ver-
70 lauf voneinander separiert und dann einzeln analysiert
71 werden können. Dass das möglich ist, ist nämlich weder
72 selbstverständlich noch in der Praxis einfach durchführ-
73 bar.

74 I Was soll, will und kann die Probabilistische 75 Analyse?

76 Wir gehen von einem Rechenverfahren aus, das zur Lö-
77 sung konkreter Einzelprobleme (sogenannten Instanzen)
78 / einer Problemart Π eingesetzt werden kann und das
79 nachweislich für jedes I aus Π nach einer gewissen, von
80 I abhängigen, Rechenzeit die gewünschte Lösung liefert.
81 Wie üblich wollen wir im Folgenden anstelle der Rechen-
82 zeit die Anzahl der elementaren Rechenbefehle betrach-
83 ten. Die in der Komplexitätstheorie oft beachtete Dar-
84 stellungsgröße der darin verarbeiteten Zahlen (also die
85 Ziffernzahl oder die Kodierungslänge) spielt hier keine
86 Rolle.

87 Natürlich geht man davon aus, dass in der Regel
88 mit Zunahme der Anzahl von zu verarbeitenden Zahlen
89 der Rechenaufwand wächst. Deshalb teilt man die Menge
90 der Probleminstanzen I vom Typ Π ein in Kategorien, die
91 nach Größe der Datenmenge sortieren, wie $\Pi_n = \{I \mid I$
92 ist Instanz von Π und enthält n Zahlen}.

93 Innerhalb einer solchen n -Kategorie können nun In-
94 stanzen mit höchst unterschiedlichem Rechenaufwand
95 $R(A, I)$ auftreten (A steht für Algorithmus, I für die In-
96 stanz).

97 Der klassische, von extremer Vorsicht geprägte Ansatz
 98 zur Beurteilung des Algorithmus A ist die *Worst-Case-*
 99 *Analyse*. Man versucht dabei, für jedes n das Maximum
 100 über $\{R(A, I) \mid I \in \Pi_n\}$ zu erkennen, und beurteilt nach
 101 dem Verlauf dieser Maximalwerte in Abhängigkeit von n .
 102 Aber diese Kennzahl bzw. der Verlauf dieser Kennzah-
 103 len gibt oft keinen stichhaltigen Eindruck vom normalen
 104 Verhalten des Algorithmus. Um Informationen darüber
 105 erhalten zu können, müsste man erstmal eine klare Vor-
 106 stellung davon haben, welche Instanzen $I \in \Pi_n$ real zu
 107 lösen sind und wie oft diese auftreten. In der Sprache der
 108 Wahrscheinlichkeitstheorie müsste man also wissen, wie
 109 die realen I über Π_n verteilt sind. Dies weiß niemand,
 110 aber viele haben dazu individuelle Vorstellungen. Basie-
 111 rend auf einer klaren Vorstellung wäre man vielleicht in
 112 der Lage, über Zufallsexperimente den Rechenaufwand
 113 statistisch zu evaluieren. Jedoch entzieht sich diese Me-
 114 thodik der empirischen Austestung einer Ausdehnung in
 115 höhere Dimensionen und das so gewonnene Datener-
 116 gebnismaterial liefert meist keinen Einblick in die wirk-
 117 lichen Zusammenhänge und erlaubt deshalb auch keine
 118 zutreffende qualitative Deutung.

119 Für den Mathematiker stellt sich nun die folgende Her-
 120 ausforderung:

121 Kann ich bei Unterstellung der von mir angenommenen
 122 Verteilung der Instanzen in Π_n eine Analyse von pro-
 123 babilistischen Kenngrößen (Erwartungswerte, Varianzen
 124 usw.) analytisch gewinnen?

125 Dass dies meist nur unter außergewöhnlichen Umstän-
 126 den mit Ja beantwortet werden kann, ist offensichtlich.
 127 Und deshalb setzt ein rekursiver (eigentlich so nicht ge-
 128 wollter) Selektionsprozess ein:

129 Unser Mathematiker sucht nach Verteilungen oder Ver-
 130 teilungsmodellen, die er analysieren kann. Also tritt die
 131 Frage nach der „realen Verteilung“ (die ja niemand kennt)
 132 in den Hintergrund. Das Ergebnis der – wenn überhaupt
 133 – erfolgreichen Analyse basiert also auf einer Unterstel-
 134 lung, die sowohl vom Produzent als auch vom Abnehmer
 135 akzeptiert sein muss.

136 Hat man sich auf ein einheitliches stochastisches Mo-
 137 dell geeinigt, dann kann ein Wettbewerb zwischen ver-
 138 schiedenen Analysen um eine genauere Approximation
 139 der gewünschten Größen (z. B. Verkleinerung von Ober-
 140 schranken) stattfinden. Hüten muss man sich aber vor
 141 einem Herumbasteln am stochastischen Modell mit dem
 142 Ziel, die Oberschranken zu senken. Allzuoft prägt dann
 143 das stochastische Modell selber das Ergebnis und man
 144 untersucht in Wirklichkeit nicht mehr die Güte des Al-
 145 gorithmus. Wie eminent sich solche stochastischen An-
 146 nahmen auswirken können, wird in den folgenden Ab-
 147 schnitten verdeutlicht. Aber selbst bei Festhalten an ei-
 148 nem stochastischen Modell ist die Akzeptanz noch nicht
 149 gesichert. Nun setzt nämlich die Kritik an dem gewählten
 150 Modell aus Sicht der Anwender ein. Zu rechtfertigen ist
 151 nun, dass das verwendete Modell etwas mit Realproble-
 152 men und deren Auftrittshäufigkeiten zu tun hat.

153 Hinzu kommt noch ein anderer Effekt, der auch in den
 154 folgenden Abschnitten noch konkret verdeutlicht wird.
 155 Man tut sich oft schwer damit, eine Analyse basierend
 156 auf dem Einsatz von Algorithmus A in Reinform erfolg-
 157 reich durchzuführen. Oft wird nun der Ausweg gewählt,
 158 den Algorithmus so zu variieren, dass der variierte Al-
 159 gorithmus \tilde{A} tatsächlich analysierbar wird. Dessen Über-
 160 einstimmung mit A liegt dann vor allem (und manchmal
 161 nur noch) darin, dass er ebenfalls das Problem $I \in \Pi_n$
 162 erfolgreich löst.

163 Bei all diesen Kritikpunkten gegen die genauen qualitati-
 164 ven und quantitativen Ergebnisse der Analyse sollten aber
 165 deren Verdienste nicht kleingeredet werden. Durch die
 166 intensive mathematische Analyse gelingt es oft, die Ur-
 167 sachen für ein typischerweise auftretendes oder ein kri-
 168 tisch verlaufendes Verhalten des Algorithmus zu erken-
 169 nen. Abseits von quantitativen Auswertungen liefert dies
 170 eine Fülle von nachvollziehbaren Einsichten und Plausibi-
 171 litäten. Zumeist steht am Ende von vielerlei Bemühun-
 172 gen der Eindruck, dass der Algorithmus unter all den ver-
 173 schiedenen unterstellten Verteilungen immer ganz passa-
 174 bel funktioniert.

175 2 Lineare Optimierung mit dem 176 Simplexverfahren

177 Der Autor hat sich in den vergangenen Jahrzehnten in-
 178 tensiv mit der probabilistischen Analyse des Rechenauf-
 179 wands beim Lösen Linearer Optimierungsprobleme be-
 180 fasst. Deshalb soll das im vorigen Abschnitt in genereller
 181 Form Angesprochene nun in seiner Bedeutung für dieses
 182 Forschungsfeld konkretisiert und verdeutlicht werden.

Man will also folgendes Problem lösen:

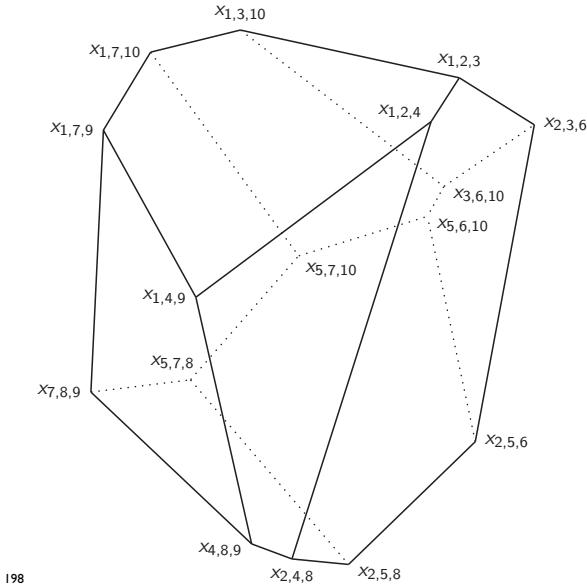
$$\begin{array}{lll}
 \text{maximiere} & v^T x & \text{als Zielfunktion} \\
 \text{unter} & \text{den Restriktionen} & \\
 \text{unter} & a_1^T x \leq b^1 & \\
 & \vdots & \text{bzw. } Ax \leq b \\
 (LP) & a_m^T x \leq b^m & \\
 \text{wobei} & v, x, a_1, \dots, a_m \in \mathbb{R}^n, \quad b \in \mathbb{R}^m. & \\
 \end{array}$$

185 Also ist hier n die Dimension oder die Anzahl der Ent-
 186 scheidungsvariablen, m ist die Anzahl der Restriktionen,
 187 und grundsätzlich sei $m \geq n$.

188 So entsteht ein Polyeder $P(A, b) = \{x \in \mathbb{R}^n \mid Ax \leq b\}$
 189 als Zulässigkeitsbereich für die Wahl des optimalen Va-
 190 riablenvektors x_{opt} .

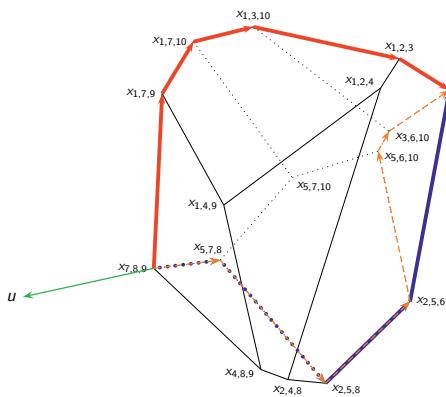
191 Es können folgende Fälle vorkommen:

- 192 – das Polyeder ist leer, weil das Restriktionensystem
 193 nicht erfüllbar ist;
- 194 – $v^T x$ (die Zielfunktion) lässt sich in $P(A, b)$ unbegrenzt
 195 steigern;
- 196 – in $P(A, b)$ gibt es Punkte x_{opt} mit maximalem $v^T x$ -
 197 Wert.



198

199 Jede Ecke ist bestimmt durch drei Restriktionen, die dort straff werden.



200

201 Von der Optimalecke zur Richtung u führen mehrere Simplexpfade (unterschiedlicher Länge) zur v -Optimalecke. Der rote Pfad entspricht dem Schatteneckenalgorithmus.

202 Es wird von einem Lösungsverfahren erwartet, dass es in 203 den ersten beiden Fällen diesen Sachverhalt meldet und 204 im letzten Fall einen der Optimalpunkte benennt.

205 Wir unterstellen die Nichtentartung der Problemstellung 206 (in (A, b) seien alle quadratischen Untermatrizen von vol- 207 lem Rang und die a_i sowie die Zielfunktion v seien in 208 allgemeiner Lage).

209 Wenn nun noch gilt $P(A, b) \neq \emptyset$, dann besitzt $P(A, b)$ 210 wegen $m \geq n$ eine Ecke.

211 Und dann verläuft ausgehend von dieser Ecke x_0 ein 212 (oder mehrere) Kantenpfad, wobei auf jeder Kante der 213 Zielfunktionswert $v^T x$ strikt ansteigt und der

214 – entweder mit einer ins Unendliche verlaufenden Kante 215 endet (Unbeschränktheit).

216 Die Ecke, an der diese Kante beginnt, kann als Ab- 217 bruchcke genutzt werden.

218 – oder bei einer Ecke (x_{opt}) endet, für die $v^T x$ optimal 219 ist.

220 Damit man diese Methode einsetzen kann, ist entschei- 221 dend, dass sich bei Polyedern $P(A, b)$, die eine Ecke be- 222 sitzen und auf denen die $v^T x$ -Werte beschränkt bleiben, 223 unter den Optimalpunkten eine Ecke befindet.

224 Diese iterierte Wanderung von Ecke über Kante zu einer 225 Nachbarecke bis hin zur Optimal- oder Abbruchcke ist 226 das Wesen des Simplexalgorithmus.

227 Jeder Wechsel von einer Ecke zu einer Nachbarecke (also 228 über eine Kante) hat arithmetisch folgende Bedeutung:

229 Die erste Ecke x_Δ werde bestimmt durch eine n - 230 elementige Indexmenge $\{\Delta^1, \Delta^2, \dots, \Delta^n\} \subset \{1, \dots, m\}$ 231 als Lösung des Gleichungssystems

$$232 a_{\Delta^1}^T x = b^{\Delta^1}, \dots, a_{\Delta^n}^T x = b^{\Delta^n}.$$

233 Für die Zulässigkeit von x_Δ ist entscheidend, dass auch 234 für alle $k \notin \Delta$ gilt $a_k^T x_\Delta \leq b^k$.

235 Nun wird ein Eckenwechsel dadurch ausgelöst, dass für 236 ein einzelnes Δ^i die Bedingung $a_{\Delta^i}^T x = b^{\Delta^i}$ gelockert wird 237 zu $a_{\Delta^i}^T x < b^{\Delta^i}$ unter Beibehaltung der anderen $n-1$ Glei- 238 chungen. Dadurch wird eine Bewegung auf einer Kante 239 weg von x_Δ initialisiert.

240 Diese Bewegung muss aber dort enden, wo $P(A, b)$ ver- 241 lassen werden würde. Dies ist erkennbar daran, dass nun 242 eine vorher lockere Restriktion $a_j^T x_\Delta \leq b^j$ mit $j \notin \Delta$ 243 die Gleichung $a_j^T x_{neu} = b^j$ erfüllt. x_{neu} ist dann die Ecke 244 am anderen Kantenende und erfüllt die Gleichungen

$$a_{\Delta^1}^T x_{neu} = b^{\Delta^1}, \dots, a_{\Delta^{i-1}}^T x_{neu} = b^{\Delta^{i-1}}, a_j^T x_{neu} = b^j, \\ a_{\Delta^{i+1}}^T x_{neu} = b^{\Delta^{i+1}}, \dots, a_{\Delta^n}^T x_{neu} = b^{\Delta^n}.$$

245 Für das Weitere wird zur Vereinfachung der Darstellung 246 jeweils $\Delta = \{1, 2, \dots, n\}$ unterstellt.

247 Diese Eckenwechsel (Pivotschritte) werden fortgesetzt, 248 bis (wie vorher beschrieben) eine Optimal- oder Ab- 249 bruchcke erreicht ist. Da von jeder Ecke (bei Nichtent- 250 artung) n Kanten ausgehen und ein Teil davon die Ziel- 251 funktion verbessert, steht diese Kantenmenge zur Aus- 252 wahl und es erfordert eine feste Regel (eine Variante), 253 um systematisch jeweils eine Kante zum Weiterkommen 254 zu benutzen.

255 Nun sei noch klargestellt, dass zum Auffinden der er- 256 wählten Startecke des Hauptproblems ein leicht modi- 257 fiziertes Problem (statt LP) gelöst werden muss. Des- 258 sen Lösung geschieht nach dem gleichen erwähnten Prin- 259 zip. Das modifizierte Problem besitzt aber den Vorzug, 260 dass man dafür eine Ausgangscke kennt. Es besitzt je- 261 doch den Nachteil, dass seine Lösung nur irgendeine Ecke 262 des Originalpolyeders liefert, nicht etwa die Optimal- 263 ecke. Diese so durch die „Phase I“ gelieferte Ecke wird 264 dann als Startecke für die beschriebene „Phase II“ be- 265 nutzt.

266 Die gebräuchlichsten Varianten (der Auswahl der nächs- 267 ten Kante) orientieren sich an ganz unterschiedlichen 268 Kriterien, wie z. B.

269 – am Zielfunktionsfortschritt, den die Benutzung dieser 270 Kante bringen wird (größte Verbesserung)

- 271 – am Abweichungswinkel zwischen der Zielrichtung v
 272 und der Kantenrichtung (steilster Anstieg)
 273 – an der Darstellung der Zielrichtung v durch die an der
 274 aktuellen Ecke straffen (d.h. mit Gleichheit erfüllten)
 275 Restriktionen
 276 $v = \sum_{i=1}^n \xi^i a_{\Delta^i}$ („Regel von Dantzig“)
 277 – an den Originalindizes der straffen Restriktionen („Re-
 278 gel von Bland“).

279 Eine für die probabilistische Analyse entscheidende Son-
 280 derrolle spielt hier die parametrische Variante von Gass
 281 und Saaty [9] (mittlerweile auch bekannt unter dem Na-
 282 men Schatteneckenalgorithmus). Diese soll wegen ihrer
 283 für das Hiesige wesentlichen Bedeutung näher erläutert
 284 werden. Die Bedeutung fußt auf dem sogenannten Polar-
 285 kegelsatz bzw. dem Lemma von Farkas:

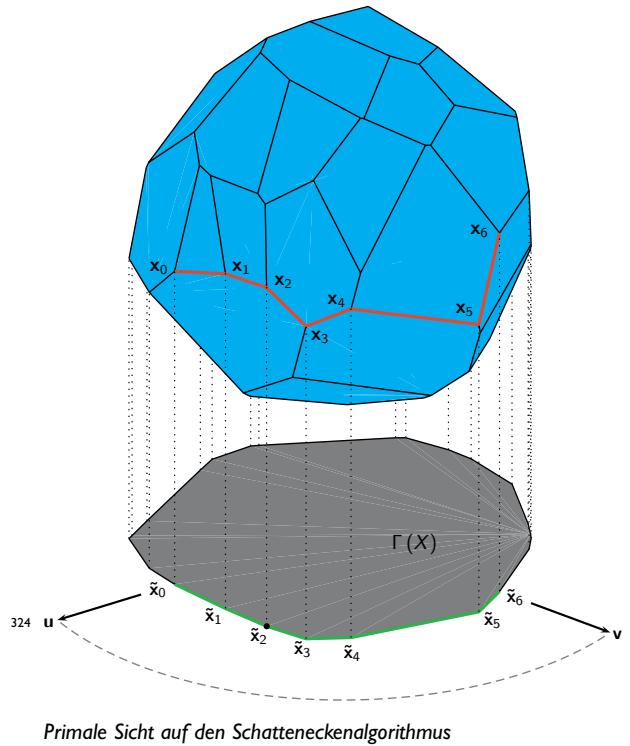
- 286 Eine Ecke x_{Δ} ist dann optimal, wenn die Restriktionsvek-
 287 toren $a_{\Delta^1}, \dots, a_{\Delta^n}$ zu den dort straffen Restriktionen die
 288 Zielrichtung konisch erzeugen, d.h. wenn gilt:
 289 v liegt im Kegel dieser Vektoren bzw. in $\sum_{i=1}^n \xi^i a_{\Delta^i} = v$
 290 sind alle $\xi^i \geq 0$.

291 Nun betrachtet man die vorliegende Startecke als Opti-
 292 malecke zu einer Zielrichtung u , die man als konische
 293 Kombination aus $a_{\Delta^1}, \dots, a_{\Delta^n}$ gewinnen kann, also $u =$
 294 $\sum_{i=1}^n \xi^i a_{\Delta^i}$ mit $\xi^i \geq 0$ für $1 \leq i \leq n$.

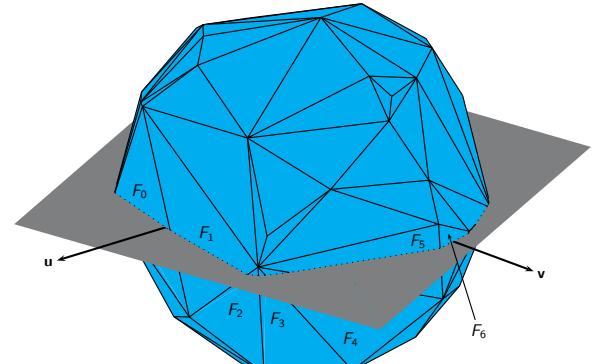
295 Wir sind uns aber bewusst, dass nicht $u^T x$, sondern $v^T x$
 296 optimiert (maximiert) werden sollte.

297 Deshalb betrachtet man nun Mischzielrichtungen $w_{\lambda} =$
 298 $(1 - \lambda)u + \lambda v$ mit $\lambda \in [0, 1]$. Während bei $\lambda = 0$ und
 299 $w_0 = u$ die Startecke für w_0 optimal ist, wird diese Ecke
 300 ihre Optimalität allmählich einbüßen, wenn wir λ stei-
 301 gern. Ab einem $\bar{\lambda} \geq 0$ ist dann eine andere Ecke optimal
 302 und dieser Wechsel vollzieht sich etliche Male, bis $\lambda = 1$
 303 und $w_1 = v$ erreicht werden kann. Es zeigt sich, dass
 304 jeder solche Wechsel der Optimalecke der Wanderung
 305 von der bisherigen Optimalecke zu einer Nachbarecke
 306 entspricht, so dass diese Wanderung einen Simplexpfad
 307 von der $u^T x$ -Optimalecke zur $v^T x$ -Optimalecke ergibt.
 308 Der Aufwand für einen solchen Wechsel ist leicht be-
 309 stimmbar ($O(m \cdot n)$). Deshalb ist man brennend daran
 310 interessiert, wie viele Eckenwechsel diese Wanderung er-
 311 fordert. Das kann man somit auch daran ablesen, wie vie-
 312 le Ecken denn bei der λ -Steigerung temporär für gewis-
 313 se w_{λ} die Optimalstellung innehaben. Der parametrische
 314 Algorithmus folgt mit seiner Eckenwanderung genau die-
 315 ser Folge von Optimalecken.

316 Der Autor hat in seiner Dissertation 1977 eine geometri-
 317 sche Bedeutung dieses Prozesses erkannt. Jede erreichte
 318 Ecke ist ja bezüglich einer Zielrichtung $w_{\lambda} = (1 - \lambda)u + \lambda v$
 319 extremal. Projiziert man nun das Polyeder $P(A, b)$ auf die
 320 Ebene $\text{Span}(u, v)$, so bildet sich ein zweidimensionaler
 321 Schatten. Die erwähnte Ecke wird wegen ihrer Extrema-
 322 lität auf eine Ecke des Schattens projiziert. Für diese Art
 323 von Ecken hat der Autor deshalb den Namen Schatten-



Primale Sicht auf den Schatteneckenalgorithmus



Duale Sicht auf den Schatteneckenalgorithmus

326 ecken und für den obigen Algorithmus den Namen Schat-
 327 teneckenalgorithmus geprägt [3, 4].

328 Folglich würde eine entdeckte Oberschranke für die Zahl
 329 der Schattenecken auch eine Oberschranke für die Sim-
 330 plexschritte auf dem Weg von Start- zu Zielecke abgeben.

331 Dem Beobachter dieser Bemühungen um probabilisti-
 332 sche Resultate, die in den folgenden Kapiteln beschrie-
 333 ben werden, stellt sich die Frage, weshalb eine solche
 334 Analyse ausgerechnet für den Schatteneckenalgorithmus
 335 und nicht auch für andere Varianten funktioniert. Denn es
 336 ist schon auffällig, dass beim Rotationssymmetrie-Modell
 337 (Kapitel III) beim Umklapp-Modell (Kapitel IV) und bei
 338 der Glättungsanalyse (Kapitel V) tatsächlich nur Ergeb-
 339 nisse über den Schatteneckenalgorithmus vorliegen.

340 Der Grund liegt im erkennbar klaren geometrischen
 341 Konzept des Schatteneckenalgorithmus, das sich mühe-
 342 los auf die Betrachtung der Einzelkandidaten (das sind in
 343 diesem Fall die Ecken bzw. Basislösungen) herunterbre-
 344 chen lässt.

345 Hier wird nämlich jeder Kandidat (das ist etwa die Kombination aus a_1, a_2, \dots, a_n bzw. die Lösung x_Δ des entsprechenden Gleichungssystems) daraufhin abgefragt, ob

348 simultan

349 – x_Δ Ecke von $P(A, b)$ ist bzw. ob $\text{Conv}(a_1, \dots, a_n)$ Facette von $\text{Conv}(a_1, \dots, a_m, 0)$ ist

350 – x_Δ eine Schattenecke ist bzw. ob $\text{Conv}(a_1, \dots, a_n)$ und $\text{Span}(u, v)$ sich schneiden.

353 Damit ergibt sich die Schatteneckenzahl direkt aus der Anzahl der Kandidaten, die beides gleichzeitig erfüllen. Also setzt sich diese „Summe“ aus den Einzelbeobachtungen zusammen. Den Erwartungswert bekommt man entsprechend, wenn man die Wahrscheinlichkeit dafür ermittelt, dass unser typischer Kandidat $\Delta = \{1, \dots, n\}$ beide Bedingungen gleichzeitig erfüllt.

360 Bei anderen Varianten des Simplexalgorithmus gibt es nur eine dynamische Charakterisierung mit Hilfe der Beobachtung des ganzen Pfades. Ob hier eine Ecke oder ein Kandidat (a_1, \dots, a_n) vom Simplexpfad besucht wird, kann nicht allein aus der Beobachtung der Daten dieses Kandidaten erkannt werden. Vielmehr hängt dies entscheidend davon ab, wie der Verlauf des Simplexpfares war, bevor das $v^T x$ -Niveau unseres Kandidaten erreicht war. Wurden dabei etwa keine Nachbarecken unseres Kandidaten erfasst, dann bleibt auch unser Kandidat außen vor. Es türmen sich also hier enorme Abhängigkeiten zwischen allen Daten auf, die viel zu komplex sind, als dass man sie (bisher) in den Griff hätte bekommen können.

3 Probabilistische Analyse für den Schatteneckenalgorithmus unter dem Rotationssymmetrie-Modell

377 Wenn wir also bei der Variante Schatteneckenalgorithmus die Frage nach der Zahl der Ecken auf dem Simplexpfad schlüssig ersetzen können durch die Frage nach der Anzahl von Schattenecken, dann bietet sich die Stochastische Geometrie zur Beantwortung an.

382 Um stochastische Informationen über diese Anzahl zu gewinnen, müssen wir uns zunächst auf ein stochastisches Modell festlegen. Dazu setzen wir (in diesem Abschnitt) das sogenannte Rotationssymmetrie-Modell ein.

Rotationssymmetrie-Modell

Zur Erzeugung der Probleminstanzen von

(LP1) maximiere $v^T x$ unter $a_1^T x \leq 1, \dots, a_m^T x \leq 1$

386 seien die Zufallsvektoren v, a_1, \dots, a_m und ein Hilfsvektor u jeweils über $\mathbb{R}^n \setminus \{0\}$

- stochastisch unabhängig
- identisch
- rotationssymmetrisch erzeugt.

388 (Die Festlegung der rechten Seite b^i auf 1 ist bedeutsam, solange $b^i \geq 0$ gilt. In diesem Fall kann diese Normierung durch Anpassung der a_i erfolgen.)

391 Will man stochastische Geometrie betreiben, dann empfiehlt es sich, den Dualraum zu betrachten, in dem die Zufallsvektoren v, a_1, \dots, a_m direkt erzeugt werden. Dies ist günstiger als im Primalraum, wo das Polyeder $P(A, b)$ und die Vektoren x leben, Studien anzustellen, die nur indirekt durch die Wahl der v, a_1, \dots, a_m geprägt werden.

397 Die folgenden beiden Zusammenhänge besorgen die erforderliche Übersetzung

399 (a) x_Δ (bestimmt durch $a_1^T x_\Delta = 1, \dots, a_n^T x_\Delta = 1$) ist genau dann Ecke von

400 $P(A, b) = \{x \mid a_1^T x \leq 1, \dots, a_m^T x \leq 1\}$, wenn $\text{Conv}(a_1, \dots, a_n)$ Facette (Randsimplex) von $\text{Conv}(0, a_1, \dots, a_m)$ ist.

404 (b) Die Ecke x_Δ von $P(A, b)$ ist genau dann Schattenecke bei Projektion auf $\text{Span}(u, v)$, wenn $\text{Conv}(a_1, \dots, a_n) \cap \text{Span}(u, v) \neq \emptyset$.

Da es nur $\binom{m}{n}$ Auswahlmengen der Art $\{a_{\Delta 1}, \dots, a_{\Delta n}\} \subset \{a_1, \dots, a_m\}$ gibt, lässt sich die erwartete Anzahl von Schattenecken errechnen als

$EW(\text{Anzahl Schattenecken}) =$

$$\binom{m}{n} P(\text{Conv}(a_1, \dots, a_n) \text{ ist Facette}$$

und $\text{Conv}(a_1, \dots, a_n) \cap \text{Span}(u, v) \neq \emptyset$.

407 Da hier a_1, \dots, a_n zufällig erzeugt sind, berücksichtigen wir deren Verteilungsfunktionen F in der folgenden Integraldarstellung :

$$EW(S) = \binom{m}{n} \int_{\mathbb{R}^n} \dots \int_{\mathbb{R}^n} P([\text{Conv}(a_1, \dots, a_n) \text{ Facette}] \wedge [\text{Conv}(a_1, \dots, a_n) \cap \text{Span}(u, v) \neq \emptyset]) dF(a_1) \dots dF(a_n)$$

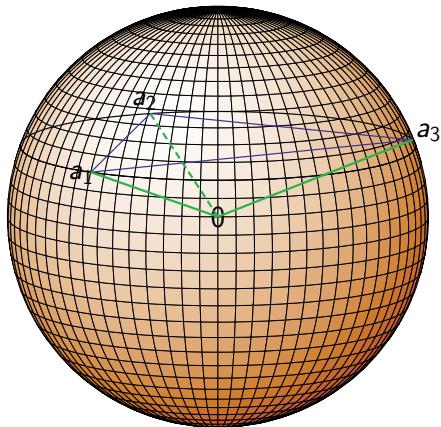
410 Es stellt sich heraus, dass das erste Ereignis, nur noch von 411 der Lage von a_{n+1}, \dots, a_m abhängt, sobald a_1, \dots, a_n ein- 412 mal festliegen. Genauso ist für das zweite Ereignis dann 413 nur noch entscheidend, wie u und v liegen. Somit können 414 wir in der Integraldarstellung aufspalten

$$EW(S) = \binom{m}{n} \int_{\mathbb{R}^n} \dots \int_{\mathbb{R}^n} P(\text{Conv}(a_1, \dots, a_n) \text{ Facette}) \cdot P(\text{Conv}(a_1, \dots, a_n) \cap \text{Span}(u, v) \neq \emptyset) dF(a_1) \dots dF(a_n).$$

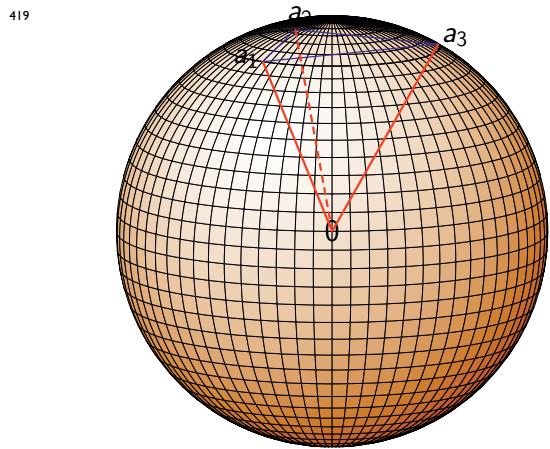
415 Man beachte aber, dass, solange sich a_1, \dots, a_n ändern, 416 sicher nicht gilt:

$$P([\text{Conv}(a_1, \dots, a_n) \text{ def. Facette}] \wedge [\text{Span}(u, v) \cap \text{Conv}(a_1, \dots, a_n) \neq \emptyset]) = P(a_1, \dots, a_n \text{ def. Facette}) \cdot P(\text{Span}(u, v) \cap \text{Conv}(a_1, \dots, a_n) \neq \emptyset).$$

417 Um dies plausibel zu machen, betrachte man die Gleich- 418 Verteilung auf der Vollkugel des \mathbb{R}^n .



Bei tiefer Lage des Dreiecks aus a_1, a_2, a_3 öffnet sich $\text{Cone}(a_1, a_2, a_3)$ stark, die Schnittwahrscheinlichkeit ist demnach groß. Gleichzeitig gibt es oberhalb extrem viel Platz für Punkte a_i , die die Facetteneigenschaft zerstören könnten. Die beiden Ereignisse sind also negativ korreliert.



Bei hoher Lage des Dreiecks aus a_1, a_2, a_3 öffnet sich $\text{Cone}(a_1, a_2, a_3)$ nur schwach, die Schnittwahrscheinlichkeit ist demnach klein. Gleichzeitig gibt es oberhalb kaum Platz für Punkte a_i , die die Facetteneigenschaft zerstören könnten. Die beiden Ereignisse sind also negativ korreliert.

a_1, \dots, a_n legen einen Simplex fest, der wiederum als affine Hülle eine Hyperebene definiert. Die Facetteneigenschaft erfordert, dass sich alle übrigen Vektoren a_{n+1}, \dots, a_m unterhalb dieser Hyperebene ansiedeln, d. h. in dem Halbraum, der auch den Ursprung enthält. Es liegt auf der Hand, dass bei Gleichverteilung aller Punkte auf der Vollkugel dieses Ereignis hochwahrscheinlich ist, wenn der Ursprung einen sehr großen Abstand zur Hyperebene hat, und sehr unwahrscheinlich wird, wenn der Abstand zwischen beiden klein ist. Aber genau umgekehrt ist es bei der Schnittbedingung. Hier muss ja der Kegel $\text{Cone}(a_1, \dots, a_n)$ von $\text{Span}(u, v)$ geschnitten werden. Liegt nun die Hyperebene als Träger von a_1, \dots, a_n nahe beim Nullpunkt, dann können a_1, \dots, a_n sehr weit voneinander entfernt sein. Das heißt, der betrachtete Kegel kann sich sehr weit öffnen, was die Schnittwahrscheinlichkeit in die Höhe treibt. Hat aber die Hyperebene einen großen Abstand vom Nullpunkt, dann müssen sich die n Punkte auf einer sehr kleinen $n - 1$ -dimensionalen Kreisscheibe ansiedeln. Infolgedes-

sen kann sich nur ein kleiner Kegel öffnen und die Schnittwahrscheinlichkeit wird extrem klein.

Dieser gegenläufige Effekt (diese negative Korrelation) sorgt nun dafür, dass das genannte Integral ziemlich klein bleibt. (Hätte man hier die oben beschriebene Produktform bzw. die Unabhängigkeit, dann wären die Oberschranken fulminant höher ausgefallen). Obige Unterscheidung hat zu Entstehungszeiten zu erheblichen Fehlinterpretationen, Missverständnissen und Anzweiflungen geführt. Wir werden im Späteren über andere Verteilungsmodelle reden, bei denen die hier zuweilen fälschlich unterstellte Unabhängigkeit wirklich zutrifft, was psychologisch einige Erklärungen für diese damaligen Auseinandersetzungen liefert. Nun ergab die korrekte Auswertung dieser negativen Korrelation für die Integraldarstellung folgendes Ergebnis

Fundamentalsatz (Schattenecken im RSM) [7]

Die erwartete Anzahl von Schattenecken bei m Restriktionen und n Entscheidungsvariablen bei Projektion auf $\text{Span}(u, v)$ ist für alle rotationssymmetrischen Verteilungen nicht größer als

$$m^{\frac{1}{n-1}} n^2 \cdot \text{Const.}$$

Man mache sich an dieser Stelle klar, dass bisher nur der Eckenwanderungsweg von der Startecke zur Abbruchcke unter die Lupe genommen worden ist. Aber zunächst verfügen wir ja noch gar nicht über die Startecke bzw. die Optimalecke zu einer frei gewählten Zielrichtung u .

Und hier kommt man in ein Dilemma. Im obigen Fundamentalsatz über die Schatteneckenanzahl wird nämlich verlangt, dass u unabhängig von den übrigen Problemdaten zu wählen ist. Um dann eine Schattenecke zu bekommen, muss man eine Ecke x_0 haben, die $u^T x$ maximiert. Würde man aber zuerst das $u^T x$ -Optimum suchen, dann wäre das Problem eigentlich nur verlagert worden von v auf u und keineswegs leichter geworden. Würde man u umgekehrt bestimmen aus $\text{Cone}(a_{\Delta 1}, \dots, a_{\Delta n})$, wobei $a_{\Delta 1}, \dots, a_{\Delta n}$ die straffen Restriktionsvektoren an einer irgendwie gewonnenen Startecke sind, dann hätte man die Erfordernis der unabhängigen Wahl von u gegenüber a_1, \dots, a_m verletzt und deshalb hätte man keinen rigoros richtigen Beweis mehr. Der Leser sei darauf hingewiesen, dass sich diese Komplikation durch alle folgenden Kapitel dieses Aufsatzes schlept.

Hier eröffnete sich ein Ausweg, der zwar vom effizientesten (Praktiker-)Verhalten abweicht, aber für einen rigorosen Beweis sorgt. Der sogenannte Dimensionssteigerungsalgorithmus löst das $v^T x$ -Maximierungsproblem, indem er nacheinander den Schatteneckenalgorithmus zur Lösung von Problemen mit wachsender Dimension einsetzt. Dazu lösen wir sukzessive die folgenden Probleme in den Stufen k , wobei k von 1 bis n läuft.

$$\begin{aligned} & \text{maximiere} && v^T x \\ & (\text{Stufe } k) \text{ unter} && a_1^T x \leq 1, \dots, a_m^T x \leq 1 \\ & && x_{k+1} = 0, \dots, x_n = 0 \end{aligned}$$

488 Man kann auf diese Weise den Optimalpunkt der k -ten
 489 Stufe dazu verwenden, um die $(k+1)$ -te Stufe zu starten.
 490 Denn dieser Optimalpunkt der k -ten Stufe liegt auf einer
 491 Schattenkante des $(k+1)$ -ten Problems bei Projektion auf
 492 $\text{Span}(e_{k+1}, v)$. Man beachte, dass diese beiden Vektoren
 493 das Gebot der Unabhängigkeit von den anderen Eingabe-
 494 daten erfüllen. Deshalb kann man dort wie gehabt den
 495 Schatteneckenalgorithmus einsetzen, um das $(k+1)$ -te
 496 Ergebnis zu erhalten. Die letzte Stufe führt uns (wenn
 497 nicht vorher schon abgebrochen worden war) zum Opti-
 498 malpunkt des Hauptproblems. Dadurch, dass man nun
 499 n Stufen durchläuft, schwächt die Oberschranke, die ja alle
 500 Schattenecken berücksichtigt, zunächst einmal um einen
 501 Faktor n an.

Man hat nach [7] und [6] folglich für alle Dimensions-
 paare (m, n) :

502 $EW(\text{Schatteneckenzahl bei Dimensionssteigerung})$

$$\leq m^{\frac{1}{n-1}} \cdot n^3 \cdot \text{Const.}$$

504 Jedoch kann man zumindest asymptotisch (d.h. bei $m \gg n$) nachweisen, dass der Korrekturbedarf in den höheren
 505 Stufen immer kleiner wird. Weil nämlich die Nachbesse-
 506 rung ja nur die letzte Komponente v^{k+1} der Zielrichtung
 507 betrifft, wird nur noch ein mit k immer kleiner werden-
 508 der Anteil der vorhandenen Schattenecken für Simplex-
 509 schritte genutzt, so dass man bei $m \gg n$ eine Abschät-
 510 zung folgender Art hat [12]

511 $EW(\text{benutzte Anzahl Ecken bei Dimensionssteigerung})$

$$\leq m^{\frac{1}{n-1}} \cdot n^{\frac{5}{2}} \cdot \text{Const.}$$

512 Es bleibt zu klären, wie man mit Problemen umgeht, bei
 513 denen die rechten Seiten b^i auch negativ sein können und
 514 bei denen die b^i unabhängig von den a_i gewählt werden.
 515 Leider lassen sich unsere Transformationen in den Dual-
 516 raum und unsere Skalierungen dann nicht mehr direkt
 517 durchziehen. Wir können aber einen rigorosen Beweis
 518 für folgendes Vorgehen führen.

519 Löse zunächst mit dem beschriebenen Dimensionssteige-
 520 rungsalgorithmus das Problem in der (bis jetzt bearbeit-
 521 baren) Einheits-Form, also

522 $(LP1) \quad \max v^T x \text{ unter } a_i^T x \leq 1 \text{ für } i = 1, \dots, m$

523 und bewahre dessen Optimaldecke.

Dann können wir durch Einführung einer zusätzlichen Va-
 524 riablen x^{n+1} eine Stufe an den Dimensionssteigerungsalgo-
 525 rithmus anhängen, bei der wir verlangen

$$\max x^{n+1} \text{ unter } a_i^T x + x^{n+1} (1 - b^i) \leq 1 \quad \forall i = 1, \dots, m$$

526 Dies führt zu folgender Beobachtung

527 – bei $x^{n+1} = 0$ liegt immer noch das Original mit Re-
 528 striktionen $a_i^T x \leq 1$ vor
 529 – bei $x^{n+1} = 1$ erfüllen wir die Anforderungen des er-
 530 weiterten Problems mit allgemeinem b^i .

529 Das passt aber in unser Schema vom Dimensionssteige-
 530 rungsalgorithmus als Stufe $n+1$. Als letztes soll sich also
 531 x^{n+1} von 0 lösen dürfen. Ist durch Steigerung von x^{n+1}
 532 der Wert 1 nicht erreichbar, dann ist das Originalpro-
 533 blem nicht zulässig. Ist aber durch Steigerung von x^{n+1}
 534 der Wert 1 erreichbar, dann ist man mit (x^1, \dots, x^n) zu-
 535 lässig. Hat man auch für diese letzte Stufe den Schatten-
 536 eckenalgorithmus benutzt, dann ist man beim Erreichen
 537 von $x^{n+1} = 1$ gleich beim optimalen x angelangt.

538 Eine Verteilungsannahme für die Werte b^1, \dots, b^m , für die
 539 auf diese Art ein Resultat gewonnen wurde, ist z. B.

540 – a_1, \dots, a_m sowie v seien nach dem Rotations-
 541 symmetrie-Modell verteilt.
 542 – die b^1, \dots, b^m seien davon und untereinander unabhän-
 543 gig jeweils gleichverteilt auf $[-1, 1]$.

544 Daraus entsteht dann eine Art Zylinderverteilung für die
 545 Vektoren (a_i, b^i) (Rotationssymmetrie in den ersten n
 546 Komponenten, Gleichverteilung in der letzten).

547 Während für die ersten n Stufen die obigen Ergebnis-
 548 se (Rotationssymmetrie) Bestand haben, muss die letzte
 549 Stufe wegen ihrer andersartigen Verteilung anders ana-
 550 lysiert werden. Aber auch für diese letzte Stufe wurde
 551 bestätigt, dass (bei $m \gg n$) gilt [10]

552 $EW(\text{durchlaufene Anzahl von Schattenecken})$

$$\leq m^{\frac{1}{n-1}} \cdot n^{\frac{5}{2}} \cdot \text{Const.}$$

Unter Hinzunahme dieser letzten Stufe lassen sich des-
 553 halb auch die Probleme mit beliebigen b^i unter unserem
 554 Modell lösen mit einem Gesamtaufwand von

$$555 \quad EW(\text{Gesamtaufwand}) \leq m^{\frac{1}{n-1}} \cdot n^{\frac{5}{2}} \cdot \text{Const.}$$

556 Nun wird der Praktiker entgegenhalten: Ich suche mir
 557 in einer ersten Phase eine Startecke (das ist irgendwie
 558 vergleichbar mit einem Durchlauf des Schatteneckenal-
 559 gorithmus). Zu dieser so erhaltenen Startecke bestim-
 560 me ich ein u , das diese Ecke $u^T x$ -optimal macht. Danach
 561 optimiere ich in Phase II mit dem Schatteneckenalgorith-
 562 mus die Zielfunktion $v^T x$. Da beides vergleichbare Op-
 563 timierungsprobleme sind, schätze ich den Aufwand hier-
 564 für auf ca. $m^{\frac{1}{n-1}} n^2$. Obwohl mannigfache Testläufe diese
 565 Hypothese irgendwie bestätigen, kann man nicht davon
 566 sprechen, dass diese Oberschranke bewiesen wäre. Wir
 567 stellen zwei essentielle Verletzungen der Voraussetzun-
 568 gen für den Beweis fest:

569 – Das vom Praktiker aufgestellte Phase-I-Problem ent-
 570 spricht nicht den Anforderungen des Rotations-
 571 symmetrie-Modells.
 572 – Die nachträgliche Bestimmung der Zielrichtung u aus
 573 dem Polarkegel der gewonnenen Startecke (also aus
 574 $\text{Cone}(a_{\Delta^1}, \dots, a_{\Delta^n})$) verletzt offensichtlich die essenti-
 575 ellen Unabhängigkeits-Anforderungen an das Problem
 576 im Rotationssymmetrie-Modell.

577 Also befindet man sich hier tatsächlich in einer Situation,
 578 wo man – wie angedeutet – einen um einen Faktor $n^{\frac{1}{2}}$
 579 (bei $m \gg n$) erhöhten Laufzeitbedarf für einen rigorosen
 580 Beweis in Kauf nehmen muss.

577 4 Stochastische Unabhängigkeit der
578 Ungleichungsausrichtungen

579 Parallel zum Rotationssymmetriemodell wurde ein gänz-
580 lich verschiedenes Modell, das Umklapp-Modell (Sign-
581 Invariance-Modell) für Average-Case-Analysen herange-
582 zogen [1, 2, 11, 14, 17].

583 Hier unterstellt man, dass der Zielvektor v und die Re-
584 striktionsvektoren a_1, \dots, a_m zunächst einmal festgelegt
585 sind (mit der kleinen Einschränkung der Nichtentartung).
586 Nun besteht die Variation der Probleme darin, dass für
587 jede Restriktionsungleichung und unabhängig voneinan-
588 der entschieden wird, ob diese Restriktion von der Form
589 $a_i^T x \leq b^i$ oder $a_i^T x \geq b^i$ sein soll. Beide Versionen sol-
590 len mit Wahrscheinlichkeit $\frac{1}{2}$ auftreten. Die untersuchten
591 und zu lösenden Probleme sind also von der Art

$$\begin{array}{ll} \text{maximiere} & v^T x \\ \text{unter} & a_1^T x \leq b^1, \dots, a_m^T x \leq b^m \\ \text{sowie} & P(\leq) = \frac{1}{2} = P(\geq) \text{ für jede} \\ & \text{Einzelrestriktion.} \end{array}$$

593 594 Auf diese Weise generiert eine Festlegung der Restrikti-
595 onsvektoren schon 2^m gleichberechtigte Partnerproble-
596 me. Bei der Average-Case Analyse werden jetzt fiktiv alle
597 diese Probleme gelöst, ihre Schrittzahlen werden aufad-
598 duiert und danach wird durch 2^m geteilt.

599 Bestechend an diesem Ansatz ist die Tatsache, dass (so-
600 lange die a_i nichtentartet festgelegt wurden) sich für je-
601 de Festlegung der gleiche Mittelwert ergibt: Man erspart
602 sich also hier alle Berücksichtigungen der Variation die-
603 ser a_i, b^i . Diese angesprochene Gleichheit ist eine Folge
604 der altbekannten Ergebnisse von Buck über Hyperebe-
605 nenarrangements im \mathbb{R}^n . Danach ergibt sich bei Vorga-
606 be von m Hyperebenen (nichtentartet) eine Partition des
607 Raumes in Zellen, deren Anzahl immer gleich ist, nämlich
608 $\binom{m}{0} + \binom{m}{1} + \binom{m}{2} + \dots + \binom{m}{n}$ [8].

609 Diese Zellen entsprechen genau denjenigen Umklapp-
610 Probleminstanzen (unter den 2^m möglichen), die über-
611 haupt zulässige Punkte x besitzen. Man beachte aber, dass
612 $\binom{m}{0} + \binom{m}{1} + \dots + \binom{m}{n} \leq 2^m$ bei $m \geq n$ und dass der
613 Quotient

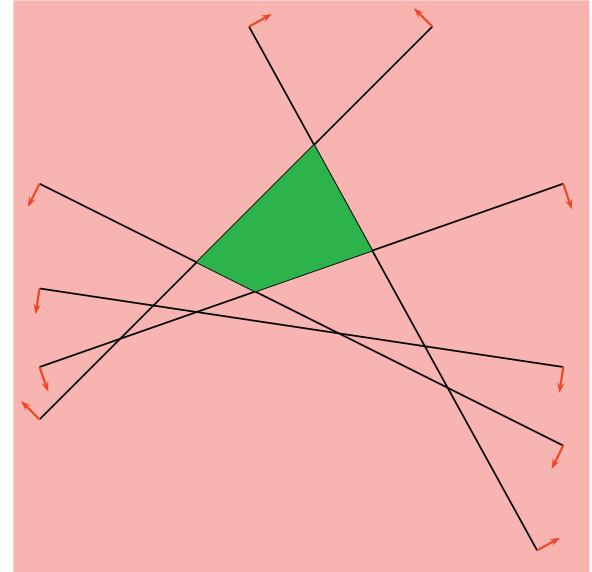
$$\frac{1}{2^m} \cdot [\binom{m}{0} + \dots + \binom{m}{n}]$$

614 dramatisch gegen 0 driftet, wenn bei festem n die Re-
615 striktionszahl m gegen ∞ geht.

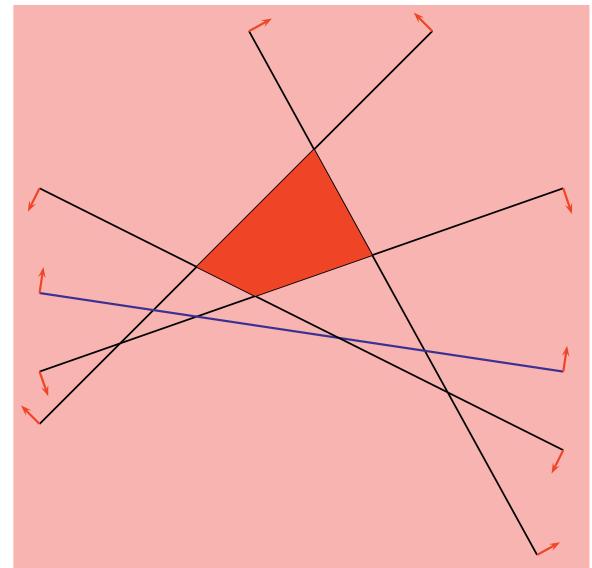
616 Dies bedeutet, dass allein schon bedingt durch das ange-
617 setzte stochastische Modell nur in einem verschwindend
618 kleinen Anteil der Probleminstanzen überhaupt noch ei-
619 ne Phase II stattfindet.

620 Der entsprechende Effekt wird auch noch dann erzielt,
621 wenn man auf die zulässigen Probleme konditioniert.

622 Denn es lässt sich zeigen, dass die Gesamtzahl der Ecken



Bei 5 Hyperebenen entstehen 16 Zellen. Wenn die Ausrichtungen alle passen, wird eine Zelle zulässig.



Dreht man bei einer vorher redundanten Restriktion die Ausrichtung um, dann gibt es gar keine Zulässigkeit mehr.

624 625 aller Zellen nicht größer ist als $2^n \binom{m}{n}$. Dann entwickelt
626 sich die Relation

$$\frac{2^n \binom{m}{n}}{\binom{m}{0} + \binom{m}{1} + \dots + \binom{m}{n}}$$

627 und dies wird auch klein bei $m \gg n$.

628 Der Grund hierfür liegt im Wesentlichen in der Tatsache
629 begründet, dass bei $m \gg n$ eine zusätzliche Hyperebe-
630 ne eine vorher bestehende Zelle zumeist verfehlt, also
631 nicht mehr teilt. Dann gibt es in diesem Fall zwei gleich-
632 wahrscheinliche Möglichkeiten aufgrund der Ausrichtung
633 dieser Restriktion

634 – die Zulässigkeit dieser Zelle wird vernichtet,

635 – die Zelle bleibt zulässig und komplett unversehrt (die
 636 neue Hyperebene hat also nichts zur Eckenvermehr-
 637 ung bei dieser Zelle und insgesamt beigetragen).

638 So weit hat unser stochastisches Modell also die Zeichen
 639 bereits gesetzt. Man beachte: wir haben bisher noch gar
 640 nicht festgelegt, welche Variante des Simplexalgorithmus
 641 eingesetzt wird.

642 Auch hier erwies sich der Schatteneckenalgorithmus als
 643 einzige auswertbare Option.

644 Es ergab sich als Resultat der Mittelwertbildung über
 die in Phase II (der Optimierungsphase) durchgeführten
 645 Pivotschritte (= Zahl der Schattenecken) ein Durch-
 schnitt von nicht mehr als n Schritten pro zulässiger
 Zelle [1, 2, 11, 17].

646 Noch haben wir aber nicht erörtert, wie man jeweils ei-
 647 ne Startecke (in Phase I) findet. Dazu haben [1, 2] einen
 648 sogenannten *Constraint-By-Constraint*-Algorithmus einge-
 649 setzt.

650 Dieser Algorithmus verläuft in $m-n$ Stufen. In jeder Stufe
 651 wird dafür gesorgt, dass eine weitere Restriktion erfüllt
 652 ist. Zur Erklärung seien gerade die \leq -Richtungen aktuell.
 653 Zunächst sind nur n Restriktionen erfüllt, wir agieren also
 654 auf $X^{(n)} = \{x \mid a_1^T x \leq b^1, \dots, a_n^T x \leq b^n\}$.

655 Nun versucht man ausgehend von der einzigen Ecke von
 656 $X^{(n)}$ die Restriktion $a_{n+1}^T x \leq b^{n+1}$ auch noch einzuhal-
 657 ten. Ist dies nicht erreichbar, dann kann man schon wegen
 658 Unzulässigkeit abbrechen.

659 Andernfalls gehen wir zur nächsten Restriktion über und
 660 wiederholen das Verfahren.

661 So wird also in $X^{(k)} = \{x \mid a_1^T x \leq b^1, \dots, a_k^T x \leq b^k\}$ in
 662 Zielrichtung a_{k+1} minimiert, bis $a_{k+1}^T x \leq b^k$ erreicht ist.
 663 Bezeichnend ist, dass dazu in jeder Stufe der parametri-
 664 sche Algorithmus eingesetzt wird. Ausgehend von einer
 665 einmal geschickt lexikographisch gewählten Zielrichtung
 666 $u \in \text{Cone}(a_1, \dots, a_n)$, die an der einzigen Ecke von $X^{(n)}$
 667 optimiert wird, und mit Hilfe der Zielrichtung a_{k+1} kann
 668 dann für jede Stufe eine Schatten-Projektionsebene fest-
 669 gelegt und damit jeweils der Schatteneckenalgorithmus
 670 eingesetzt werden.

671 Man nutzt hier die Cooptimalitätseigenschaft des para-
 672 metrischen Algorithmus aus. Dies bedeutet, dass für je-
 673 des erreichte Niveau bzgl. $a_i^T x$ der auf dem Schattene-
 674 ckenpfad erreichte Punkt den besten $u^T x$ -Wert aufweist.
 675 Deshalb erreicht dieses Verfahren also im Fall der Zuläs-
 676 sigkeit gleich die $u^T x$ -optimale Ecke von $X^{(m)} = P(A, b)$.
 677 Und von dort aus kann die Phase II starten. Im Fall der
 678 Unzulässigkeit bricht eine der Stufen ab.

679 In [1, 2] konnten die Autoren so nachweisen, dass auf
 diese Weise der Gesamtaufwand im Erwartungswert
 680 die Schranke von $2(n+1)^2$ nicht überschreitet.

681 Also verhindert hier eigentlich das Modell (und nicht die
 682 Qualität des Algorithmus) das Anwachsen der Schrittzah-
 683 len mit m . Die Autoren konstatieren: „These results can-
 684 not show a good behaviour of the algorithm for $m \gg n$,

685 but they are a consequence of the stochastic model per
 686 se“ [1].

687 Ein anderer interessanter Aspekt ist aber bei diesem Mo-
 688 dell zu beobachten (ganz im Gegensatz zum Rotations-
 689 symmetrie-Modell). Hier sind nämlich die Ereignisse
 690 – a_1, \dots, a_n definieren über $a_1^T x = b^1, \dots, a_n^T x = b^n$ ei-
 691 ne Ecke von $X^m = P(A, b)$
 692 – der Polarkegel dieser Basislösung enthält die Zielrich-
 693 tung u , also $u \in \text{Cone}(a_1, \dots, a_n)$

694 tatsächlich stochastisch unabhängig voneinander (denn
 695 hier liegen a_1, \dots, a_n ja von vornherein fest).

696 Hat man nämlich \bar{x} ermittelt als Lösung von $a_1^T \bar{x} =$
 697 $b^1, \dots, a_n^T \bar{x} = b^n$, dann ist für die Eigenschaft der Ba-
 698 sislösung unwichtig, ob $a_i^T x \leq b^i$ oder $a_i^T x \geq b^i$ ($i \leq n$)
 699 gefordert wurde.

700 Nun aber entscheiden über die primale Zulässigkeit von \bar{x}
 701 die Ausrichtungen der restlichen Ungleichungen, nämlich
 702 $a_i^T x \leq b^i$ oder $a_i^T x \geq b^i$ für $i = n+1, \dots, m$. Um aber zu
 703 wissen, ob diese Basislösung dual zulässig ist, muss man
 704 für eine Richtung v (die Zielrichtung) wissen, ob sie im
 705 Polarkegel von \bar{x} liegt. Dieser hat formal die Beschreibung
 706 $\text{Cone}(\pm a_1, \dots, \pm a_n)$, wobei + zu wählen ist bei Aus-
 707 richtung \leq und – bei Ausrichtung \geq . Da im Nichtenttar-
 708 tungsfall genau eine Ausrichtungskombination zum Erfolg
 709 führt, entscheiden also die Ausrichtungen zu a_1, \dots, a_n
 710 über die duale Zulässigkeit. Dagegen wird über die pri-
 711 male Zulässigkeit von den Ausrichtungen zu a_{n+1}, \dots, a_m
 712 entschieden. Die Definition des Modells sorgt dann für
 713 die stochastische Unabhängigkeit, da ja die Ausrichtun-
 714 gen unabhängig voneinander zu wählen sind.

715 5 Glättungsanalyse des Simplexverfahrens

716 Anfang des neuen Jahrhunderts kam eine ganz anders-
 717 artige Philosophie oder Sicht der Dinge bezüglich der
 718 normalen Laufzeit-Bewertung auf. Unter dem Namen
 719 „Smoothed Analysis“ (Glättungsanalyse) sollte die folgen-
 720 de Frage geklärt werden:

721 Wenn alle Daten des Problems (also a_1, \dots, a_m) gering-
 722 fügig Gauß-verteilt gestört werden und das Problem je-
 723 weils unter allen solchen Störungskonstellationen gelöst
 724 würde, wird sich dann eine signifikante Glättung erge-
 725 ben? Das heißt, werden diese Erwartungswerte erkenn-
 726 bar tiefer liegen als die extrem hohen Aufwandswerte,
 727 die man bei festen Problemen in Kauf nehmen muss
 728 (Klee–Minty-Probleme)? Wenn man will, kann man dies
 729 als eine andere Art von Average-Case-Analyse ansehen.

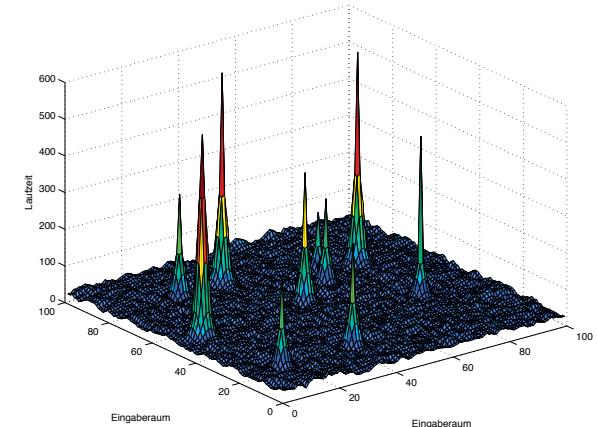
730 Eigentlich zu lösen ist also das feste Originalproblem

$$\begin{array}{ll} \text{maximiere} & \bar{v}^T x \\ (\overline{LP}) \text{ unter} & \bar{a}_1^T x \leq \bar{b}^1, \dots, \bar{a}_m^T x \leq \bar{b}^m \\ \text{mit} & \bar{a}_1, \dots, \bar{a}_m \in \mathbb{R}^n, \bar{b}^1, \dots, \bar{b}^m \in \mathbb{R}. \end{array}$$

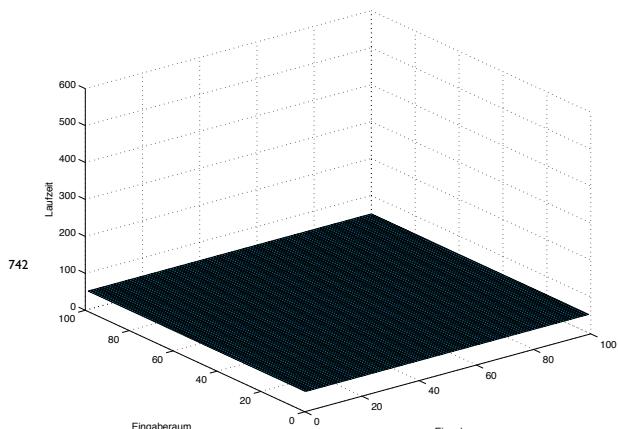
731 Nun sollen an den Vektoren $\bar{a}_1, \dots, \bar{a}_m$ und an den \bar{b}^i
 732 Störungen der Form $N(\bar{a}_i, \sigma)$ (also normal verteilt um
 733 \bar{a}_i bzw. \bar{b}^i mit Streuung σ) angebracht werden, so dass
 734 jeweils tatsächlich gelöst wird

735 maximiere $\bar{v}^T x$
 unter $a_1^T x \leq b^1, \dots a_m^T x \leq b^m$
 736 mit $(a_i, b^i) \in N((\bar{a}_i, \bar{b}^i), \sigma)$.

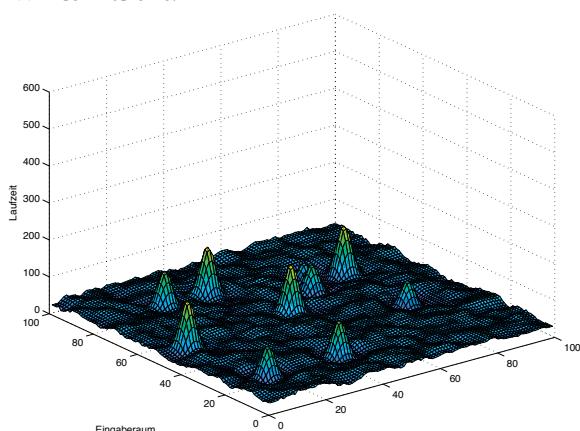
737 Der erwartete Aufwand, der bei dieser Art Mittelung
 738 entsteht, orientiert sich also nicht mehr an einer Gesamt-
 739 Verteilung, sondern konkret an diesem festen Original-
 740 problem. Es besteht die Hoffnung, dass man mit gering-



Bei einzelnen Problemen ergeben sich extreme Ausschläge der Rechenzeit.



Eine Average-Case Analyse ergibt einen Mittelwert über dem ganzen Raum der Probleme.



Die Glättungsanalyse mittelt in der Nähe eines jeden Festproblems und behebt so die extremen Ausschläge.

743 fügig wachsendem σ schon die extremen Ausschläge von
 744 Festproblemen vermeiden (besänftigen) kann. Und man
 745 hofft, dass für diesen Glättungseffekt eine gleichmäßige
 746 Oberschranke für alle Originalprobleme gefunden wer-
 747 den kann. Es ist deshalb angemessen und unvermeidlich,
 748 dass man bei der Gütemessung der Glättung neben den
 749 üblichen Dimensionen m und n auch die Störungsinten-
 750 sität einbezieht. Dies sollte konsequenterweise mit σ^{-1} ,
 751 also indirekt proportional zur Streuung geschehen. Die
 752 Verfechter dieser Sicht auf die Bewertung der Algorith-
 753 muslaufzeit führen folgende Argumente ins Feld:

- 754 – Im Gesamtbereich aller Problemstellungen wird der
 755 einzelne Anwender es oft mit einem oder einzelnen
 756 Typen zu tun haben. Viele denkbare Problemstellun-
 757 gen bzw. Datensetzungen haben keinen realen Anwen-
 758 dungshintergrund (auch wenn eine Abgrenzung zwi-
 759 schen realen und fiktiven nicht möglich ist).
- 760 – Das eigentlich gleiche Problem wird oft mit unter-
 761 schiedlicher Technologie und unterschiedlicher Model-
 762 lierungsgenauigkeit gelöst bzw. beschrieben.
- 763 – Es soll vermieden werden, dass die Effekte bei „ernst-
 764 zunehmenden“ realen Problemstellungen ausgeglichen
 765 oder verwässert werden durch fiktive Probleme, die
 766 rein aus Verteilungs- oder Symmetriegründen den ers-
 767 ten formal gleichgestellt sind.

768 Die Basis für alle im Folgenden besprochenen Aufwands-
 769 untersuchungen für die Lösung des Gesamtproblems bil-
 770 det das Fundamentaltheorem, das eine Oberschranke da-
 771 für angibt, wie viele Ecken des Zufallspolyeders $\{x \mid$
 772 $Ax \leq 1\}$ bei Projektion auf eine (wirklich feste) Ebene
 773 $\text{Span}(\bar{u}, \bar{v})$ zu Ecken des Bildes dieser Projektion werden.
 774 Also sind wir wieder bei der Anzahl der Schattenecken.

Fundamentaltheorem bei Glättung [15, 16]

Seien alle \bar{a}_i von der Norm $\|\bar{a}_i\| \leq 1$ und alle $\bar{b}^i = 1$.
 Dann ist der Erwartungswert der Schatteneckenanzahl
 bei Projektion auf $\text{Span}(\bar{u}, \bar{v})$ durch $D(m, n, \sigma)$
 gleichmäßig nach oben beschränkt, wobei

$$D(m, n, \sigma) := \frac{(58.888.678) \cdot m \cdot n^3}{\left[\min \left\{ \sigma, \frac{1}{3\sqrt{n} \ln m} \right\} \right]^6}.$$

775 776 777 Setzt man beispielsweise σ so an, dass sich im Nenner
 778 Gleichheit der beiden Größen in der Minimierungsklam-
 779 mer ergibt, dann hat man $D(m, n, \sigma) \approx \text{Const.} \cdot mn^3 \cdot$
 780 $n^3 \cdot (\ln m)^3$.

781 782 783 784 785 Die Herleitung dieses Ergebnisses ist enorm kompliziert
 786 und spielt in immenser Weise mit dem Instrumentarium
 787 der Eigenschaften von verknüpften und approximierten
 788 Normalverteilungen. Eine Verdeutlichung davon kann da-
 789 her hier nicht vorgenommen werden.

786 787 788 789 790 Die einschränkende Bedingung $\|\bar{a}_i\| \leq 1$ erweist sich bei
 791 der Algorithmusanalyse als sehr hinderlich, weil man ja ei-
 792 ne gleichmäßige Oberschranke für alle Probleme und alle
 793 Spielarten der a_i sucht. Jedoch lassen sich nach diesbe-
 794 züglicher Relaxierung durch Skalierung auch noch dann

791 Beschränkungen retten, wenn die obige Bedingung ver-
 792 letzt ist. Allerdings fallen dann die Oberschranken ent-
 793 sprechend höher aus. Erlaubt man auch höhere Normen
 794 der \bar{a}_i , dann ist in obiger Formel σ zu ersetzen durch

$$\frac{\sigma}{\max\{\|\bar{a}_i\| \mid i = 1, \dots, m\}}.$$

795 Erlaubt man auch Vektoren Mengen $\{a_1, \dots, a_m\}$, die mit
 796 unterschiedlichen Covarianzmatrizen M_i verteilt sind,
 797 dann braucht man zur grundsätzlichen Erhaltung der obi-
 798 genen Abschätzung, dass die Eigenwerte der M_i zwischen
 799 σ^2 und $\frac{1}{9n \ln m}$ liegen. Erlaubt man, dass die b^i positiv
 800 bleiben, aber von 1 abweichen, dann ist σ nun zu erset-
 801 zen durch die (nun noch kleinere) Größe

$$\frac{\sigma \cdot \min\{b^1, \dots, b^m\}}{\max\{\|a_1\|, \dots, \|a_m\|\} \cdot \max\{b^1, \dots, b^m\}}$$

802 Wie zu erkennen ist, treibt jede Relaxierung der Vor-
 803 aussetzungen aus dem Fundamentaltheorem die Ober-
 804 schranken nach oben. Und da eine gleichmäßige Ober-
 805 schranke angestrebt wird, muss man alle denkbaren Kon-
 806 stellationen abdecken. Außerdem ergeben sich die be-
 807 schriebenen „Überschreitungen“ selbst aus günstigen An-
 808 fangsdaten, wenn man zur Algorithmusanalyse des Pha-
 809 senwechsels bzw. des Gesamtalgorithmus kommt.

810 Betrachten wir nun die Problem-Gesamtbearbeitung und
 811 deren Analyse unter Zuhilfenahme des Fundamentaltheo-
 812 rems. Denn dazu braucht man zunächst eine Startecke
 813 und dazu ein festes \bar{u} . Im Gegensatz zum früher beschrie-
 814 benen Dimensionssteigerungsalgorithmus kommt man in
 815 diesem Fall nicht automatisch an der \bar{u} -optimalen Ecke
 816 vorbei und kann auch nicht mehr davon ausgehen, dass
 817 die Ankunftsecke nach Phase I unabhängig von (A, b)
 818 ist – es ist nicht einmal mehr selbstverständlich, dass
 819 die erzeugten Störungen dieser Ankunftsecke nach Ab-
 820 schluss der Phase I dem normalverteilten Störungsbild
 821 entsprechen, wie es im Fundamentaltheorem von den
 822 Daten verlangt wird. So wird u auch noch abhängig
 823 von a_1, \dots, a_n , da die a_1, \dots, a_n ja stochastisch zum Wa-
 824 ckeln gebracht werden. Es kann z.B. geschehen, dass
 825 $\bar{u} \in \text{Cone}(\bar{a}_1, \dots, \bar{a}_n)$, aber $u \notin \text{Cone}(a_1, \dots, a_n)$.

826 Natürlich könnte man sich, ausgehend von der Phase-I-
 827 Ecke, durch Optimierung von $\bar{u}^T x$ die „richtige“ Starte-
 828 cke besorgen. Dies hätte die Problematik aber nur von
 829 $\bar{v}^T x$ auf $\bar{u}^T x$ verlagert.

830 Deshalb bereinigen Spielman und Teng die obigen Kom-
 831 plikationen durch angemessene Erhöhung ihrer Ober-
 832 schranken. Um einen analysierbaren Lösungsalgorithmus
 833 zu bekommen, werden zwei Schritte ausgeführt. Zu-
 834 nächst wird ein zulässiges Hilfsproblem künstlich so kon-
 835 struiert, dass man für eine gewünschte Indexmenge (z.B.
 836 $\Delta = \{1, \dots, n\}$) mit Hilfe der Restriktionen a_1, \dots, a_n ei-
 837 ne Startecke erhält. Da Δ in aller Regel von vornherein
 838 noch keine Ecke liefert, werden alle anderen Restrikto-
 839 ren $a_{n+1}^T x \leq b^{n+1}, \dots, a_m^T x \leq b^m$ relaxiert, d.h. ihre
 840 Oberbegrenzungen werden erhöht oder die a_j werden

841 skaliert. Die künstlichen b^i werden positiv gewählt, um
 842 vom Fundamentaltheorem profitieren zu können.

843 Weil die notwendige Stärke dieser Relaxierungen von
 844 vornherein nicht bekannt ist, werden hierzu vorsichti-
 845 gerweise Größen verwendet, die aus den Eigenwerten/
 846 Konditionszahlen der Ausgangsmatrix entnommen wer-
 847 den können, was wiederum die Oberschranken anhebt.

848 In einem ersten Schritt löst man dieses eben geschaffene
 849 Hilfsproblem, bedenkt aber, dass eigentlich das Haupt-
 850 problem (evtl. auch mit negativen b^i) zu lösen wäre.
 851 Unbeschränktheit – auch für das Hauptproblem – wird
 852 schon in diesem ersten Schritt sichtbar.

853 Nun parametrisiert / interpoliert man zwischen den re-
 854 laxierten Versionen der a_i, b^i mit positiven b^i und den
 855 originalen Versionen. Auch dafür eignet sich der parame-
 856 trische Algorithmus (in einer dualen Form). (Dies ent-
 857 spricht der Vorgehensweise, die schon am Ende von Ka-
 858 pitel 3 beschrieben worden war). Vorteilhaft dafür ist,
 859 dass auf diese Weise das Fundamentaltheorem greift.
 860 Dies ist dann der zweite Schritt, der die Originallösung
 861 liefert.

862 Hierbei hing im ersten Schritt zwar nicht die Zielrichtung
 863 v , wohl aber die Anfangsrichtung u von (A, b) ab. Im Ge-
 864 gensatz dazu entspricht der zweite Schritt durchaus den
 865 Anforderungen des Fundamentaltheorems. Jedoch führen
 866 die Skalierungen im ersten Schritt und der Phasenwech-
 867 sel zum Anschwellen der Schranken aus dem Fundamen-
 868 taltheorem. Auch das Wackeln von u lässt sich auffangen,
 869 aber dazu müssen wiederum die Schwellen angehoben
 870 werden.

871 Als Alternative zu diesem mehrmaligen Schranken-Heben
 872 käme auch eine Randomisierung in Frage. Die Hoffnung
 873 ist dabei, dass man hier nur die Schranke so weit wie
 874 nötig anheben muss. Insbesondere legt man sich hier
 875 nicht auf eine Wunschbasis wie a_1, \dots, a_n fest, sondern
 876 probiert eine (komplexitätstheoretisch noch vertretba-
 877 re) Vielzahl von Anfangsbasen aus. Davon verwendet man
 878 diejenige, die die geringste Erhöhung der Schranken ver-
 879 ursachen würde. Es wird dabei stochastisch ausgenutzt,
 880 dass durch das Wackeln der Vektoren a_i (im Gegensatz
 881 zu den festen \bar{a}_i) eine relativ hohe Wahrscheinlichkeit da-
 882 für besteht, dass der kleinste Eigenwert der A -Teilmatrix
 883 von 0 erkennbar abweicht und dass die Konditionszahl
 884 relativ harmlos wird. Schließlich werden auch die b^i be-
 885 liebig aus \mathbb{R} gewählt und gestört. Legt man die Ergebnisse
 886 der stochastischen Glättungsanalyse alle betont pessimis-
 887 tisch aus (also so, dass alle Oberschrankenerhöhungen
 888 notwendig sind), dann kommt Vershynin für die Spielman-
 889 Teng-Laufzeitschranke auf eine Größenordnung von

$$m^{86} n^{55} \sigma^{-30}.$$

890 Vershynin hat in [18] die Abschätzungsmethodik für sein
 891 Fundamentaltheorem neu gestaltet. Darin erreicht er für
 892 $\sigma \leq \frac{1}{6\sqrt{n \ln m}}$ eine Oberschranke von $Cd^3 \sigma^{-4}$.

893 Zusätzlich hat er die Algorithmusdurchführung modifi-
 894 ziert. Zunächst einmal werden die Übergänge vom Hilfs-

problem zum Originalproblem so gestaltet, dass die parametrische Interpolation simultan auf die Zielfunktion und auf die rechten Seiten b^i zugreift. Die Variation der rechten Seite wird dabei aufgefangen durch eine zusätzliche Variable. Man vermeidet dadurch Verzerrungen von gaußverteilten Störungen beim Phasenwechsel.

Um das Startproblem in den Griff zu bekommen, geht Vershynin einen anderen Weg als Spielman/Teng. Er verzichtet auf die Auswahl einer Wunschbasis und damit auch auf die Relaxierung der restlichen Restriktionen. Stattdessen fügt er n Restriktionen

$$a_{m+1}^T x \leq b^{m+1}, \dots, a_{m+n}^T x \leq b^{m+n}$$

hinzu und macht diese so restriktiv, dass diese eine Ecke generieren. Diese Ecke dient als Startecke und das Baryzentrum dieser Vektoren wird die Startrichtung \tilde{u} . Von dieser fiktiven Ecke aus wird nun der parametrische Algorithmus gestartet.

Man gelangt dann zu einer Optimal- oder Abbruchcke. Diese aus dem modifizierten Problem gewonnene Abbruchcke wird dann daraufhin überprüft, ob sie nichts mehr mit den künstlichen (hinzugefügten) Restriktionen zu tun hat (ob diese also dort alle locker sind). Ist dies der Fall, dann ist das Problem ja gelöst.

Andernfalls (also wenn die erreichte Abbruchcke noch straffe hinzugefügte Restriktionen impliziert) muss der Lösungsprozess als misslungener Test verworfen werden. Nun wird ein erneuter Test durchgeführt, mit anders gewählten Vektoren a_{m+1}, \dots, a_{m+n} . Um zu erreichen und um zu zeigen, dass diese Tests nicht oft scheitern, geht er wie folgt vor:

Er bestimmt gleichverteilt über der Einheitskugel eine Richtung \tilde{u} . Nun bestimmt er a_{m+1}, \dots, a_{m+n} so, dass $\|a_i - \tilde{u}\|$ für alle i sehr klein ist und dass gilt $\tilde{u} \in \text{Cone}(a_{m+1}, \dots, a_{m+n})$. Zum Nachweis der Erfolgswahrscheinlichkeit betrachtet er den sogenannten *numbset* des Problems. Dieser besteht aus dem Bereich/Anteil der Einheitskugel, in den man Restriktionen hinzufügen kann, ohne dass dies auf das Ergebnis Einfluss haben würde. Es ist klar, dass dieser *numbset* immer mindestens die Hälfte der Einheitskugel umfasst (er enthält einen Halbraum).

Man kann also sicher sein, dass man mit Wahrscheinlichkeit nahe bei $\frac{1}{2}$ alle a_{m+i} im *numbset* untergebracht hat.

Und dies würde zum Erfolg des Tests führen: Eine tatsächliche Unterschranke dafür, dass all diese Vektoren im *numbset* liegen, ist $P = \frac{1}{4}$.

Wiederholt man diese Prozedur, so ergibt sich eine Oberschranke für die erwartete Anzahl der nötigen Testläufe (in diesem Falle 4). Somit wurde ein randomisierter Algorithmus eingesetzt, der (im schlechten Falle) nicht funktionieren muss, dessen nachhaltige Wiederholung aber mit hoher Wahrscheinlichkeit zum Ziel führt. Dadurch, dass es hier zu weniger Systembrüchen (Phasenwechseln) und zu weniger Verteilungsanpassungen kommt, verläuft hier die Kombination der verschiedenen Gaußanalysen und die Anwendung des Fundamentaltheorems über Zusatzabschätzungen viel harmloser.

Satz [18]

Für diesen beschriebenen randomisierten Algorithmus gilt eine gleichmäßige Oberschranke von der Größenordnung

$$O(\ln^2 m \cdot \ln(\ln m)(n^3 \sigma^{-4} + n^5 \ln^2 m + n^9 \ln^4 n)).$$

Hier hat man angesichts der Aufgabenstellung die Randomisierung akzeptiert.

Ohne Randomisierung (dafür aber mit Hilfe des umständlicheren Dimensionssteigerungsalgorithmus) hat Emanuel Schnalzger aufbauend auf den Ergebnissen von Spielman & Teng sowie von Vershynin nachweisen können, dass mit diesem Algorithmus ein erwarteter geglätteter Aufwand von

$$\text{Const} \left(\frac{1}{\sigma^2} + \frac{(n+1)^4}{\sigma^4} + (\ln m)^2 (n+1)^6 \right)$$

verbunden ist [13].

Diskutabel ist nun, ob man neben Datenzufälligkeit auch noch Verfahrenszufälligkeit (Randomisierung) akzeptiert. Und hier schließt sich der Kreis: Was wollte ich eigentlich – was soll ich – was kann ich?

Literatur

- [1] Adler, I., R. M. Karp, R. Shamir (1987). A Simplex Variant Solving an $m \times d$ Linear Program in $O(\min(m^2, d^2))$ Expected Number of Steps. *Journal of Complexity* **3** 372–387.
- [2] Adler, I., N. Megiddo (1985). A Simplex Algorithm where the Average Number of Steps is Bounded Between two Quadratic Functions of the Smaller Dimension. *Journal of the ACM* **32** 871–895.
- [3] Borgwardt, K.H.(1977). Untersuchungen zur Asymptotik der mittleren Schrittzahl von Simplexverfahren in der linearen Optimierung, Dissertation Universität Kaiserslautern.
- [4] Borgwardt, K.H.(1982a). Some Distribution-Independent Results About the Asymptotic Order of the Average Number of Pivot Steps of the Simplex Method. *Mathematics of Operations Research* **7** 441–462.
- [5] Borgwardt, K.H.(1982b). The Average Number of Pivot Steps Required by the Simplex-Method is Polynomial, *Zeitschrift für Operations Research* **26** 157–177.
- [6] Borgwardt, K.H. (1987). *The Simplex Method, A Probabilistic Analysis*, Springer Verlag, Heidelberg.
- [7] Borgwardt, K.H. (1999). A sharp upper bound for the expected number of shadow-vertices in LP-polyhedra under orthogonal projection on two-dimensional planes, In *Mathematics of Operations Research*, Vol.24, Nr. 4, 925–984
- [8] Buck, R.C. (1943). Partition of Space, In *American Mathematical Monthly* **50** 541–544
- [9] Gass, S. & Saaty, Th. (1955). The Computational Algorithm for the Parametric Objective Function, *Naval Research Logistics, Quarterly* **2**, 39–45
- [10] Göhl, M. (2013). Der durchschnittliche Rechenaufwand des Simplexverfahrens unter einem verallgemeinerten Rotationssymmetrie-Modell, Dissertation, Universität Augsburg.
- [11] Haimovich, M.(1983).*The Simplex Algorithm is Very Good! - On the Expected Number of Pivot Steps and Related Properties of Random Linear Programs*. Report Columbia University, New York.

- 1002 [12] Höfner, G. (1995). *Lineare Optimierung mit dem Schatteneckenalgoritmus –Untersuchungen zum mittleren Rechenaufwand und Entartungsverhalten*, Dissertation Universität Augsburg.
- 1003 [13] Schnalzger, E. (2014). Interner Report Universität Augsburg: *Smoothed Analysis des Dimensionssteigerungsalgorithmus*.
- 1004 [14] Smale, S. (1983). On the Average Speed of the Simplex Method of Linear Programming. *Mathematical Programming* **27** 241–262.
- 1005 [15] Spielman, D. & Tng, Sh.H. (2004). Smoothed Analysis of Algorithms: Why the Simplex Algorithm usually takes Polynomial Time, In *Journal of the ACM* **51**(3), 385–463.
- 1006 [16] Spielman, D. & Tng, Sh.H. (2008). Smoothed Analysis of Algorithms: Why the Simplex Algorithm usually takes Polynomial Time, 7. Version, Februar 2008, Report.
- 1007 [17] Todd, M.J. (1986). Polynomial Expected Behavior of a Pivoting Algorithm for Linear Complimentarity and Linear Programming Problems. *Mathematical Programming* **35** 173–192.
- 1008 [18] Vershynin, R. (2009). Beyond Hirsch conjecture: walls on random polytopes and smoothed complexity of the simplex method, In *SIAM Journal on Computing* **39**(2), 646–678.

1021 Prof. Dr. Karl Heinz Borgwardt, Lehrstuhl für Diskrete Mathema-
1022 tik, Optimierung und Operations Research, Institut für Mathema-
1023 tik, Universität Augsburg, 86135 Augsburg
1024 borgwardt@math.uni-augsburg.de

1025 1949 in Landstuhl/Pfalz geboren, Studium
1026 der Mathematik und Betriebswirtschaftslehre
1027 in Saarbrücken, 1973–1979 Wissenschaftlicher
1028 Mitarbeiter an der TU Kaiserslautern dort Pro-
1029 motion 1977 und Habilitation 1985. 1979–1984
1030 Planungsabteilung der Deutsche Bank AG Zen-
1031 trale (Frankfurt), 1982 Lanchester Prize für
1032 den Nachweis der Polynomialität der mittleren
1033 Schrittzahl beim Simplexverfahren, 1984–2014 Professor für Opti-
1034 mierung und Operations Research am Institut für Mathematik der
1035 Universität Augsburg. Forschungsgebiete: Probabilistische Analyse
1036 von Algorithmen, Lineare Optimierung, Heuristiken, Stochastische
1037 Polyedertheorie



1038

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Professor of Insurance Mathematics

The Department of Mathematics (www.math.ethz.ch) at ETH Zurich invites applications for a full professor position in Insurance Mathematics.

We are seeking candidates with an internationally recognized research record and with proven ability to direct research of highest quality in the field of insurance mathematics including mathematical finance or mathematical economics. We expect the successful candidate to integrate scientifically into the Department as well as to take a leading role in communication between academia, insurance and financial industry.

The successful candidate is expected to lead the education program for actuarial mathematics at ETH Zurich and will be expected to teach undergraduate level courses (mainly German) and graduate level courses (English).

**Please apply online at
www.facultyaffairs.ethz.ch**

Applications should include a curriculum vitae, a list of publications, and a statement of future research and teaching interests. The letter of application should be addressed **to the President of ETH Zurich, Prof. Dr. Ralph Eichler**. **The closing date for applications is 15 September 2014**. ETH Zurich is an equal opportunity and family friendly employer and is further responsive to the needs of dual career couples. We specifically encourage women to apply.