

11 Generierung von Szenariobäumen und Szenarioreduktion für stochastische Optimierungsprobleme in der Energiewirtschaft

Heitsch, H. ; Henrion, R. ; Küchler, C. ; Römisch, W.

Kurzfassung: Ein wesentlicher Teil der Eingangsdaten für stochastische Optimierungsprobleme sind Szenarien, die die mögliche Entwicklung der stochastischen Parameter beschreiben. In diesem Kapitel wird die Reduktion von Mengen von Szenarien als Inputs für stochastische energiewirtschaftliche Optimierungsprobleme diskutiert. Die Methodik und die vorgeschlagenen Algorithmen zur Szenarioreduktion beruhen auf der Verwendung angepasster Abstände von (multivariaten) Wahrscheinlichkeitsverteilungen bzgl. derer sich die stochastischen Optimierungsmodelle stabil verhalten. Solche Abstände sind für lineare bzw. gemischt-ganzzahlige zweistufige Optimierungsmodelle Fortet-Mourier Metriken bzw. gemischte Diskrepanz- und Fortet-Mourier Metriken. Weiter wird die Anwendung von Methoden der Szenarioreduktion zur Generierung von Szenariobäumen diskutiert.

Schlagworte: *Stochastische Optimierung, Szenarien, Szenariobaum, Szenarioreduktion, zweistufig, mehrstufig, gemischt-ganzzahlig*

11.1 Einleitung

Optimierungsprobleme in der Energiewirtschaft enthalten häufig unsichere Parameter, die in der Zielfunktion oder in den Gleichungen bzw. Ungleichungen, die die Restriktionen beschreiben, auftreten. Solche unsicheren Parameter sind beispielsweise zukünftige Strompreise am Spot- und Terminmarkt, Preise von Stromderivaten an den Strombörsen, Primärenergiepreise (für thermische Kraftwerke), die elektrische Last, Zuflüsse zu Reservoirs hydraulischer Kraftwerke, oder das Dargebot an erneuerbaren Energien für entsprechende Erzeugungsanlagen. Alle diese unsicheren Parameter erlauben eine Interpretation als Zufallsgrößen oder stochastische Prozesse. In den meisten Fällen existieren historische Daten bzw. ausgereifte stochastische Modelle. Als Beispiele dafür seien die Modelle für Strompreise [2, 3, 17, 25, 27, 29], die elektrische Last [5, 27] oder die Windgeschwindigkeit [6] genannt.

Bezeichnet $\xi = \{\xi_t\}_{t=1}^T$ den stochastischen Prozess mit d -dimensionalen Zufallsvariablen ξ_t für jeden Zeitpunkt t , der in ein energiewirtschaftliches Modell über einem diskreten Zeithorizont mit T Zeitpunkten eingeht, so kann ein solches Modell häufig in der Form

$$\min \left\{ \mathbb{E} \left(\sum_{t=1}^T f_t(x_t, \xi_t) \right) : x_t \in X_t, \sum_{\tau=0}^{t-1} A_{t\tau} x_{t-\tau} = h_t(\xi_t), t = 1, \dots, T \right\} \quad (11.1)$$

geschrieben werden. Dabei sind die Entscheidungen $x = (x_1, \dots, x_T)$ stochastische Prozesse, die den in (11.1) enthaltenen Restriktionen genügen müssen. Einerseits sollen sie zu jedem Zeitpunkt t zu einer gewissen Menge X_t gehören, die durch endlich viele lineare Gleichungen und Ungleichungen sowie möglicherweise durch Ganzzahligkeitsbedingungen beschrieben wird. Überdies liegen zeitübergreifende Restriktionen und Bilanzgleichungen für den Prozess x an jedem Zeitpunkt vor. Erstere beschreiben z.B. die

zeitliche Entwicklung der Zustände eines Energiesystems. Die Matrizen $A_{t\tau}$ besitzen dabei geeignete Dimension und $h_t(\xi_t)$ sind evtl. vom stochastischen Prozess ξ abhängige Vektoren, die z.B. Zuflüsse, die elektrische Last und Windenergie beschreiben. Die Funktionen $f_t(\cdot, \xi_t)$ beschreiben Kostenfunktionen, die von unsicheren Preisen abhängen können, und z.B. (stückweise) lineare oder quadratische Funktionen von Komponenten von x_t sind. Ziel ist hier die Minimierung der erwarteten Kosten (das Funktional \mathbb{E} liefert den Erwartungswert).

Der Zeitpunkt $t = 1$ repräsentiert die Gegenwart oder unmittelbar bevorstehende Zukunft, d.h. der Wert von ξ_1 ist bekannt oder kann hinreichend gut vorhergesagt werden. Daher wird im Modell (11.1) angenommen, dass die Entscheidung x_1 der ersten Zeitstufe deterministisch ist. Dies wird durch die Restriktion $x_1 = \mathbb{E}(x_1)$ beschrieben. Dann läßt sich das stochastische Optimierungsmodell (11.1) in das Optimierungsproblem

$$\min \{ f_1(x_1, \xi_1) + \mathbb{E}(\Phi(x_1, \check{\xi})) : x_1 = \mathbb{E}(x_1), x_1 \in X_1, A_{10}x_1 = h_1(\xi_1) \} \quad (11.2)$$

bzgl. der Erststufen-Entscheidung x_1 umformulieren, wobei $\check{\xi} := (\xi_2, \dots, \xi_T)$ der unsicheren Zukunft entspricht und die Funktion Φ definiert ist durch

$$\Phi(x_1, \check{\xi}) := \inf \left\{ \mathbb{E} \left(\sum_{t=2}^T f_t(x_t, \xi_t) \right) : x_t \in X_t, \sum_{\tau=0}^{t-1} A_{t\tau} x_{t-\tau} = h_t(\xi_t), t = 2, \dots, T \right\}.$$

Im Problem (11.2) ist dann die kostenminimale Erststufenentscheidung gesucht, die auch die erwarteten zukünftigen Kosten bei unsicherer Zukunft minimiert. Hierbei handelt es sich um ein *zweistufiges stochastisches Optimierungsproblem*.

Während die Erststufenentscheidung x_1 deterministisch ist, hängt die Zweitstufenentscheidung $\tilde{x} := (x_2, \dots, x_T)$ und auch jede ihrer Komponenten vom gesamten stochastischen Prozess ξ ab.

Will man überdies modellieren, dass die Entscheidung x_t zum Zeitpunkt t nur von der bis dahin vorhandenen Information (ξ_1, \dots, ξ_t) abhängt, dass also der Entscheidungsprozess rekursiv ist und zu jedem neuen Zeitpunkt auf der Basis neuer Information erfolgt, so kann dies über neue Restriktionen der Form

$$x_t = \mathbb{E}(x_t | \xi_1, \xi_2, \dots, \xi_t) \quad (t = 1, \dots, T) \quad (11.3)$$

in (11.1) formuliert werden. Dabei stellt $\mathbb{E}(\cdot | \xi_1, \xi_2, \dots, \xi_t)$ die bedingte Erwartung bzgl. des Zufallsvektors $(\xi_1, \xi_2, \dots, \xi_t)$ dar und die Restriktion bedeutet, dass für jedes $t = 2, \dots, T$, die Entscheidung x_t eine Funktion von $(\xi_1, \xi_2, \dots, \xi_t)$ darstellt. Die neuen Nebenbedingungen (11.3) heißen auch *Nichtantizipativitätsrestriktionen* und Problem (11.1) mit (11.3) *mehrstufiges stochastisches Optimierungsproblem*. Die Restriktion für $t = 1$ in (11.3) entspricht der bisherigen Bedingung $x_1 = \mathbb{E}(x_1)$ in (11.2).

Bezeichnet man die Restriktionsmenge der Erststufenentscheidung x_1 mit

$$\hat{X}_1 := \{x_1 \in X_1 : x_1 = \mathbb{E}(x_1), A_{10}x_1 = h_1(\xi_1)\},$$

so läßt sich eine Funktion f_0 wie folgt einführen

$$f_0(x_1, \xi) := \begin{cases} f_1(x_1, \xi_1) + \Phi(x_1, \check{\xi}) & , \text{ falls } x_1 \in \hat{X}_1 \text{ und } \Phi(x_1, \check{\xi}) \text{ endlich ist,} \\ +\infty & , \text{ sonst.} \end{cases}$$

Dabei sichern geeignete Annahmen an den Prozess ξ , dass $\Phi(x_1, \xi) > -\infty$ gilt. Die Endlichkeit von $\Phi(x_1, \xi)$ bedeutet dann, dass eine Entscheidung (x_1, \dots, x_T) existiert, die alle Restriktionen erfüllt. Die Existenz solcher zulässigen Entscheidungen muss nachgewiesen werden.

Mit diesen Festlegungen hat das Erststufen-Optimierungsproblem (11.2) die äquivalente Form

$$\min \left\{ \mathbb{E}(f_0(x_1, \xi)) = \int_{\mathbb{R}^s} f_0(x_1, \xi) P(d\xi) : x_1 \in \hat{X}_1 \right\}, \quad (11.4)$$

wobei P die multivariate Wahrscheinlichkeitsverteilung von ξ auf \mathbb{R}^s mit $s := Td$ ist und \mathbb{E} den Erwartungswert bzgl. P darstellt.

Um zwei- oder mehrstufige stochastische Optimierungsprobleme in Anwendungen näherungsweise numerisch zu lösen, muss der stochastische Prozess ξ durch einen Prozess mit endlich vielen Szenarien (oder *Szenarios*) oder, anders formuliert, die Wahrscheinlichkeitsverteilung P durch eine (diskrete) Verteilung mit endlich vielen Atomen ersetzt werden. Es seien nun ξ^i diese Szenarien mit den zugehörigen Wahrscheinlichkeiten p_i , $i = 1, \dots, N$. Alle Szenarien besitzen die Form

$$\xi^i = (\xi_1^i, \xi_2^i, \dots, \xi_T^i) \quad \text{mit} \quad \xi_1^i = \xi_1 \quad (i = 1, \dots, N).$$

Das stochastische Optimierungsproblem (11.4) nimmt dann die Form

$$\min \left\{ \sum_{i=1}^N p_i f_0(x_1, \xi^i) = f_1(x_1, \xi_1) + \sum_{i=1}^N p_i \Phi(x_1, \hat{\xi}^i) : x_1 \in \hat{X}_1 \right\} \quad (11.5)$$

an. Bei der Lösung von (11.5) stellt sich das folgende Problem. Einerseits ermöglicht eine große Zahl von Szenarien eine gute Darstellung der stochastischen Parameter. Andererseits wächst mit der Zahl der Szenarien der Aufwand zur Lösung des Optimierungsproblems. Daher ist es mitunter notwendig, sich bei der Lösung von (11.5) auf eine deutlich reduzierte Auswahl aus den N "Ausgangsszenarien" zu beschränken.

In den Abschnitten 11.3.1 und 11.4 werden die in [10, 15, 13, 14] entwickelten Ansätze und Algorithmen zur Szenarioreduktion für lineare bzw. gemischt-ganzzahlige zweistufige stochastische Optimierungsprobleme beschrieben. Der Zugang beruht auf Stabilitätsaussagen für stochastische Programme bzgl. geeigneter Abstände von Wahrscheinlichkeitsverteilungen, die im nächsten Abschnitt 11.2 diskutiert werden.

Mehrstufige stochastische Optimierungsprobleme modellieren einen Prozess sukzessiver Beobachtungen und Entscheidungen. Die im Verlauf der Zeit sukzessive erfolgende Beobachtung der stochastischen Parameter kann durch die Anordnung der Szenarien in einer Baumstruktur (vgl. Abbildung 11.5) modelliert werden. Bezeichnet k_t die Anzahl der möglichen Werte von ξ bzw. x zum Zeitpunkt t , so wächst diese Anzahl mit t , wobei $k_1 = 1$ der Wurzel und $k_T = N$ den Blättern des Baums entsprechen. Die Restriktion (11.3) nimmt dann die Gestalt (szenarienverkoppelnder) Gleichungsrestriktionen

$$x_t^i = \mathbb{E}(x_t | \xi_1, \xi_2, \dots, \xi_t)^i = \left(\sum_{j \in C_t^i} p_j \right)^{-1} \sum_{j \in C_t^i} p_j x_t^j \quad (i = 1, \dots, k_t, t = 1, \dots, T) \quad (11.6)$$

an, wobei die Mengen C_t^i , $i = 1, \dots, k_t$, eine Zerlegung der Menge $\{1, \dots, N\}$ darstellen und (wegen $k_1 = 1$ bzw. $k_T = N$) $C_1^1 = \{1, \dots, N\}$ bzw. $C_T^i = \{i\}$, $i =$

$1, \dots, k_T = N$ gilt. Sämtliche Mengen C_t^i werden mit wachsendem t kleiner. Unterteilt sich die Menge C_t^i ($i = 1, \dots, k_t$) beim Übergang von t zu $t + 1$ in die von i abhängigen Mengen C_{t+1}^j , $j = j_i, \dots, j_{i+1} - 1$ (mit $j_1 = 1$ und $j_{k_{t+1}} - 1 = k_{t+1}$), so bedeutet dies, dass sich Szenario i in t in $j_{i+1} - j_i$ Szenarien *verzweigt*. Allerdings weisen aus historischen oder simulierten Daten gewonnene Szenarien eine solche Baumstruktur im Allgemeinen nicht auf. Im Abschnitt 11.3.2 wird beschrieben, wie durch sukzessive Anwendung der Szenarioreduktion eine Baumstruktur hergestellt werden kann.

11.2 Stabilität und Metriken für Wahrscheinlichkeitsverteilungen

Für eine numerische Lösung stochastischer Optimierungsprobleme ist die ursprüngliche Wahrscheinlichkeitsverteilung P durch eine Verteilung Q mit endlich vielen Atomen bzw. Szenarien zu ersetzen. Dabei ist es naheliegend, die Wahrscheinlichkeitsverteilung Q derart zu wählen, dass die Ergebnisse der Optimierung (d.h. Optimalwerte und Lösungen) sich bei dieser Ersetzung möglichst wenig verändern, oder anders ausgedrückt, sich *stabil* verhalten.

Wir verwenden im Folgenden die Bezeichnungen

$$v(P) := \inf \left\{ \int_{\mathbb{R}^s} f_0(x_1, \xi) P(d\xi) : x_1 \in \hat{X}_1 \right\}$$

$$S(P) := \left\{ x_1 \in \hat{X}_1 : \int_{\mathbb{R}^s} f_0(x_1, \xi) P(d\xi) = v(P) \right\}$$

für den Optimalwert bzw. die Menge aller Lösungen von (11.4). In den meisten Anwendungen ist die Menge \hat{X}_1 beschränkt und es gelten die in [22] hergeleiteten Abschätzungen:

$$|v(P) - v(Q)| \leq \sup_{x_1 \in \hat{X}_1} \left| \int_{\mathbb{R}^s} f_0(x_1, \xi) (P - Q)(d\xi) \right| \quad (11.7)$$

$$\sup_{y_1 \in S(Q)} d(y_1, S(P)) \leq \psi_P^{-1} \left(\sup_{x_1 \in \hat{X}_1} \left| \int_{\mathbb{R}^s} f_0(x_1, \xi) (P - Q)(d\xi) \right| \right). \quad (11.8)$$

Hierbei bezeichnet $d(y_1, S(P))$ den Abstand eines Elementes y_1 aus der Menge $S(Q)$ aller Näherungslösungen zur ursprünglichen Lösungsmenge. Die Abbildung ψ_P^{-1} ist die Inverse der Wachstumsfunktion ψ_P der Zielfunktion $x_1 \mapsto \int_{\mathbb{R}^s} f_0(x_1, \xi) P(d\xi)$ in der Nähe der Lösungsmenge $S(P)$, d.h.

$$\psi_P(t) := \inf \left\{ \int_{\mathbb{R}^s} f_0(x_1, \xi) P(d\xi) - v(P) : d(x_1, S(P)) \geq t, x_1 \in \hat{X}_1 \right\}.$$

Die Wachstumsfunktion ψ_P ist monoton wachsend auf $[0, +\infty)$ und es gilt $\psi_P(0) = 0$. Die Abschätzungen (11.7) und (11.8) machen deutlich, dass

$$d(P, Q) := \sup_{x_1 \in \hat{X}_1} \left| \int_{\mathbb{R}^s} f_0(x_1, \xi) P(d\xi) - \int_{\mathbb{R}^s} f_0(x_1, \xi) Q(d\xi) \right| \quad (11.9)$$

der entscheidende Abstand der beiden Wahrscheinlichkeitsverteilungen ist. Dieser Abstand ist aber i.a. schwer numerisch auszuwerten, da f_0 mittels der Infimum-Funktion

Φ definiert ist. Für wichtige Klassen von Modellen ist es möglich, Abschätzungen des Abstandes (11.9) mit einfacher auszuwertenden Metriken herzuleiten. Dazu wird das Stetigkeits- und Wachstumsverhalten der in (11.4) auftretenden Integranden $f_0(x_1, \cdot)$ auf einer Menge Ξ untersucht, die die Träger der beiden Wahrscheinlichkeitsverteilungen P und Q enthält. Anschließend wird eine Menge \mathcal{F} von Funktionen bestimmt, die diesem Stetigkeits- und Wachstumsverhalten entspricht und jede der Funktionen $C f_0(x_1, \cdot)$ ($x_1 \in \hat{X}_1$) mit einer geeignet gewählten positiven Konstante $C > 0$ enthält. Dann gilt

$$d(P, Q) \leq \frac{1}{C} \sup_{f \in \mathcal{F}} \left| \int_{\mathbb{R}^s} f(\xi) P(d\xi) - \int_{\mathbb{R}^s} f(\xi) Q(d\xi) \right|.$$

Für die beiden Fälle (i) linearer und (ii) gemischt-ganzzahliger linearer zweistufiger Programme ergeben sich dabei folgende Funktionenmengen für $r \in \{1, 2\}$ und eine Menge \mathcal{B} polyedrischer Teilmengen von Ξ (vgl. [22, 24, 26]):

$$(i) \mathcal{F}_r(\Xi) = \{f : \Xi \rightarrow \mathbb{R} : f(\xi) - f(\tilde{\xi}) \leq c_r(\xi, \tilde{\xi}), \forall \xi, \tilde{\xi} \in \Xi\}.$$

$$(ii) \mathcal{F}_{r,\mathcal{B}}(\Xi) = \{f \mathbf{1}_B : f \in \mathcal{F}_r(\Xi), |f(\xi)| \leq \max\{1, |\xi|^r\}, \forall \xi \in \Xi, B \in \mathcal{B}\}.$$

Hierbei ist $|\cdot|$ eine Norm in \mathbb{R}^s , $\mathbf{1}_B$ die charakteristische Funktion der Menge B , d.h., $\mathbf{1}_B(\xi) = 1$ für $\xi \in B$ und $\mathbf{1}_B(\xi) = 0$ für $\xi \in \Xi \setminus B$, und

$$c_r(\xi, \tilde{\xi}) := \max\{1, |\xi|^{r-1}, |\tilde{\xi}|^{r-1}\} |\xi - \tilde{\xi}| \quad (\xi, \tilde{\xi} \in \Xi).$$

Man wählt $r = 1$, wenn ξ entweder in der Zielfunktion oder in der rechten Seite von Restriktionen auftritt. Tritt ξ in Zielfunktion *und* rechten Seiten auf, so setzt man $r = 2$. Die Funktionenmenge $\mathcal{F}_{r,\mathcal{B}}(\Xi)$ ist relevant für die Stabilität eines gemischt-ganzzahligen linearen stochastischen Programms, falls die Integranden $f_0(x_1, \cdot)$ Unstetigkeiten auf den Rändern von Mengen aus \mathcal{B} besitzen, aber im Inneren dieser Mengen zu \mathcal{F}_r gehören. Ein wichtiger Spezialfall für die Klasse \mathcal{B} ist die Menge aller Rechtecke in \mathbb{R}^s

$$\mathcal{B}_{\text{rect}} = \{I_1 \times I_2 \times \dots \times I_s : \emptyset \neq I_j \text{ ist abgeschlossenes Intervall in } \mathbb{R}, j = 1, \dots, s\},$$

die relevant für rein-ganzzahlige Optimierungsprobleme in der zweiten Stufe ist. Im Fall (i) entstehen die sog. *Fortet-Mourier Metriken* der Ordnung r

$$\zeta_r(P, Q) = \sup_{f \in \mathcal{F}_r(\Xi)} \left| \int_{\Xi} f(\xi) P(d\xi) - \int_{\Xi} f(\xi) Q(d\xi) \right| \quad (11.10)$$

und im Fall (ii) die verallgemeinerten Fortet-Mourier Metriken mit Ordnung r und Klasse \mathcal{B}

$$\zeta_{r,\mathcal{B}}(P, Q) = \sup_{f \in \mathcal{F}_r(\Xi), B \in \mathcal{B}} \left| \int_B f(\xi) P(d\xi) - \int_B f(\xi) Q(d\xi) \right|. \quad (11.11)$$

Da eine Folge von Wahrscheinlichkeitsverteilungen bzgl. der Metrik $\zeta_{r,\mathcal{B}}$ genau dann konvergiert, wenn sie sowohl bzgl. ζ_r als auch bzgl. der \mathcal{B} -Diskrepanz

$$\alpha_{\mathcal{B}}(P, Q) := \sup_{B \in \mathcal{B}} |P(B) - Q(B)|$$

konvergiert (vgl. [14]), kann man im Fall (ii) anstelle der schwierig zu handhabenden Metrik (11.11) auch den gemischten Abstand

$$d_{\lambda,r,B}(P,Q) = \lambda \alpha_B(P,Q) + (1-\lambda) \zeta_r(P,Q) \quad (11.12)$$

für ein $\lambda \in [0, 1]$ verwenden. Sind P und Q diskrete Wahrscheinlichkeitsverteilungen, so sind die Abstände (11.10) Optimalwerte linearer Optimierungsprobleme und lassen sich mit Algorithmen zur linearen Optimierung berechnen. In Abschnitt 11.4 wird gezeigt, wie in diesem Fall auch der Abstand (11.12) durch lineare Optimierung berechnet werden kann.

Ist die Menge Ξ beschränkt und abgeschlossen, so besitzen die Fortet-Mourier Metriken ζ_r eine duale Darstellung als Monge-Kantorovich Transport-Funktional. Solche Transport-Funktionale haben allgemein die Form

$$\mu_c(P,Q) := \inf \left\{ \int_{\Xi \times \Xi} c(\xi, \tilde{\xi}) \eta(d\xi, d\tilde{\xi}) : \eta \in M(P,Q) \right\}, \quad (11.13)$$

wobei $M(P,Q)$ die Menge aller Wahrscheinlichkeitsverteilungen auf $\Xi \times \Xi$ ist, deren Marginalverteilungen auf Ξ gerade P und Q sind, und c eine symmetrische und nichtnegative Kostenfunktion darstellt (vgl [21, 23]). Damit μ_c endlich ist, muss c bzgl. P und Q integrierbar sein. Für die spezielle Wahl der Kostenfunktion c

$$c(\xi, \tilde{\xi}) := |\xi - \tilde{\xi}|^r \quad (\forall \xi, \tilde{\xi} \in \Xi)$$

mit $r \in \{1, 2\}$ entsteht dabei die L_r -minimale Metrik ℓ_r definiert durch

$$\ell_r(P,Q)^r = \inf \left\{ \int_{\Xi \times \Xi} |\xi - \tilde{\xi}|^r \eta(d\xi, d\tilde{\xi}) : \eta \in M(P,Q) \right\} \quad (11.14)$$

auf der Menge $\mathcal{P}_r(\Xi)$ aller Wahrscheinlichkeitsverteilungen mit endlichem r -ten Moment (vgl. [23, Chapter 2.6]). Eine Folge von Wahrscheinlichkeitsverteilungen konvergiert bzgl. der L_r -minimalen Metrik ℓ_r genau dann, wenn sie bzgl. der Fortet-Mourier Metrik ζ_r konvergiert. Ebenso wie ζ_r ist daher ℓ_r eine Wahrscheinlichkeits-Metrik, bezüglich derer sich zweistufige lineare stochastische Programme stabil verhalten. Ist Ξ beschränkt und abgeschlossen, so ist die duale Darstellung

$$\zeta_r(P,Q) = \inf \left\{ \int_{\Xi \times \Xi} \hat{c}_r(\xi, \tilde{\xi}) \eta(d\xi, d\tilde{\xi}) : \eta \in M(P,Q) \right\} \quad (11.15)$$

für alle $P, Q \in \mathcal{P}_r(\Xi)$ gültig. Dabei ist \hat{c}_r die sog. *reduzierte* Kostenfunktion von c_r und durch

$$\hat{c}_r(\xi, \tilde{\xi}) := \inf \left\{ \sum_{j=1}^{n+1} c_r(z_{j-1}, z_j) : z_0 = \xi, z_{n+1} = \tilde{\xi}, z_j \in \Xi, n \in \mathbb{N} \right\} \quad (11.16)$$

definiert. Die Funktion \hat{c}_r ist eine Metrik auf Ξ und es gilt $\hat{c}_r \leq c_r$ (vgl. [23, Chapter 4.1.3]). Im Fall von diskreten Wahrscheinlichkeitsverteilungen P und Q sind Transport-Funktionale Optimalwerte linearer Optimierungsprobleme.

Liegt ein lineares mehrstufiges stochastisches Optimierungsproblem vor (mit der zusätzlichen Restriktion (11.3)), so genügt es i.a. für die Stabilität nicht, Metriken für Wahrscheinlichkeitsverteilungen vom Typ ℓ_r oder ζ_r zu betrachten (vgl. [12, 11]). Letzteres ist nur möglich unter weiteren Voraussetzungen an die Wahrscheinlichkeitsverteilung P und die Approximationen Q (vgl. [18]).

11.3 Lineare stochastische Optimierungsprobleme

11.3.1 Szenarioreduktion für zweistufige Modelle

Die Idee der Szenarioreduktion beruht auf der Verwendung von Metriken für Wahrscheinlichkeitsverteilungen bzgl. derer sich die stochastischen Optimierungsprobleme sich stabil verhalten (vgl. Abschnitt 11.2). In der Arbeit [10] wurden die Methodik und Algorithmen zur Szenarioreduktion aus [4, 8] unter direkter Verwendung von Fortet-Mourier Metriken (vgl. (11.10)) weiterentwickelt. Dies wird im Folgenden beschrieben. Die Methodik beruht auf dem Prinzip der optimalen Szenarioreduktion bzgl. einer Metrik d für Wahrscheinlichkeitsverteilungen, die mittels eines Monge-Kantorovich Transportproblems gegeben ist. Es sei P eine diskrete Wahrscheinlichkeitsverteilung mit Szenarien ξ^i und Wahrscheinlichkeiten p_i für $i \in \{1, \dots, N\}$. Weiter sei Q_J eine Wahrscheinlichkeitsverteilung, deren Träger $\text{supp}(Q_J)$ nur die folgende Teilmenge der Szenarien von P umfasst:

$$\text{supp}(Q_J) = \{\xi^i \mid i \in \{1, \dots, N\} \setminus J\} \quad \text{und} \quad J \subset \{1, \dots, N\}.$$

Die Wahrscheinlichkeiten von Q_J werden mit q_i ($i \notin J$) bezeichnet. Motiviert durch die Stabilitätsresultate in Abschnitt 11.2 besteht das Ziel nun darin, die Wahrscheinlichkeitsverteilung Q_J so zu bestimmen, dass die Abstände $\zeta_r(P, Q_J)$ bzw. $\ell_r(P, Q_J)$ möglichst klein werden.

Dies bedeutet, dass, bei fixiertem $n \in \mathbb{N}$ mit $n < N$, die Indexmenge J und die Wahrscheinlichkeiten q_i , $i = 1, \dots, n$, als Lösung des Minimumproblems

$$\min \left\{ d(P, Q_J) : J \subset \{1, \dots, N\}, \#J = N - n, q_i \geq 0, i \notin J, \sum_{i \notin J} q_i = 1 \right\} \quad (11.17)$$

mit $d = \zeta_r$ bzw. $d = \ell_r$ bestimmt werden (*optimale Szenarioreduktion*).

Das Minimumproblem (11.17) läßt sich durch

$$\min_J \left\{ \min_q \left\{ d(P, Q_J) : q_i \geq 0, i \notin J, \sum_{i \notin J} q_i = 1 \right\} : J \subset \{1, \dots, N\}, \#J = N - n \right\}$$

in ein inneres Optimierungsproblem zur Bestimmung von q_i , $i \notin J$, und ein äußeres Minimierungsproblem für die Indexmenge J aufspalten. Bei der inneren Minimierung ergibt sich das von J abhängige Infimum D_J durch

$$D_J := \min \left\{ d(P, Q_J) : q_i \geq 0, i \notin J, \sum_{i \notin J} q_i = 1 \right\}. \quad (11.18)$$

Besitzt die Metrik d eine Darstellung als Transport-Funktional, so läßt sich das innere Optimierungsproblem (11.18) explizit lösen (vgl. [4, Theorem 2]). Für die Metrik $d = \ell_r$ ergibt sich

$$D_J = \min \left\{ \ell_r(P, Q_J) : q_i \geq 0, i \notin J, \sum_{i \notin J} q_i = 1 \right\} = \left(\sum_{j \in J} p_j \min_{i \notin J} |\xi^i - \xi^j|^r \right)^{\frac{1}{r}}. \quad (11.19)$$

Darüber hinaus wird das Minimum in (11.19) durch die Wahrscheinlichkeitsverteilung Q_J^* mit

$$q_i^* = p_i + \sum_{\substack{j \in J \\ i(j)=i}} p_j \quad \text{und} \quad i(j) \in \arg \min_{i \notin J} |\xi^i - \xi^j|^r, \quad i \notin J \quad (11.20)$$

realisiert. Im Fall $d = \zeta_r$ gilt analog (vgl. [10])

$$D_J = \min \left\{ \zeta_r(P, Q_J) : q_i \geq 0, i \notin J, \sum_{i \notin J} q_i = 1 \right\} = \sum_{j \in J} p_j \min_{i \notin J} \hat{c}_r(\xi^i, \xi^j). \quad (11.21)$$

Das Minimum wird durch die Wahrscheinlichkeitsverteilung Q_J^* mit

$$q_i^* = p_i + \sum_{\substack{j \in J \\ i(j)=i}} p_j \quad \text{und} \quad i(j) \in \arg \min_{i \notin J} \hat{c}_r(\xi^i, \xi^j), \quad i \notin J, \quad (11.22)$$

realisiert, wobei die reduzierten Kosten \hat{c}_r sich aus der gegenüber (11.16) vereinfachten Formel

$$\hat{c}_r(\xi^i, \xi^j) := \min \left\{ \sum_{k=1}^{n-1} c_r(\xi^{l_k}, \xi^{l_{k+1}}) : n \in \mathbb{N}, l_k \in \{1, \dots, N\}, l_1 = i, l_n = j \right\}.$$

berechnen. Die Berechnung von \hat{c}_r kann durch Anwendung eines einfachen *Dijkstra-Algorithmus* erfolgen (vgl. hierzu auch [19, 7]).

Die Formeln (11.20) und (11.22) liefern in beiden Fällen eine Vorschrift zur *optimalen Neuverteilung*. Diese besteht darin, die Wahrscheinlichkeit eines gestrichenen Szenarios zu der Wahrscheinlichkeit eines (bzgl. der Abstände $|\cdot|^r$ bzw. \hat{c}_r) nächstgelegenen Szenarios zu addieren.

Mit Hilfe des Funktionals D_J kann das Problem (11.17) der optimalen Szenarioreduktion nun in der Form

$$\min \{D_J : J \subset \{1, \dots, N\}, \#J = N - n\} \quad (11.23)$$

geschrieben werden. Das Funktional D_J wird dabei je nach Metrik ℓ_r bzw. ζ_r mit Hilfe der Formeln (11.19) bzw. (11.21) berechnet. Die Optimierungsaufgabe besteht in der Bestimmung einer Indexmenge J mit vorgeschriebener Kardinalität $1 \leq N - n = \#J < N$ und minimalen Kosten D_J . Solche kombinatorischen Optimierungsprobleme sind im Allgemeinen \mathcal{NP} -schwer. Das Problem (11.23) stellt insbesondere ein sog. *n-median-Problem* dar (vgl. [28, Chapter 25]).

Aufgrund ihrer hohen Komplexität können kombinatorische Probleme dieser Art nur selten exakt gelöst werden (vgl. [1]) und man greift in der Regel auf Näherungsalgorithmen zurück. Speziell für das Problem der optimalen Szenarioreduktion wurden in [4] und [8] zwei heuristische Algorithmen zur näherungsweise Lösung von (11.23) vorgeschlagen, der Vorwärts- und Rückwärtsalgorithmus.

Heuristische Algorithmen

Zum Zweck der einfacheren Darstellung definieren wir

$$c_{ij} := \begin{cases} \hat{c}_r(\xi^i, \xi^j), & \text{falls } d = \zeta_r, \\ |\xi^i - \xi^j|^r, & \text{falls } d = \ell_r. \end{cases}$$

Die Basis für den Vorwärtsalgorithmus der Szenarioreduktion bildet die einfache Gestalt des Problems im Spezialfall $n = 1$. Es läßt sich kurz in der Form

$$\min_{u \in \{1, \dots, N\}} \sum_{\substack{j=1 \\ j \neq u}}^N p_j c_{uj}$$

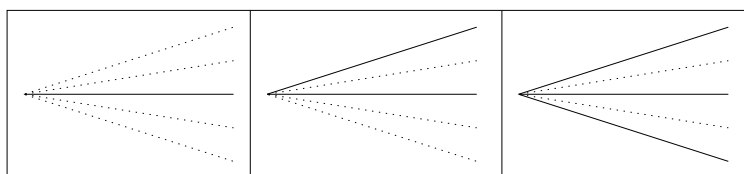
darstellen. Wird das Minimum in u^* angenommen, erhält man mit $J = \{1, \dots, N\} \setminus \{u^*\}$ eine Lösung von (11.23) im Fall $n = 1$. Damit ist das erste Element des Trägers von Q_J gefunden. Unter Verwendung der additiven Gestalt von D_J in (11.19) bzw. (11.21) bestimmt man bei fixiertem ersten Element analog das zweite Element des Trägers von Q_J und setzt diesen Prozess bis $\#J = N - n$ fort. Dies führt zu folgender Heuristik zur approximativen Lösung von (11.23) für beliebiges n .

Algorithmus 11.1. (Vorwärtsalgorithmus Szenarioreduktion)

Step 0: $J^{[0]} := \{1, \dots, N\}$.

Step k: $u_k \in \arg \min_{u \in J^{[k-1]}} \sum_{j \in J^{[k-1]} \setminus \{u\}} p_j \min_{i \in J^{[k-1]} \setminus \{u\}} c_{ij}$,
 $J^{[k]} := J^{[k-1]} \setminus \{u_k\}$.

Step n+1: Optimale Neuverteilung mit $J := J^{[n]}$.



Der Algorithmus kann in folgender Weise interpretiert werden. Der Szenarioindex u_k im Schritt k wird so gewählt, dass der Wert des Functionals $D_{J^{[k]}}$ minimiert wird, d.h.

$$D_{J^{[k-1]} \setminus \{u_k\}} = \min_{u \in J^{[k-1]}} D_{J^{[k-1]} \setminus \{u\}}.$$

Die Menge der verbleibenden Szenarien und damit der Träger der Wahrscheinlichkeitsverteilung Q_J ist am Ende durch die Indexmenge $\{u_1, \dots, u_n\}$ bestimmt.

Die Idee des Rückwärtsalgorithmus der Szenarioreduktion basiert auf dem zweiten Spezialfall des Problems (11.23). Dieser entsteht für $n = N - 1$ und besitzt die Form

$$\min_{l \in \{1, \dots, N\}} p_l \min_{i \neq l} c_{il}.$$

Wird das Minimum in l^* angenommen, ergibt sich $J = \{l^*\}$ als eine Lösung von (11.23) im Fall $n = N - 1$. Auch dieser Spezialfall läßt sich schrittweise zu einem heuristischen Algorithmus zur approximativen Lösung von (11.23) für beliebige n fortsetzen.

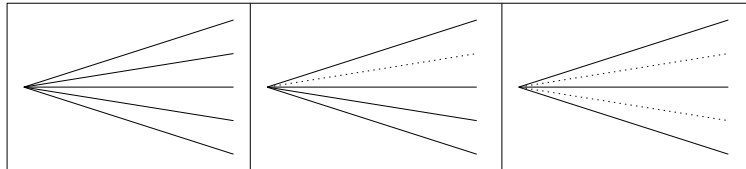
Algorithmus 11.2. (Rückwärtsalgorithmus Szenarioreduktion)

Step 0: $J^{[0]} := \emptyset$.

$$\text{Step } k: \quad l_k \in \arg \min_{l \notin J^{[k-1]}} \sum_{j \in J^{[k-1]} \cup \{l\}} p_j \min_{i \notin J^{[k-1]} \cup \{l\}} c_{ij},$$

$$J^{[k]} := J^{[k-1]} \cup \{l_k\}.$$

Step N-n+1: Optimale Neuverteilung mit $J := J^{[N-n]}$.



Die Interpretation des Algorithmus ist ähnlich wie zuvor. Der Szenarioindex l_k im Schritt k wird erneut so gewählt, dass der Wert des Funktionals $D_{J^{[k]}}$ minimiert wird, d.h.

$$D_{J^{[k-1]} \cup \{l_k\}} = \min_{l \notin J^{[k-1]}} D_{J^{[k-1]} \cup \{l\}}.$$

Die Menge der verbleibenden Szenarien nach Ablauf des Algorithmus und damit der Träger des Maßes Q_J ist hier durch die Indexmenge $\{1, \dots, N\} \setminus \{l_1, \dots, l_{N-n}\}$ bestimmt.

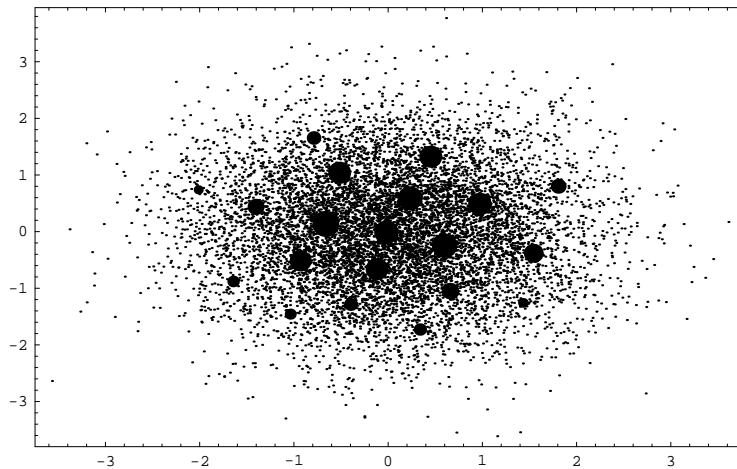


Abbildung 11.1: Ergebnis der Szenarioreduktion am Beispiel der zweidimensionalen Standardnormalverteilung. Ausgehend von 10 000 simulierten Szenarien, wurde mit dem Vorwärtsalgorithmus eine Szenarioreduktion auf 20 Szenarien bzgl. ℓ_2 durchgeführt. Dabei sind die Durchmesser der Kreise für die verbleibenden 20 repräsentativen Szenarien proportional zu ihren neuen Wahrscheinlichkeiten.

11.3.2 Anwendung: Szenariobaumkonstruktion für mehrstufige Modelle

Die vorgestellten Techniken zur Szenarioreduktion können als Grundlage für die Konstruktion von Szenariobäumen im Fall mehrstufiger stochastischer Optimierungsprobleme

me genutzt werden. In diesem Fall modellieren die Szenarien einen Prozess sukzessiver Beobachtungen und Entscheidungen und besitzen deshalb Baumstruktur. Diese beschreibt den fortschreitenden Informationszuwachs bzw. die zunehmenden Entscheidungsvarianten innerhalb des Optimierungszeitraums.

Jedoch weisen Szenarien, welche aus historischen Daten oder durch Simulation aus angepassten Zeitreihen- oder Regressionsmodellen erhält, im Allgemeinen keine Baumstruktur auf. Dies erfordert eine Methodik, die aus diesen Ausgangsszenarien einen geeigneten Szenariobaum ξ_{tr} generiert, der den stochastischen Prozess ξ approximiert.

Im Folgenden bezeichnen wir den stochastischen Prozess, der aus den Ausgangsszenarien $\xi^i = (\xi_1^i, \dots, \xi_T^i)$ mit den Wahrscheinlichkeiten p_i , $i = 1, \dots, N$, besteht, mit $\hat{\xi}$. Um eine deterministische erste Stufe zu realisieren, sei zusätzlich die Bedingung $\xi_1^1 = \dots = \xi_1^N =: \xi_1$ erfüllt (siehe Abbildung 11.2 zur Veranschaulichung).

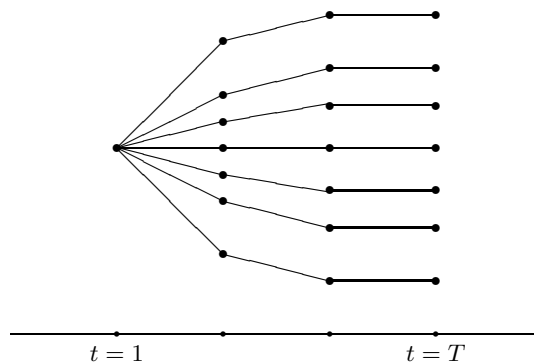


Abbildung 11.2: Prozess $\hat{\xi}$ mit Ausgangsszenarien für $T = 4$ und $N = 7$

Ein naheliegendes Konzept zur Erzeugung der Baumstruktur besteht in der rekursiven Anwendung der Szenarioreduktion unter Berücksichtigung variabler Zeitintervalle. Zieht man dabei nacheinander Zeitintervalle der Form $\{1, \dots, t\}$ für entweder wachsende oder fallende Zeitparameter t in Betracht, lassen sich erneut zwei Algorithmen ableiten, die je nach Ausrichtung als Vorwärts- und Rückwärtsalgorithmus in [9] und später in [11] ausführlich diskutiert werden. Beide Vorgehensweisen ermöglichen die schrittweise Rekonstruktion der implizit vorliegenden Informationsstruktur (vgl. [7]). Im Folgenden wird der Vorwärtsalgorithmus vorgestellt.

Vorwärtsalgorithmus zur Baumkonstruktion

Die Strategie des Vorwärtsalgorithmus basiert auf einer rekursiven Szenarioreduktion für wachsende Zeiträume $\{1, \dots, t\}$ für $t = 2, \dots, T$. Das Resultat sind aufsteigende Partitionen der Indexmenge $I := \{1, \dots, N\}$, die zum Zeitpunkt t die Gestalt

$$C_t := \{C_t^1, \dots, C_t^{k_t}\}, \quad k_t \in \mathbb{N}.$$

annehmen. Die Partitions Mengen erfüllen dabei für jedes t die Partitionseigenschaft

$$C_t^k \cap C_t^{k'} = \emptyset, \quad \text{falls } k \neq k', \quad \text{und} \quad \bigcup_{k=1}^{k_t} C_t^k = I.$$

Die Elemente der Partition C_t werden mitunter auch Szenario-Cluster genannt. Der nun folgende Algorithmus erlaubt die Konstruktion eines Szenariobaumprozesses ξ_{tr} , dessen Gestalt über sog. Toleranzparameter ε_t , $t = 2, \dots, T$, innerhalb des Algorithmus gesteuert werden kann.

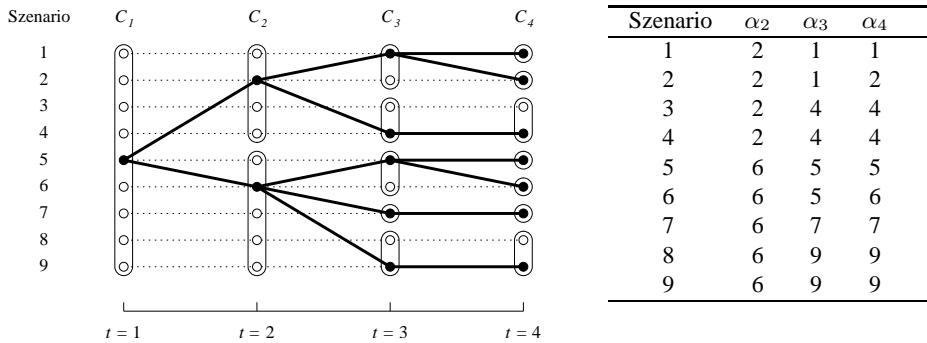


Abbildung 11.3: Schematische Darstellung der Clusterung durch den Vorwärtsalgorithmus 11.3 bei der Szenariobaumkonstruktion (links) und Veranschaulichung der zugehörigen Zuordnungsabbildungen α_t (rechts).

Algorithmus 11.3. (Vorwärtsalgorithmus Baumkonstruktion)

[Initialisierung]

Setze $C_1 = \{I\}$ und $t := 2$.

[Cluster-Berechnung]

Es sei $C_{t-1} = \{C_{t-1}^1, \dots, C_{t-1}^{k_{t-1}}\}$. Führe für jede Szenarioteilmenge $\{\xi_t^i\}_{i \in C_{t-1}^k}$ separat für jedes $k \in \{1, \dots, k_{t-1}\}$ eine Szenarioreduktion durch (d.h. nur bezüglich der t -ten Komponente). Definiere Indexmengen J_t^k und I_t^k der jeweils gestrichenen bzw. verbleibenden Szenarien sowie partielle Zuordnungen $i_t^k : J_t^k \rightarrow I_t^k$, so dass

$$i_t^k(j) \in \arg \min_{i \in I_t^k} |\xi_t^i - \xi_t^j|^r, \quad j \in J_t^k,$$

entsprechend der Szenarioreduktion (vgl. Abschnitt 11.3.1). Definiere eine Abbildung $\alpha_t : I \rightarrow I$, so dass

$$\alpha_t(j) = \begin{cases} i_t^k(j), & j \in J_t^k \text{ für ein } k = 1, \dots, k_{t-1}, \\ j, & \text{sonst.} \end{cases} \quad (11.24)$$

Die Partition zum Zeitpunkt t wird damit durch

$$C_t := \{ \alpha_t^{-1}(i) \mid i \in I_t^k, k = 1, \dots, k_{t-1} \}$$

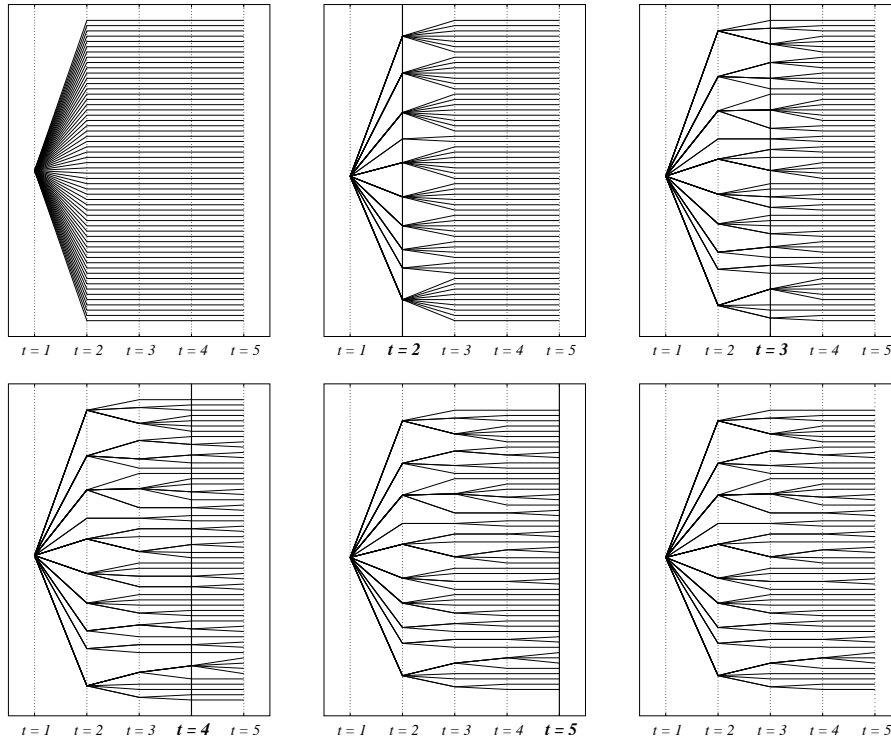


Abbildung 11.4: Illustration der schrittweisen Szenariobaumkonstruktion von Algorithmus 11.3 an einem Beispiel beginnend mit $N = 58$ (individuellen) Szenarien und $T = 5$.

definiert, was einer Verfeinerung der Partition C_{t-1} entspricht. Im Fall $t < T$ führe eine weitere Cluster-Berechnung mit $t := t + 1$ durch. Andernfalls gehe zur abschließenden Baumgenerierung.

[Baumgenerierung]

Entsprechend der Partition C_T und der Abbildungen (11.24) definiere den Szenariobaumprozess ξ_{tr} über die Szenarien

$$\xi_{\text{tr}}^k = \left(\xi_1^*, \xi_2^{\alpha_2(i)}, \dots, \xi_t^{\alpha_t(i)}, \dots, \xi_T^{\alpha_T(i)} \right) \quad \text{für ein } i \in C_T^k,$$

versehen mit den Wahrscheinlichkeiten $q_k := \sum_{i \in C_T^k} p_i$ für $k = 1, \dots, k_T$.

Für die Umsetzung der Cluster-Berechnung innerhalb des Algorithmus können beide Algorithmen der Szenarioreduktion aus Abschnitt 11.3.1 herangezogen werden. Der Fehler im Schritt t läßt sich durch

$$\text{err}_t := \sum_{k=1}^{k_{t-1}} \sum_{j \in J_t^k} p_j \min_{i \in I_t^k} |\xi_t^i - \xi_t^j|^r$$

bestimmen. Analog zu [7, Theorem 6.4] kann damit der gesamte Approximationsfehler

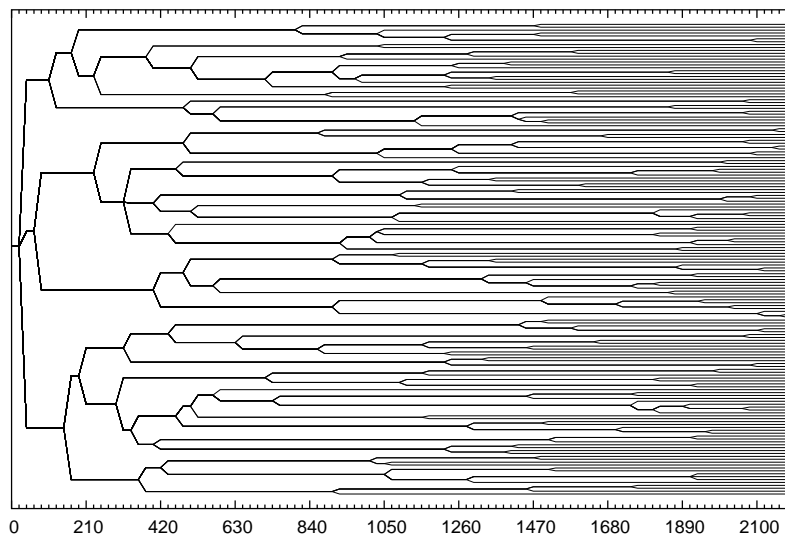


Abbildung 11.5: Last-Zufluss Szenariobaum mit 163 Szenarien über einem zweijährigen Zeithorizont

durch

$$\ell_r(P_{\hat{\xi}}, P_{\xi_{\text{tr}}}) \leq \left(\sum_{t=2}^T \text{err}_t \right)^{\frac{1}{r}}$$

abgeschätzt werden, wobei $P_{\hat{\xi}}$ und $P_{\xi_{\text{tr}}}$ die Wahrscheinlichkeitsverteilungen von $\hat{\xi}$ bzw. ξ_{tr} bezeichnen. Die letzte Abschätzung erlaubt die Steuerung des Konstruktionsprozesses durch die vorgegebenen Toleranzen ε_t für err_t und jedes $t = 2, \dots, T$.

Der Vorwärtsalgorithmus wurde im Rahmen verschiedener energiewirtschaftlicher Modelle angewendet. Zum Beispiel wurden Last-Spotpreis Szenariobäume für einen jährlichen Zeithorizont und Wind-Spotpreis Bäume für mehrere Monate und jeweils stündlicher Diskretisierung erzeugt. Abbildung 11.5 zeigt einen Last-Zufluss Szenariobaum für einen zweijährigen Zeithorizont mit drei täglichen Diskretisierungspunkten, bei dem der Vorwärtsalgorithmus ausgehend von $N = 456$ gleichwahrscheinlichen Szenarien bei $T = 2184$ Diskretisierungspunkten und nur wöchentlich erlaubten Verzweigungen einen Baum mit $k_T = 163$ Szenarien erzeugt hat. Dabei wurden die Toleranzen ε_t linear monoton fallend bzgl. t gewählt.

11.4 Gemischt-ganzzahlige zweistufige stochastische Programme

11.4.1 Optimale Neuverteilung

In diesem Abschnitt entwickeln wir einen Ansatz zur Lösung des Problems der optimalen Szenarioreduktion (11.17) für den Abstand $d = d_{\lambda, 2, \mathcal{B}_{\text{rect}}}$ (vgl. (11.12)) gegeben durch

$$d_{\lambda, 2, \mathcal{B}_{\text{rect}}}(P, Q) := \lambda \alpha_{\mathcal{B}_{\text{rect}}}(P, Q) + (1 - \lambda) \zeta_2(P, Q). \quad (11.25)$$

Hierbei ist ζ_2 die Fortet-Mourier Metrik (11.10) für $r = 2$ und $\alpha_{\mathcal{B}_{\text{rect}}}$ die Rechtecksdiskrepanz

$$\alpha_{\mathcal{B}_{\text{rect}}}(P, Q) = \sup_{B \in \mathcal{B}_{\text{rect}}} |P(B) - Q(B)|.$$

Im Folgenden benutzen wir der Einfachheit halber die Schreibweise $\mathbf{d}_\lambda = d_{\lambda, 2, \mathcal{B}_{\text{rect}}}$. Die in Abschnitt 11.3.1 vorgestellten Verfahren beruhen auf der wiederholten Evaluation des Wertes

$$D_J := \min \left\{ \mathbf{d}_\lambda(P, Q_J) : q_i \geq 0, i \notin J, \sum_{i \notin J} q_i = 1 \right\}$$

für verschiedene Mengen $J \subset \{1, \dots, N\}$. Im Gegensatz zu den in Abschnitt 11.3.1 betrachteten Fällen, gibt es für den Fall $d = \mathbf{d}_\lambda$ keine explizite, leicht auszuwertende Darstellung von D_J . Im Folgenden zeigen wir, wie eine numerische Evaluation von D_J und damit eine Lösung des Szenarioreduktionsproblems für $d = \mathbf{d}_\lambda$ möglich ist. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $J = \{n+1, \dots, N\}$, d.h. der Träger des Maßes Q_J ist die Menge $\{\xi^1, \dots, \xi^n\}$.

Wegen der Definition von \mathbf{d}_λ hat D_J die Form

$$D_J = \min \left\{ \lambda \alpha_{\mathcal{B}_{\text{rect}}}(P, Q_J) + (1 - \lambda) \zeta_2(P, Q_J) : q_i \geq 0, i \notin J, \sum_{j \notin J} q_j = 1 \right\} \quad (11.26)$$

Im Folgenden zeigen wir, wie das durch (11.26) gegebene Minimierungsproblem als lineares Programm geschrieben werden kann. Zunächst scheint der Term $\alpha_{\mathcal{B}_{\text{rect}}}(P, Q_J)$ nicht direkt zugänglich. Da die Träger der Maße P und Q_J jedoch in der endlichen Menge $\{\xi^1, \dots, \xi^N\}$ enthalten sind, können wir für jedes Rechteck $B \in \mathcal{B}_{\text{rect}}$ die dazugehörige Indexmenge $I(B)$ durch

$$I(B) := \{i \in \{1, \dots, N\} : \xi^i \in B\}$$

definieren und erhalten

$$|P(B) - Q_J(B)| = \left| \sum_{i \in I(B)} p_i - \sum_{j \in I(B) \cap \{1, \dots, n\}} q_j \right|. \quad (11.27)$$

Definieren wir nun das System der Indexmengen $\mathcal{I}_{\mathcal{B}_{\text{rect}}}$ durch

$$\mathcal{I}_{\mathcal{B}_{\text{rect}}} := \{I(B) : B \in \mathcal{B}_{\text{rect}}\},$$

so kann die Diskrepanz zwischen P und Q_J wie folgt geschrieben werden:

$$\alpha_{\mathcal{B}_{\text{rect}}}(P, Q_J) = \max_{I \in \mathcal{I}_{\mathcal{B}_{\text{rect}}}} \left| \sum_{i \in I} p_i - \sum_{j \in I \cap \{1, \dots, n\}} q_j \right| \quad (11.28)$$

$$= \min \left\{ t_\alpha \mid \begin{array}{l} -\sum_{j \in I \cap \{1, \dots, n\}} q_j \leq t_\alpha - \sum_{i \in I} p_i, I \in \mathcal{I}_{\mathcal{B}_{\text{rect}}} \\ \sum_{j \in I \cap \{1, \dots, n\}} q_j \leq t_\alpha + \sum_{i \in I} p_i, I \in \mathcal{I}_{\mathcal{B}_{\text{rect}}} \end{array} \right\}. \quad (11.29)$$

Insbesondere besteht das System $\mathcal{I}_{\mathcal{B}_{\text{rect}}}$ aus nur endlich vielen Indexmengen. Aus der Darstellung (11.29) folgt, dass die Bestimmung von $\alpha_{\mathcal{B}_{\text{rect}}}(P, Q_J)$ (für festes q) ein lineares Optimierungsproblem ist. Ebenso ist die Minimierung der Diskrepanz über dem Standard-Simplex

$$S := \left\{ q \in \mathbb{R}^n : q_i \geq 0, i = 1, \dots, n, \sum_{i=1}^n q_i = 1 \right\}$$

ein lineares Programm (in t_α und q).

Allerdings kann die Menge $\mathcal{I}_{\mathcal{B}_{\text{rect}}}$ sehr groß sein und bis zu 2^N Elemente beinhalten. Die entsprechende Anzahl von Nebenbedingungen kann daher eine numerische Lösung von (11.29) sehr schwierig machen. Betrachtet man allerdings die Ungleichungen in (11.29) genauer, so stellt man fest, dass für charakteristische Indexmengen, deren Durchschnitte mit der Menge $\{1, \dots, n\}$ übereinstimmen, die linken Seiten der Ungleichungen identisch sind. Daher genügt es, nur jeweils diejenige Ungleichung mit der kleinsten rechten Seite zu berücksichtigen, die es zu einer bestimmten linken Seite gibt. Um dies zu formalisieren, betrachten wir das folgende *reduzierte System der Indexmengen*

$$\mathcal{I}_{\mathcal{B}_{\text{rect}}}^* := \{I(B) \cap \{1, \dots, n\} : B \in \mathcal{B}_{\text{rect}}\}.$$

Jede Menge $I^* \in \mathcal{I}_{\mathcal{B}_{\text{rect}}}^*$ entspricht einer linken Seite einer Ungleichung in (11.29). Die entsprechenden kleinsten rechten Seiten der Ungleichungen ergeben sich aus den folgenden Größen:

$$\begin{aligned} \gamma^{I^*} &:= \max \left\{ \sum_{i \in I} p_i \mid I \in \mathcal{I}_{\mathcal{B}_{\text{rect}}}, I \cap \{1, \dots, n\} = I^* \right\} \\ \gamma_{I^*} &:= \min \left\{ \sum_{i \in I} p_i \mid I \in \mathcal{I}_{\mathcal{B}_{\text{rect}}}, I \cap \{1, \dots, n\} = I^* \right\} \end{aligned}$$

Das Optimierungsproblem (11.29) hat damit die folgende Form:

$$\alpha_{\mathcal{B}_{\text{rect}}}(P, Q_J) = \min \left\{ t_\alpha \mid \begin{array}{l} -\sum_{j \in I^*} q_j \leq t_\alpha - \gamma^{I^*}, I^* \in \mathcal{I}_{\mathcal{B}_{\text{rect}}}^* \\ \sum_{j \in I^*} q_j \leq t_\alpha + \gamma_{I^*}, I^* \in \mathcal{I}_{\mathcal{B}_{\text{rect}}}^* \end{array} \right\}. \quad (11.30)$$

Da die Anzahl der Elemente in $\mathcal{I}_{\mathcal{B}_{\text{rect}}}^*$ höchstens 2^n ist, kann der Übergang von (11.29) nach (11.30) tatsächlich die Anzahl von Nebenbedingungen deutlich reduzieren. Andererseits müssen nun das System $\mathcal{I}_{\mathcal{B}_{\text{rect}}}^*$ und die Koeffizienten γ^{I^*} und γ_{I^*} für $I^* \in \mathcal{I}_{\mathcal{B}_{\text{rect}}}^*$ bestimmt werden können. Dies erfolgt im nächsten Abschnitt.

Der Wert D_J ist dann der Optimalwert eines linearen Optimierungsproblems:

$$D_J = \min \left\{ \lambda t_\alpha + (1 - \lambda) t_\zeta \mid \begin{array}{l} t_\alpha, t_\zeta \geq 0, q_i \geq 0, \sum_{i=1}^n q_i = 1, \\ \eta_{i,j} \geq 0, i = 1, \dots, N, j = 1, \dots, n, \\ t_\zeta \geq \sum_{j=1, \dots, n} \hat{c}_2^{i,j} \eta_{i,j}, \\ \sum_{j=1}^n \eta_{i,j} = p_i, i = 1, \dots, N, \\ \sum_{i=1}^N \eta_{i,j} = q_j, j = 1, \dots, n, \\ -\sum_{j \in I^*} q_j \leq t_\alpha - \gamma^{I^*}, I^* \in \mathcal{I}_{\mathcal{B}_{\text{rect}}}^* \\ \sum_{j \in I^*} q_j \leq t_\alpha + \gamma_{I^*}, I^* \in \mathcal{I}_{\mathcal{B}_{\text{rect}}}^* \end{array} \right\} \quad (11.31)$$

11.4.2 Bestimmung der Koeffizienten

In diesem Abschnitt zeigen wir, wie das System $\mathcal{I}_{\mathcal{B}_{\text{rect}}}^*$ und die Koeffizienten $\gamma^{I^*}, \gamma_{I^*}$ für $I^* \in \mathcal{I}_{\mathcal{B}_{\text{rect}}}^*$, deren Kenntnis für die Lösung von (11.31) notwendig ist, bestimmt werden können.

In [13] wurde gezeigt, dass es genügt, dazu eine bestimmte Menge von *stützenden* Rechtecken zu untersuchen. Ein Rechteck ist stützend, wenn jede seiner Seitenflächen in ihrem relativen Inneren einen Punkt aus $\{\xi^1, \dots, \xi^n\}$ enthält, siehe auch Abbildung 11.6. Um dies zu formalisieren, betrachten wir die folgenden Mengen:

$$R_j := \{\xi_j^i, i = 1, \dots, n\} \cup \{-\infty, +\infty\}, j = 1, \dots, s.$$

Zu jedem $\underline{b}_j \in R_j$ finden wir einen Index $\underline{i}_j \in \{1, \dots, n, -\infty, +\infty\}$, so dass $\xi_j^{\underline{i}_j} = \underline{b}_j$ falls $|\underline{b}_j| < \infty$, und $\underline{i}_j = \underline{b}_j$ sonst.

Definition 11.4. Ein Rechteck $B = \times_{j=1}^s [\underline{b}_j, \bar{b}_j] \in \mathcal{B}_{\text{rect}}$ heißt *stützend*, wenn die folgenden Bedingungen erfüllt sind:

- (i) Es gilt $\underline{b}_j, \bar{b}_j \in R_j$ für $j = 1, \dots, s$.
- (ii) Es seien $\underline{i}_j, \bar{i}_j$ die Indizes zu $\underline{b}_j, \bar{b}_j \in R_j$. Dann gilt für alle $j, l \in \{1, \dots, s\}, j \neq l$

$$\begin{aligned} \underline{b}_j < \xi_j^{\underline{i}_l} < \bar{b}_j & \text{ falls } \underline{i}_l \neq \pm\infty, \\ \underline{b}_j < \xi_j^{\bar{i}_l} < \bar{b}_j & \text{ falls } \bar{i}_l \neq \pm\infty. \end{aligned}$$

Die Menge aller stützenden Rechtecke bezeichnen wir mit \mathcal{P} .

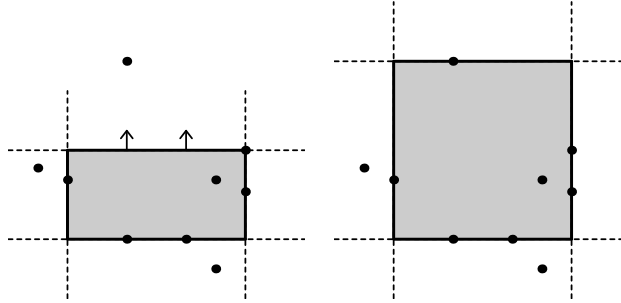


Abbildung 11.6: Nicht stützendes Rechteck (links) und stützendes Rechteck (rechts). Die Punkte stehen für die verbleibenden Szenarien ξ^1, \dots, ξ^n .

Wir beschreiben jetzt, wie $\mathcal{I}_{\mathcal{B}_{\text{rect}}}^*$ und $\gamma^{I^*}, \gamma_{I^*}$ mit Hilfe des Systems \mathcal{P} bestimmt werden können. Entsprechend den Propositionen 1 und 2 sowie Korollar 1 in [13] gelten die folgenden Darstellungen:

$$\mathcal{I}_{\mathcal{B}_{\text{rect}}}^* = \bigcup_{B \in \mathcal{P}} \{I^* \subseteq \{1, \dots, n\} : \cup_{j \in I^*} \{\xi^j\} = \{\xi^1, \dots, \xi^n\} \cap \text{int } B\}$$

und für jede reduzierte Indexmenge $I^* \in \mathcal{I}_{\mathcal{B}_{\text{rect}}}^*$ überdies

$$\gamma^{I^*} = \max \left\{ \mathbb{P}(\text{int } B) : B \in \mathcal{P}, \cup_{j \in I^*} \{\xi^j\} = \{\xi^1, \dots, \xi^n\} \cap \text{int } B \right\},$$

$$\gamma^{I^*} = \sum_{i \in \underline{I}} p_i \quad \text{mit} \quad \underline{I} := \{i \in \{1, \dots, N\} : \min_{j \in I^*} \xi_l^j \leq \xi_l^i \leq \max_{j \in I^*} \xi_l^j, l = 1, \dots, s\}.$$

Dies zeigt, dass man das System $\mathcal{I}_{\mathcal{B}_{\text{rect}}}^*$ bestimmen kann, indem man nacheinander alle stützenden Rechtecke $B \in \mathcal{P}$ untersucht; der Durchschnitt $\{\xi^1, \dots, \xi^n\} \cap \text{int } B$ ergibt jeweils eine Menge $I^* \in \mathcal{I}_{\mathcal{B}_{\text{rect}}}^*$. Der Wert γ^{I^*} ist dabei die maximale Wahrscheinlichkeit einer solchen Menge B die J erzeugt.

Da es $\binom{n+2}{2}^s$ viele Rechtecke mit Eigenschaft (i) von Definition 11.4 gibt, scheint es nicht vorteilhaft, diese alle auf Eigenschaft (ii) zu überprüfen. Statt dessen konstruieren wir für $l = 1, \dots, s$, rekursiv l -dimensionale Rechtecke mit Eigenschaft (i), wobei wir in jedem Schritt Eigenschaft (ii) von Definition 11.4 überprüfen. Die geschieht in der (rekursiv definierten) Funktion **Iterate** des Algorithmus 11.5.

Numerische Resultate zeigen, dass die Laufzeit von Algorithmus 11.5 zum größten Teil durch die Berechnung der Parameter bestimmt wird, wogegen das Lösen des linearen Optimierungsproblems selbst kaum ins Gewicht fällt. Weiter hängt die Laufzeit etwa linear von N ab, und von n und s durch die Beziehung $\binom{n+1}{2}^s$. Während damit eine große Anzahl von Eingangsszenarien kein Problem für den Algorithmus darstellt, sind bei beschränkter Laufzeit nur moderate Werte von n und s realistisch.

11.4.3 Bestimmen optimaler Szenarien

Da nun ein Verfahren zur Berechnung von D_J zur Verfügung steht, lassen sich prinzipiell die in Abschnitt 11.3.1 vorgestellten heuristischen Vorwärts- und Rückwärtsalgorithmen verwenden, um eine möglichst gute Indexmenge J zu ermitteln. Dazu wird der entsprechende D_J Term im k -ten Schritt minimiert, d.h. für den Vorwärtsalgorithmus wird u_k bestimmt durch $u_k \in \arg \min_{u \in J^{[k-1]}} D_{J^{[k-1]} \setminus \{u\}}$, für den Rückwärtsalgorithmus l_k durch $l_k \in \arg \min_{l \notin J^{[k-1]}} D_{J^{[k-1]} \cup \{l\}}$.

Da im Fall $d = \mathbf{d}_\lambda$ die Auswertung des Funktionals D_J jedoch numerisch aufwändiger ist, erhöhen sich die Rechenzeiten deutlich. Da insbesondere der Aufwand für eine Lösung des inneren Problems sich annähernd wie $\binom{n+1}{2}^s$ verhält, ist der Rückwärtsalgorithmus für größere Werte von N nicht geeignet. Daher erscheint es notwendig, Heuristiken für das kombinatorische Optimierungsproblem zu entwickeln, die nur wenige Berechnungen von D_J verlangen. Eine solche Heuristik kann beispielsweise auf Quasi-Monte Carlo Methoden (vgl. [20]) beruhen. Quasi-Monte Carlo Methoden liefern Folgen von (gleichgewichteten) Punkten, die die Gleichverteilung auf dem Einheitswürfel $[0, 1]^s$ approximieren. A priori ist nicht klar, wie diese zur Approximation von *beliebigen* Verteilungen in \mathbb{R}^s verwendet werden können. Allerdings ist dies möglich, wenn das Maß P eine Dichte besitzt und die s Komponenten der Zufallsvariablen *statistisch unabhängig* sind. Dann lassen sich die aus Quasi-Monte Carlo gewonnenen Punkte mittels der inversen Verteilungsfunktionen der Marginalverteilungen geeignet transformieren, vgl. [16]. In dem in diesem Kapitel betrachteten Fall *diskreter* Verteilungen ergäbe der so transformierte Punkt nicht unbedingt einen Punkt des Trägers $\{\xi^1, \dots, \xi^N\}$ des Maßes P . Daher suchen wir also einen solchen Punkt $\xi^i, i \in \{1, \dots, N\}$, der möglichst nahe liegt. Dies ist formalisiert durch Algorithmus 11.6. Abbildung 11.7 zeigt das Ergebnis

Algorithmus 11.5. (Berechnung von D_J / Optimale Neuverteilung bezüglich \mathbf{d}_λ)

0. Setze $\mathcal{I}_{\mathcal{B}_{\text{rect}}}^* = \{\emptyset\}$, $\gamma^{I^*} = 0$ für alle $I^* \subseteq \{1, \dots, n\}$.
1. Setze $I^* = \{1, \dots, n\}$ und $\bar{I} = \{1, \dots, N\}$.
2. Rufe Funktion **Iterate**(0, 0, 0, I^* , \bar{I}) auf.
3. Löse das lineare Programm (11.31).

Funktion **Iterate**(l , $(\underline{i}_j)_{j=1}^l, (\bar{i}_j)_{j=1}^l, I^*, \bar{I}$):

Falls $l = s$:

Rufe Funktion **UpdateData**((\underline{i}_j)_{j=1}^l, (\bar{i}_j)_{j=1}^l, J , \bar{I}) auf.

Funktionsende.

Setze $l = l + 1$.

Für $\underline{i}_l = 1, \dots, n + 1$ und $\bar{i}_l = 1, \dots, n + 1$:

Falls Bedingung (ii) für alle $j \in \{1, \dots, l - 1\}$, $i = l$,
und $i \in \{1, \dots, l - 1\}$, $j = l$ erfüllt ist:

$$\text{Setze } \underline{b}_l = \begin{cases} \xi_l^{\underline{i}_l} & \text{falls } \underline{i}_l \in \{1, \dots, n\}, \\ -\infty & \text{falls } \underline{i}_l = n + 1, \end{cases} \quad \bar{b}_l =$$

$$\begin{cases} \xi_l^{\bar{i}_l} & \text{falls } \bar{i}_l \in \{1, \dots, n\}, \\ +\infty & \text{falls } \bar{i}_l = n + 1. \end{cases}$$

Setze $\bar{I} = \bar{I} \cap \{i \in \{1, \dots, N\} : \underline{b}_l < \xi_l^i < \bar{b}_l\}$.

Falls $\bar{I} \neq \emptyset$:

Setze $I^* = I^* \cap \bar{I}$.

Rufe Funktion **Iterate**(l , $(\underline{i}_j)_{j=1}^l, (\bar{i}_j)_{j=1}^l, I^*, \bar{I}$) auf.

Funktionsende.

Funktion **UpdateData**((\underline{i}_j)_{j=1}^s, (\bar{i}_j)_{j=1}^s, I^* , \bar{I}):

Falls $I^* \notin \mathcal{I}_{\mathcal{B}_{\text{rect}}}^*$:

Setze $\mathcal{I}_{\mathcal{B}_{\text{rect}}}^* = \mathcal{I}_{\mathcal{B}_{\text{rect}}}^* \cup \{I^*\}$.

Setze $\gamma_{I^*} = \sum_{i \in \underline{I}} p_i$ mit $\underline{I} = \{i \in \bar{I} : \min_{j \in J} \xi_l^j \leq \xi_l^i \leq \max_{j \in J} \xi_l^j\}$, für $l = 1, \dots, s$.

Setze $\gamma^{I^*} = \max\{\gamma^{I^*}, \sum_{i \in \bar{I}} p_i\}$.

Funktionsende.

einer solchen auf Quasi-Monte Carlo basierten Heuristik. Die linke Seite von Abbildung 11.8 zeigt die Diskrepanzen zwischen $n = 1, \dots, 50$ Quasi-Monte Carlo Punkten und den 1 000 gleichverteilten Punkten sowie die benötigte Zeit von Algorithmus 11.6. Die rechte Seite von Abbildung 11.8 zeigt die aus dem Vorwärtsalgorithmus resultierenden Abstände $\alpha_{\mathcal{B}_{\text{rect}}}$ und ζ_2 für verschiedene Werte von $\lambda \in [0, 1]$, dabei wurden $n = 10$ aus $N = 100$ Punkten einer Gleichverteilung auf $[0, 1]^2$ ausgewählt.

11.5 Zusammenfassung

Nach einer Einführung in zwei- und mehrstufige stochastische Optimierungsprobleme mit energiewirtschaftlichem Hintergrund wird das Verhalten der optimalen Verluste (bzw. Gewinne) und der Menge der optimalen Erststufenentscheidungen bei Approximation der zugrunde liegenden Wahrscheinlichkeitsverteilung diskutiert. Es werden Abstände

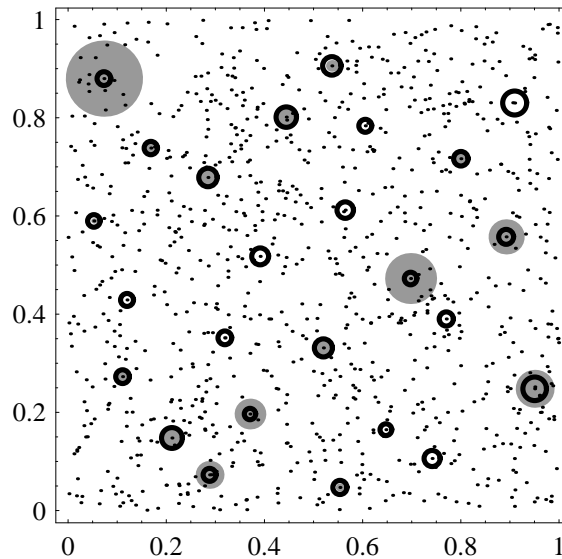


Abbildung 11.7: $N = 1000$ Punkte, simuliert von einer Gleichverteilung auf $[0, 1]^2$. Dann wurden die ersten $n = 25$ Punkte der *Halton-Folge* zu den Basen 2 und 3 berechnet und mittels Algorithmus 11.6 transformiert. Die Wahrscheinlichkeiten der n Punkte wurde bezüglich des Abstandes \mathbf{d}_λ optimiert, einmal mit $\lambda = 1$ (graue Kreise), und einmal mit $\lambda = 0.9$ (schwarze Kreise). Dabei sind die Durchmesser der Kreise proportional zu den entsprechenden Wahrscheinlichkeiten.

Algorithmus 11.6. (Transformation einer gleichverteilten Folge von Punkten)

Bestimme die inversen Verteilungsfunktionen $F_j^{-1}(t)$ für $j = 1, \dots, s$, gegeben durch

$$F_j^{-1}(t) := \inf\{z \in \mathbb{R} \mid \sum_{i:\xi_i^j \leq t} p_i \geq t\} \text{ für } t \in [0, 1].$$

Bestimme n Quasi-Monte Carlo Punkte $z^k = (z_1^k, \dots, z_s^k)$, $k = 1, \dots, n$.

Setze $J := \{1, \dots, N\}$.

Für $k = 1, \dots, n$:

Setze $y^k := (F_1^{-1}(z_1^k), \dots, F_s^{-1}(z_s^k))$.

Setze $u_k \in \arg \min_{u \in \{1, \dots, N\}} \|\xi^u - y^k\|$.

Setze $J = J \setminus u_k$.

von Wahrscheinlichkeitsverteilungen identifiziert bzgl. derer sich die stochastischen Optimierungsprobleme stabil verhalten.

Für solche metrischen Abstände d wurde in (11.17) das Problem der optimalen Szenario-reduktion formuliert. Dieses besteht darin, eine Wahrscheinlichkeitsverteilung mit einer großen Anzahl N von Szenarien durch eine Verteilung mit einer kleineren Anzahl n bzgl. d bestmöglich zu approximieren. Dieses Problem läßt sich für die Fälle, dass d durch ein Transport-Funktional gegeben ist, auf ein rein kombinatorisches Optimierungsproblem vom n -median Typ (vgl. (11.23)) reduzieren. Für letztere wurden Heuristiken zur

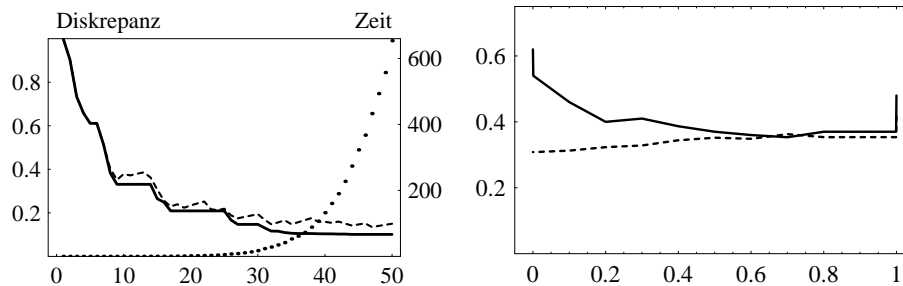


Abbildung 11.8: Linke Seite: Abstand d_λ ($\lambda = 1$) zwischen gleichgewichteten QMC-Punkten (gestrichelt) und QMC-Punkten, deren Wahrscheinlichkeiten mit Algorithmus 11.5 optimiert wurden (Linie), sowie Laufzeit in Sekunden von Algorithmus 11.5. Rechte Seite: Abstände α_{rect} und ζ_2 (gestrichelt), resultierend aus dem Vorwärtsalgorithmus für verschiedene Werte von $\lambda \in [0, 1]$.

näherungsweise Lösung vorgeschlagen (Algorithmen 11.1 und 11.2). Für zweistufige Modelle kann für d die Fortet-Mourier Metrik ζ_r verwendet werden. Für mehrstufige Modelle ist es möglich, die Technik zur Szenarioreduktion sukzessive in jeder Stufe einzusetzen und auf diesem Weg Szenariobäume zu erzeugen (Algorithmus 11.3).

Für gemischt-ganzzahlige zweistufige Optimierungsmodelle wird eine Konvexkombination der sog. Rechtecks-Diskrepanz α_{rect} und einer Fortet-Mourier Metrik als geeignetes Abstandsmaß zur optimalen Szenarioreduktion vorgeschlagen. Es wird gezeigt, wie dieses Problem mittels linearer Optimierung gelöst werden kann (vgl. (11.31)). Da der numerische Aufwand in diesem Fall deutlich größer ist, ist der zugehörige Algorithmus 11.5 nur für moderat große Werte von n und der Dimension s des Zufallsvektors ξ einsetzbar. Zur Aufwandsreduktion wird eine Heuristik zur Lösung des kombinatorischen Optimierungsproblems verwendet. Numerische Beispiele illustrieren die in diesem Kapitel diskutierten Algorithmen.

Literatur

- [1] AVELLA, P. ; SASSANO, A. ; VASIL'EV, I.: Computational study of large-scale p-median problems. In: *Mathematical Programming* 109 (2007), S. 89–114
- [2] BURGER, M. ; KLAR, B. ; MÜLLER, A. ; SCHINDLMAYR, G.: A spot market model for pricing derivatives in electricity markets. In: *Quantitative Finance* 4 (2004), S. 109–122
- [3] CLEWLOW, L. ; STRICKLAND, C.: *Energy Derivatives: Pricing and Risk Management*. London : Lacima, 2000
- [4] DUPAČOVÁ, J. ; GRÖWE-KUSKA, N. ; RÖMISCH, W.: Scenarios reduction in stochastic programming: An approach using probability metrics. In: *Mathematical Programming* 95(A) (2003), S. 493–511

- [5] EICHHORN, A. ; RÖMISCH, W. ; WEGNER, I.: Mean-risk optimization of electricity portfolios using multiperiod polyhedral risk measures. In: *IEEE St. Petersburg PowerTech Proceedings*, 2005
- [6] EWING, B.T. ; KRUSE, J.B. ; SCHROEDER, J.L.: Time series analysis of wind speed with time-varying turbulence. In: *Environmetrics* 16 (2005), S. 1–9
- [7] HEITSCH, H.: *Stabilität und Approximation stochastischer Optimierungsprobleme*, Humboldt-Universität zu Berlin, Diss., 2007
- [8] HEITSCH, H. ; RÖMISCH, W.: Scenario reduction algorithms in stochastic programming. In: *Computational Optimization and Applications* 24 (2003), S. 187–206
- [9] HEITSCH, H. ; RÖMISCH, W.: Generation of multivariate scenario trees to model stochasticity in power management. In: *IEEE St. Petersburg PowerTech Proceedings*, 2005
- [10] HEITSCH, H. ; RÖMISCH, W.: A note on scenario reduction for two-stage stochastic programs. In: *Oper. Res. Lett.* 35 (2007), S. 731–738
- [11] HEITSCH, H. ; RÖMISCH, W.: Scenario tree modeling for multistage stochastic programs. In: *Mathematical Programming*, to appear (2008)
- [12] HEITSCH, H. ; RÖMISCH, W. ; STRUGAREK, C.: Stability of multistage stochastic programs. In: *SIAM Journal on Optimization* 17 (2006), S. 511–525
- [13] HENRION, R. ; KÜCHLER, C. ; RÖMISCH, W.: Discrepancy distances and scenario reduction in two-stage stochastic integer programming. In: *Journal of Industrial and Management Optimization* 4 (2008), S. 363–384
- [14] HENRION, R. ; KÜCHLER, C. ; RÖMISCH, W.: A scenario reduction heuristic for two-stage stochastic integer programs. (in Vorbereitung)
- [15] HENRION, R. ; KÜCHLER, C. ; RÖMISCH, W.: Scenario reduction in stochastic programming with respect to discrepancy distances. In: *Computational Optimization and Applications* (to appear)
- [16] HLAWKA, E. ; MÜCK, R.: Über eine Transformation von gleichverteilten Folgen II. In: *Computing* 9 (1972), S. 127–138
- [17] HOCHREITER, R. ; PFLUG, G. C. ; WOZABAL, D.: Multi-stage stochastic electricity portfolio optimization in liberalized energy markets. In: *System Modeling and Optimization*. Boston : Springer, 2006, S. 219–226
- [18] KÜCHLER, C.: On stability of multistage stochastic programs. In: *SIAM Journal on Optimization* (to appear)
- [19] MAGNANTI, T.L. ; WOLSEY, L.A.: Optimal Trees. In: *Network Models*. Elsevier, 1995, Kapitel 9, S. 503–616
- [20] NIEDERREITER, H.: *Random Number Generation and Quasi-Monte Carlo Methods*. Philadelphia : SIAM, 1992 (CBMS-NSF Regional Conference Series in Applied Mathematics)

- [21] RACHEV, S.T.: *Probability Metrics and the Stability of Stochastic Models*. Chichester : Wiley, 1991
- [22] RACHEV, S.T. ; RÖMISCH, W.: Quantitative stability in stochastic programming: The method of probability metrics. In: *Mathematics of Operations Research* 27 (2002), S. 792–818
- [23] RACHEV, S.T. ; RÜSCHENDORF, L.: *Mass Transportation Problems*. Bd. I. Berlin : Springer, 1998
- [24] RÖMISCH, W. ; VIGERSKE, S.: Quantitative stability of fully random mixed-integer two-stage stochastic programs. In: *Optimization Letters* 2 (2008), S. 377–388
- [25] SCHMÖLLER, H. K. ; HARTMANN, T. ; KRUCK, I. ; HAUBRICH, H.-J.: Modeling power price uncertainty for midterm generation planning. In: BORGHETTI, A. (Hrsg.) ; NUCCI, C. A. (Hrsg.) ; PAOLONE, M. (Hrsg.): *IEEE Bologna PowerTech Proceedings*, 2003
- [26] SCHULTZ, R.: Rates of convergence in stochastic programs with complete integer recourse. In: *SIAM Journal of Optimization* 6 (1996), S. 1138–1152
- [27] SEN, S. ; YU, L. ; GENC, T.: A stochastic programming approach to power portfolio optimization. In: *Operations Research* 54 (2006), S. 55–72
- [28] VAZIRANI, V. V.: *Approximation Algorithms*. Springer, Berlin, 2001
- [29] WERON, R. ; BIERBRAUER, M. ; TRÜCK, S.: Modeling electricity prices: jump diffusion and regime switching. In: *Physica A* 336 (2004), S. 39–48