Konvexitätsaussagen zum linearen stochastischen Optimierungsproblem

KURT MARTI

1. Problemstellung

In der vorliegenden Arbeit¹ betrachten wir die beiden stochastischen Optimierungsprobleme

$$\begin{aligned} & \min F(x) \\ \text{bezüglich} & \\ & P(Y^1(x, \xi) \leq 0, \dots, Y^m(x, \xi) \leq 0) \geq \alpha \\ & x \geq \theta \end{aligned} \quad (1.1)$$

$$\begin{aligned} & \min F(x) \\ \text{bezüglich} & \\ & P(Y^i(x, \xi) \leq 0) \geq \alpha_i, \quad i = 1, \dots, m \\ & x \geq \theta. \end{aligned} \quad (1.2)$$

Dabei ist x ein n -Vektor, ξ ein r -Zufallsvektor, Y^i eine Abbildung $Y^i: \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}$, P ein Wahrscheinlichkeitsmaß und F das zu minimierende Funktional. Ferner ist $0 \leq \alpha, \alpha_i \leq 1, i = 1, \dots, m$ und $x \geq \theta$ bedeutet $x_k \geq \theta_k, k = 1, \dots, n$.

Im weiteren befassen wir uns mit dem linearen stochastischen Optimierungsproblem, d.h. der Zufallsvektor $y_x \equiv Y(x, \xi)$ wird durch

$$y_x = -Ax + b \quad (1.3)$$

definiert, wobei die Elemente der $(m \times n)$ -Matrix A sowie des m -Vektors b reelle Zufallsvariable sind, deren gemeinsame Verteilung als bekannt vorausgesetzt wird.

2. Konvexe Probleme

Unter einem konvexen Optimierungsproblem soll die Minimierung einer konvexen oder die Maximierung einer konkaven Funktion über einem konvexen Bereich verstanden werden.

Setzen wir die Zielfunktion F als konvex voraus, reduziert sich die Frage nach der Konvexität der Probleme (1.1) und (1.2) auf diejenige nach der Konvexität der zulässigen Bereiche

$$X = X(\alpha) = \{x \mid x \geq \theta, P(y_x \leq \theta) \geq \alpha\}$$

von (1.1) und

$$X = X(\alpha_1, \dots, \alpha_m) = \{x \mid x \geq \theta, P(y_x^i \leq 0) \geq \alpha_i, i = 1, \dots, m\}$$

¹ Wir verwenden die Notation von [5].

von (1.2). Diese sind, wie man sich an Beispielen (siehe z.B. § 3) klar macht, für $\alpha < 1$ bzw. $\alpha_i < 1$ im allgemeinen nicht konvex. Hingegen erhalten wir (siehe [1]) durch eine leichte Überlegung die Konvexität von X in einem wichtigen Spezialfall.

Lemma 1. Für $\alpha = 1$ bzw. $\alpha_i = 1, i = 1, \dots, m$ ist der zulässige Bereich $X(\alpha)$ von (1.1) bzw. $X(\alpha_1, \dots, \alpha_m)$ von (1.2) konvex.

Damit gleichwertig ist das folgende

Lemma 2. Zu jedem Problem (1.1) bzw. (1.2) gibt es eine Zahl α^0 bzw. m Zahlen α_i^0 mit $0 \leq \alpha^0, \alpha_i^0 \leq 1$, so daß für alle $\alpha \geq \alpha^0$ bzw. $\alpha_i \geq \alpha_i^0$ $X(\alpha)$ bzw. $X(\alpha_1, \dots, \alpha_m)$ konvex ist.

Es stellt sich nun sofort die Frage, ob α^0 bzw. α_i^0 nur den Wert 1 oder aber auch solche < 1 annehmen kann. Durch Einsetzen spezieller W -Masse für die gemeinsame Verteilung der Elemente von A und b suchen wir wenigstens eine teilweise Antwort auf die obige Frage.

3. Konvexitätsaussagen zum Problem (1.2)

Der zulässige Bereich $X(\alpha_1, \dots, \alpha_m)$ läßt sich durch

$$X(\alpha_1, \dots, \alpha_m) = \bigcap_{i=1}^m X_i(\alpha_i)$$

mit

$$X_i(\alpha_i) = \{x \mid x \geq \theta, P_{y_k^i}((-\infty, 0]) \geq \alpha_i\}$$

darstellen, wobei $P_{y_k^i}$ die Verteilung von y_k^i ist. Da der Durchschnitt konvexer Mengen wieder konvex ist, untersuchen wir die Mengen $X_i(\alpha_i)$.

Dazu sei x ein beliebiger, aber fester n -Vektor und

$$\{f_{ik} \mid i = 1, \dots, m; k = 1, \dots, n+1\} \quad \text{mit } f_{ik} = a_{ik}$$

für $i = 1, \dots, m, k = 1, \dots, n$ und $f_{in+1} = b_i$ für $i = 1, \dots, m$ die Familie der in (1.2) auftretenden Zufallsvariablen $f_{ik}: (\Omega, \mathfrak{A}, P) \rightarrow (\mathbb{R}, \mathfrak{B})$. Wie bereits einleitend bemerkt, setzen wir deren gemeinsame Verteilung als bekannt voraus, womit dann auch für jedes $i, i = 1, \dots, m$ die gemeinsame Verteilung $P_{f_i}(f_i(\omega) = (f_{i1}(\omega), \dots, f_{in+1}(\omega)), \omega \in \Omega)$ der $f_{i1}, \dots, f_{in+1}$ gegeben ist.

Ist $H_x: \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ durch

$$H_x(t_1, \dots, t_{n+1}) = - \sum_{k=1}^n x_k \cdot t_k + t_{n+1}$$

definiert, dann gilt

$$P_{y_k^i} = H_x(P_{f_i}),$$

wobei $H_x(P_{f_i})$ das Bild von P_{f_i} bei der Abbildung H_x ist.

Wir beschränken uns nun auf Zufallsvariable (f_{ik}) mit (I) diskreter und (II) stetiger Verteilung:

I. Es sei für ein $i \in I = \{1, \dots, m\}$

$$P_{f_i} = \sum_{r=1}^{N_i} \alpha_{ir} \varepsilon_{z_{ir}}$$

$$\left(0 < \alpha_{ir} \leq 1, \sum_{r=1}^{N_i} \alpha_{ir} = 1, N_i < \infty, z_{ir} \in \mathbb{R}^{n+1}, \right.$$

$$\left. \varepsilon_z \text{ ist das durch die Einheitsmasse in } z \text{ definierte Borelmaß} \right)$$

die gemeinsame Verteilung der $f_{i1}, \dots, f_{in+1}$.

Unter der Voraussetzung (welche im ganzen § gelten soll), daß die Zielfunktion F konvex ist, gilt der

Satz 1. *Haben die Zufallsvariablen $a_{i1}, \dots, a_{in}, b_i$ jeder i -ten Zeile eine gemeinsame diskrete Verteilung, so ist das Optimierungsproblem (1.2) für*

$$\alpha_i > \alpha_i^0 = 1 - \min_{1 \leq r \leq N_i} \alpha_{ir}$$

konvex.

Beweis. Nach dem Obigen gilt

$$P_{y_i}(B) = P_{f_i}(H_x^{-1}(B))$$

für eine Borelmenge $B \subset \mathbb{R}$. Setzen wir darin $B = (-\infty, 0]$, so ergibt sich

$$\sum_{r=1}^{N_i} \alpha_{ir} \varepsilon_{z_{ir}}(H_x^{-1}(-\infty, 0]) \geq \alpha_i \quad (3.1)$$

als i -te Restriktion der Aufgabe (1.2).

Nun ist aber nach Definition des Maßes ε_z

$$\varepsilon_{z_{ir}}(H_x^{-1}(-\infty, 0]) = \varepsilon_{H_x(z_{ir})}((-\infty, 0]),$$

d.h. die Restriktion (3.1) geht über in

$$\sum_{r=1}^{N_i} \alpha_{ir} \varepsilon_{H_x(z_{ir})}((-\infty, 0]) \geq \alpha_i. \quad (3.2)$$

Soll jetzt die Ungleichung (3.2) für ein $\alpha_i > \alpha_i^0$ erfüllt sein, dann müssen die Gleichungen

$$\varepsilon_{H_x(z_{ir})}((-\infty, 0]) = 1, \quad r = 1, \dots, N_i$$

oder damit äquivalent die Ungleichungen

$$H_x(z_{ir}) \leq 0, \quad r = 1, \dots, N_i \quad (3.3)$$

gelten, denn die aus (3.2) und $\alpha_i > \alpha_i^0$ resultierende Ungleichung

$$\sum_{r=1}^{N_i} \alpha_{ir} \varepsilon_{H_x(z_{ir})}((-\infty, 0]) + \min_r \alpha_{ir} > 1$$

gilt genau dann, falls sämtliche Koeffizienten der $\alpha_{ir}, r = 1, \dots, N_i$ gleich 1 sind.

Aus (3.3) liest man aber unter Berücksichtigung der Definition des Operators H_x sofort die Konvexität von $X_i(\alpha_i)$ für alle $\alpha_i > \alpha_i^0$ ab. Daraus folgt die Behauptung.

Bemerkungen. Für $\alpha_i > \alpha_i^0, i = 1, \dots, m$ folgt aus (3.3) noch die Gleichung

$$X(\alpha_1, \dots, \alpha_m) = X(1, \dots, 1).$$

Das folgende Beispiel zeigt, daß im Falle $0 < \alpha_i \leq \alpha_i^0, i = 1, \dots, m$ das Optimierungsproblem i. a. nicht konvex ist.

Wir setzen dazu $n = m = 1$. In der Zufallsvariablen $y_x = -ax + b$ (x ist jetzt eine reelle Zahl) habe das Paar (a, b) die diskrete Verteilung

$$P_{(a,b)} = \frac{1}{2} \varepsilon_{z_1} + \frac{1}{2} \varepsilon_{z_2}$$

mit $z_1 = (-3, -1), z_2 = (3, 2)$. Es folgt $\alpha^0 = \frac{1}{2}$. Mittels einer Skizze macht man sich leicht klar, daß von den beiden Punkten z_1 und z_2 für

$$\begin{array}{ll} 0 \leq x \leq \frac{1}{3} & \text{bloß } z_1, \\ \frac{1}{3} < x < \frac{2}{3} & \text{keiner,} \\ \frac{2}{3} \leq x & \text{bloß } z_2 \end{array}$$

in $H_x^{-1}(-\infty, 0]$ liegt. Also gilt

$$P(y_x \leq 0) = \begin{cases} \frac{1}{2}, & 0 \leq x \leq \frac{1}{3} \\ 0, & \frac{1}{3} < x < \frac{2}{3} \\ \frac{1}{2}, & \frac{2}{3} \leq x \end{cases}$$

und der zulässige Bereich $X(\alpha)$ ist für alle Werte von α mit $0 < \alpha \leq \alpha^0$ offensichtlich nicht konvex, während $X(\alpha)$ für $\alpha > \alpha^0$ leer, also konvex ist.

II. Wir betrachten nun die folgenden beiden stetigen Verteilungen für die Variable f_i :

A. Es habe f_i für ein beliebiges $i \in I$ eine $(n+1)$ -dimensionale Normalverteilung.

Nach Definition dieser Verteilung existieren ein $(n+1)$ -Zufallsvektor z mit der W -Dichte

$$\phi(t) = (2\pi)^{-\frac{n+1}{2}} \exp(-\frac{1}{2}\langle t, t \rangle), \quad t \in \mathbb{R}^{n+1}$$

($\langle \cdot, \cdot \rangle$ bezeichnet das gewöhnliche Skalarprodukt im $\mathbb{R}^{n+1}$), eine lineare Abbildung C^i des $\mathbb{R}^{n+1}$ in sich und ein fester $(n+1)$ -Vektor m^i , so daß die Zufallsvariable f_i durch die Gleichung

$$f_i = C^i z + m^i \quad (3.4)$$

definiert wird.

Nun gilt der folgende

Satz 2. Haben die Zufallsvariablen $a_{i1}, \dots, a_{in}, b_i$ jeder i -ten Zeile eine $(n+1)$ -dimensionale Normalverteilung, so ist das Optimierungsproblem (1.2) für

$$\alpha_i \geq \frac{1}{2}, \quad i = 1, \dots, m$$

konvex.

Beweis. Im folgenden sei $x = (x_1, \dots, x_n)'$ wieder ein beliebiger aber fester n -Vektor und $\hat{x} = (x_1, \dots, x_n, -1)'$. Aus (3.4) folgt, daß die Variable

$$y_x^i = -\langle \hat{x}, f_i \rangle$$

gemäß

$$N\left(-\langle m^i, x \rangle, \sum_{r=1}^{n+1} (C_r^{i'} x)^2\right)$$

normalverteilt ist. Für die Varianz von y_x^i gilt dann

$$V y_x^i = \|C^{i'} \hat{x}\|^2,$$

wenn

$$\|t\| = \left(\sum_{k=1}^{n+1} t_k^2\right)^{\frac{1}{2}}, \quad t \in \mathbb{R}^{n+1}.$$

Ist Φ bzw. Φ^{-1} die Verteilungsfunktion bzw. deren Inverse der $N(0, 1)$ -Normalverteilung, so ergibt sich die Wahrscheinlichkeit $P(y_x^i \leq 0)$ zu

$$\Phi\left(\frac{-E y_x^i}{\sqrt{V y_x^i}}\right) = \Phi\left(\frac{\langle m^i, \hat{x} \rangle}{\|C^{i'} \hat{x}\|}\right),$$

und als i -te Restriktion folgt

$$-\langle m^i, \hat{x} \rangle + \Phi^{-1}(\alpha_i) \cdot \|C^{i'} \hat{x}\| \leq 0. \quad (3.5)$$

Da $\|\cdot\|$ eine Norm ist, ist die linke Seite von (3.5) konvex in $\hat{x}$, falls noch

$$\Phi^{-1}(\alpha_i) \geq 0$$

oder

$$\alpha_i \geq \Phi(0) = \frac{1}{2}$$

ist. Dies besagt aber, daß $X_i(\alpha_i)$ für alle $\alpha_i \geq \alpha_i^0 = \frac{1}{2}$ eine konvexe Menge ist, was zu beweisen war.

Das folgende Beispiel zeigt, daß im Fall $0 < \alpha_i < \frac{1}{2}$, $i = 1, \dots, m$ das Optimierungsproblem i. a. nicht konvex ist.

Wir setzen $n=2$, $m=1$ und a_{11} , a_{12} sowie b_1 seien unabhängige, normalverteilte Zufallsvariable mit den Parametern

$$\begin{aligned} \mu_{11} = \mu_{12} = 0, \quad \mu_{13} > 0, \\ \sigma_{11} = \sigma_{12} = 1, \quad \sigma_{13} = 0. \end{aligned}$$

Die Abbildung C^1 ergibt sich daraus zu

$$C^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

und für die Varianz von y_x^1 gilt

$$V y_x^1 = x_1^2 + x_2^2 \equiv \|x\|^2.$$

Damit lautet die einzige Restriktion von (1.2)

$$\mu_{13} + \Phi^{-1}(\alpha_1) \cdot \|x\| \leq 0.$$

Ist $0 < \alpha_1 < \frac{1}{2}$, also $\Phi^{-1}(\alpha_1)$ negativ, so folgt daraus

$$\|x\| \geq \frac{\mu_{13}}{|\Phi^{-1}(\alpha_1)|}$$

und $X = X(\alpha_1)$ ist offensichtlich nicht konvex für diese Werte von α_1 .

B. Es habe f_i für ein beliebiges $i \in I$ eine $(n+1)$ -dimensionale Cauchy-Verteilung.

Wir definieren die $(n+1)$ -dimensionale Cauchy-Verteilung analog der $(n+1)$ -dimensionalen Normalverteilung. Es existiert also ein $(n+1)$ -Zufallsvektor z mit der Wahrscheinlichkeitsdichte

$$\phi(t) = \pi^{-n-1} \prod_{i=1}^{n+1} \frac{1}{1+t_i^2}, \quad t \in \mathbb{R}^{n+1},$$

eine lineare Abbildung C^i des $\mathbb{R}^{n+1}$ in sich und ein fester $(n+1)$ -Vektor m^i , so daß die Zufallsvariable f_i durch die Gleichung

$$f_i = C^i z + m^i$$

definiert wird.

Daraus folgt zunächst, daß $z_1, \dots, z_{n+1}$ eine Familie unabhängiger Zufallsvariablen mit je einer $C(0, 1)$ -Cauchyverteilung, d.h. einer Verteilung mit der W -Dichte

$$v(t) = \frac{1}{\pi} \frac{1}{1+t^2}, \quad t \in \mathbb{R}$$

ist. Entsprechend Satz 2 gilt

Satz 3. Haben die Zufallsvariablen $a_{i1}, \dots, a_{in}, b_i$ jeder i -ten Zeile eine $(n+1)$ -dimensionale Cauchyverteilung, so ist das Optimierungsproblem (1.2) für

$$\alpha_i \geq \frac{1}{2}, \quad i = 1, \dots, m$$

konvex.

Beweis. Aus $y_x^i = -\langle \hat{x}, f_i \rangle$ folgt

$$y_x^i = -\langle z, C^{i'} \hat{x} \rangle - \langle m^i, \hat{x} \rangle.$$

Jedes $z_r, r = 1, \dots, n+1$ hat eine $C(0, 1)$ -Verteilung, also ist $-(C_r^{i'} \hat{x}) \cdot z_r$

$C(0, |C_r^{i'} \hat{x}|)$ -cauchyverteilt.

Aus der Unabhängigkeit der $z_1, \dots, z_{n+1}$ folgt diejenige der Variablen $\{-(C_r^{i'} \hat{x}) \cdot z_r, r = 1, \dots, n+1\}$. Da die Summe zweier unabhängiger $C(\mu_i, \lambda_i)$ -cauchyverteilter Zufallsgrößen ($\lambda_i \geq 0$) eine $C(\mu_1 + \mu_2, \lambda_1 + \lambda_2)$ -verteilte Variable ist, folgt, daß y_x^i gemäß

$$C\left(-\langle m^i, \hat{x} \rangle, \sum_{r=1}^{n+1} |C_r^{i'} \hat{x}|\right)$$

cauchyverteilt ist. Mit

$$\|t\| = \sum_{k=1}^{n+1} |t_k|, \quad t \in \mathbb{R}^{n+1}$$

lautet die i -te Restriktion von (1.2)

$$\Psi \left(\frac{\langle m^i, \hat{x} \rangle}{\|C^{i'} \hat{x}\|} \right) \geq \alpha_i,$$

wenn Ψ die Verteilungsfunktion der $C(0, 1)$ -Verteilung ist, oder

$$-\langle m^i, \hat{x} \rangle + \Psi^{-1}(\alpha_i) \cdot \|C^{i'} \hat{x}\| \leq 0, \quad (3.6)$$

wenn Ψ^{-1} wieder die Inverse von Ψ ist.

Gilt nun

$$\Psi^{-1}(\alpha_i) \leq 0$$

d.h.

$$\alpha_i \geq \Psi(0) = \frac{1}{2},$$

so ist die linke Seite von (3.6) ein konvexes Funktional in $\hat{x}$. Damit ist aber $X_i(\alpha_i)$ für alle $\alpha_i \geq \frac{1}{2}$ eine konvexe Menge, was zu zeigen war.

Wiederum ergibt sich durch ein Beispiel, daß das Optimierungsproblem im Falle $0 < \alpha_i < \frac{1}{2}$, $i = 1, \dots, m$ i. a. nicht konvex ist.

Es sei dazu $n = 2$, $m = 1$, $m^1 = (0, 0, m_3^1)$ mit $m_3^1 > 0$ sowie

$$C^1 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Damit lautet die Restriktion von (1.2)

$$m_3^1 + \Psi^{-1}(\alpha_1) |x_1 - x_2| \leq 0.$$

Ist $0 < \alpha_1 < \frac{1}{2}$, also $\Psi^{-1}(\alpha_1)$ negativ, so folgt daraus

$$|x_1 - x_2| \geq \frac{m_3^1}{|\Psi^{-1}(\alpha_1)|}$$

und $X = X(\alpha_1)$ ist offensichtlich nicht konvex für diese Werte von α_1 . Hingegen gilt das folgende

Korollar 1. Ist für jedes $i \in I$ $f_{i1}, \dots, f_{in+1}$ eine Familie unabhängiger $C(\mu_{ik}, \lambda_{ik})$ -verteilter Zufallsvariablen, so ist das Optimierungsproblem (1.2) für

$$\alpha_i \geq 0, \quad i = 1, \dots, m$$

konvex.

Beweis. Die Abbildung C^i ist in diesem Fall durch

$$C^i = (\lambda_{ik} \cdot \delta_{kk'}) \quad (\delta_{kk'} = \text{Kroneckersymbol})$$

gegeben, ferner gilt

$$m^i = (\mu_{i1}, \dots, \mu_{in+1}).$$

Da $\lambda_{ik} \geq 0$ ist, lautet mit $x \geq 0$ die Restriktion (3.6)

$$-\sum_{k=1}^n \mu_{ik} x_k + \mu_{in+1} + \Psi^{-1}(\alpha_i) \left(\sum_{k=1}^n \lambda_{ik} x_k + \lambda_{in+1} \right) \leq 0$$

oder

$$\sum_{k=1}^n (\lambda_{ik} \Psi^{-1}(\alpha_i) - \mu_{ik}) x_k + \lambda_{in+1} \Psi^{-1}(\alpha_i) + \mu_{in+1} \leq 0$$

was eine lineare Bedingung darstellt. Dies war zu zeigen.

Literatur

1. Kall, P.: Qualitative Aussagen zu einigen Problemen der stochastischen Programmierung. Z. Wahrscheinlichkeitstheorie verw. Geb. 6, 246–272 (1966).
2. Symonds, G.H.: Stochastic scheduling – part I (1964), Tech. Memo No. 22, – part IV (1965), Tech. Memo No. 39. Case Institute of Technology, Cleveland, Ohio 44106.
3. Charnes, A.: (Northwestern University, Evanston, Illinois), Kirby M. (Research Analysis Corporation, McLean, Virginia): Optimal decision rules for the *E*-model of chance-constrained programming. Presented at the 12th International Meeting of the Institute of Management Sciences, Vienna, Austria, 1965.
4. Kirby, M.: The current state of chance-constrained programming. Systems Research Memorandum No. 181, Northwestern University, Evanston, Illinois.
5. Bauer, H.: Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie I. Band 1216/1216a, Sammlung Götschen. Berlin 1964.

Kurt Marti
Universität Mannheim
Lehrstuhl für Mathematik II
BRD-6800 Mannheim 1
Schloß
Deutschland

(Eingegangen am 6. Juli 1969)