(Lineare) stochastische Optimierung

Bsp:

Aus zwei Sorten Rohöl wird Benzin und Heizöl erzeugt. Die Produktivität sowie der Mindestbedarf (pro Woche) und die Kosten sind in folgender Tabelle angegeben:

	Rohöl 1	Rohöl 2	Bedarf
Benzin	2	6	180
Heizöl	3	3	162
Kosten	2	3	

Weiters besteht eine Kapazitätsbeschränkung, dass maximal 100 Einheiten Rohöl pro Woche verarbeitet werden können.

Man ist also mit folgendem linearen Optimierungsproblem konfrontiert:

$$\begin{array}{rcl} 2x_1 + 3x_2 & \Rightarrow & \min \\ x_1 + x_2 & \leq & 100 \\ 2x_1 + 6x_2 & \geq & 180 \\ 3x_1 + 3x_2 & \geq & 162 \\ x_1, x_2 & \geq & 0 \end{array}$$

Als kostengünstigsten Produktionsplan erhält man $x_1 = 36, x_2 = 18$ und die minimalen Kosten betragen 126 GE.

Angenommen, sowohl der Mindestbedarf als auch die Produktivitäten unterliegen Zufallsschwankungen und sind bei Erstellung des Produktionsplans noch nicht bekannt.

	Rohöl 1	Rohöl 2	Bedarf
Benzin	$2 + \tilde{\eta}_1$	6	$180 + \tilde{\zeta}_1$
Heizöl	3	$3.4 - \tilde{\eta}_2$	$162 + \tilde{\zeta}_2$
Kosten	2	3	

wobei $\tilde{\zeta}_1 \sim \mathcal{N}(0, 12)$, $\tilde{\zeta}_2 \sim \mathcal{N}(0, 9)$, $\tilde{\eta}_1 \sim \mathcal{U}(-0.8, 0.8)$, $\tilde{\eta}_2 \sim \mathcal{EXP}(\lambda = 2.5)$.

Weiters wird angenommen, dass die unbeschränkten Zufallsvariablen $\tilde{\zeta}_1, \tilde{\zeta}_2$ und $\tilde{\eta}_2$ nur Werte innerhalb folgender 99% Konfidenzintervalle annehmen:

$$\zeta_1 \in [-30.91, 30.91], \qquad \zeta_2 \in [-23.18, 23.18], \qquad \eta_2 \in [0.0, 1.84]$$

Man ist also mit folgendem stochastischen linearen Optimierungsproblem konfrontiert:

$$\begin{array}{rcrcl} 2x_1 & + & 3x_2 & \Rightarrow & \min \\ x_1 & + & x_2 & \leq & 100 \\ (2+\tilde{\eta}_1)x_1 & + & 6x_2 & \geq & 180 + \tilde{\zeta}_1 \\ 3x_1 & + & (3.4 - \tilde{\eta}_2)x_2 & \geq & 162 + \tilde{\zeta}_2 \\ & & x_1, x_2 & \geq & 0 \end{array}$$

- Dieses Optimierungsproblem ist nicht wohldefiniert, da völlig unklar ist, was eine "optimale" Lösung bedeuten soll, solang die Realisierungen der Zufallsvariablen unbekannt sind.

- Änderung von ζ_i bewirkt eine Parallelverschiebung der Begrenzungsgeraden des zulässigen Bereichs.

- Änderung von η_i bewirkt eine Drehung.

Fat Solution: Eine Lösung, die für alle Realisierungen der Zufallsvariablen (in den oben angegebenen Intervallen) zulässig ist.

In unserem Beispiel tritt der "schlechteste" Fall für $\zeta_1 = 30.91, \zeta_2 = 23.18, \eta_1 = -0.8, \eta_2 = 1.84$ ein. Als "Fat Solution" erhält man: $x_1 = 48.49, x_2 = 25.45$ und die Kosten betragen 173.34 GE.

<u>Andere Möglichkeit:</u>

Für nicht erfüllte Nachfrage sind von der Raffinerie Strafzahlungen zu entrichten, und zwar 7 bzw. 12 Geldeinheiten je nicht gelieferter Mengeneinheit Benzin bzw. Heizöl.

Gesucht ist ein Produktionsplan, der die Summe aus Produktionskosten und erwarteten Regress-kosten minimiert.

Notation:

$\tilde{\xi} := (\tilde{\zeta}_1, \tilde{\zeta}_2, \tilde{\eta}_1, \tilde{\eta}_2)^t$

$y_i(\tilde{\xi}) \cdots$ Regress-variable, gibt die nicht-erfüllte Nachfrage an.

Weiters:

$$\begin{aligned} \alpha(\tilde{\xi}) &:= 2 + \tilde{\eta}_1, & \beta(\tilde{\xi}) &:= 3.4 - \tilde{\eta}_2 \\ h_1(\tilde{\xi}) &:= 180 + \tilde{\zeta}_1, & h_2(\tilde{\xi}) &:= 162 + \tilde{\zeta}_2. \end{aligned}$$

Man erhält folgendes stochastisches Optimierungsproblem:

$$\begin{array}{rcrcrcrcr} 2x_1 & + & 3x_2 & + & \mathrm{E}_{\tilde{\xi}}[7y_1(\tilde{\xi}) & + & 12y_2(\tilde{\xi})] & \Rightarrow & \min \\ x_1 & + & x_2 & & & & & \leq & 100 \\ \alpha(\tilde{\xi})x_1 & + & 6x_2 & + & y_1(\tilde{\xi}) & & & \geq & h_1(\tilde{\xi}) \\ 3x_1 & + & \beta(\tilde{\xi})x_2 & & & + & y_2(\tilde{\xi}) & \geq & h_2(\tilde{\xi}) \\ & & x_1, x_2 & \geq & 0 & & & & \end{array}$$

Probleme:

- multivariate Integration bei Erwartungswertberechnung

- Funktionen $y_i(\xi)$ sind implizit gegeben.

Falls die Zufallsvariable $\tilde{\xi}$ eine endliche diskrete Verteilung besitzt, i.e. $\tilde{\xi}$ nimmt mit Wahrscheinlichkeiten p_i die Werte ξ^i an, $i = 1, \cdots, r$, so erhält man das folgende lineare Optimierungsproblem

$$
\begin{array}{rcrcrcrcll}
2x_1 & + & 3x_2 & + & \sum_{i=1}^r p_i[7y_1(\xi^i) & + & 12y_2(\xi^i)] & \Rightarrow & \min & \\
x_1 & + & x_2 & & & & & \leq & 100 & \\
\alpha(\xi^i)x_1 & + & 6x_2 & + & y_1(\xi^i) & & & \geq & h_1(\xi^i) & \forall i = 1, \cdots r \\
3x_1 & + & \beta(\xi^i)x_2 & & & + & y_2(\xi^i) & \geq & h_2(\xi^i) & \forall i = 1, \cdots r \\
x_1 & & & & & & & \geq & 0 & \\
 & & x_2 & & & & & \geq & 0 & \\
 & & & & y_1(\xi^i) & & & \geq & 0 & \forall i = 1, \cdots, r \\
 & & & & & & y_2(\xi^i) & \geq & 0 & \forall i = 1, \cdots, r
\end{array}
$$

d.h. man erhält ein lineares Optimierungsproblem mit 2r + 2 Variablen $(x_1, x_2, y_1(\xi^1), y_2(\xi^i))$ und 2r + 1 Nebenbedingungen.

Approximation stetiger Verteilungen

Die Verteilung der stetigen Zufallsvariable $\tilde{\xi}$ kann folgendermaßen durch eine diskrete endliche Zufallsvariable approximiert werden:

- generiere eine große (z.B. $K = 10000$) Zufallsstichprobe $\xi_i, i = 1, \cdots, K$
- Zerlege das 99%-Intervall in r Teil-"Intervalle" $I_j, j = 1, \cdots, r$.
- Wähle als mögliche Ausprägungen ξ^j der disketen Approximation die Mittelpunkte dieser Teilintervalle
- Schätze die Wahrscheinlichkeit p_j der Ausprägung ξ^j durch die relative Häufigkeit der Elemente der Zufallsstichprobe, die im Intervall I_j liegen.

Duale Dekompositionsmethode

Betrachten das Problem

$$
\begin{array}{rrcl}
 & \{\vec{c}^t\vec{x} + \vec{q}^t\vec{y}\} & \rightarrow & \min \\
\text{s.t.} & A\vec{x} & = & \vec{b} \\
 & T\vec{x} + W\vec{y} & = & \vec{h} \\
 & \vec{x} \geq 0, \vec{y} \geq 0 & &
\end{array}
$$

Annahmen:

- Das Problem ist lösbar
- Die Menge $\{\vec{x} | A\vec{x} = \vec{b}, \vec{x} \geq \vec{0}\}$ ist beschränkt.

Dieses Problem kann zu folgendem äquivalenten Problem umformuliert werden:

$$
\begin{array}{rrcl}
 & \{\vec{c}^t\vec{x} + f(\vec{x})\} & \rightarrow & \min \\
\text{s.t.} & A\vec{x} & = & \vec{b} \\
 & \vec{x} & \geq & \vec{0}
\end{array}
$$

wobei

$$f(\vec{x}) = \min\{\vec{q}^t\vec{y} | W\vec{y} = \vec{h} - T\vec{x}, \vec{y} \geq \vec{0}.\}$$

Weiters kann dieses Problem umformuliert werden zu

$$\begin{array}{rrcl} & \{\vec{c}^t\vec{x} + \theta\} & \rightarrow & \min \\ \text{s.t.} & A\vec{x} & = & \vec{b} \\ & \theta - f(\vec{x}) & \geq & \vec{0} \\ & \vec{x} & \geq & \vec{0} \end{array}$$

Dualer Dekompositions-Algorithmus

- **Schritt 1:** Sei θ_0 eine untere Schranke für

  $$min\{\vec{q}^t\vec{y} | A\vec{x} = \vec{b}, T\vec{x} + W\vec{y} = \vec{h}, \vec{x} \geq \vec{0}, \vec{y} \geq \vec{0}\}$$

  Löse das Problem

  $$min\{\vec{c}^t\vec{x} + \theta | A\vec{x} = \vec{b}, \theta \geq \theta_0, \vec{x} \geq \vec{0}\}$$

  Dies führt zu einer Lösung $(\hat{\vec{x}}, \hat{\theta})$. Sei $\mathcal{B}_1 := \{\mathbb{R}^n \times \{\theta\} | \theta \geq \theta_0\}$.

- **Schritt 2:** Evaluiere die Regress Funktion

  $$\begin{array}{rcl} f(\hat{\vec{x}}) & = & min\{\vec{q}^t\vec{y} | W\vec{y} = \vec{h} - T\hat{\vec{x}}, \vec{y} \geq 0\} = \\ & = & \max\{(\vec{h} - T\hat{\vec{x}})^t\vec{u} | W^t\vec{u} \leq \vec{q}\} \end{array}$$

  Man muss nun zwischen den folgenden zwei Fällen unterscheiden:

  - $f(\hat{\vec{x}}) = +\infty$
    dann ist $\hat{\vec{x}}$ nicht zulässig bezüglich aller Nebenbedingungen des Optimierungsproblems.
    $\Rightarrow$ Es existiert ein $\tilde{\vec{u}}$ sodass $W^t\tilde{\vec{u}} \leq \vec{0}$ und $(\vec{h} - T\hat{\vec{x}})^t\tilde{\vec{u}} > 0$ ist.
    Andererseits muss für jedes zulässige $\vec{x}$ ein $\vec{y}$ existieren, sodass

    $$W\vec{y} = \vec{h} - T\vec{x} \tag{1}$$

    gilt. Skalarmultiplikation von (1) mit $\tilde{\vec{u}}$ ergibt

    $$\tilde{\vec{u}}^t(\vec{h} - T\vec{x}) = \underbrace{\tilde{\vec{u}}^t W}_{\leq 0} \underbrace{\vec{y}}_{\geq 0} \leq 0.$$

    Daher kann durch die zusätzliche Nebenbedingung

    $$\tilde{\vec{u}}(\vec{h} - T\vec{x}) \leq 0$$

    die unzulässige Lösung $\hat{\vec{x}}$ "weggeschnitten" werden (feasibility cut). Schrenken nun den Parameterraum ein zu $\mathcal{B}_1 = \mathcal{B}_1 \cap \{(\vec{x}^t, \theta) | \tilde{\vec{u}}^t(\vec{h} - T\vec{x}) \leq 0\}$. Gehe zu Schritt 3.

– $f(\hat{\vec{x}})$ ist endlich. Dann gibt es zu $\hat{\vec{x}}$ eine primale Lösung $\hat{\vec{y}}$ und eine duale Lösung $\hat{\vec{u}}$.

Die Dualität ergibt:

$$f(\hat{\vec{x}}) = (\vec{h} - T\hat{\vec{x}})^t \hat{\vec{u}},$$

für alle anderen Vektoren $\vec{x}$ gilt jedoch

$$\begin{aligned} f(\vec{x}) &= sup\{(\vec{h} - T\vec{x})^t \vec{u} | W^t \vec{u} \leq \vec{q}\} \\ &\geq (\vec{h} - T\vec{x})^t \hat{\vec{u}} \\ &= \hat{\vec{u}}^t (\vec{h} - T\vec{x}\} \end{aligned}$$

Falls die aktuelle Lösung $(\hat{\vec{x}}, \hat{\theta})$ die Nebenbedingung $f(\vec{x}) \leq \theta$ erfüllt, (i.e. falls $f(\hat{\vec{x}}) \leq \hat{\theta}$) hat man eine optimale Lösung gefunden.

Anderenfalls impliziert die Nebenbedingung $\theta \geq f(\vec{x})$ die lineare Nebenbedingung

$$\theta \geq \hat{\vec{u}}^t (\vec{h} - T\vec{x})$$

Durch die zusätzliche Nebenbedingung

$$\theta \geq \hat{\vec{u}}^t (\vec{h} - T\vec{x})$$

wird die sub-optimale Lösung $(\hat{\vec{x}}, \hat{\theta})$ "weggeschnitten" (optimality cut) Schrenken nun den Parameterraum ein zu $\mathcal{B}_1 = \mathcal{B}_1 \cap \{(\vec{x}^t, \theta) | \hat{\vec{u}}^t (\vec{h} - T\vec{x}) \leq \theta\}$. Gehe zu Schritt 3.

Schritt 3: Löse das modifizierte Problem

$$\min\{\vec{c}^t \vec{x} + \theta | (\vec{x}^t, \theta) \in \mathcal{B}_1\}$$

Dies ergibt eine Lösung $(\tilde{\vec{x}}, \tilde{\theta})$. Setze nun $(\hat{\vec{x}}, \hat{\theta}) := (\tilde{\vec{x}}, \tilde{\theta})$ und gehe zu Schritt 2.

Unter den obigen Voraussetzungen führt dieser Algorithmus in endlich vielen Schritten zur optimalen Lösung.

References

[1] Kall, Peter und Wallace, Stein W., Stochastic Programming, J. Wiley and Sons, Chichester, 1995.