Metody robustní statistiky ve stochastické optimalizaci

Martin Branda

Katedra pravděpodobnosti a matematické statistiky
Matematicko-fyzikální fakulta
Univerzita Karlova v Praze

Robust 2010
1. – 5. února, Králíky, Poutní dům

Content

1. Stochastické programování
2. Influenční funkce (Influence function)
3. Bod selhání (Breakdown point)
4. Závěr a reference

Content

1 Stochastické programování

2 Influenční funkce (Influence function)

3 Bod selhání (Breakdown point)

4 Závěr a reference

Motivace
Stochastické programování

Obecná úloha stochastického programování:

$$\min_{x \in X} F(x, P), \tag{1}$$

kde $\emptyset \neq X \subseteq \mathbb{R}^n$ uzavřená a $F : \mathbb{R}^n \times \mathcal{P} \to (-\infty, \infty]$, P je **ZNÁMÉ.**

- Finance: výnos, riziko, cash-flow balance.
- Vodohospodářství: přítok, odtok, stav vody.
- Dopravní problém: nabídka, poptávka.
- Marketing: náhodná sledovanost.
- ...

Úloha stochastického programování

Optimální hodnota

$$\varphi(P) = \min_{x \in X} F(x,P) := \int_\Omega f(x,\omega)P(d\omega),\ P \in \mathcal{P} \tag{2}$$

a **množina optimálních řešení** závislé na P

$$\Psi(P) = \{x \in X :\ F(x,P) = \varphi(P)\},\ P \in \mathcal{P}, \tag{3}$$

kde $\mathcal{P}$ je množina pravděpodobnostních měr na $(\Omega, \mathcal{F})$,
$\emptyset \neq X \subseteq \mathbb{R}^n$ uzavřená a $f : \mathbb{R}^n \times \mathbb{R}^s \to (-\infty, \infty]$ měřitelná v ω.

Motivace

Úloha s náhodným parametrem

Obecná **úloha s náhodným parametrem**:

$$\min \left\{ g_0(x,\omega) \;:\; x \in X, g_k(x,\omega) \le 0,\ k = 1,\dots,K \right\}, \tag{4}$$

$\omega \in \mathbb{R}^s$ je reálný náhodný vektor na $(\Omega, \mathcal{F}, P)$ a $g_k : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}$, $k = 0,\dots,K$.

Penalizovaná účelová funkce

Integrand $f : \mathbb{R}^n \times \Omega \to \mathbb{R}$ může být například, M.B. (2008):

$$f_\mu(x,\omega) = g_0(x,\omega) + \mu \cdot \vartheta(g_1(x,\omega),\ldots,g_K(x,\omega)),\ \mu > 0,$$

kde ϑ je **penalizační funkce** $\vartheta : \mathbb{R}^K \to \mathbb{R}_+,\ j = 1,\ldots,m$, např.

$$\vartheta^1(u) = \sum_{k=1}^{K} [u_k]_+, \tag{5}$$

$$\vartheta^2(u) = \max_{1\le k\le K} [u_k]_+. \tag{6}$$

Další úlohy stochastického programování

- Úlohy s užitkovými funkcemi (očekávaný užitek).
- Některé ”mean-risk” investiční problémy.
- Dvoustupňové úlohy.

Robustifikace úlohy SP

Robustifikace úlohy stochastického programování:

1. **Exogenní**: robustní odhad parametrů rozdělení P.
2. **Endogenní**: ”robustifikace” účelové funkce a optimalizačního postupu. Míry (kritéria) robustnosti...

Exogenní robustifikace úlohy SP

Známe-li rodinu rozdělení náhodných částí:

1. Robustní odhad parametrů rozdělení P.
2. Monte-Carlo simulace $\{\omega^s\}_{s=1}^S$.
3. Řešení deterministické SAA úlohy pro *dostatečně velký* výběr S

$$\min_{x \in X} \frac{1}{S} \sum_{s=1}^{S} f(x, \omega^s) \tag{7}$$

Shapiro (2003).

Příklad: prodavač novin (Newsboy problem)

- x .. počet objednaných výtisků k prodeji,
- r .. nákupní cena,
- c .. prodejní cena, $0 < r < c$,
- ξ .. náhodná poptávka.

Prodavač S dní prodává, pozoruje poptávku ξ^s a poté se rozhodne řešit

$$\min_{x \in \mathbb{N}} (r - c)x + \frac{c}{S} \sum_{s=1}^{S} [x - \xi^s]_+. \tag{8}$$

Má nebo nemá vyloučit některé pozorování ξ^s? Má obecně používat například useknutý průměr?

Co bude zjednodušeno...

- Nebudeme se zabývat **optimálními řešeními**, ale pouze účelovou funkcí a optimální hodnotou.
- Nebudeme uvádět **empirické (výběrové) verze** problémů a měr robustnosti.

Content

1 Stochastické programování

2 Influenční funkce (Influence function)

3 Bod selhání (Breakdown point)

4 Závěr a reference

Nechť T je funkcionál na množině pravděpodonostních měr $\mathcal{P}$

$$T(\cdot) \in \left\{ \varphi(\cdot), \{F(x,\cdot)\}_{x\in X} \right\} \tag{9}$$

Nechť Δ_δ značí **Diracokovu míru** pozorování δ. Uvažujeme kontaminované rozdělení

$$(1-t)P + t\Delta_\delta,\ \ t \in [0,1].$$

Definition

Influenční funkce (IF) funkcionálu T v bodě P definujeme

$$IF(\delta; T, P) = \lim_{t \to 0_+} \frac{T((1-t)P + t\Delta_\delta) - T(P)}{t} \tag{10}$$

v takových δ, pro které limita existuje.

IF popisuje vliv nekonečně malé kontaminace rozdělení pomocí realizace δ náhodného vektoru.

Influenční funkce ve SP

Influenční funkce účelové funkce

$$IF\big(\delta, F(x,\cdot), P\big) = f(x,\delta) - F(x,P),\ \forall x \in X. \tag{11}$$

Za předpokladů, J.F. Bonnans, A. Shapiro (2000):

- sdružená spojitost F,
- kompaktnost,
- existence optimálních řešení $\Psi(P) \neq \emptyset$

platí následující vztah (”záměna minimalizace a derivace”)

$$IF(\delta,\varphi,P) = \min_{x\in\Psi(P)} f(x,\delta) - \varphi(P). \tag{12}$$

Influenční funkce ve SP

Influenční funkce účelové funkce

$$IF\big(\delta, F(x,\cdot), P\big) = f(x,\delta) - F(x,P),\ \forall x \in X. \tag{11}$$

Za předpokladů, J.F. Bonnans, A. Shapiro (2000):

- sdružená spojitost F,
- kompaktnost,
- existence optimálních řešení $\Psi(P) \neq \emptyset$

platí následující vztah ("záměna minimalizace a derivace")

$$IF(\delta, \varphi, P) = \min_{x \in \Psi(P)} f(x,\delta) - \varphi(P). \tag{12}$$

Globální citlivost (Gross-error sensitivity) T v bodě P

$$\gamma^*(T,P) = \sup_{\delta} |IF(\delta;T,P)|, \tag{13}$$

kde supremum bereme přes všechna δ, pro které $IF(\delta;T,F)$ existuje.

Lokální citlivost (Local shift sensitivity) T v bodě P

$$\lambda^*(T,P) = \sup_{\delta\neq\delta'} \frac{|IF(\delta;T,P) - IF(\delta';T,P)|}{|\delta-\delta'|}. \tag{14}$$

Globální citlivost (Gross-error sensitivity) T v bodě P

$$\gamma^*(T,P) = \sup_{\delta} |IF(\delta; T, P)|, \tag{13}$$

kde supremum bereme přes všechna δ, pro které $IF(\delta; T, F)$ existuje.

Lokální citlivost (Local shift sensitivity) T v bodě P

$$\lambda^*(T,P) = \sup_{\delta \neq \delta'} \frac{|IF(\delta; T, P) - IF(\delta'; T, P)|}{|\delta - \delta'|}. \tag{14}$$

Příklad: Lineární regrese

Teoretický problém lineární regrese může být formulován následovně

$$\min_{\beta} \mathbb{E}(Y - X^T\beta)^2,$$

kde Y a X mají konečné druhé momenty. Nechť je varianční matice $var(X)$ regulární, potom pro optimální řešení a optimální hodnotu platí, Anděl (2007):

$$\hat{\beta} = (var(X))^{-1} cov(X, Y)$$
$$\varphi = var(Y) - cov(Y, X)(var(X))^{-1} cov(X, Y).$$

Snadno poté spočteme influenční funkce

$$\begin{aligned} IF(\delta, F(\beta, \cdot)P) &= (\delta_Y - \delta_X^T\beta)^2 - \mathbb{E}(Y - X^T\beta)^2, \\ IF(\delta, \varphi, P) &= (\delta_Y - \delta_X^T\hat{\beta})^2 - \varphi(P) \end{aligned}$$

Příklad: L-odhady ve SP

Nechť $\omega \in \mathbb{R}$ (úroková míra) s distribuční funkcí G_P a funkce h splňuje

$$\int_0^1 h(y)dy = 1,$$

např. $h(y) = \frac{1}{1-2\alpha} I_{[\alpha,1-\alpha]}(y)$. Potom definujeme L-robustní účelovou funkci

$$F^L(x,P) = \int_{-\infty}^{\infty} f(x,\omega)h(G_P(\omega))dG_P(\omega) \tag{15}$$

Příklad: L-odhady ve SP

Za předpokladů

(A1) Distribuční funkce G_P je *striktně rostoucí*,

(A2) $f(x,\cdot)$ je *diferencovatelná* pro všechna $x \in X$ (problematický ve SP)

dostáváme, Jurečková (2001): $IF(\delta; T, P) =$

$$= \int_{-\infty}^{\infty} G_P(\omega) \frac{d\ f(x,\omega)}{d\omega} h(G_P(\omega)) dy - \int_{\delta}^{\infty} \frac{d\ f(x,\omega)}{d\omega} h(G_P(\omega)) d\omega.$$

Content

1. Stochastické programování
2. Influenční funkce (Influence function)
3. Bod selhání (Breakdown point)
4. Závěr a reference

Bod selhání

P.L. Davies, U. Gather (2005): Nechť pro **pravděpodobnostní metriku** d_{PM} platí

$$\sup_{P,Q\in\mathcal{P}} d_{PM}(P,Q) = 1$$

Potom **bod selhání** je definován jako

$$\varepsilon^*(T,P) = \inf\left\{\varepsilon > 0 : \sup_{d_{PM}(P,Q)<\varepsilon} |T(P) - T(Q)| = \infty\right\}$$

Prohorovova metrika

Prohorovova vzdálenost dvou pravděpodobnostních měr P and Q in $\mathcal{P}(\mathcal{X})$:

$$d_P(P, Q) = \inf\{\varepsilon : \; P(A) \leq Q(A^\varepsilon) + \varepsilon \text{ for all events } A\},$$

kde $A^\varepsilon = \{x \in \mathbb{R} : d(A, x) < \varepsilon\}$. Důvody vedoucí na Prohorovovu metriku

- Zaokrouhlování pozorování (zaokrouhlovací chyba).
- Část dat může pocházet ze zcela jiných rozdělení.
- Model je pouze aproximací skutečného problému (slabá konvergence).

Omezená lipschitzovská metrika

Další metrika, která metrizuje slabou konvergenci, je **omezená lipschitzovská metrika**

$$d_{BL}(P,Q) = \sup\left\{ \left| \int_\Omega f(\omega)P(d\omega) - \int_\Omega f(\omega)Q(d\omega) \right| : \right.$$
$$\left. |f(\omega_1) - f(\omega_2)| \leq \tilde{d}(\omega_1,\omega_2), \forall \omega_1,\omega_2 \in \Omega \right\},$$

kde

$$\tilde{d}(\omega_1,\omega_2) = \frac{d(\omega_1,\omega_2)}{1 + d(\omega_1,\omega_2)} \leq 1,\ \forall \omega_1,\omega_2 \in \Omega,$$

kde d je metrika na $\mathbb{R}^s$.

Metrická stabilita

Pro $P, Q \in \mathcal{P}$

$$|\varphi(P) - \varphi(Q)| \leq L \cdot d_{MI}(P, Q) \leq \tilde{L} \cdot d_I(P, Q) \tag{16}$$

pro nějaké $L, \tilde{L} > 0$ a vhodnou "minimal information" (pseudo)metriku d_{MI} a **ideální metriku** d_I, např. **Kantorovichovu metriku**

$$\zeta_1(P, Q) = \sup \left\{ \left| \int_\Omega f(\omega) P(d\omega) - \int_\Omega f(\omega) Q(d\omega) \right| : \right.$$
$$\left. |f(\omega_1) - f(\omega_2)| \leq d(\omega_1, \omega_2), \forall \omega_1, \omega_2 \in \Omega \right\}$$

Pro $P, Q \in \mathcal{P}$ platí následující vztahy, A.W. van der Vaart (1998):

$$d_P(P,Q)^2 \leq d_{BL}(P,Q) \leq 2d_P(P,Q).$$

a zřejmě

$$d_{BL}(P,Q) \leq \zeta_1(P,Q).$$

Content

1. Stochastické programování
2. Influenční funkce (Influence function)
3. Bod selhání (Breakdown point)
4. Závěr a reference

Co už je hotové...

J. Dupačová (...):

- Asymptotika.
- Chování M-odhadů z hlediska (parametrické a stochastické) optimalizace.
- Kontaminační techniky.
- Minimax

$$\min_{x \in X} \max_{P \in \tilde{\mathcal{P}}} F(x, P), \tag{17}$$

kde $\tilde{\mathcal{P}} \subset \mathcal{P}$.

Co dále...

- Výběrové verze a výpočty.

Reference

- **J. Anděl** (2007). *Základy matematické statistiky* Matfyzpress, Praha.
- **J.F. Bonnans, A. Shapiro** (2000). *Perturbation Analysis of Optimization Problems*. Springer Series in Operations Research and Financial Engineering, Springer-Verlag, New York.
- **M. Branda** (2008). *Recourse reformulation of chance constrained problems*. Proceedings of 15th Summer School ROBUST 2008 (J. Antoch, G. Dohnal eds).
- **P.L. Davies, U. Gather** (2005). *Breakdown and Groups*. The Annals of Statistics, Vol. 33, No. 3, 977-1035.
- **P.L. Davies, U. Gather** (2005). *The Breakdown Point - Examples and Counterexamples*. REVSTAT - Statistical Journal, Vol. 5, No. 1, 1-17.

Reference

- **J. Dupačová** (1986). *Stability in stochastic programming with recourse. Contaminated distributions*. Mathematical Programming Studies, Volume 27, 133-144.
- **J. Dupačová** (1990). *Stability and sensitivity-analysis for stochastic programming*. Annals of Operations Research 27, Issue 1-4, 115-142.
- **P.J. Huber** (1981). Robust Statistics. John Wiley and Sons.
- **F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw and W.A. Stahel** (1986). *Robust Statistics - The Approach Based on Influence Functions*. John Wiley and Sons.
- **J. Jurečková** (2001). *Robustní statistické metody*. Nakladatelství Karolinum.

Reference

- **W. Römisch** (2003). *Stability of Stochastic Programming Problems*. In Stochastic Programming (A. Ruszczynski and A. Shapiro eds.), Handbook in Operations Research and Management Science Vol. 10, Elsevier, Amsterdam, 483-554.
- **A. Shapiro** (2003). *Monte Carlo Sampling Methods*. In Stochastic Programming (A. Ruszczynski and A. Shapiro eds.), Handbook in Operations Research and Management Science Vol. 10, Elsevier, Amsterdam, 483-554.
- **A.W. van der Vaart** (1998). *Asymptotic statistics*. Cambridge University Press.

Děkuji za pozornost.