The Moment Problem

Jiayi Guo

Shanghai University of Finance and Economics

Distributionally Robust Newsvendor

A newsvendor needs to decide on the number of units of an item to order before the actual demand is observed.

- c: unit purchase cost
- $p > c > 0$: unit revenue
- zero salvage value
- $\tilde{d}$: random demand which ambiguous distribution $F(\cdot)$ lie a set of possible distributions $\mathcal{F}$.
- Solve:
$$\max_{q\in\mathbb{R}_+} \inf_{F\in\mathcal{F}} \left(p\mathbb{E}_F[\min(q,\tilde{d})] - cq\right)$$
- Let $\eta = 1 - c/p \in [0,1)$ denote the critical ratio. Solve:
$$\min_{q\in\mathbb{R}_+} \sup_{F\in\mathcal{F}} \left(\mathbb{E}_F[\tilde{d} - q]^+ + (1-\eta)q\right)$$

Scarf's Model

Set of demand distributions in Scarf's model (1958) is:

$$\mathcal{F}_{1,2} = \left\{ F \in \mathbb{M}(\Re_+) : \int_0^\infty dF(w) = 1, \int_0^\infty w dF(w) = m_1, \int_0^\infty w^2 dF(w) = m_2 \right\}$$

- Given an order quantity $q > m_2/2m_1$, the worst-case demand distribution for $\sup_{F \in \mathcal{F}_{1,2}} \mathbb{E}_F[\tilde{d} - q]_+$ is two-point:

$$\tilde{d}_q = \begin{cases} q - \sqrt{q^2 - 2m_1 q + m_2}, & \text{w.p. } \frac{1}{2}\left(1 + \frac{q - m_1}{\sqrt{q^2 - 2m_1 q + m_2}}\right) \\ q + \sqrt{q^2 - 2m_1 q + m_2}, & \text{w.p. } \frac{1}{2}\left(1 - \frac{q - m_1}{\sqrt{q^2 - 2m_1 q + m_2}}\right) \end{cases}$$

- Given an order quantity in the range $0 \le q \le m_2/2m_1$, the worst-case distribution is two-point but fixed and given by $\tilde{d}_{m_2/2m_1}$, where

$$\tilde{d}_{m_2/2m_1} = \begin{cases} 0, & \text{w.p. } 1 - \frac{m_1^2}{m_2}, \\ \frac{m_2}{m_1}, & \text{w.p. } \frac{m_1^2}{m_2} \end{cases}$$

Scarf's Model

- The worst-case bound is given as:

$$\sup_{F \in \mathcal{F}_{1,2}} \mathbb{E}_f[\tilde{d} - q]_+ = \begin{cases} \frac{1}{2}\left(\sqrt{q^2 - 2m_1 q + m_2} - (q - m_1)\right), & \text{if } q > \frac{m_2}{2m_1} \\ m_1 - \frac{q m_1^2}{m_2}, & \text{if } 0 \leq q \leq \frac{m_2}{2m_1} \end{cases}$$

- A closed form solution for the optimal order quantity is:

$$q_\eta^{\text{scarf}} = \begin{cases} m_1 + \frac{\sqrt{m_2 - m_1^2}}{2} \frac{2\eta - 1}{\sqrt{\eta(1-\eta)}}, & \text{if } \frac{m_2 - m_1^2}{m_2} < \eta < 1 \\ 0, & \text{if } 0 \leq \eta < \frac{m_2 - m_1^2}{m_2} \end{cases}$$

Optimization under Distribution Uncertainty

- Stochastic Optimization: $\min\{E_\xi f(x,\xi) \mid x \in X\}$
- Distributions in practice are often not known!
- Robustness under Distribution Uncertainty: $\min_{x\in X} \max_{\xi\in\mathbb{F}} E_\xi f(x,\xi)$
- Research Objective: to solve $\max_{\xi\in\mathbb{F}} E_\xi f(x,\xi)$ and $\min_{x\in X} \max_{\xi\in\mathbb{F}} E_\xi f(x,\xi)$?
- Analytical form solution for $\max_{\xi\in\mathbb{F}} E_\xi f(x,\xi)$ is useful in solving $\min_{x\in X} \max_{\xi\in\mathbb{F}} E_\xi f(x,\xi)$
- Numerical solution for $\max_{\xi\in\mathbb{F}} E_\xi f(x,\xi)$ can help solving $\min_{x\in X} \max_{\xi\in\mathbb{F}} E_\xi f(x,\xi)$, when x is univariate, or the function is convex or continuous in x

The Moment Problem

Let's first focus on $\max_{\xi \in \mathbb{F}} E_\xi f(x, \xi)$, where the distribution class $\mathbb{F}$ has been described by expectations of functions:

$$\max\{E_\xi f(\xi) \mid E_\xi f_i(\xi) = M_i, \text{ for all } i \in I\}$$

where ξ is the random variable, $f(\xi)$ is defined as $f(x, \xi)$ for the given x, and I is the index set of functions f_i.

The Moment Problem: the distribution class $\mathbb{F}$ in $\min_{\xi \in \mathbb{F}} E_\xi f(x, \xi)$ is governed by the moments information:

$$\max\{E_\xi f(\xi) \mid E_\xi \xi^i = M_i, \text{ for all } i \in I\}$$

s

Distributional Robust as Infinite LP

Let's focus on how to reform each "for-all" type constraint for given x, t, using the f part as example:

$$
\begin{array}{rcll}
Z_P^1 & = & \max_{F(\cdot)} & \int_y f(y) \cdot dF(y) \\
 & & \text{s.t.} & \int_y 1 \cdot dF(y) = 1 \\
 & & & \int_y f_i(y) \cdot dF(y) = M_i \text{ for all } i \in I \\
 & & & dF(y) \geq 0
\end{array} \tag{1}
$$

Observation: This is an infinite dimensional LP! (By treating $dF(y)$ as variables with y as indices.)

The Dual Form and Strong Duality

The dual problem is

$$
\begin{aligned}
Z_D^1 \quad = \quad \min \quad & z_0 + \textstyle\sum_{i \in I} M_i z_i \\
\text{s.t.} \quad & h(y) := z_0 + \textstyle\sum_{i \in I} z_i f_i(y) \geq f(y), \quad \forall y \in \mathbb{R}.
\end{aligned}
\tag{2}
$$

Theorem (Bertsimas and Popescu,05)

Strong Duality holds when slater condition holds, i.e., there exists a distribution with positive density everywhere satisfying the expectation (moment) conditions!

Complementary Slackness Condition

Suppose F is a primal feasible solution, and z is a dual feasible solution. If strong duality holds for the problem, then (F, z) is a pair of primal-dual optimal solution if and only if the complementary slackness condition holds:

$$dF(y)\left[z_0 + \sum_{i \in I} z_i f_i(y) - f(y)\right] = 0, \quad \forall y \in \mathbb{R}.$$

Further Analysis

Intuition: Optimal primal-dual solutions for non-degenerate linear programming problems can always be obtained at **Basic Feasible Solutions.**

Theorem (Bertsimas and Popescu,05)

Suppose strong duality holds for the problem. Therefore, optimal solution for the primal problem can be obtained at a discrete distribution with at most $k+1$ points.

Sum of Squares: Numerical Approach for Dual Moment Problem

The dual problem (4)

$$
\begin{aligned}
Z_D^1 \quad = \quad & \min \quad z_0 + \textstyle\sum_{i \in I} M_i z_i \\
& \text{s.t.} \quad h(y) := z_0 + \textstyle\sum_{i \in I} z_i f_i(y) \geq f(y), \quad \forall y \in \mathbb{R}.
\end{aligned}
$$

1. SOS(Sum of Squares): Any univariate nonnegative polynomial function is the sum of square of two real polynomial functions.
2. The SOS condition can be presented by SDP constraints.

SOS

Theorem (Sum of Squares)

Any univariate nonnegative polynomial function is the sum of squares of two real polynomial functions.

1. Fundamental Theorem of Algebra:

   $f(x) = \prod_{j=1}^{n-2k}(x - r_j)^{m_j} \prod_{l=1}^{k}\left[(x - z_l)^{q_l}(x - \bar{z}_l)^{q_l}\right]$

2. $f(\cdot)$ is nonnegative if and only if all multipliers (m_j) for the real roots are even numbers

3. Complex terms are sum of squares: $(x - z_l)(x - \bar{z}_l) = (x - a_l)^2 + b_l^2$

4. Multiple of sum of two squares

$$(u^2 + v^2)(w^2 + y^2) = (uw + vy)^2 + (uy - vw)^2$$

SDP Representation

1. One can representing real function $f(x) = \sum_{j=1}^{n+1} c_{j-1}x^{j-1}$ as $c^T v_x$, where $v_x = (1, x, x^2, \cdots, x^n)^T$.
2. Since $f(x)^2 = v_x^T cc^T v_x$, a square of order n real function can always be represented as $v_x^T S v_x$ with $S \succeq 0$.
3. So is the sum of squares of real function!

Theorem (SDP Representation)

$$\sum_{i=0}^{2n} a_i x^i \geq 0 \forall x \in R \Longleftrightarrow \exists S \in \mathbb{S}_+^{n+1}, \sum_{j=1}^{i+1} S_{j,i+2-j} = a_i \forall i = 0, 1, \cdots, 2n.$$

Multivariate SOS Approach

The multivariate problem is related to Hilbert's 17th problem:

Theorem (Arkin, 1927)

Nonnegative multivariate real polynomial function is always the sum of squares of fraction of real polynomials:

$$f(x) \geq 0 \text{ for all } x \in \Re \Longleftrightarrow f(x) = \sum_{i=1} \left[\frac{f_i(x)}{q(x)} \right]^2$$

For numerical approach of polynomial constrained polynomial optimization problems, please refer to the SOS hierarchies designed by J. Lasserre., P. Parillo., J. Nie et. al.

Applications in OM

1. (Delage and Ye, 10) Portfolio Selection
2. (Du, Han and Zuluaga, 12) Newsvendor: Improved Scarf Bound
3. (MaK, Rong and Zhang, 14) Appointment Scheduling
4. (Wang and Zhang, 15) Process Flexibility
5. (Li, Jiang and Mathieu, 16) Power Control
6. (Natarajan, Sim and Uichanco, 18) Newsvendor: asymmetric distribution
7. (Das, Dhara, Natarajan, 18) Newsvendor: heavy tail demand distribution
8. (Xin and Goldberg, 20) Inventory Control – time (in)consistency

Methodologies: Closed-form solution, or Semi-definite Programming

Applications: Probability and Distributional Inequalities

1. Markov Inequality: $P(\xi \geq a) \leq \frac{a}{E\xi}$ if $\xi \geq 0$.
2. (one-sided) Chebyshev inequality: $P(\xi \geq E\xi + t\sigma_\xi) \leq \frac{1}{1+t^2}$.
3. Zelen's Inequality: $Prob(\xi \geq E\xi + t\sigma) \leq \left(1 + t^2 + \frac{(t^2 - t\kappa_3 - 1}{\kappa_4 - \kappa_3^2 - 1}\right)^{-1}$ for $t \geq \frac{\kappa_3 + \sqrt{\kappa^2 + 4}}{2}$, where $\kappa_m = M_m/\sigma^m$.
4. Cantelli's Inequality: $P(|\xi - E\xi| \geq a) \leq \frac{v_{2m} - v_m^2}{v_{2m} - v_m^2 + (a^m - v_m)^2}$ for $a^m \geq \frac{v_{2m}}{v_m}$, where $v_m = E(\xi - E\xi)^m$.
5. Scarf's Bound: $E(r - \xi)_+ \leq \frac{r - \mu + \sqrt{\sigma^2 + (r-\mu)^2}}{2}$

Moment Problem Approach: Improved Inequalities

1. (Bertismas and Popescu, 05): Tight bounds for $P(\xi \geq (1+\delta)E\xi)$ and $P(|\xi| \geq (1+\delta)E\xi)$ with different kinds of information.
2. (Berge, 09): $E|\xi| \geq \frac{3\sqrt{3}}{2\sqrt{q}}\left(E\xi^2 - \frac{E\xi^4}{q}\right)$ for all $q > 0$.

Probability Bound with 4-th Order Moment Information

$$
\begin{aligned}
Z_P^1 = \max \quad & \int_{x \geq 0} dF(x) \\
\text{s.t.} \quad & \int_{x \in \mathbb{R}} dF(x) = 1, \\
& \int_{x \in \mathbb{R}} x^i dF(x) = M_i, i = 1, 2, 4. \\
& dF(x) \geq 0.
\end{aligned}
$$

$$
\begin{aligned}
Z_D^1 = \min \quad & y_0 + y_1 M_1 + y_2 M_2 + y_4 M_4 \\
\text{s.t.} \quad & g(x) \doteq y_0 + y_1 x + y_2 x^2 + y_4 x^4 \geq 1_{x \geq 0} \forall x \in \mathbb{R}.
\end{aligned}
$$

Closed Form Dual Feasible Solution and Upper Bound

Observing from the dual constraint, the optimal distribution is of at most 3 point. By constructing the primal-dual optimal pair, we obtain the following:

Theorem (He, Zhang and Zhang,09)

$$P(\xi \geq 0) \leq 1 - \frac{4}{9}(2\sqrt{3} - 3) \max_{v>0} \left\{ (-\frac{2E\xi}{v} + \frac{3E\xi^2}{v^2} - \frac{E\xi^4}{v^4} \right\}.$$

$$P(\xi \geq E\xi) \leq 1 - (2\sqrt{3} - 3)\frac{1}{\kappa_4}, \ \textit{where}\ \kappa_4 = \frac{E(\xi - E\xi)^4}{\sigma_\xi^4}$$

The Tight Result

THEOREM 2.2.

$$\operatorname{Prob}\{X \geq 0\} \leq \begin{cases} 1 - \dfrac{M_1^2}{M_2}, & \text{if } \dfrac{M_4}{M_2^2} \geq \dfrac{M_2}{M_1^2} \text{ and } M_1 < 0, \\ 1, & \text{if } \dfrac{M_4}{M_2^2} \geq \dfrac{M_2}{M_1^2} \text{ and } M_1 > 0, \\ 1 - \dfrac{4(2\sqrt{3}-3)}{9} \sup\limits_{v>0} \left(-\dfrac{2M_1}{v} + 3\dfrac{M_2}{v^2} - \dfrac{M_4}{v^4} \right), & \text{if } \dfrac{M_4}{M_2^2} < \dfrac{M_2}{M_1^2} \text{ and } \alpha \geq \dfrac{\sqrt{3}-1}{2} V_{\min}, \\ \dfrac{1}{2} + \dfrac{1}{2} \dfrac{\alpha + 2M_1}{\sqrt{4M_2 + \alpha^2 + 4M_1\alpha}}, & \text{if } \dfrac{M_4}{M_2^2} < \dfrac{M_2}{M_1^2} \text{ and } \alpha < \dfrac{\sqrt{3}-1}{2} V_{\min}, \end{cases} \tag{13}$$

where $\alpha \triangleq \sqrt{(M_4 - M_2^2)/(M_2 - M_1^2)}$, $V_{\min} \triangleq \sqrt{(\sqrt{3}-1)M_2 + ((7-4\sqrt{3})/4)M_1^2} + ((2-\sqrt{3})/2)M_1$. *Furthermore, the bound is tight; i.e., there exists an X such that the inequality (13) holds as an equality.*

Bivariate Moment Problem

$$v_P(\theta; q) = \sup_{\mathbb{P} \in \mathcal{F}(\theta)} \mathbb{E}_{\mathbb{P}} \left[(X_1 + X_2 - q)_+ \right]$$

where the ambiguity set is

$$\mathcal{F}(\theta) = \left\{ \mathbb{P} \in \mathbb{M}(\mathbb{R}^2_+) \;\middle|\; \begin{array}{ll} \mathbb{E}_{\mathbb{P}}[X_1] = \mu_1, & \mathbb{E}_{\mathbb{P}}[X_2] = \mu_2, \\ \mathbb{E}_{\mathbb{P}}\left[X_1^2\right] = a\mu_1^2, & \mathbb{E}_{\mathbb{P}}\left[X_2^2\right] = b\mu_2^2, \\ \mathbb{E}_{\mathbb{P}}[X_1 X_2] = c\mu_1\mu_2 & \end{array} \right\}.$$

Assumption

We assume $a > 1$, $b > 1$, $c \geq 0$ *and* $(a-1)(b-1) \geq (c-1)^2$.

Closed-form Optimal Value

Theorem

Given a non-empty ambiguity set $\mathcal{F}(\theta)$ and $q > 0$, the optimal value $v_P(q;\theta)$ can be characterized as

$$
v_P(q;\theta) = \begin{cases}
\mu_1 + \mu_2 - q \cdot \dfrac{a+b-2c}{ab-c^2} & \text{if Condition 1 holds;} \\[2mm]
\dfrac{b-1}{2b}\left((q+Q_b) - \dfrac{b-c}{b-1}\mu_1\right) + \mu_1 + \mu_2 - q & \text{if Condition 2 holds;} \\[2mm]
\dfrac{a-1}{2a}\left((q+Q_a) - \dfrac{a-c}{2a}\mu_2\right) + \mu_1 + \mu_2 - q & \text{if Condition 3 holds;} \\[2mm]
\dfrac{b-1}{2b}\left(\dfrac{b-c}{b-1}\mu_1 - (q-Q_b)\right) & \text{if Condition 4 holds;} \\[2mm]
\dfrac{a-1}{2a}\left(\dfrac{a-c}{a-1}\mu_2 - (q-Q_a)\right) & \text{if Condition 5 holds;} \\[2mm]
\dfrac{1}{2}(Q_c - q + \mu_1 + \mu_2) & \text{if Condition 6 holds.}
\end{cases}
$$