分布不确定集下的DRO问题转化思路

舒晴

2025 年 8 月 5 日

1 分布不确定集

首先介绍三种不同形式的分布不确定集P：

1. 最一般的概率分布集合：$P = \{P \in \mathcal{P}(Z)\}$
2. 具有给定期望的分布集合：$P = \{P \in \mathcal{P}(Z) : \mathbb{E}[Z] = \mu\}$
3. 具有给定一阶和二阶矩的分布集合：$P = \{P \in \mathcal{P}(Z) \mid \mathbb{E}_P[Z] = \mu, \mathbb{E}_P[ZZ^T] = M\}$

2 矩分布不确定集（Moment Ambiguity Sets）

矩分布不确定集是通过有限个（广义）矩条件定义的分布集合。其数学形式为：

$$\mathcal{P} = \{P \in \mathcal{P}(Z) : \mathbb{E}_P[f(Z)] \in \mathcal{F}\},$$

其中：

- $f : Z \to \mathbb{R}^m$ 是一个Borel可测的矩函数（例如，一阶矩是均值，二阶矩是方差）
- $\mathcal{F} \subseteq \mathbb{R}^m$ 是一个不确定性集合，用于约束矩的取值（例如，均值和方差的取值范围）
- $\mathbb{E}_P[f(Z)]$ 是分布 P 下 $f(Z)$ 的期望（即广义矩）

意义：
矩模糊集包含所有支撑在 Z 上且满足矩条件 $\mathbb{E}_P[f(Z)] \in \mathcal{F}$ 的概率分布。例如：

- 若 $f(Z) = Z$ 且 $\mathcal{F} = [\mu_{\min}, \mu_{\max}]$，则分布不确定集包含所有均值在 $[\mu_{\min}, \mu_{\max}]$ 内的分布

3 仅支撑分布不确定集（Support-Only Ambiguity Sets）

这是矩分布不确定集的一个特例，定义为：

$$\mathcal{P} = \mathcal{P}(Z),$$

即所有支撑在 Z 上的概率分布。它对应矩分布不确定集中：

- 矩函数 $f(z) = 1$（常数函数）
- 不确定性集合 $\mathcal{F} = \{1\}$

意义：
此时，DRO问题退化为经典的鲁棒优化（Robust Optimization）问题：

$$\inf_{x \in \mathcal{X}} \sup_{P \in \mathcal{P}(Z)} \mathbb{E}_P[\ell(x, Z)] = \inf_{x \in \mathcal{X}} \sup_{z \in Z} \ell(x, z).$$

- 左边是DRO问题：在最坏分布 P 下最小化期望损失
- 右边是鲁棒优化问题：直接在最坏情况 $z \in Z$ 下最小化损失 $\ell(x, z)$

4 DRO问题的一般形式

考虑矩分布不确定集下的DRO问题：

$$\inf \sup \mathbb{E}_P[\ell(z; \xi)] \tag{1}$$

可以转化为：

$$\inf_{x \in \mathcal{X}} \sup_{p \in \mathcal{P}(Z)} \mathbb{E}_p[\ell(x, Z)] = \inf_{x \in \mathcal{X}} \sup_{z \in Z} \ell(x, z) \tag{2}$$

5 内层问题的转化

引入辅助函数h定义如下：

$$h(u_0, u) = \inf_{\nu \in \mathcal{M}_+(Z)} \left\{ \int \ell(z) d\nu(z) : \int d\nu(z) = u_0, \int f(z) d\nu(z) = u \right\} \quad (3)$$

考虑将内层最大化问题转化为最小化问题，从而进行求解：

$$\sup_{P \in P} \mathbb{E}_P[\ell(Z)] = \sup_{u \in F} \sup_{P \in \mathcal{P}(Z)} \{\mathbb{E}_P[\ell(Z)] : \mathbb{E}_P[f(Z)] = u\} \quad (4)$$

$$= \sup_{u \in F} -h(1, u) \quad (5)$$

6 关键引理和定理

引理 1 (最优值函数的凸性). 如果H是凸函数，则$h(u) = \inf_{v \in V} H(u, v)$是凸函数。

引理 2 (Fenchel-Moreau定理). 对于任何凸函数$h : \mathbb{R}^d \to \mathbb{R}$，有$h \geq h^{**}$，且在$rint(dom(h))$上等式成立。

引理 3 (h函数的定义域). 若对任意 $P \in \mathcal{P}(Z)$ 均有 $\mathbb{E}_P[l(Z)] > -\infty$，则成立

$$dom(h) = cone(\{1\} \times C)$$

命题 1 (h的双共轭). 在$\{1\} \times rint(C)$生成的锥上（除原点外）h^{**}与h一致。

$$h^{**}(\mu_0, \mu) = \sup\{-\mu_0\lambda_0 - \mu^T\lambda : \lambda_0 + f(z)^T\lambda \leq \ell(z)\ \forall z\} \quad (6)$$

命题 2 (推论). 对任意 $u \in rint(C)$，有 $h(1, u) = h^{**}(1, u)$。

命题1的详细证明如下：

证明. 对于任意固定的$(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m$，由定义$h$的凸共轭：

$$h^*(-\lambda_0, -\lambda) = \sup_{u_0 \in \mathbb{R}, u \in \mathbb{R}^m} \left\{ -u_0\lambda_0 - u^\top \lambda - h(u_0, u) \right\} \tag{7}$$

由$h(u_0, u) = \inf_{\nu \in \mathcal{M}_+(Z)} \int \ell(z) d\nu(z)$，上式

$$= \sup_{\substack{u_0 \in \mathbb{R}, u \in \mathbb{R}^m \\ v \in \mathcal{M}_{f,+}(\mathcal{Z})}} \left\{ -u_0\lambda_0 - u^\top \lambda + \int_{\mathcal{Z}} \ell(z) dv(z) \right\} \tag{8}$$

其中约束条件：

$$\int_{\mathcal{Z}} dv(z) = u_0 \tag{9}$$

$$\int_{\mathcal{Z}} f(z) dv(z) = u \tag{10}$$

将其代入(8)式并整理有：

$$= \sup_{v \in \mathcal{M}_{f,+}(\mathcal{Z})} \int_{\mathcal{Z}} \left(\ell(z) - \lambda_0 - f(z)^\top \lambda \right) dv(z) \tag{11}$$

由于$\mathcal{M}_{f,+}(Z)$ 包含 Z 上所有加权的Dirac测度，v在此处可以视为一个正向的向量。因此当$\ell(z) - \lambda_0 - f(z)^\top \lambda > 0$时，上式可取值到无穷大，故上式：

$$= \begin{cases} 0, & \text{若 } \ell(z) - \lambda_0 - f(z)^\top \lambda \leq 0, \ \forall z \in \mathcal{Z} \\ \infty, & \text{否则} \end{cases} \tag{12}$$

□

接下来我们再计算h^*的共轭函数h^{**}。对于任意固定的$(u_0, u) \in \mathbb{R} \times \mathbb{R}^m$，$h^*$的共轭函数满足

$$h^{**}(u_0, u) = \sup_{\lambda_0 \in \mathbb{R}, \, \lambda \in \mathbb{R}^m} \left\{ -u_0\lambda_0 - u^\top \lambda - h^*(-\lambda_0, -\lambda) \right\} \tag{13}$$

注意这里取上确界时只能考虑(12)式中的第一种情况，第二种情况中无穷的上确界不存在。

$$\begin{cases} h^*(-\lambda_0, -\lambda) = 0 \\ \ell(z) - \lambda_0 - f(z)^\top \lambda \le 0 \end{cases}$$

将条件带入(13)式得到：

$$= \sup_{\lambda_0 \in \mathbb{R},\, \lambda \in \mathbb{R}^m} \left\{ -u_0 \lambda_0 - u^\top \lambda : \lambda_0 + f(z)^\top \lambda \ge \ell(z), \forall z \in Z \right\} \tag{14}$$

7 问题转化

接下来我们将(4)式左侧的极大化问题转化为等价的极小化问题：

定理 1 (矩分布不确定集的对偶理论). *若$\mathcal{P}$是矩分布不确定集(2.1)，则成立如下弱对偶关系：*

$$\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}[\ell(Z)] \le \begin{cases} \inf & \lambda_0 + \delta^*_{\mathcal{F}}(\lambda) \\ s.t. & \lambda_0 \in \mathbb{R},\ \lambda \in \mathbb{R}^m \\ & \lambda_0 + f(z)^\top \lambda \ge \ell(z), \quad \forall z \in \mathcal{Z} \end{cases} \tag{15}$$

证明. 为简化表述，引入对偶可行集的简记：

$$\mathcal{L} = \left\{ (\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m : \lambda_0 + f(z)^\top \lambda \ge \ell(z),\ \forall z \in Z \right\} \tag{16}$$

基于(5)式，可得：

$$\sup_{P \in \mathcal{P}} \mathbb{E}_P[\ell(Z)] = \sup_{u \in \mathcal{F}} -h(1, u) \tag{17}$$

由命题二， $h(1,u) = h^{**}(1,u)$ ，带入得到(17)式：

$$\le \sup_{u \in \mathcal{F}} \inf_{(\lambda_0, \lambda) \in \mathcal{L}} \left(\lambda_0 + u^\top \lambda \right) \tag{18}$$

由于“先取上确界再取下确界”的值大于等于“先取下确界再取上确界的值”，(18)式：

$$\le \inf_{(\lambda_0, \lambda) \in \mathcal{L}} \sup_{u \in \mathcal{F}} \left(\lambda_0 + u^\top \lambda \right) \tag{19}$$

此处引入支撑函数 $\delta^*_{\mathcal{F}}(\lambda)$ ，定义如下：

$$\delta_{\mathcal{F}}(u) = \begin{cases} 0, & \text{如果 } \lambda \in \mathcal{F}, \\ +\infty, & \text{否则}. \end{cases}$$

其共轭函数

$$\delta^*_{\mathcal{F}}(\lambda) = \sup_{u \in \mathcal{F}} u^T \lambda - \delta_{\mathcal{F}}(u) = \sup_{u \in \mathcal{F}} u^T \lambda \tag{20}$$

带入(19)得到：

$$= \inf_{(\lambda_0, \lambda) \in \mathcal{L}} (\lambda_0 + \delta^*_{\mathcal{F}}(\lambda)) \tag{21}$$

□

这一系列转化使得原本难以直接求解的DRO问题可以转化为更易处理的对偶问题，为求解提供了理论保证。