

压缩感知中的分布鲁棒优化模型及其求解方法

李达臣

辽宁师范大学数学学院, 辽宁 大连

收稿日期: 2024年2月27日; 录用日期: 2024年3月21日; 发布日期: 2024年3月27日

摘要

压缩感知(CS)理论作为当前先采样后压缩方法的替代方法引起了很多的关注, 相关学者已经提出很多CS恢复算法如CoSaMP等, 而带噪声的压缩感知模型可以表示为 l_1 -范数问题, 由于不易判断矩阵的有限等距性(RIP)以及观测矩阵的不确定性, 考虑将该 l_1 -范数问题建模为分布鲁棒优化问题, 再借助KL-散度、函数变换以及内部最大化期望函数等方法得到分布鲁棒优化问题的等价形式。

关键词

KL-距离散度, 分布鲁棒优化, 机会约束优化, 压缩感知

Distributed Robust Optimization Model and Its Solution Method in Compressed Sensing

Dachen Li

School of Mathematics, Liaoning Normal University, Dalian Liaoning

Received: Feb. 27th, 2024; accepted: Mar. 21st, 2024; published: Mar. 27th, 2024

Abstract

Compressed sensing (CS) theory has attracted a lot of attention as an alternative to the current sampling-then-compression method. Relevant scholars have proposed many CS recovery algorithms such as CoSaMP, etc., and the noisy compressed sensing model can be expressed as an l_1 -norm problem. Since it is difficult to judge the finite isometric property (RIP) of the matrix and the uncertainty of the observation matrix, consider modeling the l_1 -norm problem as a distributed robust optimization problem, and then use KL-divergence, function transformation and internal maximization methods such as the expectation function obtain equivalent forms of distributed robust optimization problems.

Keywords

KL-Distance Divergence, Distributed Robust Optimization, Chance-Constrained Optimization, Compressed Sensing

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

压缩感知, 也被称为压缩采样或者稀疏采样, 其主要思想是利用信号中存在的冗余性, 同时进行采样和压缩。压缩感知理论从比 Nyquist 采样理论所建议的要少得多的采集测量中证明, 当信号在某些域中表现出稀疏性时, 它可以以很高的概率重建。在此领域研究的相关学者、已经提出了许多压缩感知恢复算法, 例如: Tropp [1]等提出的正交匹配追踪算法, 主要应用于语音处理、图像压缩、信号处理等领域, 具有简单易实现、在稀疏信号重构中表现良好的特点, 但对于非稀疏信号可能性能有所下降[2]; Wright [3]等提出的基追踪算法, 主要应用于信号处理、图像压缩、医学成像等领域, 可以处理非稀疏信号, 通过求解凸优化问题, 能够得到全局最优解, 但计算复杂[4]; Candès [5]等提出的压缩采样匹配追踪算法, 主要应用于通信系统、成像、生物医学工程等领域, 在高度稀疏信号的情况下表现良好, 相对于正交匹配追踪算法有更好的收敛性, 但对于非稀疏信号仍存在问题[6]; Goldstein [7]等提出的迭代硬阈值算法, 主要应用于图像重构、信号处理、通信系统等领域, 但收敛速度相对较快, 算法简单, 但受噪声干扰较大[8]。

在压缩感知理论中, 通常通过与信号相互独立的随机投影对信号进行采样, 并通过最小化 l_0 -范数或者 l_1 -范数优化问题重建, 建立该优化问题的前提是信号在某些变换域中是稀疏的[8] [9]。一般地, 使用 $m \times n$ 观测矩阵 $\phi \in R^{m \times n}$ 的 m 个行向量来投影稀疏系数为 s 的向量 $x \in R^n$, 得到观测向量 $y \in R^m$ 。压缩感知的一般模型可以写成:

$$y = \phi x, \quad (1.1)$$

其中, 当稀疏系数为 s 时可令

$$x = \psi s,$$

也就是说, 当信号稀疏时, 压缩感知模型可转化为如下形式:

$$y = \phi x = \phi \psi s. \quad (1.2)$$

考虑带噪声的稀疏系数为 s 的压缩感知问题, 可建模为最小化 l_0 -范数问题:

$$\begin{aligned} \min \quad & \|\psi s\|_0 \\ \text{s.t.} \quad & y = \phi \psi s. \end{aligned} \quad (1.3)$$

问题(1.3)是不连续的, 是一个 NP-hard 问题。

Donoho [10]等证明了在一定条件下 l_0 -范数问题与下述 l_1 -范数问题是等价的,

$$\begin{aligned} \min \quad & \|\psi s\|_1 \\ \text{s.t.} \quad & y = \phi \psi s. \end{aligned} \quad (1.4)$$

在实际测量过程中, 噪声容易干扰观测信号, 造成重建信号的过程中出现偏差, 假设观测到的带噪

声的测量向量为 y ，噪声为 e 。那么观测向量可以表示为 $y = \phi x + e$ ，由于希望噪声 e 非常小，而且噪声是客观存在无法忽视的，所以假设 e 小于或等于一个正数 ε ，则问题(1.4)可转化为如下带噪声的压缩感知信号重建模型[10]:

$$\begin{aligned} \min \quad & \|\psi s\|_1 \\ \text{s.t.} \quad & \|\phi \psi s - y\|_2 \leq \varepsilon. \end{aligned}$$

一般地，带噪声的压缩感知模型为如下形式:

$$\begin{aligned} \min \quad & \|x\|_1 \\ \text{s.t.} \quad & \|\phi x - y\|_2 \leq \varepsilon. \end{aligned} \tag{1.5}$$

其中， $\varepsilon > 0$ 代表噪声， ϕ 代表 $m \times n$ 的观测矩阵， $x \in R^n$ 表示原始信号， $y \in R^m$ 代表观测值。

关于 l_1 -范数问题的求解，近年来有很多学者对其进行研究做出了很多贡献:

1) 聂操男[11]首先将 l_1 -范数问题转化为概率约束优化模型，再通过构造特征函数 $1_{[0,+\infty)}(z)$ 的一个光滑近似函数

$$\varphi(z, t) = \begin{cases} \frac{\pi}{2}z - \frac{t}{2} + \frac{\pi t}{4} & z > 0 \\ \frac{\pi}{2}z - \frac{t}{2} \cos \frac{\pi}{t}z + \frac{\pi t}{4} & -\frac{t}{2} < z < 0 \\ 0 & z \leq -\frac{t}{2} \end{cases}$$

然后建立压缩感知中的概率约束优化问题的的光滑 D.C.近似问题(P_δ)

$$\begin{aligned} \min_{x \in X} \quad & \|x\|_1 \\ \text{s.t.} \quad & p(x, \varepsilon, \delta) \leq \alpha, \end{aligned} \tag{1.6}$$

并对(1.6)进行了收敛性分析，最后分析并求解压缩感知中的概率约束优化问题的光滑 D.C.近似问题。

2) 陈苗给[12]出求解含有 l_1 -范数问题的光滑 FR 共轭梯度法。先将 l_1 -范数问题等价转化为无约束优化问题，

$$\min_{z \in R^{2n}} f(z) = \frac{1}{2} \left\| \begin{pmatrix} (H+I)z + c \\ (H-I)z + c \end{pmatrix} \right\|^2 \tag{1.7}$$

其中 $H \in R^{n \times n}$ ， $c \in R^n$

然后对此无约束优化问题进行光滑化处理，利用绝对值函数的光滑逼近函数:

$$\phi_i(z, \mu) = \sqrt{\left((H-I)z + c \right)_i^2 + \mu^2}, \tag{1.8}$$

其中， $\mu \in R_+$ ， $i = 1, 2, \dots, 2n$ 。

且(1.8)满足

$$\lim_{\mu \rightarrow 0} \phi_i(z, \mu) = \left| \left((H-I)z + c \right)_i \right|,$$

其中， $i = 1, 2, \dots, 2n$ 。

基于(1.8)，得到(1.7)的目标函数的光滑函数

$$\tilde{f}(z, \mu) = \frac{1}{2} \sum_{i=1}^{2n} \tilde{f}^2(z, \mu),$$

其中

$$\tilde{f}(z, \mu) = ((H + I)z + c)_i - \phi_i(z, \mu), \quad i = 1, 2, \dots, 2n.$$

这样就将含有 l_1 -范数问题转化为光滑无约束优化问题，并且在一般条件下给出了 FR 共轭梯度法的全局收敛结果。

3) 屈彪等[13]将 l_1 -范数问题转化为一个分裂可行问题

$$\text{求 } x^* \in C, \text{ 使得 } Ax^* \in Q, \tag{1.9}$$

其中 $C = \{x \in R^n \mid \|x\|_1 - s \leq 0\}$, $Q = \{y \in R^m \mid y = b\}$ 。

同时也可以将 l_1 -范数问题转化为一个凸可行问题

$$\text{求 } x^* \in \left(\bigcap_{j=1}^m Q_j \right) \cap C \tag{1.10}$$

其中, $Q_j = \{x \in R^n \mid \langle a^j, b \rangle = b_j\}$, $j = 1, 2, \dots, m$, a^j 是观测矩阵第 j 列, b_j 是观测向量 b 的第 j 个分量。

在式(1.9)和(1.10)基础上, 设计了几种松弛投影算法与经典投影算法简化了投影计算量和步长计算步骤。

2. 压缩感知中的分布鲁棒约束优化模型

对于带噪声的压缩感知模型, 由于不易判断矩阵的有限等距性(RIP)以及观测矩阵的不确定性, 因此可将 l_1 -范数问题建模为分布鲁棒优化问题:

$$\begin{aligned} \min \quad & \|x\|_1 \\ \text{s.t.} \quad & \Pr_{\sim P} \left\{ \|\phi(\xi)x - y\|_2 \leq \varepsilon \right\} \geq 1 - \beta, \\ & \forall P \in \mathbb{P}. \end{aligned} \tag{2.1}$$

其中, $x \in X \subseteq R^n$ 是 n 维的决策向量, ξ 是支撑集 $\Xi \in R^k$ 上的具有概率分布 P 的随机向量, \Pr 为概率, $y \in R^m$, $\beta \in (0, 1)$ 是预先给定的置信水平。 \mathbb{P} 表示任意分布, $\varepsilon > 0$, $\phi \in R^{m \times n}$ 。

考虑特征函数

$$1_{(0, +\infty)}(z) = \begin{cases} 1, & z \in (0, +\infty), \\ 0, & z \notin (0, +\infty). \end{cases}$$

由概率约束条件 $\Pr_{\sim P} \left\{ \|\phi(\xi)x - y\|_2 \leq \varepsilon \right\} \geq 1 - \beta$ 得

$$\begin{aligned} 1 - \Pr_{\sim P} \left\{ \|\phi(\xi)x - y\|_2 \leq 0 \right\} &= \Pr_{P \in \mathbb{P}} \left\{ \|\phi(\xi)x - y\|_2 - \varepsilon \geq 0 \right\} \\ &= E_P \left[1_{(0, +\infty)} \left(\|\phi(\xi)x - y\|_2 - \varepsilon \right) \right] \leq \beta. \end{aligned}$$

从而问题(2.1)可表示成:

$$\begin{aligned} \min \quad & \|x\|_1 \\ \text{s.t.} \quad & \max_{P \in \mathbb{P}} E_P \left[1_{(0, +\infty)} \left(\|\phi(\xi)x - y\|_2 - \varepsilon \right) \right] \leq \beta \end{aligned} \tag{2.2}$$

3. 内部最大化问题的求解

假设这里已经获得一个额定分布 P_0 , 通过考虑 P_0 与未知分布 P 之间的距离来构造模糊集

$$\mathbb{P} = \{P \in D : D(P \| P_0) \leq \eta\}, \tag{3.1}$$

其中, D 表示所有概率分布的集合, 常数 η 是模糊度指数, 它控制模糊度集 P 的大小, $D(P \| P_0)$ 表示分布 P 到额定分布 P_0 的距离散度。

假设 P 相对于 P_0 是绝对连续的(记为 $P \ll P_0$), 也就是说对于每个可测集合 A , 若 $P(A) = 0$ 也就意味着 $P_0(A) = 0$ 。设分布 P 和 P_0 在支撑集 $\Xi \in R^k$ 上有概率密度函数 $p(\xi)$ 和 $p_0(\xi)$, 定义

$$L(\xi) = \frac{p(\xi)}{p_0(\xi)},$$

则 $L(\xi)$ 称为似然比, $L(\xi) > 0$, 且 $E_p[L(\xi)] = 1$, 显然, 对于测量分布 P 与额定分布 P_0 扰动或偏差问题, 似然比是一个很好的工具[14]。

考虑 KL-距离散度函数:

$$\phi_{KL}(t) = t \log t, t \geq 0,$$

则从 P 到 P_0 的 KL-距离散度为

$$D(P \| P_0) = \int_{\Xi} p(z) \log \frac{p(z)}{p_0(z)} dz,$$

显然, $D(P \| P_0) \geq 0$, 当且仅当在 P_0 上有 $p(z) = p_0(z)$ 几乎一定成立时, 等式成立。

定义一个不确定集为:

$$\mathbb{L} = \{L(\xi) \in \mathbb{L}(P_0) : E_{P_0}[L(\xi)] = 1, L(\xi) \geq 0 \text{ a.s.}\},$$

使用测度变换方法, 得

$$D(P \| P_0) = \int_{\Xi} \frac{p(z)}{p_0(z)} \log \frac{p(z)}{p_0(z)} p(z) dz = E_{P_0}[L(\xi) \log L(\xi)].$$

因此, 可以将约束函数转换为期望形式, 其中期望是关于额定分布 P_0 的。

考虑问题(2.2)的约束函数

$$\max_{P \in \mathbb{L}} E_p \left[\mathbf{1}_{(0, +\infty)} (\|\phi(\xi)x - y\|_2 - \varepsilon) \right],$$

令 $H(x, \xi) = \mathbf{1}_{(0, +\infty)} (\|\phi(\xi)x - y\|_2 - \varepsilon)$, 有

$$E_p[H(x, \xi)] = \int_{\Xi} H(x) p(\xi) d\xi = \int_{\Xi} H(x, \xi) \frac{p(\xi)}{p_0(\xi)} p(\xi) d\xi = E_{P_0}[H(x, \xi)L(\xi)].$$

问题(2.2)中的内部最大化问题可以重新表述为

$$\begin{aligned} \max \quad & E_{P_0}[H(x, \xi)L(\xi)] \\ \text{s.t.} \quad & E_{P_0}[L(\xi) \log L(\xi)] \leq \beta, \\ & L(\xi) \in \mathbb{L}. \end{aligned} \tag{3.2}$$

易证, 问题(3.2)是一个凸优化问题。

事实上, $\forall \lambda \in [0, 1]$ 和 $\forall L_1(\xi), L_2(\xi) \in \mathbb{L}$, 可以得到

$$L_\lambda(\xi) = L_1(\xi) + (1 - \lambda)L_2(\xi) \in \mathbb{L}.$$

此外, 令 $h(y) = y \log y$, 由于 $h(y)$ 是在 R^+ 上关于 y 的凸函数, 对于每个 ξ , 有

$$L_\lambda(\xi) \log L_\lambda(\xi) \leq L_1(\xi) \log L_1(\xi) + (1 - \lambda)L_2(\xi) \log L_2(\xi),$$

由此可得

$$\begin{aligned} \mathbb{E}_{P_0} [L_\lambda(\xi) \log L_\lambda(\xi)] &= \mathbb{E}_{P_0} [\lambda L_1(\xi) + (1-\lambda)L_2(\xi)] \\ &\leq \lambda \mathbb{E}_{P_0} [L_1(\xi) \log L_1(\xi)] + (1-\lambda) \mathbb{E}_{P_0} [L_2(\xi) \log L_2(\xi)] \end{aligned}$$

也就是说, $\mathbb{E}_{P_0} [L(\xi) \log L(\xi)]$ 是凸的。

类似地, 令 $g(y) = H(x, \xi)y$, 由于 $g(y)$ 是在 \mathbb{R}^+ 上关于 y 的凸函数, 那么

$$H(x, \xi)L_\lambda(\xi) \leq \lambda H(x, \xi)L_1(\xi) + (1-\lambda)H(x, \xi)L_2(\xi),$$

由此可见

$$\begin{aligned} \mathbb{E}_{P_0} [H(x, \xi)L_\lambda(\xi)] &= \mathbb{E}_{P_0} [\lambda H(x, \xi)L_1(\xi) + (1-\lambda)H(x, \xi)L_2(\xi)] \\ &\leq \lambda \mathbb{E}_{P_0} [H(x, \xi)L_1(\xi)] + (1-\lambda) \mathbb{E}_{P_0} [H(x, \xi)L_2(\xi)] \end{aligned}$$

故 $\mathbb{E}_{P_0} [H(x, \xi)L(\xi)]$ 关于 $L(\xi)$ 也是凸的, 因此问题(2.4)也是个凸优化问题。

对于每个 $x \in X$, 记 $M_H(t) = \mathbb{E}_{P_0} [e^{tH(x, \xi)}]$ 表示 $H(x, \xi)$ 关于分布 P_0 的矩母生成函数, 令 $S = \{s \in \mathbb{R} : s > 0, M_H(s) < +\infty\}$.

假设 1 [15] 对于每个 $x \in X$, S 是一个非空集。

令

$$\ell_0(\alpha, L(\xi)) = \mathbb{E}_{P_0} [H(x, \xi)L(\xi)] - \alpha [\mathbb{E}_{P_0} L(\xi) \log L(\xi) - \eta],$$

问题(3.2)等价于

$$\underset{L(\xi) \in \mathbb{L}}{\text{maximize}} \underset{\alpha \geq 0}{\text{minimize}} \ell_0(\alpha, L(\xi)). \tag{3.3}$$

交换极大极小算子的顺序, 得到(3.3)的 Lagrangian 对偶为

$$\underset{\alpha \geq 0}{\text{minimize}} \underset{L(\xi) \in \mathbb{L}}{\text{maximize}} \ell_0(\alpha, L(\xi)). \tag{3.4}$$

假设问题(3.2)满足 Slater 条件[16], 即存在 $L(\xi)$, 使得

$$\mathbb{E}_{P_0} [L(\xi) \log L(\xi)] < \eta,$$

又由于(3.2)是凸问题, 所以问题(3.3)和问题(3.4)的强对偶性成立, 因此要求解问题(3.2), 只需求解问题(3.4)即可。忽略 $\alpha\eta$, 问题(3.4)可以表示为

$$\begin{aligned} \max_{L(\xi) \in \mathbb{L}} \quad & \mathbb{E}_{P_0} [H(x, \xi)L(\xi) - \alpha L(\xi) \log L(\xi)] \\ \text{s.t.} \quad & \mathbb{E}_{P_0} [L(\xi)] = 1. \end{aligned} \tag{3.5}$$

问题(3.5)是一个凸优化问题, 设 $v(\alpha)$ 是问题(3.5)的最优值函数。

考虑两种情况: 1) $\alpha = 0$; 2) $\alpha \neq 0$ 并且 $\frac{1}{\alpha} \in S$;

1) 设 $H_u(x)$ 是测度 P_0 上 $H(x, \xi)$ 的本质确界, 即

$$H_u(x) = \inf \{t \in \mathbb{R} : \Pr_{P_0} \{H(x, \xi) > t\} = 0\},$$

然后构造一个分布序列 $P_j \ll P_0$ 集中趋于 $H_u(x)$, 进而构造 $L_j(\xi) \in \mathbb{L}^0$, 使得当 $j \rightarrow +\infty$ 时, $\mathbb{E}_{P_0} [H(x, \xi)L_j(\xi)]$ 趋于 $H_u(x)$ 。因此, $v(\alpha) = H_u(x)$ 。

2) 定义函数

$$F(L(\xi)) := E_{P_0} [H(x, \xi)L(\xi) - \alpha L(\xi) \log L(\xi)],$$

$$G(L(\xi)) := E_{P_0} [L(\xi)] - 1,$$

其中, $F(L(\xi))$ 关于 $L(\xi)$ 是凸的, $G(L(\xi))$ 关于 $L(\xi)$ 是线性的。因此可以计算函数的导数。对于在 $L(\xi)$ 任意可行方向 $V(\xi)$,

$$\begin{aligned} DF(L(\xi))V(\xi) &= \frac{F(L(\xi) + tV(\xi)) - F(L(\xi))}{t} \\ &= \lim_{t \rightarrow 0} \frac{E_{P_0} [H(x, \xi)(L(\xi) + tV(\xi)) - \alpha(L(\xi) + tV(\xi)) \log(L(\xi) + tV(\xi))] - E_{P_0} [H(x, \xi)L(\xi) - \alpha L(\xi) \log L(\xi)]}{t} \\ &= \lim_{t \rightarrow 0} E_{P_0} [H(x, \xi)V(\xi)] \\ &\quad - \alpha \lim_{t \rightarrow 0} E_{P_0} \left[\frac{E_{P_0} [(L(\xi) + tV(\xi)) \log(L(\xi) + tV(\xi)) - L(\xi) \log L(\xi)]}{t} \right] \end{aligned}$$

对于任意可行的 y 和 $V(\xi)$, 函数 $k(t) = \frac{y + tV(\xi) \log(y + tV(\xi)) - y \log y}{t}$ 是关于 t 单调的, 因此, 通过单调收敛定理[16], 可以得到

$$\begin{aligned} DF(L(\xi))V(\xi) &= E_{P_0} [H(x, \xi)V(\xi)] - E_{P_0} \left[\lim_{t \rightarrow 0} \frac{L(\xi) + tV(\xi) - \log(L(\xi) + tV(\xi)) - L(\xi) \log L(\xi)}{t} \right] \\ &= E_{P_0} [H(x, \xi)V(\xi)] - \alpha E_{P_0} [(\log L(\xi) + 1)V(\xi)] \\ &= E_{P_0} [(H(x, \xi) - \alpha(\log L(\xi) + 1))V(\xi)] \end{aligned}$$

类似地, 可以得到

$$DF(L(\xi))V(\xi) = E_{P_0} [V(\xi)],$$

问题(3.5)的拉格朗日函数如下:

$$\begin{aligned} \ell(L(\xi), \lambda) &= E_{P_0} [H(x, \xi)L(\xi)] - \alpha L(\xi) \log L(\xi) + \lambda E_{P_0} [L(\xi) - 1] \\ &= E_{P_0} [H(x, \xi)L(\xi) - \alpha L(\xi) \log L(\xi) + \lambda L(\xi)] - \lambda \end{aligned}$$

命题 3.1 [16] 如果存在 $(L^*(\xi), \lambda^*)$, 使得 $L^*(\xi) \in \mathbb{L}^0$, $G(L^*(\xi)) = 0$ 并且

$$L^*(\xi) \in \arg \max_{L(\xi) \in \mathbb{L}^0} \ell(L(\xi), \lambda^*) \tag{3.6}$$

因此这个问题简化为解决问题(3.6), 则 $L^*(\xi)$ 是问题(3.5)的最优解。问题(3.6)是一个本质上没有约束的凸优化问题。因此, 它的最优解应该是某种意义上的驻点, 即它应该强制 $\ell(L(\xi), \lambda)$ 的导数在某种意义上为零。由导数的表达式和导数算子的线性, 得到

$$D\ell(L(\xi), \lambda)V = E_{P_0} [(H(x, \xi) - \alpha(\log L(\xi) + 1) + \lambda)V]$$

命题 3.2 [15] 假设 $L(\xi) = L^*(\xi, \lambda)$ 满足 $H(x, \xi) - \alpha(\log L(\xi) + 1) + \lambda = 0$, 也就是 $D\ell(L(\xi), \lambda) = 0$ (即

$D\ell(L(\xi), \lambda)$ 是零线性算子)那么 $\ell(L^*(\xi, \lambda), \lambda) < +\infty$, 并且

$$L^*(\xi, \lambda) \in \arg \max_{L(\xi) \in \mathbb{L}^0} \ell(L(\xi), \lambda),$$

证明: 由 $H(x, \xi) - \alpha(\log L(\xi) + 1) + \lambda = 0$ 可得, $e^{\frac{H(x, \xi) + \lambda}{\alpha} - 1} = L(\xi)$, 将其带入

$$G(L^*(\xi)) = E_{P_0} [L(\xi)] - 1 = 0,$$

得到

$$E_{P_0} \left[e^{\frac{H(x, \xi) + \lambda}{\alpha} - 1} \right] - 1 = 0,$$

该期望函数可变形为

$$e^{E_{P_0} \left[\frac{H(x, \xi) + \lambda}{\alpha} - 1 \right]} = 1,$$

即

$$E_{P_0} [H(x, \xi)] = \alpha - \lambda$$

得到

$$\lambda = \alpha - E_{P_0} [H(x, \xi)],$$

再将 $H(x, \xi) - \alpha(\log L(\xi) + 1) + \lambda = 0$ 变形

$$L(\xi) = e^{\frac{\lambda - \alpha}{\alpha}} \cdot e^{\frac{H(x, \xi)}{\alpha}},$$

又由 $L(\xi) = L^*(\xi, \lambda)$, 也就是最优解的形式为

$$L^*(\xi, \lambda) = L(\xi) = e^{\frac{\lambda - \alpha}{\alpha}} \cdot e^{\frac{H(x, \xi)}{\alpha}}.$$

命题 3.2 表明优化问题(3.5)的最优目标值是有限的, 还得出最优解的形式为

$$L^*(\xi, \lambda) = e^{\frac{\lambda - \alpha}{\alpha}} \cdot e^{\frac{H(x, \xi)}{\alpha}},$$

设 $\lambda^* = -\alpha \log E_{P_0} \left[e^{\frac{H(x, \xi)}{\alpha}} \right] + \alpha$, 可以得到 $G(L^*(\xi), \lambda^*) = 0$.

$$L^*(\xi) = L^*(\xi, \lambda^*) = \frac{e^{\frac{H(x, \xi)}{\alpha}}}{E_{P_0} \left[e^{\frac{H(x, \xi)}{\alpha}} \right]}$$

由上述给出的 $L^*(\xi)$ 与 λ^* , 满足 Bonnans 等人提出的命题 3.3 的条件[16], 这表明 $L^*(\xi)$ 可以求解问题(3.5)。将 $L^*(\xi)$ 代入问题(3.5), 得到问题(3.5)的最优值函数:

$$v(\alpha) = \alpha \log E_{P_0} \left[e^{\frac{H(x, \xi)}{\alpha}} \right] + \alpha \eta \quad (3.7)$$

当假设 1 成立时, 对于 $\forall x \in X$, 总存在 $\alpha > 0$, 使得问题(3.7)有限, 这说明问题(3.5)的最优值函数是

有限的。进一步结合两种情况，得到以下定理

定理 1 若假设 1 成立且不确定集为 \mathbb{L} ，则问题(3.2)等价于

$$\min_{\alpha \geq 0} f(\alpha) = \alpha \log E_{P_0} \left[e^{\frac{1_{(0,+\infty)}(\|\phi(\xi)x-y\|_2 - \varepsilon)}{\alpha}} \right] + \alpha \eta \quad (3.8)$$

综上问题(2.1)可以等价转化为

$$\begin{aligned} \min_{x \in X} & \|x\| \\ \text{s.t.} & \inf_{\alpha \geq 0} f(\alpha) \leq \beta \end{aligned} \quad (3.9)$$

其中， $f(\alpha)$ 由式(3.8)定义。

4. 总结及展望

这篇文章的主要内容是将一个 l_1 -范数问题建模为分布鲁棒优化问题，并通过使用 KL-散度、函数变换以及内部最大化期望函数等方法，得到了分布鲁棒优化问题的等价形式。

文章可能首先介绍了压缩感知模型的背景，包括 l_1 -范数问题模型，但其由于不易判断矩阵的有限等距性(RIP)以及观测矩阵的不确定性，需要将其建模为分布鲁棒优化问题等。接着，文章阐述建模的过程，包括如何利用 KL-散度和函数变换来对 l_1 -范数问题进行转换，并通过内部最大化期望函数等方法得到等价形式。

应用分布鲁棒优化问题模型方法对于解决 l_1 -范数问题，可以进一步转化为凸优化问题，进而应用凸优化问题解决方法，未来可以探索例如转化为 CVaR 模型，并考虑使用内点法或梯度下降法等算法解决。相信以后在这个领域会有更深入的理解和探索。

在问题适用性方面，该方法建模为分布鲁棒优化问题的适用性可能受到特定问题结构和数据分布的限制；在计算复杂性角度方面，因为在实际应用中，由于使用了 KL-散度、函数变换等方法，可能导致算法的计算复杂性较高；在参数敏感性方面且方法中可能存在一些参数，对于这些参数的选择可能对最终结果产生较大影响。

今后研究方向和改进方向的改进，可以研究如何优化算法，降低计算复杂性，使方法更适用于大规模数据和实时应用；可以探索方法中参数的自适应性，减少对目标的参数调整需求，提高方法的易用性；可以加强对方法理论基础的研究，确保方法在不同场景下的有效性，并探索更一般化的问题建模方法。

参考文献

- [1] Tropp, J.A. and Gilbert, A.C. (2007) Signal Recovery from Random Measurements via Orthogonal Matching Pursuit. *IEEE Transactions on Information Theory*, **53**, 4655-4666. <https://doi.org/10.1109/TIT.2007.909108>
- [2] Wang, J., Kwon, S. and Shim, B. (2012) Generalized Orthogonal Matching Pursuit. *IEEE Transactions on Signal Processing*, **60**, 6202-6216. <https://doi.org/10.1109/TSP.2012.2218810>
- [3] Wright, S.J., Nowak, R.D. and Figueiredo, M.A.T. (2009) Sparse Reconstruction by Separable Approximation. *IEEE Transactions on Signal Processing*, **57**, 2479-2493. <https://doi.org/10.1109/TSP.2009.2016892>
- [4] Chen, S.S., Donoho, D.L. and Saunders, M.A. (2001) Atomic Decomposition by Basis Pursuit. *SIAM Review*, **43**, 129-159. <https://doi.org/10.1137/S003614450037906X>
- [5] Candès, E.J., Romberg, J. and Tao, T. (2006) Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information. *IEEE Transactions on Information Theory*, **52**, 489-509. <https://doi.org/10.1109/TIT.2005.862083>
- [6] Gui, G., Wan, Q., Peng, W., et al. (2010) Sparse Multipath Channel Estimation Using Compressive Sampling Matching Pursuit Algorithm. arXiv:1005.2270.

-
- [7] Goldstein, T. and Osher, S. (2009) The Split Bregman Method for L1-Regularized Problems. *SIAM Journal on Imaging Sciences*, **2**, 323-343. <https://doi.org/10.1137/080725891>
- [8] Zhang, J., Zhao, D., Zhao, C., *et al.* (2012) Image Compressive Sensing Recovery via Collaborative Sparsity. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, **2**, 380-391. <https://doi.org/10.1109/JETCAS.2012.2220391>
- [9] Zhang, J., Zhao, D., Zhao, C., *et al.* (2012) Compressed Sensing Recovery via Collaborative Sparsity. 2012 *Data Compression Conference, Snowbird*, 10-12 April 2012, 287-296. <https://doi.org/10.1109/DCC.2012.71>
- [10] Donoho, D.L. (2006) Compressed Sensing. *IEEE Transactions on Information Theory*, **52**, 1289-1306. <https://doi.org/10.1109/TIT.2006.871582>
- [11] 聂操男. 压缩感知中的概率约束优化模型的光滑近似[D]: [硕士学位论文]. 大连: 辽宁师范大学, 2019.
- [12] 陈苗. 一类含有 l_1 范数优化问题的梯度类算法[D]: [硕士学位论文]. 青岛: 青岛大学, 2020.
- [13] 屈彪, 张文伟, 于丽超. 线性方程组 l_1 范数问题的松弛投影算法及其应用[J]. *运筹学学报*, 2017, 21(2): 57-65.
- [14] 王榆. 基于修正的 χ^2 -距离散度的不确定概率优化问题[D]: [硕士学位论文]. 大连: 辽宁师范大学, 2016.
- [15] Hu, Z. and Hong, L.J. (2013) Kullback-Leibler Divergence Constrained Distributionally Robust Optimization. *Optimization Online*, **1**, 9.
- [16] Shapiro, A., Dentcheva D. and Ruszczyński, A. (2009) Lectures on Stochastic Programming. *Society for Industrial and Applied Mathematics*. <https://doi.org/10.1137/1.9780898718751>