

# 稀疏优化模型及其正则化方法

刘思凡<sup>\*</sup>, 王浩<sup>†</sup>, 胡耀华<sup>‡</sup>

**摘要** 稀疏优化模型是目前最优化领域中非常热门的研究前沿课题, 在压缩感知、图像处理、机器学习、统计建模等领域都获得了成功的应用。本文将光谱分析技术、数字信号处理、推荐系统等多个应用问题为例, 阐述稀疏优化模型的建模过程与核心思想。稀疏优化模型属于组合优化模型, 非常难以求解(NP-难)。正则化方法是稀疏优化模型的一类常用的求解方法。我们将介绍正则化方法的原理与几类常见的正则化模型, 并阐述正则化模型的稳定性理论与多种先进算法。数值实验表明这些算法都具有快速、高效、稳健等显著优点。稀疏正则化模型将在大数据时代中发挥更显著的计算优势与应用价值。

**关键字** 稀疏优化模型; 正则化方法; 稳定性理论; 机器学习思想; 贪婪算法; 一阶迭代算法

## 1 引言

最优化理论和方法是应用数学的一个重要分支, 为诸多科学和应用问题提供了统一的建模框架和研究方法, 相关研究成果在工农业生产、经济管理、金融投资、交通运输、通信控制、图像处理、生命科学等领域已获得广泛应用。

稀疏优化模型是最优化领域中非常热门的研究课题, 它旨在寻找一个欠定线性系统的稀疏解, 即只有极少数的分量不为零。稀疏优化模型最早是由美国科学院院士David Donoho等人[9]于1998年提出来的, 它的本质思想是结合解的稀疏性结构来构建数学模型, 克服欠定线性反问题的不稳定性, 进而提升模型的稳定性和准确性。特别地, 在2005年, 数学家Emmanuel Candès与陶哲轩[7]给出了稀疏优化模型/压缩感知的数学理论, 证明了在已知信号的稀疏性的情况下, 稀疏优化模型能够利用极少数的采样数(显著优于奈奎斯特采样定理)来重建原信号。此文奠定了稀疏优化模型的理论根基。在过去的十年中, 稀疏优化模型吸引了学术界与业界的大量关注, 并且在很多领域都取得了成功的应用, 诸如压缩感知[9, 12]、图像科学[3, 14]、机器学习[1, 18]、统计建模[15, 28]、基因组学数据分析[24, 31]。

本文将简单介绍稀疏优化模型的建模思想与求解稀疏优化模型的正则化方法。在第二节, 我们将以光谱分析技术、数字信号处理应用问题为例, 阐述稀疏优化模型的建模过程与原

---

<sup>\*</sup>深圳大学 数学与统计学院, 深圳市现代机器学习与应用重点实验室, 广东 深圳 518060

<sup>†</sup>上海科技大学 信息科学与技术学院, 上海 201210

<sup>‡</sup>深圳大学 数学与统计学院, 深圳市现代机器学习与应用重点实验室, 广东 深圳 518060

<sup>§</sup>基金项目: 国家自然科学基金项目(11601343, 11601344), 广东省自然科学基金项目(2016A030310038)

<sup>¶</sup>通讯作者: 胡耀华, Email: mayhhu@szu.edu.cn

理。正则化方法是稀疏优化模型的一类常用的求解方法，其主旨思想是用可解的正则化模型来逼近稀疏优化模型，进而提供可行的数值解法。在第三节中，我们将介绍正则化方法的原理与几类常见的正则化模型。在第四节中，我们将阐述正则化模型的稳定性理论，给出其逼近稀疏优化模型的量化估计，从而建立正则化方法的可靠性。在第五节中，我们将介绍稀疏优化模型/正则化模型的两类流行的先进算法：贪婪算法与一阶迭代算法，并且运用这些算法进行了光谱分析模拟实验以及算法稳定性比较，实验结果显示这些稀疏优化算法都具有良好的实际应用效果、计算效率和鲁棒性能。最后，我们对全文做出简要的总结。

## 2 稀疏优化建模

稀疏优化模型是大数据背景下最优化领域中非常热门的模型工具与研究课题，它在过去十年中井喷式发展，在很多领域已获得广泛应用。本节首先以光谱分析技术为例，阐述稀疏优化模型的建模过程与原理，然后启发更多的高新技术应用，比如信号复原、基因调控网络预测、推荐系统、视频监控等。

**光谱分析** 光谱是复色光经过色散系统(如棱镜、光栅)分光后，被色散开的单色光按波长大小依次排列的曲线。光谱是各种物质的“指纹”，不同物质的原子内部的电子运动情况不同，它们所发射/吸收的光波也不同，从而会形成不同的特征光谱。光谱分析就是利用特征光谱来鉴别物质以及确定它的组成成分和相对含量的定量方法。铷、铯、氦等新元素就是通过光谱分析技术发现的。吸收光谱是一类典型的光谱(包括红外光谱、紫外光谱等)，它是光在通过物质时，某些波长的光被物质吸收后产生的特征光谱。下文以吸收光谱分析为例，介绍光谱分析技术的数学模型。

假设待测物质为纯净物，根据朗伯-比尔定律[23]，我们可建立出射光强度 $I(\lambda)$ 与入射光强度 $I_0(\lambda)$ 的关系：

$$I(\lambda) = I_0(\lambda) e^{-xa(\lambda)}, \quad (2.1)$$

其中 $\lambda$ 是波长， $a(\lambda)$ 是待测纯净物的特征吸收光谱， $x := \rho L$ ， $\rho$ 是待测纯净物的浓度， $L$ 是光路长度。若待测物质含有多种有机物/无机物，根据吸光度的可加性，待测物质的光谱 $xa(\lambda)$ 可由这些物质的特征吸收光谱线性表示，即 $xa(\lambda) = \sum_j x_j a_j(\lambda)$ ，其中 $x_j$ 和 $a_j(\lambda)$ 分别为其组成成分的浓度与特征吸收光谱。另外，考虑到光的散射会导致的出射光强度减小，我们可将光谱噪音 $\epsilon(\lambda)$ 加入到模型中表示散射项(仪器误差等也计入 $\epsilon(\lambda)$ )。此时，朗伯-比尔定律(2.1)演变为

$$I(\lambda) = I_0(\lambda) e^{-\sum_j x_j a_j(\lambda) - \epsilon(\lambda)}.$$

对上式两边取对数，则有

$$\ln(I(\lambda)) = \ln(I_0(\lambda)) - \sum_j x_j a_j(\lambda) - \epsilon(\lambda).$$

入射光强度 $I_0(\lambda)$ 与出射光强度 $I(\lambda)$ 均可由光谱仪观测，记 $b(\lambda) := \ln(I_0(\lambda)) - \ln(I(\lambda))$ ，则上式演变为

$$b(\lambda) = \sum_j x_j a_j(\lambda) + \epsilon(\lambda), \quad (2.2)$$

也就是说，光谱仪观测所得的待测物质的光谱可以近似表示成其组成成分的特征光谱的线性组合(权重为各组成成分的浓度)。所以，理想状况的光谱分析问题可以建模成在特征光谱的标准数据库\*中变量选择来线性逼近光谱仪的观测光谱。

在实际应用中，光谱分析问题要更为复杂一些。一个常见的困难是波长偏移现象，即由于机器测量误差，光谱仪观测的光谱的波长与标准光谱的波长相比，往往会有一些波长的偏移。光谱仪测量的波长偏移通常是线性偏移的情形： $\nu_j = p_j\lambda + q_j$ ， $p_j \in \{p_1, \dots, p_t\}$ ， $q_j \in \{q_1, \dots, q_t\}$ 。图1 描绘了几种常见有机物的紫外光谱图及其波长偏移可能。

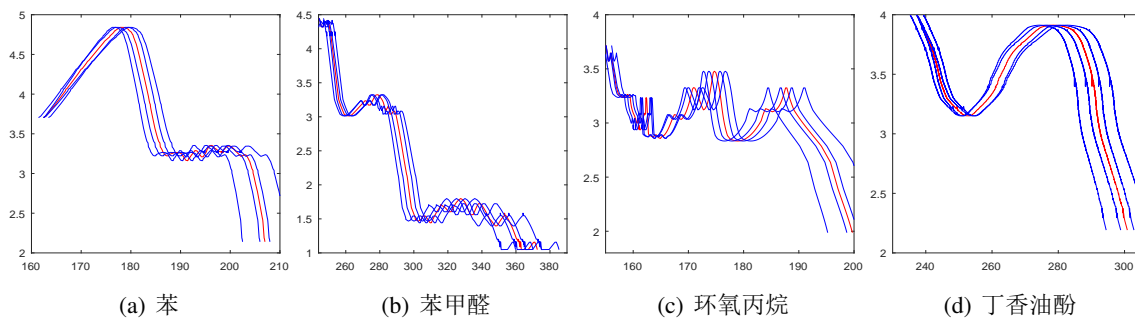


Figure 1: 横轴为波长，纵轴为强度。图中的红色曲线表示NIST提供的标准光谱，蓝色曲线表示发生波长偏移现象的光谱。

由于波长偏移现象往往会造成整条光谱的强度产生系统性的误差，无法忽略不计；每次测量的波长偏移也是未知的。所以，实际情况的光谱分析是分析待测物质的组成成分和相对含量，并且同时预测波长偏移的距离。这样说来，实际情况的光谱分析问题应该建模成在标准光谱及所有可能的波长偏移光谱的扩展数据库中变量选择来线性逼近光谱仪的观测光谱。根据(2.2)，光谱分析问题可表示成如下的线性反问题

$$b = Ax + \varepsilon, \quad (2.3)$$

其中  $b \in \mathbb{R}^m$  表示光谱仪观测所得的待测物质的光谱向量； $A \in \mathbb{R}^{m \times n}$  表示标准光谱及波长偏移光谱的扩展数据库矩阵，某一行  $A_j$  表示一种有机物/无机物的标准光谱(或者波长偏移光谱)； $x \in \mathbb{R}^n$  表示各种组成成分的浓度， $x_i = 0$  说明不含有这种物质； $\varepsilon \in \mathbb{R}^m$  表示光谱的散射噪音与仪器误差。

就紫外/红外吸收光谱而言，通常需要考虑的有机物/无机物(可能的成分)大约有200 ~ 500种，光谱的波长偏移通常有4 ~ 8种可能性，因而，可供变量选择的标准/偏移光谱数量为  $n = 800 \sim 4000$ 。但是，光谱仪实际所观测的光谱通常只有  $m = 50 \sim 300$  个样本点。所以，光谱分析问题(2.3)是一个欠定的线性方程组( $m \ll n$ )，它是不适定问题，有无数个解。这使得经典的方程组求解方法和线性拟合的最小二乘法不适用于光谱分析问题。我们进一步考虑光谱分析的一个重要的结构特点：通常待测物质往往只有少量几种物质，大约2 ~ 20种，这相比变量选择的可能性800 ~ 4000小很多。这就是光谱分析问题的稀疏性结构，即光谱分析问题的解只有少量分量不为零，大部分分量都为零。从数学语言的角度，我们用符号  $\ell_0$  范数

\*光谱的标准数据源于美国国家标准与技术局(NIST): <http://webbook.nist.gov/chemistry/>

来刻画稀疏性结构，即 $\|x\|_0$  定义为向量 $x$  的非零分量的个数：

$$\|x\|_0 := \text{card}(\{i : x_i \neq 0\}).$$

因此，基于(2.3)，光谱分析问题可以建模成以系统误差为约束的稀疏度最小化模型：

$$\begin{aligned} \min \quad & \|x\|_0 \\ \text{s.t.} \quad & \|Ax - b\|^2 \leq \epsilon. \end{aligned} \quad (2.4)$$

此模型就是典型的稀疏优化模型。我们已经通过光谱分析技术为例，详细阐述了稀疏优化模型的建模过程与原理。接下来，我们将启发更多的高新技术应用，简单介绍他们的稀疏表示原理与建模，从而展示稀疏优化模型的广泛应用与价值。

**信号复原/图像处理** 在信号/图像处理领域，一个主要的问题就是如何将连续信号(比如图像)转化为离散的数字信号，以便计算机对其进行处理，然后将离散的数字信号复原为连续信号。第一个过程称为采样，第二个过程称为信号处理，第三个过程为信号复原。信号复原的传统方法是基于奈奎斯特采样定理[22]的信号采样技术。根据奈奎斯特采样定理，我们可知当信号的频率小于采样频率的一半，便能够从采样数据中成功地恢复原本信号。但是，这个方法需要大量的空间存储采样数据，这限制了高速信号处理与实时处理技术的发展。压缩感知理论的诞生是信号复原技术的突破性进展，它建立了新的信号处理框架，能够利用极少的采样数据来重建原信号，能够极大地降低了信号采集端的复杂度和存储空间。压缩感知理论的核心思想和原理就是基于信号的稀疏表示理论：在很多应用领域，虽然一个信号的数据很庞大，但是它往往都具有某些特征且特征数量远小于信号数据，所以这些特征就构成了这个信号的稀疏表示，即一个普通信号能够在某个域变换下转变成稀疏(特征)信号。例如，自然中的图像往往包含许多特征，但是对于一幅图像来说，往往只具有少数的主要特征，这幅图像可以通过这些特征近似表示。具体地，一个图像 $\Psi$  在变换域 $D$  (比如Haar小波基变换、快速傅里叶变换FFT、离散余弦变换DCT)下得到稀疏信号表示 $x = D\Psi \in \mathbb{R}^n$ 。为了便于传输和处理，我们将此稀疏信号与一个低维伪随机矩阵 $A \in \mathbb{R}^{m \times n}$  ( $m \ll n$ ) 相乘得到观测信号 $b = Ax \in \mathbb{R}^m$ 。这样，观测信号 $b$  就是图像信号 $\Psi$  的压缩数据，我们能够方便地对观测信号 $b$  进行传输或处理。在信号传输或处理之后，压缩传感的信号复原技术就是首先从观测信号 $b$  中恢复稀疏信号 $x$ ，根据稀疏表示理论，这一过程可以通过稀疏优化建模实现：

$$\begin{aligned} \min \quad & \|x\|_0 \\ \text{s.t.} \quad & \|Ax - b\|^2 \leq \epsilon. \end{aligned}$$

然后再通过变换域逆变换得到原始图像，即 $\Psi = D^{-1}x$ 。

**基因调控网络** 基因调控网络是指细胞内(或基因组内)基因和基因之间的相互作用关系所形成的网络。预测/研究基因调控网络能够帮助我们了解不同基因之间的调控关系以及它们在生物进化过程中的功能，是生命科学与医学的重要议题。通常来说，细胞中所有的基因都是在成百上千个转录因子(就是特殊功能的基因)的严格控制之下进行转录、翻译、表达等过程，进而确保其拥有正确的细胞形态与有效的细胞功能。细胞内(或基因组内)所有的转录因子与目标基因之间的调控关系(激活或抑制)就构成了一个巨大的基因调控网络[11]。

近年来高通量转录组测序技术(RNA sequencing)的兴起，使得人们能够在一次实验中完成对组织细胞内所有的基因(包括转录因子)的表达进行量化统计。具体地，通过对 $m$  个

样品基因的表达的测序与量化，我们可以得到某目标基因在所有样品中的(对数化)表达向量  $b \in \mathbb{R}^m$  与所有转录因子在所有样品中的(对数化)表达矩阵  $A \in \mathbb{R}^{m \times n}$ ，其中  $b_i$  表示该目标基因在样品  $i$  中的表达量的对数， $A_{i,j}$  表示是转录因子  $j$  在样品  $i$  中的表达量的对数。当目标基因被一个或多个转录因子调控时，目标基因在不同样本中的表达向量由调控该目标基因表达的多个转录因子的表达向量近似线性表示，即

$$b = Ax + \varepsilon,$$

其中  $x \in \mathbb{R}^n$  表示各个转录因子的调控权重，刻画了转录因子与目标基因之间的调控关系； $\varepsilon \in \mathbb{R}^m$  表示测序技术的计算误差。 $x_i$  是转录因子  $i$  对该目标基因的调控强度，正数表示该转录因子能够激活目标基因的表达，负数表示该转录因子能抑制目标基因的表达，等于零则表示该转录因子不调控此基因的表达。基因调控网络预测问题就是预测转录因子对所有目标基因的调控关系。基于物种的进化过程与生物系统的能耗最小原则，通常只有少量几个转录因子会参与调控某个目标基因的转录翻译表达过程。因此，各个转录因子对每个目标基因的调控向量具有稀疏性，进而基因调控网络预测问题可以建模成稀疏优化模型(2.4)。

**推荐系统** 推荐系统(Recommendation system)是用来预测用户对某件事物或商品的“评级”或者“偏好”，近几年被广泛应用于电影、音乐、新闻、书籍、学术论文等产品领域。我们生活中常见的购物/社交平台，比如淘宝、京东、亚马逊、优酷、推特页面等都应用了推荐系统。我们以著名的Netflix问题为例，阐述推荐系统的数学模型原理。Netflix是美国的一家电影租赁公司，他们收集了480189个用户对17770部电影的1亿个评分(点评率仅为1.2%)，希望能够开发一套推荐系统，可以通过用户对少数电影的评分预测用户的电影喜好，进而为用户推荐新的电影。如果推荐的电影越符合用户的喜好，就越能提高Netflix租赁电影的业务量与消费者的信赖。为此，Netflix于2006年设立了壹百万美元的奖金用于悬赏能够最好地提高该公司推荐系统准确度(即最准确地预测所有用户对所有电影的评分)的解决方法。这个比赛也是推动推荐系统发展的关键事件之一。

矩阵填充(Matrix completion)模型[26]是Netflix推荐系统的数学表示模型。我们可构建评分矩阵  $X \in \mathbb{R}^{480189 \times 17770}$ ，其中  $X_{i,j}$  代表第  $i$  个用户对第  $j$  部电影的评分。Netflix所提供的  $X$  仅是一个完成度为1.2%的不完全矩阵，推荐系统的任务就是要准确地填充  $X$  的每一个评分。为此，我们还需要进一步考虑Netflix推荐系统的一个重要的结构特点：评分矩阵  $X$  通常被认为是低秩的！用户数量巨大，电影数目巨大，因此矩阵  $X$  的维度也是巨大的；然而，影响用户对影片喜好的主要因素的数目是有限的(意外因素可看作评分噪声)，如电影的题材、演员、年代、导演、制作等，所以这个高维评分矩阵  $X$  本质上是一个低秩矩阵。因此，Netflix推荐系统可表示成低秩矩阵填充问题，从而建模成矩阵秩极小化模型：

$$\begin{aligned} \min \quad & \text{rank}(X) \\ \text{s.t.} \quad & X_{i,j} = r_{i,j}, \forall (i, j) \in \Omega \end{aligned} \quad (2.5)$$

其中  $\text{rank}(X)$  表示矩阵  $X$  的秩， $r_{i,j}$  表示给定的用户-电影评分数据， $\Omega$  表示已评分的用户-电影集合。秩极小化模型(2.5)与稀疏优化模型(2.4)的建模思想是异曲同工的。我们用  $\sigma(X)$  表示矩阵  $X$  的奇异值向量，根据矩阵理论的基本知识，我们可知  $\|\sigma(X)\|_0 = \text{rank}(X)$ ；(2.5)的约束条件可以看做是一个线性方程组  $AX = R$  的特殊形式。所以，矩阵秩极小化模型(2.5)可以看作奇异值向量的稀疏优化模型，是经典的稀疏结构建模的矩阵优化模型。

**视频监控** 近年来,随着计算机、图像处理、信号传输技术的快速发展,视频监控技术也有了巨大的进步。视频背景提取技术是视频监控中的一项重要技术,它源自于电影和视频产品的发展,比如在电影的制作中经常需要将电影棚内拍摄片段中的演员合成到另一个环境中,此时就需要视频背景提取技术将影片中的演员与背景分离开。同时,这项技术也促进了计算机视觉的许多应用的发展,例如监视跟踪、动作感知等。随着电影制作要求的不断提高与计算机视觉的发展,视频背景提取技术的要求也越来越高,这也促使了人们对视频背景提取技术的深入研究与技术升级。稳健主成分分析(Robust Principal Component Analysis)模型[32]是近年发展的视频背景提取技术主流研究方法之一,它的思想也是结合背景数据的低秩性与移动物体数据的稀疏性建立模型。具体来说,对于一段时间内的视频,我们将其中的每一帧画面编码成数据矩阵 $H$ 中的一个列向量,从而整段视频的所有画面数据就组成了矩阵 $H$ 。视频的每一帧画面可以看作两部分的组成:背景画面与移动物体,所以 $H$ 的每一个列向量由背景画面数据与移动物体数据组成。背景画面通常是固定不动的(或者稳定移动),每一帧画面的背景画面数据(向量)往往都具有高度的相似性(或者相关性),从而整段视频的背景画面数据(矩阵)具有低秩性;移动物体通常只含有少量的物体,比如几辆汽车或者几个行人,每一帧画面移动物体数据往往都具有高度的稀疏性,从而整段视频的移动物体数据(矩阵)也具有稀疏性。所以,整段视频数据可以分解成一个低秩矩阵 $L$ 和一个稀疏矩阵 $S$ 的和,从而建模成稳健主成分分析(稀疏-低秩)模型

$$\begin{aligned} \min \quad & \text{rank}(L) + \mu \|S\|_0 \\ \text{s.t.} \quad & H = L + S, \end{aligned}$$

其中 $\mu > 0$ 是用于平衡低秩和稀疏的参数, $L$ 是表示背景画面的低秩矩阵, $S$ 是用于捕捉画面中的移动物体的稀疏矩阵。

### 3 正则化方法

综上所述,在很多应用问题中,基础数据往往可以近似满足如下的线性系统

$$Ax = b + \varepsilon. \quad (3.1)$$

其中 $A \in \mathbb{R}^{m \times n}$ 是由一组基函数 $A_1, \dots, A_n$ 组成的矩阵, $b \in \mathbb{R}^m$ 是观测值组成的向量, $\varepsilon$ 是实际过程中所产生的观测误差向量。在大数据的背景下,很多应用问题的变量维度很高,然而实验测得的样本数据是非常有限的(仪器或成本限定)。此时一般有 $m \ll n$ ,因此(3.1)是一个欠定的线性方程组,它是高度病态的,有无数个解。在这个有限样本的高维变量选择问题中,人们往往希望仅通过关注少量的主要因素得到系统的简单而有效的理解与刻画。系统拟合的准确性通常用欧氏范数 $\|Ax - b\|^2$ 来刻画,解(变量选择)的稀疏结构可以用 $\|x\|_0$ 来度量。所以,给定允许的观测误差 $\varepsilon$ ,线性系统(3.1)的稀疏优化模型可表述为

$$\begin{aligned} \min \quad & \|x\|_0 \\ \text{s.t.} \quad & \|Ax - b\|^2 \leq \varepsilon. \end{aligned} \quad (3.2)$$

由于 $\ell_0$ 范数是非凸且非光滑的函数,模型(3.2)非常难求解,这是一个NP难问题[20]。幸运的是,求解问题(3.2)的近似局部解很容易。为了避免求解模型(3.2)的计算困难,学者提出

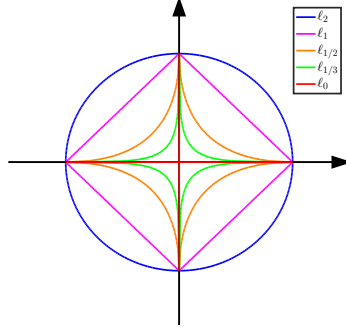


Figure 2: 不同的 $\ell_p$  范数的等高线

了它的 $\ell_1$  松弛化模型(又称作Lasso[28]或者是基追踪[9]), 并且成为解决稀疏优化模型的有效工具。 $\ell_1$  范数的定义是 $\|x\|_1 := \sum_{i=1}^n |x_i|$ , 它是 $\ell_0$  范数的最优凸逼近。 $\ell_1$  松弛化模型的思路是用 $\ell_1$  范数代替 $\ell_0$  范数来刻画稀疏结构。因此,  $\ell_1$  松弛化模型可表述为

$$\begin{aligned} \min \quad & \|x\|_1 \\ \text{s.t.} \quad & \|Ax - b\|^2 \leq \epsilon. \end{aligned} \quad (3.3)$$

Emmanuel Candès与陶哲轩[7]、David Donoho[12]建立了稀疏优化模型(3.2)与 $\ell_1$  松弛化模型(3.3)之间的关系: 在一些宽松的条件下, 比如RIP(受限正交条件)或者NSP(零空间性质), 模型(3.3)与(3.2)的解是一致的。

模型(3.2)与(3.3)都是约束优化模型, 它们的大规模问题求解仍然存在较大的计算困难。基于大规模计算的动机, 正则化方法成为处理约束优化问题的一种常见方法, 就是在目标函数中施加一个正则(惩罚)项来度量解的某种结构, 从而将稀疏优化模型转化为无约束的正则化模型。例如, 模型(3.2)与(3.3)的正则化模型分别是

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|^2 + \lambda \|x\|_0, \quad (3.4)$$

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|^2 + \lambda \|x\|_1, \quad (3.5)$$

其中 $\lambda$  是用于平衡系统的准确性与解的稀疏性的正则化参数。 $\ell_0$  正则化模型(3.4)的全局最优解仍然是NP难问题; 大量的实证研究表明 $\ell_1$  正则化模型(3.5)的稀疏诱导性与稳健性都差强人意, 而且在实际应用问题中往往只能获取次优稀疏解(参见[8, 18, 33, 36])。更一般的, 对 $p > 0$ ,  $\ell_p$  范数的定义是

$$\|x\|_p := \left( \sum_{i=1}^n |x_i|^p \right)^{1/p} \quad \text{for each } x \in \mathbb{R}^n.$$

图(2)展示了不同的 $\ell_p$  范数的等高线。如图(2)所示, 我们注意到 $\ell_p$  范数的函数近似关系为

$$\lim_{p \rightarrow 0^+} \|x\|_p^p = \|x\|_0 \quad \text{for each } x \in \mathbb{R}^n.$$

也就是说, 较之 $\ell_1$  范数,  $\ell_p$  范数( $0 < p < 1$ )是稀疏结构的更佳刻画。由此启发, 徐宗本院士[33]提出了 $\ell_{1/2}$  正则化模型, 即

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|^2 + \lambda \|x\|_{1/2}^{1/2}. \quad (3.6)$$

图3展示了 $\ell_p$  正则化模型的稀疏诱导性原理, 由图可见,  $\ell_1$ 、 $\ell_{1/2}$  与 $\ell_0$  正则化模型都能够诱导得到稀疏解, 而经典的 $\ell_2$  正则化模型无法诱导解的稀疏解。文献[8, 18, 24, 33] 均表明, 较之 $\ell_0$  与 $\ell_1$  正则化模型,  $\ell_{1/2}$  正则化模型(3.6) 在理论性质与数值性能两方面取得了较好的权衡与改进。 $\ell_{1/2}$  正则化模型(3.6)能够提升稀疏还原性, 使用更少的样本量得到更稀疏的解[8], 在稀疏信号重建中对噪声也具有更强的稳健性[33], 在基因组学数据分析中能够预测出更具有生物意义的基因调控网络[18, 24]。接下来两节内容将主要阐述正则化模型的稳定性理论和快速计算方法。

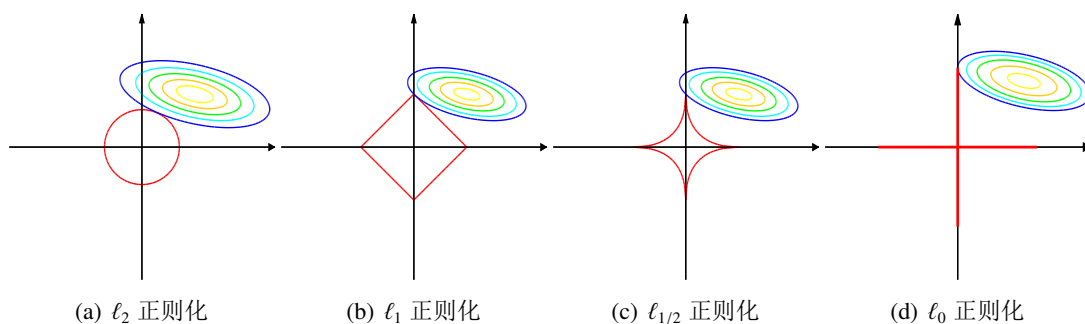


Figure 3:  $\ell_p$  范数的稀疏诱导性。红色曲线表示 $\ell_p$  范数的等高线, 彩色曲线族表示目标函数的等高线族。 $\ell_p$  正则化模型的最优点在两族等高线相切处取得。

## 4 稳定性理论

在最优化与工程应用领域中, 最小二乘法是线性拟合(回归)模型(3.1)的经典方法。最小二乘法的效用通常需要矩阵 $A^T A$  的正定性前提假设, 即

$$\min_{x \in \mathbb{R}^n} \frac{\|Ax\|}{\|x\|} > 0. \quad (4.1)$$

但是, 在高维的情况下( $m \ll n$ ), 最小二乘法往往是失效的。因为, 在这种情况下, 矩阵 $A^T A$  是严重退化的, 线性系统(3.1)是严重病态的, 它有无数个解。为了克服高维数据拟合的困难, 学者提出了稀疏表示理论, 并在多个工程应用领域中取得了巨大的成功。稀疏表示理论中的著名模型是 $\ell_1$  约束优化模型(3.3), 它在较为宽松的受限正交条件(RIP)假设下具有良好的理论性质与应用表现。RIP是稀疏表示理论中一个著名的假设条件, 它由数学家Emmanuel Candès 与陶哲轩[7]提出。

**Definition 4.1.** 令 $A \in \mathbb{R}^{m \times n}$ ,  $s + t \leq n$ 。



(i)  $A$  的  $s$ -受限等距常数的定义为

$$\eta(s) := \min \left\{ \eta : (1 - \eta)\|x_{\mathcal{J}}\|^2 \leq \|Ax_{\mathcal{J}}\|^2 \leq (1 + \eta)\|x_{\mathcal{J}}\|^2, \forall x \in \mathbb{R}^n, |\mathcal{J}| \leq s \right\}.$$

(ii)  $A$  的  $(s, t)$ -受限正交常数的定义为

$$\theta(s, t) := \min \left\{ \theta : |\langle Ax_{\mathcal{J}}, Ax_{\mathcal{T}} \rangle| \leq \theta \|x_{\mathcal{J}}\| \cdot \|x_{\mathcal{T}}\|, \forall x \in \mathbb{R}^n, |\mathcal{J}| \leq s, |\mathcal{T}| \leq t, \mathcal{J} \cap \mathcal{T} = \emptyset \right\}.$$

Emmanuel Candès与陶哲轩[7]在RIP条件假设下建立了 $\ell_1$ 约束优化模型(3.3)与稀疏优化模型(3.2)的等价性。这奠定了稀疏表示理论的理论基石。

**Theorem 4.1** ([7, 定理1.3]). 令 $\epsilon = 0$ , 记 $\bar{x}$ 为稀疏优化模型(3.2)的最优解。假设 $\|\bar{x}\|_0 \leq s$ , 且

$$\eta(s) + \theta(s, s) + \theta(s, 2s) < 1^\dagger,$$

那么,  $\bar{x}$ 是 $\ell_1$ 约束优化模型(3.3)的唯一最优解。

稀疏优化与机器学习的领域中最常见的模型是正则化模型。一般的正则化模型可表述为

$$\min_{x \in \mathbb{R}^n} \text{Loss}(x) + \lambda \text{Pen}(x), \quad (4.2)$$

其中 $\text{Loss}(\cdot)$ 是损失函数(loss function), 刻画系统或数据的拟合程度,  $\text{Pen}(\cdot)$ 是罚函数(penalty function), 度量解的(结构)规则性。适当的罚函数的引入, 往往能够有效避免机器学习中的过拟合现象, 提高模型的稳定性和泛化性能。显然, 模型(3.4)-(3.6)是模型(4.2)的特例。正则化模型(4.2)的准确性和稳定性通常由Oracle性质和 $\ell_2$ 相容性来刻画。Oracle性质是正则化模型的一种重要统计性质, 确保模型能够在先验条件的假设下成功预测出真实模型;  $\ell_2$ 相容性是估计出正则化模型与真实模型的解的误差上界, 从而确保正则化模型的准确性与稳定性。

在稀疏优化模型中, 稀疏特征值条件(SEC)是一个基本的假设条件, 即假设 $A^T A$ 的 $2s$ -稀疏最小特征值是正的:

$$\phi_{\min}(2s) := \min_{\|x\|_0 \leq 2s} \frac{\|Ax\|}{\|x\|} > 0. \quad (4.3)$$

这也就是说 $A$ 的任意 $2s \times 2s$ 维子矩阵的最小特征值是正的。SEC是比(4.1)更弱的一个假设, 也被认为是稀疏优化模型能够起效用的最起码(必要)的条件。事实上, 如果SEC满足, 那么线性系统 $Ax = b$ 的稀疏度为 $s$ 的解是唯一的; 否则, 我们可假设 $\hat{x}$ 和 $\bar{x}$ 是线性系统 $Ax = b$ 的稀疏度为 $s$ 的两个不同的解(即,  $A\hat{x} = A\bar{x}$ 且 $\|\hat{x}\|_0 = \|\bar{x}\|_0 = s$ ), 那么 $x := \hat{x} - \bar{x}$ 满足 $Ax = 0$ 且 $\|x\|_0 \leq 2s$ , 从而有 $\phi_{\min}(2s) = 0$ 。这与(4.3)矛盾。因此, 如果SEC不满足, 那么无法从带噪声的观测值中恢复真实的稀疏解。

然而, 只有SEC是不充分的, 需要更进一步的规则性条件才能保证正则化模型的良好理论性质与成果应用。Bickel[4]提出了受限特征值条件(REC)来研究 $\ell_1$ 正则化模型的Oracle性质和 $\ell_2$ 相容性, 胡耀华等[18]推广了低阶REC并用以研究低阶正则化模型的Oracle性质和 $\ell_2$

<sup>†</sup>受限正交条件(RIP)就是指这个不等式成立。文献[26]指明, 大多数独立随机矩阵都高概率地满足RIP条件。

相容性。记  $\bar{x}$  与  $x_\ell^*$  分别为稀疏优化模型(3.2)与  $\ell_p$  正则化模型的最优解，支撑集  $S := \text{supp}(\bar{x})$ 。模型的误差  $\Delta := x_\ell^* - \bar{x}$  通常满足如下性质：

$$\|\Delta_{S^c}\|_p \leq \|\Delta_S\|_p.$$

基于上述性质，REC就是要求在此限制集上满足一定的正定性，即将(4.3)的最小化限制在上式的限制集合上，同时分母改为  $x$  的一个子向量的  $\ell_2$  范数。给定  $x \in \mathbb{R}^n$  和  $I \subseteq \{1, \dots, n\}$ ，我们用  $\mathcal{I}(x; t)$  表示  $x$  在  $I^c$  ( $I$  的补集)中绝对值最大的前  $t$  个的指标集合。

**Definition 4.2.** 令  $A \in \mathbb{R}^{m \times n}$ ， $s + t \leq n$ ， $0 \leq p \leq 1$ 。  $A$  的  $p$ -受限特征值常数的定义为

$$\phi_p(s, t) := \min \left\{ \frac{\|Ax\|}{\|x_{\mathcal{T}}\|} : |\mathcal{I}| \leq s, \|x_{\mathcal{I}^c}\|_p \leq \|x_{\mathcal{I}}\|_p, \mathcal{T} = \mathcal{I}(x; t) \cup \mathcal{I} \right\}.$$

受限特征值条件(REC)<sup>‡</sup>是比SEC和RIP都弱的一类规则性条件。显然， $\phi_{\min}(2s) = \phi_0(s, s)$ ，所以

$$\text{SEC} \Leftrightarrow \phi_0(s, s) > 0.$$

**Proposition 4.1** ([4],  $\text{RIP} \Rightarrow \text{REC}$ ).

$$[\eta(t) + \theta(s, t) + \theta(t, s + t) < 1] \Rightarrow [\phi_1(s, t) > 0].$$

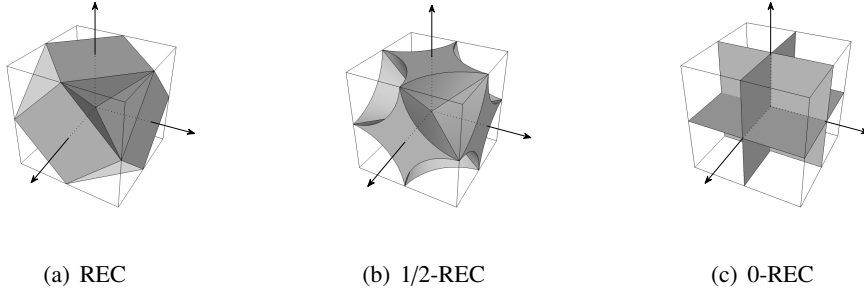


Figure 4: REC的几何解释：图中的灰色区域(锥)表示REC的受限集合，REC成立当且仅当  $A$  的零空间与灰色区域不相交。

图4展示了REC的几何意义。由于REC对应的目标函数是齐次的，REC成立当且仅当  $A$  的零空间与受限集合不相交。如图1所示，我们可知

$$1\text{-REC} \Rightarrow 1/2\text{-REC} \Rightarrow 0\text{-REC}.$$

以下命题严格地给出了REC之间的关系： $p$  越小， $p$ -REC 越弱。

**Proposition 4.2** ([18, 命题5]). 对于任意的  $0 \leq q \leq p \leq 1$ ，以下关系成立

$$[\phi_p(s, t) > 0] \Rightarrow [\phi_q(s, t) > 0].$$

<sup>‡</sup> $p$ -REC就是指  $\phi_p(s, t) > 0$ 。文献[25]指明，大多数相关(或独立)随机矩阵都高概率地满足REC条件。

综上所述，我们总结稀疏表示理论中规则性条件的关系如下：

$$\text{RIP} \Rightarrow \text{REC} \Rightarrow p\text{-REC} \Rightarrow 0\text{-REC} \Leftrightarrow \text{SEC}. \quad (4.4)$$

最后，我们在REC条件下估计正则化模型的Oracle不等式和 $\ell_2$ 相容性，从而建立展示正则化模型的稳定性理论。

**Theorem 4.2** ([16, 引理2.1与推论2.1]). 记 $\bar{x}$ 与 $x^*$ 分别为稀疏优化模型(3.2)与 $\ell_1$ 正则化模型(3.5)的最优解,  $\mathcal{S} := \text{supp}(\bar{x})$ ,  $s := \|\bar{x}\|_0$ 。假设

$$\phi_1(s, s) > 0.$$

那么,  $\ell_1$ 正则化模型(3.5)的Oracle性质与 $\ell_2$ 相容性成立:

$$\|Ax^* - b\|^2 + \lambda \|x_{\mathcal{S}^c}^*\|_1 \leq s\lambda^2 \phi_1^{-2}(s, s), \quad (4.5)$$

$$\|x^* - \bar{x}\|^2 \leq 2s\lambda^2 \phi_1^{-4}(s, s). \quad (4.6)$$

在REC前提假设下, 上述定理中的Oracle性质(4.5)估计了 $\ell_1$ 正则化模型的解在像空间(观测值)上的残差与错误预测( $\mathcal{S}^c$ 上)的惩罚值的和的上界;  $\ell_2$ 相容性(4.6)估计了 $\ell_1$ 正则化模型与稀疏优化模型的解的误差界, 即:

$$\|x^*(\ell_1) - \bar{x}\|^2 = \mathcal{O}(\lambda^2 s).$$

**Theorem 4.3** ([18, 命题8与定理9]). 记 $\bar{x}$ 与 $x^*$ 分别为稀疏优化模型(3.2)与 $\ell_0$ 正则化模型(3.4)的最优解,  $\mathcal{S} := \text{supp}(\bar{x})$ ,  $s := \|\bar{x}\|_0$ 。假设

$$\phi_0(s, s) > 0.$$

那么,  $\ell_0$ 正则化模型(3.4)的Oracle性质与 $\ell_2$ 相容性成立:

$$\|Ax^* - b\|^2 + \lambda \|x_{\mathcal{S}^c}^*\|_1 \leq s\lambda \phi_0^{-1}(s, s),$$

$$\|x^* - \bar{x}\|^2 \leq 2s\lambda \phi_0^{-2}(s, s).$$

**Theorem 4.4** ([18, 命题8与定理9]). 记 $\bar{x}$ 与 $x^*$ 分别为稀疏优化模型(3.2)与 $\ell_{1/2}$ 正则化模型(3.6)的最优解,  $\mathcal{S} := \text{supp}(\bar{x})$ ,  $s := \|\bar{x}\|_0$ 。假设

$$\phi_{1/2}(s, s) > 0.$$

那么,  $\ell_{1/2}$ 正则化模型(3.6)的Oracle性质与 $\ell_2$ 相容性成立:

$$\|Ax^* - b\|^2 + \lambda \|x_{\mathcal{S}^c}^*\|_{1/2}^{1/2} \leq s\lambda^{4/3} \phi_{1/2}^{-2/3}(s, s),$$

$$\|x^* - \bar{x}\|^2 \leq 2s\lambda^{4/3} \phi_{1/2}^{-8/3}(s, s).$$

定理4.3与4.4分别在更弱的条件下, 即0-REC(即SEC)或1/2-REC, 分别建立了 $\ell_0$ 正则化模型和 $\ell_{1/2}$ 正则化模型的模型误差界, 即:

$$\|x^*(\ell_0) - \bar{x}\|^2 = \mathcal{O}(\lambda s), \quad \|x^*(\ell_{1/2}) - \bar{x}\|^2 = \mathcal{O}(\lambda^{4/3} s).$$

根据规则性条件的关系(4.4)以及文献[25], 我们可知 $\ell_0$ 正则化模型和 $\ell_{1/2}$ 正则化模型具有更好的稳定性, 适用于更多类型的线性系统或数据拟合, 也适用于更广泛的实际应用问题。

## 5 计算方法

鉴于稀疏优化模型的广泛应用，在过去的二十年中，诸多国内外学者提出并研究了稀疏优化模型/正则化模型的许多先进计算方法。本节我们主要介绍两类流行的算法：贪婪算法与一阶迭代算法。

### 5.1 贪婪算法

贪婪算法的思想是通过最小化线性系统的残差来一步步构造近似解或者解的支撑集。它们的主要优点是简单、易于理解且快速方便；它们的常见缺陷是对噪声和高维非线性模型的影响较为敏感。目前比较流行的贪婪算法有OMP[29]、CoSaMP[21]、Foba[36]、Lar[13]。

**OMP** 正交匹配追踪算法(OMP)是一个用于逼近问题(3.2)的近似解的最基本的前向贪婪算法，它是由Tropp[29]提出来用于恢复基于少量含噪线性样本的高维稀疏信号。OMP的思想是：逐步选择与当前残差最匹配的 $A$ 的某一行，在每次迭代步中都获得最显著的更新，从而选取线性系统的稀疏表示。OMP的迭代过程为

$$r^k = b - Ax^{k-1}, \quad (5.1)$$

$$S^k = S^{k-1} \cup \arg \max_{i \notin S^{k-1}} | \langle A_i, r^k \rangle |, \quad (5.2)$$

$$x^k = \arg \min_x \{ \|Ax - b\| : \text{supp}(x) \subseteq S^k \}. \quad (5.3)$$

迭代步(5.1)计算当前迭代点的观测残差 $r^k$ ，迭代步(5.2)添加与残差 $r^k$ 相关性最高的 $A$ 的列指标到支撑集 $S^k$ 中，迭代步(5.3)采用最小二乘法在新的支撑集 $S^k$ 中更新解的估计，其解析表达式为 $(x^k)_{S^k} = (A_{S^k}^\top A_{S^k})^{-1} A_{S^k}^\top b$ 。因此，OMP的主要优点是算法思想简单易懂，迭代过程均具有显式表达，计算快速方便。而且，由于算法参数中限定了了解的稀疏性，可以有效避免出现“过拟合”现象。由于OMP中支撑集 $\text{supp}(x^k)$ 的元素只增不减，它的缺点是无法有效纠正迭代过程中出现的错误指标。文献[6]证明了在MIP条件<sup>§</sup>假设下，OMP能够高概率地准确恢复信号的支撑集。

**CoSaMP** 压缩采样匹配追踪(CoSaMP)算法是由Needell与Tropp[21]提出的一个OMP的改进算法，具有比OMP更好的数值表现。CoSaMP的思想是结合OMP思想与采样技巧，在每次迭代中，将一些随机样本加入到选定的支撑集中，并采用最小二乘法对所选支撑集进行解的估计。CoSaMP的迭代过程为

$$\begin{aligned} r^k &= b - Ax^{k-1}, \\ y &= A^\top r^k, \\ \mathcal{T} &= \text{supp}(x^{k-1}) \cup \text{supp}([y]_{2s}), \end{aligned} \quad (5.4)$$

$$z = \arg \min_x \{ \|Ax - b\| : \text{supp}(x) \subseteq \mathcal{T} \}, \quad (5.5)$$

$$x^k = [z]_s, \quad (5.6)$$

---

<sup>§</sup>MIP是比RIP更强的假设条件

其中 $[z]_s := \arg \min\{\|x - z\| : \|x\|_0 \leq s\}$ 表示向量 $z$ 的最佳 $s$ 近似。迭代步(5.4)在 $\text{supp}(x^{k-1})$ 中汇入 $2s$ 个样本得到新的支撑集 $\mathcal{T}$  (它始终不超过 $3s$ 个元素), 迭代步(5.5)采用最小二乘法在支撑集 $\mathcal{T}$ 中更新解的估计, 迭代步(5.6)仅保留最大的 $s$ 个分量进而限制支撑集 $\text{supp}(x^k)$ 不超过 $s$ 个元素。OMP每次只添加一个指标至支撑集 $\text{supp}(x^k)$ 中, 与之不同的是, CoSaMP的支撑集 $\text{supp}(x^k)$ 是从 $\mathcal{T}$ 中选择更新的, 从而具有支撑集自我矫正功能。因此, CoSaMP不仅继承了OMP的简单快速方便的优点, 而且弥补了OMP的矫正能力不足的缺点, 享有比OMP更强的理论保障。文献[21]证明了在RIP条件假设下, CoSaMP收敛到问题(3.2)的一个近似稀疏解。

**Foba** 弥补前向贪婪算法的矫正能力的另一种方法是后向贪婪算法。后向贪婪算法的思想是: 用所有的指标来做最小二乘拟合, 在每次迭代中的当前指标集内删去一个使拟合误差增加最小的指标。后向贪婪算法能够有效克服前向贪婪算法的矫正能力不足的缺点, 但是, 它从完全指标集开始依次拟合, 计算量过大, 难以实现。结合前向贪婪算法和后向贪婪算法的思想, 张潼[36]提出了前向后向自适应贪婪算法(FoBa)。FoBa执行前向贪婪步骤OMP来添加新的支撑指标, 并采用后向贪婪步骤来删去之前的前向贪婪搜索的错误指标。FoBa的迭代过程为

依次执行(5.1), (5.2), (5.3), (5.7)

$$\delta_k = \|Ax^{k-1} - b\|^2 - \|Ax^k - b\|^2,$$

$$j^k = \arg \min_{j \in \mathcal{S}^k} \|Ax_{[j]}^k - b\|^2, \quad (5.8)$$

$$\delta' = \|Ax_{[j^k]}^k - b\|^2 - \|Ax^k - b\|^2, \quad (5.9)$$

$$\text{if } \delta' \leq \frac{1}{2}\delta_k, \quad (5.10)$$

$$\mathcal{S}^k = \mathcal{S}^k \setminus \{j^k\}, \quad (5.11)$$

$$x^k = \arg \min_x \{\|Ax - b\| : \text{supp}(x) \subseteq \mathcal{S}^k\}, \quad (5.12)$$

end

其中 $x_{[j]} := (x_1, \dots, x_{j-1}, 0, x_{j+1}, \dots, x_n)^\top$ , 即 $x_j$ 变为0, 其他分量保持不变。迭代步(5.7)执行前向贪婪步骤OMP来添加新的支撑指标, 迭代步(5.8)-(5.12)采用后向贪婪步骤来删去之前的前向贪婪搜索的错误指标。(5.10)表示自适应的规则: 当后向贪婪步骤的拟合损失的增加没有超过之前前向贪婪步骤的拟合误差减少的一半时, 我们采纳后向贪婪步骤删去错误指标。此自适应原则能够确保删去前向贪婪步骤所选择的错误指标, 同时避免消除前向贪婪步骤所搜索的正确指标。FoBa避免了前向贪婪算法和后向贪婪算法的主要缺陷, 并且分别继承了他们的简单快速方便、自我矫正功能的优点。文献[36]证明了, 在RIP或者SEC(稀疏特征值条件)的假设下, FoBa能够高概率地收敛到问题(3.2)的稀疏解, 并且还推导了Oracle不等式来展示FoBa的变量选择和预测准确度。文献[18, 36]中的数值实验表明FoBa比OMP、CoSaMP表现更出色。

## 5.2 一阶迭代算法

目前, 稀疏优化与机器学习领域中比较热门的正则化模型(包括(3.4)-(3.6))都具有特定的

结构，比如损失函数和罚函数的可加形式，罚函数的可分离性。利用这些特定的结构，很多国内外学者提出和研究了多种稀疏优化与机器学习的著名算法。其中，一阶迭代算法凭借其思想简单、计算快捷、存储低和效率高等优点，已经成为稀疏优化与机器学习领域中最实用、最广泛应用的优化算法。在本小节中，我们将介绍稀疏优化模型的几种典型的一阶迭代方法。

### 5.2.1 $\ell_1$ 正则化模型

$\ell_1$  正则化模型(3.5)是稀疏优化模型(3.2)的最佳凸逼近。它已经吸引了诸多国内外学者与工程师的关注，并且成为最有效的稀疏优化工具之一。鉴于 $\ell_1$  正则化模型是一个凸优化模型，多种经典的一阶迭代凸优化算法被应用于寻求模型(3.5)的全局最优解。我们将介绍求解 $\ell_1$  正则化模型(3.5)的三种主流的一阶迭代算法：ISTA、ADMM 和SPGL1。

**ISTA** 迭代软型阈值算法(ISTA)是由Daubechies[10]提出来求解 $\ell_1$  正则化问题(3.5)。ISTA的原理是：在每次迭代中，先后对损失函数执行梯度下降运算，对罚函数执行软型阈值运算。ISTA 的迭代过程为

$$x^{k+1} = \mathbf{S}_{\nu\lambda}(x^k - \nu A^\top(Ax^k - b)).$$

这里，软型阈值算子的定义是

$$\mathbf{S}_\tau(x) := \text{sign}(x) \odot (|x| - \tau)_+ \quad \text{for each } x \in \mathbb{R}^n, \quad (5.13)$$

其中， $\text{sign}(\cdot)$  表示符号函数， $\odot$  表示向量相应分量的乘积。由于梯度下降和软型阈值运算都是显式表达，ISTA具有计算成本低与存储需求低等优点，因而得到诸多学者的关注与研究。此外，Beck和Teboulle[3]提出了一种加速版的ISTA(FISTA)。文献[10, 3]证明了ISTA和FISTA收敛到模型(3.5)的全局最优解；在FBI或者SSP的假设条件下，文献[35, 27]证明了ISTA和FISTA的线性收敛速度。

**ADMM** 交替方向乘子算法(ADMM)是由杨俊峰与张寅[34]提出来求解 $\ell_1$  正则化模型(3.5)。显然，模型(3.5)可等价表述为

$$\begin{aligned} \min \quad & \|x\|_1 + \frac{1}{2\lambda}\|r\|^2 \\ \text{s.t.} \quad & Ax + r = b. \end{aligned}$$

该模型的增广拉格朗日函数为

$$\mathcal{L}(x, r) := \|x\|_1 + \frac{1}{2\lambda}\|r\|^2 - y^\top(Ax + r - b) + \frac{\beta}{2}\|Ax + r - b\|^2.$$

ADMM的思想是应用Gauss-Seidel方法来求解增广拉格朗日函数的最小化问题。运用Gauss-Seidel方法和ISTA的近似计算，ADMM利用由投影梯度下降步、软型阈值算子(5.13) 和乘子更新准则组成的解析公式，来计算增广拉格朗日函数的最小值。ADMM的迭代过程为

$$\begin{aligned} r^{k+1} &= \frac{\lambda}{1 + \lambda\beta} (y^k - \beta(Ax^k - b)), \\ x^{k+1} &= \mathbf{S}_{\tau/\beta} (x^k - \tau A^\top(Ax^k + r^{k+1} - b - y^k/\beta)), \\ y^{k+1} &= y^k - \gamma\beta(Ax^{k+1} + r^{k+1} - b). \end{aligned}$$

文献[34]证明了ADMM收敛到模型(3.5)的全局最优解；在强凸性或者误差界的条件假设下，文献[17, 19]证明了ADMM的线性收敛速度。

**SPGL1** 谱投影梯度算法(SPGL1)是van den Berg和Friedlander[30]提出了求解 $\ell_1$  约束优化模型(3.3)。SPGL1 本质上是选用Barzilai-Borwein 谱步长[2]的投影梯度算法，即在每次迭代中，首先采用Barzilai-Borwein搜索步长的梯度下降步骤，然后将迭代点投影到 $\ell_1$  模的球上。SPGL1的迭代过程为

$$x^{k+1} = \mathbf{P}_\tau(x^k - v_k A^\top (Ax^k - b)), \quad (5.14)$$

$$\Delta x = x^{k+1} - x^k, \quad \Delta g = A^\top A(x^{k+1} - x^k),$$

$$\text{if } \langle \Delta x, \Delta g \rangle \leq 0, \quad (5.15)$$

$$v_{k+1} = v_{\max}, \quad (5.16)$$

$$\text{else} \quad (5.17)$$

$$v_{k+1} = \min\{v_{\max}, \max\{v_{\min}, \|\Delta x\|^2 / \langle \Delta x, \Delta g \rangle\}\}. \quad (5.18)$$

end

其中， $v_{\max} > v_{\min} > 0$  是设置的步长选择的上下界， $\mathbf{P}_\tau(\cdot)$  表示到 $\{x : \|x\|_1 \leq \tau\}$  的投影。迭代步(5.14)执行投影梯度步骤，文献[30] 指出 $\mathbf{P}_\tau(\cdot)$  可以通过软型阈值算子(5.13)解析表示。迭代步(5.15)-(5.18) 计算Barzilai-Borwein谱步长，诸多研究显示Barzilai-Borwein谱步长的投影梯度法在实际应用中的效果非常好。因此，SPGL1具有计算成本低与求解效率高等优点。此外，文献[30] 指明SGPL1具有 $O(n \log n)$  的计算复杂度，并且收敛到问题(3.3) 的全局最优解。

## 5.2.2 低阶正则化模型

由于算法设计方便和成功的应用经验等因素， $\ell_1$  正则化方法在学术界与业界大受欢迎。然而，在实际应用中， $\ell_1$  正则化方法尚存在某些不足。比如， $\ell_1$  正则化模型(3.5)得到的解的稀疏性远不及真实解；在压缩传感问题中， $\ell_1$  正则化模型无法在仅有极少数的测量样本的情况下恢复原信号或图像；在很多实际应用中， $\ell_1$  正则化模型常常只能得到次优稀疏解；详见文献[8, 33, 36]。

最近，低阶正则化方法被提出来以克服 $\ell_1$  正则化方法的这些缺点，并吸引了大批学者的关注与研究。例如，[8, 18]指出 $\ell_p$  ( $0 < p < 1$ ) 正则化模型可以在更弱的RIP或者REC条件下就保证良好的恢复性或者Oracle性质；[33]指出 $\ell_{1/2}$  正则化模型具有更强的稀疏诱导能力，且能够从极少数的测量样本中得到准确的稀疏解；[18, 24]表明 $\ell_0$  和 $\ell_{1/2}$  正则化方法可以更有效地从老鼠胚胎干细胞的基因表达数据中推断出可靠的基因调控网络。然而，由于低阶正则化模型是非凸非光滑的，设计求解它的有效算法解比 $\ell_1$  正则化模型具有更大的挑战与困难。在本节中，我们介绍/推广之前所提及的ISTA，来解决 $\ell_0$  和 $\ell_{1/2}$  正则化模型。

**ITA-Hard** 基于ISTA的思想，Blumensath[5]提出了求解 $\ell_0$  正则化模型(3.4)的迭代硬型阈值算法(ITA-Hard)。与ISTA类似，ITA-Hard的原理是：在每次迭代中，先后对损失函数执行梯度下降运算，对罚函数执行硬型阈值运算(替代了ISTA中的软型阈值算子(5.13))。ITA-Hard的迭代过程为

$$x^{k+1} = \mathbf{Z}_{\lambda, v}(x^k - vA^\top (Ax^k - b)).$$

这里，硬型阈值算子的定义是

$$\mathbf{Z}_{\lambda,v}(x) = \begin{cases} x, & |x| > \sqrt{2v\lambda}, \\ 0, & |x| \leq \sqrt{2v\lambda}, \end{cases} \text{ for each } x \in \mathbb{R}^n.$$

ITA-Hard算法继承了ISTA公式简洁和快速实现等优点。文献[5]证明了ITA-Hard线性收敛到模型(3.4)的局部最优解。

**ITA-Half** 徐宗本院士[33]提出了解决 $\ell_{1/2}$ 正则化模型(3.6)的半型阈值迭代算法(ITA-Half)。与ISTA唯一不同的是，ITA-Half用半型阈值算子代替了ISTA中的软型阈值算子。ITA-Half的迭代过程为

$$x^{k+1} = \mathbf{H}_{\lambda,v}(x^k - vA^\top(Ax^k - b)).$$

这里，半型阈值算子的定义是

$$\mathbf{H}_{\lambda,v}(x) = \begin{cases} \frac{16|x|^{3/2} \cos^3\left(\frac{\pi}{3} - \frac{\psi(x)}{3}\right)}{3\sqrt{3v\lambda} + 16|x|^{3/2} \cos^3\left(\frac{\pi}{3} - \frac{\psi(x)}{3}\right)} x, & |x| > \frac{3}{2}(v\lambda)^{2/3}, \\ 0, & |x| \leq \frac{3}{2}(v\lambda)^{2/3}, \end{cases}$$

其中

$$\psi(x) = \arccos\left(\frac{v\lambda}{4} \left(\frac{3}{|x|}\right)^{3/2}\right).$$

文献[33]研究了ITA-Half的全局收敛性，文献[18]证明了ITA-Half线性收敛于模型(3.6)的一个局部最优解。

我们对于光谱分析问题进行了数值模拟实验，以显示这些稀疏优化算法的实际应用效果。图5展示了不同算法对于光谱分析实验的分析结果。如图5所示，稀疏优化模型/算法都能够快速有效地预测出待测物质的真实组成成分与含量(算法运行时间约为0.01秒)。

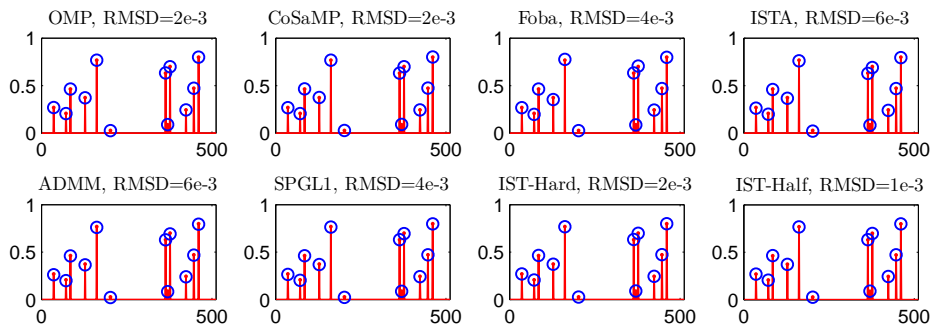


Figure 5: 光谱分析模拟实验的分析结果。红色表示待测物质的真实组成成分与含量，蓝色表示不同算法预测的结果。

我们还通过大量的模拟实验来分析这些稀疏优化算法的稳定性。如图6所示，这些稀疏优化算法在计算效率和鲁棒性能方面都具有较好的表现，对于解的不同的稀疏情况，都能具有较高的成功还原率。



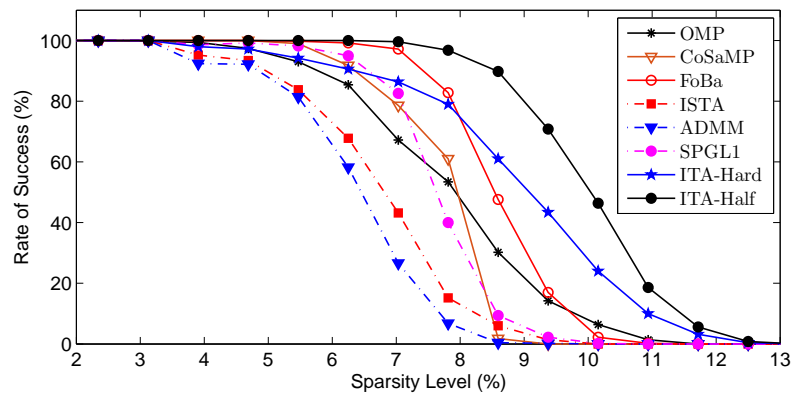


Figure 6: 500次模拟实验的稳定性分析。横轴表示解的稀疏性水平，纵轴表示不同算法的成功还原率。

## 6 结论

稀疏优化模型是目前最优化领域中非常热门的研究前沿课题，在压缩感知、图像处理、机器学习、统计建模等领域都获得了成功的应用。本文对稀疏优化模型做了简单的介绍和总结，以便读者对此模型有个清晰的认知。我们以光谱分析、信号复原、推荐系统等实际问题为例，阐述了稀疏优化模型的建模过程与核心思想。我们介绍了稀疏优化模型常见的求解方法——正则化方法，并详述了正则化模型的稳定性理论以及先进的数值算法。稳定性理论表示对于大多数独立/相关随机矩阵，正则化模型都能够在先验条件的假设下成功预测出真实模型，并且确保解的准确性。数值模拟实验表明这些数值算法都具有较好的实际应用效果、计算效率与鲁棒性能。稀疏优化模型与正则化方法将在大数据时代中更多的实际应用领域发挥更显著的计算优势与应用价值。

为了便于读者的阅读与理解，我们在这里抛砖引玉，主要介绍高维线性系统的稀疏优化模型与正则化方法。然而，大多数实际问题都是非线性问题，现今的前沿研究也延拓到非线性损失函数的稀疏优化模型，罚函数也延拓到更多的具体结构：比如非负结构、组稀疏结构等。读者不能死套公式，需要根据自己所研究的实际问题，深入理解问题的机理与结构，构造适当的损失函数和罚函数，从而建立数学模型(4.2)。

## 参考文献

### References

- [1] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski. Structured sparsity through convex optimization. *Statistical Science*, 27(4):450–468, 11 2012.
- [2] J. Barzilai and J. M. Borwein. Two-point step size gradient methods. *IMA Journal of Numerical Analysis*, 8(1):141–148, 1988.

- [3] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.
- [4] P. J. Bickel, Y. Ritov, and A. B. Tsybakov. Simultaneous analysis of lasso and dantzig selector. *Annals of Statistics*, 37(4):1705–1732, 2009.
- [5] T. Blumensath and M. E. Davies. Iterative thresholding for sparse approximations. *Journal of Fourier Analysis and Applications*, 14:629–654, 2008.
- [6] T. T. Cai and L. Wang. Orthogonal matching pursuit for sparse signal recovery with noise. *IEEE Transactions on Information Theory*, 57:4680–4688, 2011.
- [7] E. J. Candès and T. Tao. Decoding by linear programming. *IEEE transactions on Information Theory*, 51(12):4203–4215, 2005.
- [8] R. Chartrand and V. Staneva. Restricted isometry properties and nonconvex compressive sensing. *Inverse Problems*, 24(3):035020, 2008.
- [9] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Review*, 43:129–159, 2001.
- [10] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57:1413–1457, 2004.
- [11] B. Deplancke and N. Gheldof. *Gene Regulatory Networks: Methods and Protocols*. Springer, Berlin, 2012.
- [12] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(8):1289–1306, 2006.
- [13] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. *Annals of Statistics*, 32:407–499, 2004.
- [14] M. Elad. *Sparse and Redundant Representations*. Springer, New York, 2010.
- [15] J. Fan and R. Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456):1348–1360, 2001.
- [16] S. A. V. D. Geer and P. Bühlmann. On the conditions used to prove oracle results for the lasso. *Electronic Journal of Statistics*, 3:2009, 2009.
- [17] D. Han and X. Yuan. Local linear convergence of the alternating direction method of multipliers for quadratic programs. *SIAM Journal on Numerical Analysis*, 51(6):3446–3457, 2013.
- [18] Y. Hu, C. Li, K. Meng, J. Qin, and X. Yang. Group sparse optimization via  $\ell_{p,q}$  regularization. *Journal of Machine Learning Research*, 18(30):1–52, 2017.

- [19] T. Lin, S. Ma, and S. Zhang. On the global linear convergence of the ADMM with multiblock variables. *SIAM Journal on Optimization*, 25(3):1478–1497, 2015.
- [20] B. Natarajan. Sparse approximate solutions to linear systems. *SIAM Journal on Computing*, 24(2):227–234, 1995.
- [21] D. Needell and J. Tropp. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3):301–321, 2009.
- [22] H. Nyquist. Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers*, 47:617–644, 1928.
- [23] U. Platt and J. Stutz. *Differential Optical Absorption Spectroscopy: Principles and Applications*. Springer, Berlin, 2008.
- [24] J. Qin, Y. Hu, F. Xu, H. K. Yalamanchili, and J. Wang. Inferring gene regulatory networks by integrating CHIP-seq/chip and transcriptome data via LASSO-type regularization methods. *Methods*, 67(3):294–303, 2014.
- [25] G. Raskutti, M. J. Wainwright, and B. Yu. Restricted eigenvalue properties for correlated gaussian designs. *Journal of Machine Learning Research*, 11:2241–2259, 2010.
- [26] B. Recht, M. Fazel, and P. A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Review*, 52(3):471–501, 2010.
- [27] S. Tao, D. Boley, and S. Zhang. Local linear convergence of ISTA and FISTA on the LASSO problem. *SIAM Journal on Optimization*, 26(1):313–336, 2016.
- [28] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, series B*, 58(1):267–288, 1996.
- [29] J. A. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Transactions on Information Theory*, 50(10):2231–2242, 2004.
- [30] E. van den Berg and M. P. Friedlander. Probing the pareto frontier for basis pursuit solutions. *SIAM Journal on Scientific Computing*, 31(2):890–912, 2008.
- [31] J. Wang, Y. Hu, C. Li, and J.-C. Yao. Linear convergence of CQ algorithms and applications in gene regulatory network inference. *Inverse Problems*, 33(5):055017, 2017.
- [32] J. Wright, A. Ganesh, S. Rao, and Y. Ma. Robust principal component analysis: Exact recovery of corrupted low-rank matrices by convex optimization. *NIPS*, 381:2080–2088, 2009.
- [33] Z. Xu, X. Chang, F. Xu, and H. Zhang.  $L_{1/2}$  regularization: A thresholding representation theory and a fast solver. *IEEE Transactions on Neural Networks and Learning Systems*, 23:1013–1027, 2012.

- [34] J. Yang and Y. Zhang. Alternating direction algorithms for  $\ell_1$ -problems in compressive sensing. *SIAM Journal on Scientific Computing*, 33(1):250–278, 2011.
- [35] L. Zhang, Y. Hu, C. Li, and J.-C. Yao. A new linear convergence result for the iterative soft thresholding algorithm. *Optimization*, 66(7):1177–1189, 2017.
- [36] T. Zhang. Adaptive forward-backward greedy algorithm for learning sparse representations. *IEEE Transactions on Information Theory*, 57(7):4689–4708, 2011.

## 作者简介

刘思凡(1995 -), 女, 硕士生(深圳大学), 主要研究方向: 数值优化算法

王浩(1985 -), 男, 博士生(美国Lehigh大学), 助理教授, 主要研究方向: 非线性优化、机器学习

胡耀华(1984 -), 男, 博士生(香港理工大学), 副教授, 主要研究方向: 最优化理论、算法与应用