稀疏优化

给定$b \in \mathbb{R}^m$，矩阵$A \in \mathbb{R}^{m\times n}$，且向量$b$ 的维数远小于向量x的维数，即$m \ll n$. 考虑线性方程组求解问题：

$$Ax = b. \tag{2}$$

- 注意到由于$m \ll n$，方程组(2) 是欠定的，因此存在无穷多个解，重构出原始信号看似很难.
- 这些解当中大部分是不重要的，真正有用的解是所谓的“稀疏解”，即原始信号中有较多的零元素.
- 如果加上稀疏性这一先验信息，且矩阵A以及原问题的解u满足某些条件，那么可以通过求解稀疏优化问题把u与方程组(2) 的其他解区别开.
- 这类技术广泛应用于压缩感知（compressive sensing），即通过部分信息恢复全部信息的解决方案

稀疏优化

$$(\ell_0)\begin{cases}\min\limits_{x\in\mathbb{R}^n} & \|x\|_0,\\ \text{s.t.} & Ax=b.\end{cases}\quad(\ell_2)\begin{cases}\min\limits_{x\in\mathbb{R}^n} & \|x\|_2,\\ \text{s.t.} & Ax=b.\end{cases}\quad(\ell_1)\begin{cases}\min\limits_{x\in\mathbb{R}^n} & \|x\|_1,\\ \text{s.t.} & Ax=b.\end{cases}\tag{3}$$

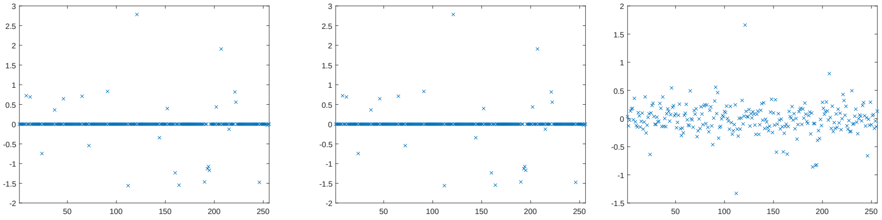
- 其中$\|x\|_0$是指x中非零元素的个数．由于$\|x\|_0$是不连续的函数，且取值只可能是整数，ℓ_0问题实际上是NP难的，求解起来非常困难．
- 若定义ℓ_1范数：$\|x\|_1=\sum_{i=1}^n|x_i|$，则得到了另一个形式上非常相似的问题，又称$\ell_1$范数优化问题，基追踪问题
- ℓ_2范数：$\|x\|_2=\left(\sum_{i=1}^n x_i^2\right)^{1/2}$

稀疏优化

在MATLAB环境里构造A, u 和b：

```
m = 128; n = 256;
A = randn(m, n);
u = sprandn(n, 1, 0.1);
b = A * u;
```

构造一个128×256矩阵A，它的每个元素都服从高斯（Gauss）随机分布．精确解u只有10%的元素非零，每一个非零元素也服从高斯分布



(a) 精确解u (b) ℓ_1问题的解 (c) ℓ_2问题的解

Figure: 稀疏优化的例子

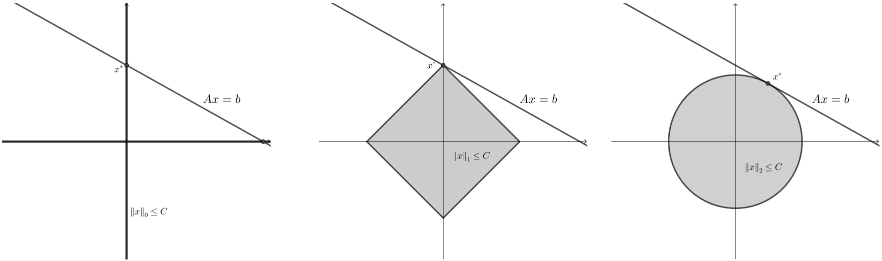
- 可以从理论上证明：若A, b满足一定的条件（例如使用前面随机产生的A和b），此时向量u也是ℓ_1范数优化问题(3)的唯一最优解．这一发现的重要之处在于，虽然问题(3) 仍没有显式解，但与零范数情形相比难度已经大大降低．
- 前面提到ℓ_0范数优化问题是NP 难问题，但ℓ_1范数优化问题的解可以非常容易地通过现有优化算法得到！从这个例子不难发现，优化学科的研究能够极大程度上帮助我们攻克现有的困难问题．但如果简单地把ℓ_1范数修改为ℓ_2范数：$\|x\|_2 = \left(\sum_{i=1}^n x_i^2\right)^{1/2}$，即求解如下优化问题：

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & \|x\|_2, \\ \text{s.t.} \quad & Ax = b. \end{aligned} \tag{4}$$

但二者是否相等是一个问题

稀疏优化

- 问题(4) 是原点到仿射集$Ax=b$的投影，可以写出它的显式表达式．但u并不是问题(4) 的解，也可从图1(a)-(c)看出该结论
- 下图可说明ℓ_1范数问题的解具有稀疏性而ℓ_2范数问题的解不具有该性质．



(a) ℓ_0范数　(b) ℓ_1范数　(c) ℓ_2范数

Figure: 三种范数优化问题求解示意图

稀疏优化

- 问题(3) 的理论和算法研究在2006年左右带来了革命性的影响．理论上研究的课题包括什么条件下问题(3) 的解具有稀疏性，如何改进这些条件，如何推广这些条件到其他应用．
- 常见的数据矩阵A一般由离散余弦变换、小波变换、傅里叶（Fourier)变换等生成．虽然这些矩阵本身并没有稀疏性，但通常具有很好的分析性质，保证稀疏解的存在性．
- 注意到绝对值函数在零点处不可微，问题(3) 是非光滑优化问题．虽然它可以等价于线性规划问题，但是数据矩阵A通常是稠密矩阵，甚至A的元素未知或者不能直接存储，只能提供Ax或$A^\mathrm{T}y$ 等运算结果．在这些特殊情况下，线性规划经典的单纯形法和内点法通常不太适用于求解大规模的问题(3)．
- 需要强调的是，问题(3) 主要特点是其最优解是稀疏向量，它是稀疏优化的一种典型形式．

LASSO问题

考虑带ℓ_1范数正则项的优化问题

$$\min_{x\in\mathbb{R}^n} \quad \mu\|x\|_1 + \frac{1}{2}\|Ax-b\|_2^2, \tag{5}$$

其中$\mu > 0$是给定的正则化参数.

- 问题(5) 又称为LASSO（least absolute shrinkage and selection operator），该问题可以看成是问题(3) 的二次罚函数形式.
- 由于它是无约束优化问题，形式上看起来比问题(3) 简单. 课件关注的大部分数值算法都将针对问题(3) 或问题(5) 给出具体形式. 因此全面掌握它们的求解方法是掌握基本最优化算法的一个标志.