

A Compressed Sense of Compressive Sensing (III)

<http://freemind.pluskid.org/machine-learning/a-compressed-sense-of-compressive-sensing-iii>

在本文第一篇的时候我们介绍了 RIP 以及相关的概念，并给出了投影矩阵 A 满足一定的 RIP 条件的时候通过 ℓ_1 -norm 最小化可以求得 CS 问题的最优解，该问题是通过凸优化来进行求解。这里我们介绍另一类主要算法 (Pursuit 类) 的一个典型，叫做 Orthogonal Matching Pursuit (OMP)。

介绍算法之前先引入一个叫做 Coherence 的概念，对于一个矩阵 $A \in \mathbb{R}^{m \times n}$ ，其 coherence 定义为

$$\mu(A) = \max_{i \neq j} \frac{|\langle A_i, A_j \rangle|}{\|A_i\|_2 \|A_j\|_2}$$

其中 A_i 表示矩阵的第 i 列。一个特殊情况是 A 的列全部正交，此时 coherence 为零，当然当 A overcomplete ($n > m$) 时这是不可能做到的，但是如果 $\mu(A)$ 很小的话，我们可以得到一些“近似正交”的结论。第一个有用的地方在于 Coherence 可以用于给出 spark 的一个下界。

我们在上次定义过，一个矩阵的 spark 是最小的数 k 使得该矩阵存在 k 列是线性相关的。我们上次介绍过，当矩阵 A 的 spark 足够大的时候，能够保证解的唯一性。但是 general 情况下计算一个矩阵的 spark 是 NP 难的，所以 Coherence 能给出一个下界就非常有用。

因为对 A 的列进行缩放并不影响 A 的 Coherence 或者 spark，不失一般性，我们考虑 A 的列全部为 unit-norm 的情况。注意 A 的任意 r 列 A 是线性无关的，等价于矩阵 $A^T A$ 的任意 $r \times r$ 的子矩阵其行列式都是 full-rank 的。另一方面，注意到 $A^T A$ 的对角线元素都是 1，而非对角元 $\langle A_i, A_j \rangle$ 都可以由 $\mu(A)$ 来 bound 住，由 Gershgorin Circle Theorem 知道，当 $r - 1 < 1/\mu$ 的时候，矩阵 $A^T A$ 的任意 $r \times r$ 子矩阵的特征值全都大于零，此时可以保证是 full-rank，亦即 $\text{spark} \geq r + 1$ 。由此可得下界： $\text{spark} \geq 1/\mu + 1$ 。类似的方法还可以将 Coherence 和 RIP 联系起来。

如果 sparsity $k < 1/2\mu$ ，此时我们有 $2k < 1/\mu < \text{spark}(A)$ ，上次我们已经证明这种情况下 k -稀疏解是唯一确定的。这里我们给出具体的 OMP 算法来将该解求出来。给出算法之前，我们再重复一下问题：已知 $A \in \mathbb{R}^{m \times n}$ 和 $b \in \mathbb{R}^m$ ，求

$$\min \|x\|_0, \quad \text{s.t. } Ax = b$$

并已知 $k < 1/2\mu$ 。OMP 算法的步骤如下：

posted on Free Mind on October 9, 2013
generated with pandoc on December 3, 2015
category: Machine Learning

tags: Compressive Sensing, Signal Processing, Algorithm

1. 初始化: $x^0 = 0, r^0 = Ax^0 - b = b, S = \emptyset$
2. For $\ell = 1, \dots, k$
 - a) 令 j 为 $\arg \max |\langle A_j, r^{\ell-1} \rangle| / \|A_j\|_2^2$
 - b) $S \leftarrow S \cup j$
 - c) $r^\ell \leftarrow \text{Proj}_{U^\perp}(b)$, 其中 $U = \text{span}(A_S)$
 - d) 如果 $r^\ell = 0$ 则退出循环
3. 求解 $A_S x_S = b$ 得出 x 的非零系数。

首先注意到每一次迭代都会有一个新的下标 j 进入集合 S , 因为如果 j 在之前的某一步中被加入到 S 中了的话, 那么 residual r 和 A_j 的内积会保持为零, 所以不会选到该下标, 亦即 j 不会进入集合两次。所以经过 k 次迭代之后将会得到 k 个下标。接下来我们要说明这 k 个下标刚好对应了 x 的 k 个非零下标。令 $T = \{j : x_j \neq 0\}$, 则我们要证明的是 $S = T$ 。

首先我们证明第一步选中的 j 是属于 T 的。令 u 是 x 的绝对值最大的元素下标, 则

$$\begin{aligned} \frac{|\langle A_u, b \rangle|}{\|A_u\|_2^2} &= \left| x_u + \sum_{j \in T/\{u\}} \frac{x_j}{\|A_u\|_2^2} \langle A_u, A_j \rangle \right| \\ &\geq |x_u| - \left| \sum_{j \in T/\{u\}} \frac{x_j}{\|A_u\|_2^2} \langle A_u, A_j \rangle \right| \\ &\geq |x_u| - \sum_{j \in T/\{u\}} \frac{|x_j|}{\|A_u\|_2^2} |\langle A_u, A_j \rangle| \\ &\geq |x_u| - (k-1)\mu|x_u| \end{aligned}$$

另一方面, 对任意的 $i \notin T$, 我们有

$$\begin{aligned} \frac{|\langle A_i, b \rangle|}{\|A_i\|_2^2} &= \left| \sum_{j \in T} \frac{x_j}{\|A_i\|_2^2} \langle A_i, A_j \rangle \right| \\ &\leq \sum_{j \in T} \frac{|x_j|}{\|A_i\|_2^2} |\langle A_i, A_j \rangle| \\ &\leq k\mu|x_u| \end{aligned}$$

当 $k < 1/2\mu$ 时, 显然前者大于后者。因此 $\arg \max$ 时一定不会选到 T 以外的下标 (虽然并不一定就是选中绝对值最大的那个 u)。也就是说, 选中的下标会是 T 中的一员。

注意到第一步之后更新 residual 之后满足 $r \in \text{span}(A_{\tilde{T}})$, 其中 $\tilde{T} = T/\{j\}$, 也就是除去了第一步所得到的下标。实际上我们可以看成是重新开始一个新的问题, 其中 \tilde{x} 是把 x 的第一步选出来的下标置零, 而 \tilde{b} 则为 residual r 。这样我们可以重复刚才的证明, 得到这一步选出来的下标还是属于 $\tilde{T} \subset T$ 。以此类推。

最后，由于我们一开始证明的 Coherence 作为 spark 下界的结果，很容易知道 A_S 的列是线性无关的，因此 OMP 最后一步的线性方程组求解也能得到 justify 了。

最后，OMP 算法虽然只需要进行 k 次迭代即可求解，但是每次迭代需要重新计算投影（第 2.3 步），普通的 Matching Pursuit (MP) 并不需要在每次选取新的下标之后对所有已有的下标重新做一次正交投影，所以计算量小了很多，但是由于在这里没有了正交性，对算法的正确性等分析也相对变得更复杂了。此外，需要注意的是，这里只讨论了最简单的没有 noise、严格 k -sparse 的情况。