

再说说稀疏信号处理

2018. 12. 22

过去数十年来，信号处理领域中各种方法层出不穷。其中有很多方法没有明确的物理背景或者意义，或者在非常特殊的情况下才有，完全只是某些数学的或者数据的折腾（英语叫 manipulation）。或者有的物理意义也只是粗线条的，很明显的等等。我认为这些算法，除了可以试一下外，很难能说出别的道道。

举个最简单的例子。在很多情况下，最大似然估计是最优的。在通信里面，最大似然估计在很多情况下是等价于最小距离判决，即哪个最靠近，哪个最有可能，就选哪个。但是，这种判决法只有在信噪比不是特别低的时候，才有道理。如果信噪比特别低了，说不定离得越远的越好呢。这个信噪比的界正是香农极限。香农极限说了，如果信噪比低于这个界，信号就检测不出来了，其实这时由于信噪比太低了，越接近的就不见得是越对的了。也就是说，在这种应用中，当信噪比低到一定程度后，所谓最优的最大似然估计法就毫无意义了。

过去十几年（2006年后）中信号处理界里最热的词可能就是“稀疏性”了。当然，稀疏性就是一种物理现象或者意义，但是我觉得只是粗线条的，从本质上就是自动假设的（英语叫 Inherited）。所以人们在求信号的优化过程中加上一个最小个数的限制。但是，信号个数这个度量（叫 L_0 ）在数学上不是一个范数，也就是说，不能确定更“接近”的一定更好（有点类似于上面的通信问题中低信噪

比的情况)，数学家说了，这时数学工具不好使。所以，人们就用了一个“最接近” L_0 的范数，“当然”就是 L_1 了。是范数了，就有距离的概念了，就可以说近就是好。问题是这个好与 L_0 好是一回事么？用 L_1 求出来的个数少是真的少么？

九十年代初的OMP法找信号是在信号库中一个一个找。先把能量最大的找到，然后减掉，再找下面的。对这种方法来说，信号库非常重要。如果信号库对了，还有啥方法可代替么？我认为没有。人们也许会问，此法利用到稀疏性了么？当然用到了，信号都是一个一个找了，还没用到稀疏性么？信号的个数总不会比一还少吧！当然，理论上可以多个多个一起找，但是这样的复杂度会太高了。

这里有两个问题。其一是信号库的问题。怎么知道信号库是对的？这个问题跟现在的稀疏信号处理的问题一样。稀疏信号处理也有类似的问题，即信号在什么域里是稀疏的，这就等价于有了正确的信号库。从这一点来说，九十年代的OMP与现在的稀疏信号处理一样。

第二个问题也许是计算复杂度。其实现在的稀疏信号处理一般优化算法的复杂度更高，为了降低计算复杂度，人们反而正是用OMP法来解。

从上面两点可以看出，九十年代初的OMP法与现在的稀疏信号处理法基本上是等价的，现在的各种方法都只是OMP法的各种变形，这点并不奇怪，从九十年代到现在都20多年了，本来也该会有对OMP法的自然变异了，而并非是因为是稀疏信号处理的推动。尽管是这

么说，但是，由于稀疏性名字的出现，所以有更多的人被吸引过来做了。遗憾的是，现在在这个领域里的很多人都让各种优化给打鸡血了，难道这些优化算法能改进 OMP 法么？

再回到稀疏性的度量 L_0 ，假如说 L_0 可以在数学上执行。因为稀疏信号处理的问题大多是不定的，即会有很多解。这时，稀疏性的限制就是说个数越少的越好。我觉得这样的思考都是在没有考虑噪声或者信噪比高的情况下的结论。哪如果信噪比不高呢？或者说信噪比要高到什么程度？

我的估计是这样的。在稀疏域中，噪声还是满域的且统计意义上是平的，叫噪声地层 (noise floor)。在噪声地层上面大于 X dB 就可以被视为信号。这个 X 就是稀疏信号处理的界，类似于上面讲的香农界。这个 X 是什么？0？还是 1.6？

也许这一点与数字通信理论完全不一样，因为在这里，信号的数值是任意的，无穷的，而数字通信理论中的信号只是有限的固定的。在这里，信号是人为定的，所以，上面的界 X 就是 X ，只知道它一定大于 0。