

文章题目：Regression shrinkage and selection via the LASSO

领域：回归模型

核心创新点：产生具有解释性的模型 LASSO，并可以进行进一步的泛化回归和拓展

论文结构与实现方法

OLS（简单线性回归）的缺点：无偏性与高方差造成准确率过低；解释性过低；

论文主要围绕如何解决上文问题（通过适当提升偏差（bias）来降低方差提升准确率，设立阈值忽略小变量，优化解释性能）

在原来的 OLS 过程的优化问题上加入了 L1 范式，来综合性的平衡上面的内容（核心就是这个，剩下的十几页主要就是把数理统计中对结果的估计（各种检验 bayes 检验、标准差分析等等）证明针对这一种情况展开分析并与其的三种回归方法进行比较（这三种都没听过，查了一下效果都很差的那种。。。），并且用各种方式的展现形式对这个原理进行展现），并如何求解、如何拓展、可能的用途等等

$$\sum_{i=1}^N \left(y_i - \sum_j \beta_j x_{ij} \right)^2 + \lambda \sum_j \beta_j^2$$

or, equivalently, minimizes

$$\sum_{i=1}^N \left(y_i - \sum_j \beta_j x_{ij} \right)^2 \quad \text{subject to } \sum_j \beta_j^2 \leq t. \quad (4)$$

优秀之处 数学推理占主导，进行了规模庞大的推导。对数学公式进行了多种形式的演绎（几何（2d， 3d）， 对比，例证，模拟等等）

可采用的点 个人觉得目前我们的水平无法信手拈来一个这么大型的数学项目，尤其是不知解寻求未知解的能力，所以从论文结构来说，我们不太可能模仿。从其他相对皮毛一些的内容来说，里面对数学公式的图形化解释可以学习；各种统计学上近乎套路式的检验可以学习。

PS：实际上来说这篇论文没看完，但是如果死扣数学公式的和检验证明的话，我觉得这一周所有科研时间都加上也不够。。。如果真的有可能的话，能做成这样真的很不错，就是有了思想之后，一连串完整的数学推导和解析。