Locally Adaptive Weighting and Screening (LAWS)

1.改进目标：本文适用于spatial(数据推断，应用LAWS去直接包含局部稀疏推断，其可以应用在很多空间数据上且是完全数据驱动，可用在2D/3D图像上。常常相依结构被认为是一个先验或从其中数据中可以较好地的估计，这样会有一些问题(1)先验是错误的，(2)估计的优劣取决于是否满足假设，(3)相依结构导致了强的正则化条件和繁冗的计算成本。

① 本文的目标：提出一种能自适应学习数据中的稀疏结构在潜在的空间过程中的FDR控制方法，主要的想法是利用辅助信息去重构多重检验问题。

② LAWS首先用screening方法估计局部的稀疏结构，其次构建空间权重去调整p值，最后选择p值去调整多重性。LAWS会在信号多的地方上调p值，反之下调p值。

③ 大范围尺度检验，若利用辅助信息，以前的文献由(1)在原数据构建辅助统计量(2)外部数据或先验。我们区别于这两种方法，用空间次序去反映局部的类和平滑性，而不是利用定性/定量变量去分析。注：grouping分析不适用于空数据分析，因为分成八个好的组需要好的先验或繁冗的计算，且分组涉及到离散一个连续性变量，会造成信息损失。

④ 优点：(1)应用广泛，2D-3D图，(2)由最优化理论导出，有扎实的理论基础，(3)完全数据驱动且提供了一种客观的方法去包含side information，(4)证明LAWS有效的时候给了一种不同的证明。

⑤ 定义稀疏程度 $\pi(s)=P(\theta(s)=1)$，其光滑性提供了关键的结构性信息，可以通过整合其邻近点的信息，进而构造更有效的方法。LAWS属于一种加权方法，根据CLfdr定义，由CAPS得其等价于一个加权分类问题。$\text{CLfdr}=\frac{\Lambda}{\Lambda+1}$，$\Lambda=\frac{1-\pi(s)}{\pi(s)}\frac{f_0}{f_1}$，关于Λ单调，而Λ中包含两种信息。(1)稀疏结构：信号有多频繁出现在s的邻域，$\frac{1-\pi(s)}{\pi(s)}$，(2)数据本身反映出该点是否为null的信息，似然比 $\frac{f_0(x|s)}{f_1(x|s)}$，因为后者很难计算，提出用p值去替代，权重为 $\frac{1-\pi(s)}{\pi(s)}$。
$p^w(s)=\min\{\frac{1-\pi(s)}{\pi(s)}p(s),1\}=\min\{\frac{p}{w},1\}$。π大，$p^w(s)\downarrow$。BH控制FDR为$(1-\pi)\alpha$。说明π可以提高Power。要注意，加权后密度会发生变化。

⑥ 要去估计一些量，如$\pi(s)$以及t_{OR}，我们首先利用核光滑的非参数方法，然后用数据估计t_{OR}。
$\pi^\tau(s)=1-\frac{P(p(s)>\tau)}{1-\tau}$，$\tau\in(0,1)$，且 $\frac{\pi^\tau(s)-\pi(s)}{\pi(s)}=-\frac{1-G_1(\tau|s)}{1-\tau}<0$。$\pi^\tau(s)$为保守估计，要去估计$\pi^\tau(s)$。
(1)smothing and screening，$\hat\pi^\tau(s)=1-\frac{\sum_{s'\in T_\tau}V_h(s,s')}{(1-\tau)\sum_{s'\in S}V_h(s,s')}$，$T_\tau=\{s\in S, p(s)>\tau\}$，$V_h(s,s')=\frac{K_h(s-s')}{K_h(0)}$，K为核函数。(2)进而，有 $\hat w(s)=\frac{\hat\pi(s)}{1-\hat\pi(s)}$，为了使算法稳定，令$\hat\pi(s)$限制在$(\nu,1-\nu)$内，$\nu$很小。(3)因为错误发现数的期望 $EFP=\sum P(p^w(s)\le t_w|\theta(s)=0)P(\theta(s)=0)=\sum\pi(s)t_w$，则若拒绝j个假设，则有大约 $\sum\hat\pi(s)p^w_{(j)}$ 个假阳，则 $FDP=\frac{1}{j}\sum\hat\pi(s)p^w_{(j)}$

2.方法步骤：Step 1：计算 $\hat\pi(s)$，$\hat w(s)$ 以及各p值和调整p值 p^w，并排序。

Step 2：令 $k^w=\max\{j:\frac{1}{j}\sum\hat\pi(s)p^w_{(j)}\le\alpha\}$

Step 3：拒绝 $H_{(1)},\cdots,H_{(k^w)}$

同濟大學
TONGJI UNIVERSITY
地址：中国上海市四平路1239号　邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21-　传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

3.假设前提.

(1). 设 $\pi^{\tau}(\cdot)$ 有连续的二阶偏导数，存在 $C>0$ s.t. $-C \le \lambda_{min}(s) \le \lambda_{max}(s) \le C$. $\forall s$. 其中 λ_{min} λ_{max} 分别表示最小和最大特征值(黑塞矩阵)

(2) 设 $Var(\sum I(p(s)>\tau)) \le C' \sum Var(I(p(s)>\tau))$. 对某个常数 $C'>1$.

(3) $(r_{i,j})_{m\times m} = R = corr(Z)$. 设 $\max_{1\le i<j\le m} |r_{i,j}| \le r < 1$, $r>0$. 进而 $\exists\, \gamma>1$ s.t. $\max_{\{i:\theta(s_i)=0\}} |\Gamma_i(\gamma)| = o(m^{K})$, 其中 $0<K<\frac{1-r}{1+r}$, 其中 $\Gamma_i(\gamma) = \{j: 1\le j\le m\}, |r_{i,j}| \ge (\log m)^{-2-\gamma}\}$

(4) $\exists\, \xi>0$ s.t. $\pi^{\tau}(s) \in [\xi, 1-\xi]$. 且 $Var[\sum I(\theta(s)=0)] = O(m^{1+\xi})$, $\xi \in [0,1)$

(5) 定义 $S_\rho = \{i: 1\le i\le m, |\mu_i| \ge (\log m)^{\frac{1+\rho}{2}}\}$. $\mu_i = E(Z_i)$. 对 $\rho>0$. $\delta>0$. $|S_\rho| \ge [\frac{1}{\sqrt{\pi}\alpha}+\delta]\sqrt{\log m}$.

NOTE: (1)是对 Alternative CDF $G_1(t|s)$ 的 mild regularity condition.
(2)假设大部分 p 值是弱相关的. (3)假设了大部分 null p 值是弱相依的，允许一部分强相依且没有对 non-nulls 施加条件. (4) $\theta(s)$ 不是完全相关，可以高度相关.

4.理论结果. $\Rightarrow$ 说明 δ_{OR}^{w} 一致优于 δ_{OR}, 由 t_{OR} 定义知 $Q^{w}(t_{OR}^{w}) \le \alpha$. 说明的是 $w(s)$ 的好!

Thm 1 (Oracle). 设 $\frac{\sum \pi(s)}{\sum(1-\pi(s))} \le 1$, $\forall s \in S$, 若 $t \to G_1(t|s)$ 凹且 $x \to G_1(\frac{t}{x}|s)$ 凸, $\forall \min w^{-1}(s) \le x \le \max w^{-1}(s)$

则有 (a) $Q^{w}(t_{OR}^{1}) \le Q^{1}(t_{OR}^{1}) \le \alpha$　(b) $\Psi^{w}(t_{OR}^{w}) \ge \Psi^{w}(t_{OR}^{1}) \ge \Psi(t_{OR}^{1})$

Prop 1 (Estimate $\pi(s)$). 在(1),(2)成立下, $\forall s \in S$, $E\{\hat{\pi}^{\tau}(s) - \pi^{\tau}(s)\}^2 \to 0$. $s \to S$

Thm 2 (Oracle LAWS) 在(3),(4),(5)下, $\forall \varepsilon>0$. $\limsup_{m\to\infty} FDR(\delta^{w}) \le \alpha$, $\lim_{m\to\infty} P(FDP(\delta^{w}) \le \alpha+\varepsilon) = 1$

Thm 3 (Data-Driven) 在 Prop 1. Thm 2 条件下, $\forall \varepsilon>0$. $\limsup_{m\to\infty} FDR(\delta^{w}) \le \alpha$. $\lim_{m\to\infty} P(FDP(\delta^{w}) \le \alpha+\varepsilon) = 1$