Weak Signal Identification & Inference in Penalized Model Selection (Shi & Qu 2017).

① Achieve better tradeoff between false discovery & weak signal detection power

② finite sample thm should be used instead of asymptotically. for weak ...

Basics about Adaptive LASSO: $\hat{\theta}_{ALASSO} = \left(|\hat{\theta}^{LS}| - \frac{\lambda}{|\hat{\theta}^{LS}|} \right)_+ \operatorname{sgn}(\hat{\theta}^{LS})$

$A_n \to \overset{true}{A}$ if $\sqrt{n}\lambda_n \to 0$, $n\lambda_n \to \infty$. (order of λ_n between $\frac{1}{n} \sim \frac{1}{\sqrt{n}}$)

noise strength: $\frac{c\sigma}{\sqrt{n}}$.

Why we should not. ~~is it hard to~~ detect weak signal asymptotically?

probability of selection (for single variable): $P_d(\theta) = P(\hat{\theta}_{ALASSO} \neq 0 \mid \theta_{true}) = \Phi\left(\frac{\theta - \sqrt{\lambda}}{\sigma/\sqrt{n}}\right) + \Phi\left(\frac{-\theta - \sqrt{\lambda}}{\sigma/\sqrt{n}}\right)$.

$\forall$ pair (γ, ν^γ): $\gamma = \Phi\left(\frac{\nu^\gamma - \sqrt{\lambda}}{\sigma/\sqrt{n}}\right) + \Phi\left(\frac{-\nu^\gamma - \sqrt{\lambda}}{\sigma/\sqrt{n}}\right)$ signal should be at least as strong as ν^γ to be detected with $p = \gamma$.

define transition phase: $\begin{cases} \theta \in \Theta^{(S)} & P_d > \gamma^s \\ \theta \in \Theta^{(W)} & \gamma^w < P_d < \gamma^s \\ \theta \in \Theta^{(N)} & P_d \le \gamma^w. \end{cases}$

Lemma: for $\gamma \in (\gamma^w, \gamma^s)$ bdd away from 0, 1, $\frac{\nu^\gamma}{\sqrt{\lambda_n}} \to 1$

good: $\nu^\gamma \to 0$

bad $\nu^\gamma \sim \sqrt{\lambda_n} \sim (n^{-\frac{1}{2}}, n^{-\frac{1}{4}})$

slower than noise strength.

in finite sample case, ν^γ is still too large to be ignored.

Detection: select $\nu_1 = z_{\frac{\tau}{2}} \frac{\sigma}{\sqrt{n}}$. $\leftarrow \tau$ FDR tolerance. $\tau > \tau_0$, more detection.

$\nu_2 = \sqrt{\lambda} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \to$ lower bound for $P_d = 1 - \alpha$.

Inference: detected strong signals: asymptotic theory based interval by Zou.

(why not apply to weak signals: no inference for undetected weak signals).

underestimated ~~sd~~ se (too conservative for heavily detected weak signal).

...... weak ... : CI based on LSE (for full model).

Similarity with PoSI: only construct CI for selected param.