

L^AT_EX Template for STAT 548 Qualifying Paper Report

Naitong Chen

September 22, 2022

1 Summary

Tibshirani and Wasserman (2015)

Stein's Lemma:

- (univariate) Let $Z \sim \mathcal{N}(0, 1)$. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be absolutely continuous, with derivative f' (and assume that $\mathbb{E}[|f'(Z)|] < \infty$). Then $\mathbb{E}[Zf(Z)] = \mathbb{E}[f'(Z)]$.
- (extesion) Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Then $\frac{1}{\sigma^2} \mathbb{E}[(x - \mu)f(x)] = \mathbb{E}[f'(X)]$.
- (multivariate) Let $X \sim \mathcal{N}(\mu, \sigma^2 I)$, where $\mu \in \mathbb{R}^n$ and $\sigma^2 I \in \mathbb{R}^{n \times n}$. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function such that, for each $i = 1, \dots, n$ and almost every $x_{-i} \in \mathbb{R}^{n-1}$, $f(\cdot, x_{-i}) : \mathbb{R} \rightarrow \mathbb{R}$ is absolutely continuous (and assume $\|f(X)\|_2 < \infty$). Then $\frac{1}{\sigma^2} \mathbb{E}[(X - \mu)f(X)] = \mathbb{E}[\nabla f(X)]$.
- (extension) Let $f = (f_1, \dots, f_n)$, then

$$\begin{aligned} & \frac{1}{\sigma^2} \mathbb{E}[(X - \mu)f_i(X)] = \mathbb{E}[\nabla f_i(X)] \\ \implies & \frac{1}{\sigma^2} \sum_{i=1}^n \text{Cov}(X_i, f_i(X)) = \frac{1}{\sigma^2} \sum_{i=1}^n \mathbb{E}[(X_i - \mu_i)f_i(X)] = \mathbb{E} \left[\sum_{i=1}^n \frac{\partial f_i}{\partial X_i}(X) \right]. \end{aligned}$$

Stein's Unbiased Risk Estimate (SURE):

Given samples $y \sim \mathcal{N}(\mu, \sigma^2 I)$, and a function $\hat{\mu} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\hat{\mu}$ is a fitting procedure that, from y , provides an estimate $\hat{\mu}(y)$ of the underlying (unknown) mean μ . Then

$$\begin{aligned} R &= \mathbb{E}_y \|\mu - \hat{\mu}(y)\|^2 \\ &= -n\sigma^2 + \mathbb{E} \|y - \hat{\mu}\|_2^2 + 2\sigma^2 \text{df}(\hat{\mu}) \\ &= -n\sigma^2 + \mathbb{E} \|y - \hat{\mu}\|_2^2 + 2 \sum_{i=1}^n \text{Cov}(y_i, \hat{\mu}_i), \end{aligned}$$

where $\text{df}(\hat{\mu}) = \frac{1}{\sigma^2} \sum_{i=1}^n \text{Cov}(y_i, \hat{\mu}_i)$. And

$$\hat{R} = -n\sigma^2 + \|y - \hat{\mu}(y)\|_2^2 + 2\sigma^2 \sum_{i=1}^n \frac{\partial \hat{\mu}_i}{\partial y_i}(y)$$

is an unbiased estimate for R .

Extending SURE to regularized estimators:

Now suppose $\hat{\mu}_\lambda$ depends on $\lambda \in \Lambda$, which controls the degree of regularization to our estimator, and assume σ is known, we can find the optimal λ , denoted $\hat{\lambda}$ by

$$\hat{\lambda} = \arg \min_{\sigma \in \Sigma} \|y - \hat{\mu}_\lambda(y)\|_2^2 + 2\sigma^2 \sum_{i=1}^n \frac{\partial \hat{\mu}_{\lambda,i}}{\partial y_i}(y).$$

2 Mini-proposals

2.1 Proposal 1: MY PROPOSAL TITLE

2.2 Proposal 2: MY OTHER PROPOSAL TITLE

3 Project report

References

Ryan Tibshirani and L Wasserman. Stein's unbiased risk estimate. *Course notes from "Statistical Machine Learning*, pages 1–12, 2015.