David Pollard: 2 November 94, Wharton

Minimax rates and distances between convex hulls

David Pollard
Yale University
pollard@stat.yale.edu

1

Collection $\mathcal{P}$ of probability measures on $(\Omega, \mathcal{A})$. Functional $\theta : \mathcal{P} \to \Theta$. Estimator $\widehat{\theta} : \Omega \to \Theta$. Loss function $L(t, \theta)$. Constant $c(\theta_0, \theta_1)$ is the largest for which $L(t, \theta_0) + L(t, \theta_1) \geq c$ for all t. For quadratic loss, $c(\theta_0, \theta_1) = \frac{1}{2}|\theta_1 - \theta_0|^2$. Minimax criterion: minimize $\mathcal{R}(\widehat{\theta}, \mathcal{P}) = \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}L(\widehat{\theta}(\omega), \theta(\mathbb{P}))$.

2

Density $q_i = d\mathbb{Q}_i/d\mu$. Define L^1 distance $\|\mathbb{Q}_0 - \mathbb{Q}_1\|_1 = \mu|q_0 - q_1|$. Affinity $\|\mathbb{Q}_0 \wedge \mathbb{Q}_1\|_1 = \mu(q_1 \wedge q_2) = \inf\{\mathbb{Q}_0 f + \mathbb{Q}_1 f_1 : f_i \geq 0,\ f_0 + f_1 \geq 1\} = 1 - \frac{1}{2}\|\mathbb{Q}_0 - \mathbb{Q}_1\|_1$. Lower bound

$$\boxed{\mathcal{R}(\widehat{\theta}, \mathcal{P}) \geq \tfrac{1}{2} \inf\{c(\theta(\mathbb{P}_0), \theta(\mathcal{P}_1)) : \mathbb{P}_i \in \mathcal{P}_i\} \alpha(\mathrm{co}\,(\mathcal{P}_0)\,, \mathrm{co}\,(\mathcal{P}_1))}$$

where $\alpha(\mathrm{co}\,(\mathcal{P}_0)\,, \mathrm{co}\,(\mathcal{P}_1)) = \sup\{\|\mathbb{Q}_0 \wedge \mathbb{Q}_1\|_1 : \mathbb{Q}_i \in \mathrm{co}\,(\mathbb{P}_i)\} = 1 - \frac{1}{2}\inf\{\|\mathbb{Q}_0 - \mathbb{Q}_1\|_1 : \mathbb{Q}_i \in \mathrm{co}\,(\mathcal{P}_i)\}$.

Proof. Put $f_i(\omega) = \inf_{\mathbb{P}\in\mathcal{P}_i} L(\widehat{\theta}(\omega), \theta(\mathbb{P}))/c$. Then

$$2\mathcal{R}(\widehat{\theta}, \mathcal{P})/c \geq \mathbb{P}_0 L(\widehat{\theta}, \theta(\mathbb{P}_0)) + \mathbb{P}_1 L(\widehat{\theta}, \theta(\mathbb{P}_1))/c \geq \mathbb{P}_0 f_0 + \mathbb{P}_1 f_1, \qquad \text{all } \mathbb{P}_i \in \mathcal{P}_i.$$

for all $\mathbb{P}_0 \in \mathcal{P}_0$ and $\mathbb{P}_1 \in \mathcal{P}_1$. Take convex combinations of measures from $\mathcal{P}_i$, giving measures $\mathbb{Q}_i$ in the convex hulls, then invoke definition of the affinity between the $\mathbb{Q}_i$. □

3

Product probability measure $\mathbb{Q}_0$ on $\mathcal{X}^n$ and $d\mathbb{Q}_\lambda/d\mathbb{Q}_0 = \prod_{i\leq n}(1 + G_i(\lambda))$, where $G_i(\lambda) = G_i(x_i, \lambda)$ is a function of the ith coordinate x_i for which $\mathbb{Q}_0 G_i(x_i, \lambda) = 0$ for each λ and i. Define $\mathbb{Q} = \sum_{\lambda\in\Lambda} w_\lambda \mathbb{Q}_\lambda$.

$$\frac{d\mathbb{Q}}{d\mathbb{Q}_0} = \Delta = \sum_{\lambda\in\Lambda} w_\lambda \Big(1 + \sum_i G_i(\lambda) + \sum_{i<j} G_i(\lambda)G_j(\lambda) + \sum_{i<j<k} G_i(\lambda)G_j(\lambda)G_k(\lambda) + \ldots\Big),$$

the sums ending with the n-fold product. Use $\|\mathbb{Q} - \mathbb{Q}_0\|_1^2 = (\mathbb{Q}_0|\Delta - 1|)^2 \leq \mathbb{Q}_0|\Delta - 1|^2$. Then get

$$\boxed{\|\mathbb{Q} - \mathbb{Q}_0\|_1^2 \leq \mathbb{P} \prod_{i\leq n}\Big(1 + \mathbb{Q}_0 G_i(x_i, \lambda) G_i(x_i, \mu)\Big) - 1 \qquad \text{independent } \lambda \text{ and } \mu \text{ from prior on } \Lambda.}$$

4

Independent $x_i \sim N(\eta_i, \sigma^2)$. Estimate $\theta(\eta) = \sum_i \beta_i \eta_i^2$ subject to constraints $|\eta_i| \leq A_i$ for each i. Use quadratic loss. Fix ξ in constraint set and $\sigma > 0$. Write $\mathbb{Q}_0$ for joint distribution of $N(0, \sigma^2)$. For $\lambda \in \Lambda = \{-1, +1\}^n$ write $\mathbb{Q}_\lambda$ for $N(\lambda_i \xi_i, \sigma^2)$. Put $\mathbb{Q} = 2^{-n}\sum_{\lambda\in\Lambda} \mathbb{Q}_\lambda$.

$$G_i(x_i, \lambda) = \{\lambda_i = +1\}\left(\exp\left(\frac{\xi_i x_i}{\sigma^2} - \frac{\xi_i^2}{2\sigma^2}\right) - 1\right) + \{\lambda_i = -1\}\left(\exp\left(-\frac{\xi_i x_i}{\sigma^2} - \frac{\xi_i^2}{2\sigma^2}\right) - 1\right),$$

giving $1 + \mathbb{Q}_0 G_i(x_i, \lambda) G_i(x_i, \mu) = \{\lambda_i = \mu_i\}\exp(\xi_i^2/\sigma^2) + \{\lambda_i \neq \mu_i\}\exp(-\xi_i^2/\sigma^2)$ and

$$\|\mathbb{Q} - \mathbb{Q}_0\|_1^2 \leq= \prod_{i\leq n}\left(\tfrac{1}{2}\exp(\xi_i^2/\sigma^2) + \tfrac{1}{2}\exp(-\xi_i^2/\sigma^2)\right) - 1 \leq \exp\left(\tfrac{1}{2}\sum_i \xi_i^4/\sigma^4\right) - 1.$$

5

Problem: Maximize $\sum_i \beta_i \xi_i^2$ subject to the constraints $\frac{1}{2}\sum_i \xi_i^4/\sigma^4 \leq 1$ and $0 \leq \xi_i \leq A_i$ all i. Solution (for σ small enough): $\xi_i = A_i \wedge \sqrt{\beta_i t}$, where $\sum_i A_i^4 \wedge (\beta_i^2 t^2) = 2\sigma^4$.