Mathematical machine learning part IV : active and online learning

Prof. Dr. Gilles Blanchard, Dr. Alexandra Carpentier*, Dr. Jana de Wiljes, Dr. Martin Wahl

1. The stochastic bandit problem

`Useful material : See Bubeck et.al (2012), and also Cesa-Bianchi et.al (2006) for a broader perspective - see also https://blogs.princeton.edu/imabandit/2016/05/11/bandit-theory-part-i/ (and part ii) for a helpful blog post.`

1.1. The problem

1.2. Upper bounds

1.3. Lower bounds

An important question is on whether the algorithm presented in the last subsection is *optimal*. But first, how can we characterise optimality? A useful tool for characterizing the efficiency of a statistical methods is the concept of *minimax lower bounds* - this framework is related to information theory.

1.3.1. Examples in a classical problem

The problems As an example let us consider a much simpler statistical setting. Consider $X_1, \ldots, X_n$ that are iid according to $\mathcal{B}(\mu)$, $\mu \in [1/4, 3/4]$. Let us write $\mathbb{E}_\mu$ and $\mathbb{P}_\mu$ for the expectation and probability when the parameter is μ.

We consider the problem of estimating μ. Consider the empirical mean

$$\hat{\mu} = \frac{1}{n} \sum_i X_i.$$

It is clear that by Hoeffding's inequality that for any $\mu \in [1/4, 3/4]$, it holds with probability larger than $1 - \delta$ that

$$|\hat{\mu} - \mu| \leq \sqrt{\frac{\log(2/\delta)}{2n}} =: c_\delta / \sqrt{n}. \quad (1)$$

So the rate of estimation for this test statistic is $n^{-1/2}$ - $\hat{\mu}$ is a $(\delta, c_\delta/\sqrt{n})$ good estimator.

Question 1 : Does there exist an estimator that performs better, i.e. such that its rate is significantly smaller than $c_\delta/\sqrt{n}$ on an event of probability $1 - \delta$?

*Contact : carpentier@math.uni-potsdam.de. **Webpage with course material TBA :** http://www.math.uni-potsdam.de/~carpentier/page3.html

imsart-generic ver. 2013/03/06 file: classHU2017_3.tex date: February 14, 2017

Consider now the testing problem

$$H_0 : \mu = 1/2 \quad \text{vs} \quad H_1 : \mu = 1/2 + u.$$

Consider the test

$$T = \mathbf{1}\{\hat{\mu} > 1/2 + u/2\}.$$

By Equation (1), the sum of its power and level is bounded as

$$\mathbb{E}_{1/2}[T] + \mathbb{E}_{1/2+u}[1-T] \leq 2\exp(-\frac{u^2 n}{2}).$$

Question 2 : Does there exist a test that performs better, i.e. such that the bound on its power and level is significantly smaller than $2\exp(-\frac{u^2n}{2})$?

1.3.2. Results

Theorem 1. *It holds that*

$$\inf_{\tilde{T} \quad test} \Big[\mathbb{E}_{1/2}[\tilde{T}] + \mathbb{E}_{1/2+u}[1-\tilde{T}]\Big] \geq \frac{1}{4e^4}\exp\Big(-10nu^2\Big).$$

This implies in particular that no estimator can be uniformly better than $(\delta, \sqrt{\frac{\log(2/\delta)}{10n}})$ *good.*

An approach using likelihood ratio and concentration Let

$$L_\mu(x_1, ..., x_n) = \mu^{\sum_i x_i}(1-\mu)^{n-\sum_i x_i} = \exp\Big(\log(\frac{\mu}{1-\mu})\sum_i x_i + n\log(1-\mu)\Big).$$

Note that L_μ is the likelihood of the data when the parameter is μ.

Let $\tilde{T}$ be a test. Let $\Omega = \{T = 0\}$. We have

$$\begin{aligned}\mathbb{P}_{1/2+u}(\Omega) &= \mathbb{E}_{1/2}\Big[\frac{L_{1/2+u}(X_1,\ldots,X_n)}{L_{1/2}(X_1,\ldots,X_n)}\mathbf{1}\{\Omega\}\Big] \\ &= \mathbb{E}_{1/2}\Bigg[\exp\Big(\log(\frac{1+2u}{1-2u})\sum_i X_i + n\log(1-2u)\Big)\mathbf{1}\{\Omega\}\Bigg].\end{aligned}$$

Consider now

$$\xi = \Big\{|\sum_i X_i - n/2| \leq \sqrt{n\frac{\log(4)}{2}}\Big\}.$$

Note that $\mathbb{P}_{1/2}(\xi) \geq 3/4$.

Let us take δ so that $\mathbb{P}_{1/2}(\Omega) = 1-\delta$ (which is required so that the test has level δ). Note that we have $\mathbb{P}_{1/2}(\Omega \cap \xi) \geq 3/4 - \delta$.

Now this implies that

$$\begin{aligned}\mathbb{P}_{1/2+u}(\Omega) &\geq \mathbb{E}_{1/2}\Bigg[\log(\frac{1+2u}{1-2u})\sum_i X_i + n\log(1-2u)\mathbf{1}\{\Omega\cap\xi\}\Bigg] \\ &\geq \mathbb{E}_{1/2}\Bigg[\exp\Big(\log(\frac{1+2u}{1-2u})\big(n/2 - \sqrt{n\frac{\log(4)}{2}}\}\big) + n\log(1-2u)\Big)\mathbf{1}\{\Omega\cap\xi\}\Bigg]\end{aligned}$$

Now note that since $0 < u \le 1/4$, we have $\log(1-2u) \ge -2u - 2u^2$ and

$$\log(\frac{1+2u}{1-2u}) \ge \log((1+2u)(1+2u)) = \log(1+4u+4u^2) \ge 4u - 8u^2.$$

So we have

$$\begin{aligned}
\mathbb{P}_{1/2+u}(\Omega) &\ge \mathbb{E}_{1/2}\Bigg[\log(\frac{1+2u}{1-2u})\sum_i X_i + n\log(1-2u)\mathbf{1}\{\Omega\cap\xi\}\Bigg] \\
&\ge \mathbb{E}_{1/2}\Bigg[\exp\Big((4u-8u^2)\big(n/2 - \sqrt{\frac{n\log(4)}{2}}\big) + n(-2u-2u^2)\Big)\mathbf{1}\{\Omega\cap\xi\}\Bigg] \\
&\ge \mathbb{E}_{1/2}\Bigg[\exp\Big(-6nu^2 - 4u\sqrt{n\frac{\log(4)}{2}}\Big)\mathbf{1}\{\Omega\cap\xi\}\Bigg] \\
&\ge \exp\Big(-6nu^2 - 4u\sqrt{n\frac{\log(4)}{2}}\Big)\mathbb{E}_{1/2}\Bigg[\mathbf{1}\{\Omega\cap\xi\}\Bigg] \\
&\ge \exp\Big(-6nu^2 - 4u\sqrt{n\frac{\log(4)}{2}}\Big)[3/4-\delta].
\end{aligned}$$

So for the test $\tilde{T}$ we know that

$$\begin{aligned}
\mathbb{E}_{1/2}[\tilde{T}] + \mathbb{E}_{1/2+u}[1-\tilde{T}] &\ge \exp\Big(-6nu^2 - 4u\sqrt{n\frac{\log(4)}{2}}\Big)[3/4-\delta] + \delta \\
&\ge \frac{1}{4}\exp(-10nu^2 - 4) \ge \frac{1}{4e^4}\exp(-10nu^2).
\end{aligned}$$

This concludes the proof.

Let $\tilde{\mu}$ be an estimator of μ. Let $0 < u \le 1/4$ and $T = \mathbf{1}\{\tilde{\mu} \in [1/2 - u/2, 1/2 + u/2]\}$. We have by the previous result that $\tilde{\mu}$ cannot be more than $(\min(2\exp(-2nu^2), 1/2), u)$ good. This concludes the proof for $u = \sqrt{\frac{\log(2/\delta)}{10n}}$.

An approach using the distance between the distributions We have

$$\mathbb{P}_{1/2}(\Omega) + \mathbb{P}_{1/2+u}(\Omega^C) = 1 + \mathbb{P}_{1/2}(\Omega) - \mathbb{P}_{1/2+u}(\Omega).$$

So we have

$$\sup_{\tilde{\mu}\ \text{estimator}}\Big[\mathbb{P}_{1/2}(|\tilde{\mu}-1/2| < u/2) + \mathbb{P}_{1/2+u}(|\tilde{\mu}-1/2-u| < u/2)\Big] \le 1 + \sup_{A\ \text{measurable}}|\mathbb{P}_{1/2}(A) - \mathbb{P}_{1/2+u}(A)|.$$

We now introduce the *total variation distance* for two measures $\mathbb{P}, \mathbb{Q}$ that are defined on the same σ-algebra

$$d_{TV}(\mathbb{P},\mathbb{Q}) = \sup_{A\ \text{measurable}}|\mathbb{P}(A) - \mathbb{Q}(A)|.$$

We have

$$\sup_{\tilde{\mu}\ \text{estimator}}\Big[\mathbb{P}_{1/2}(|\tilde{\mu}-1/2| < u/2) + \mathbb{P}_{1/2+u}(|\tilde{\mu}-1/2-u| < u/2)\Big] \le 1 + d_{TV}(\mathbb{P}_{1/2}, \mathbb{P}_{1/2+u}).$$

We now link the total variation distance to the KL divergence.

Theorem 2 (Pinsker's inequality)**.** *Let $\mathbb{P}, \mathbb{Q}$ be two measures that are defined on the same σ-algebra. It holds that*

$$d_{TV}(\mathbb{P}, \mathbb{Q}) \leq \sqrt{\frac{1}{2} d_{KL}(\mathbb{P}, \mathbb{Q})}.$$

By the chain rule this implies

$$d_{TV}(\mathbb{P}_{1/2}, \mathbb{P}_{1/2+u}) \leq \sqrt{\frac{1}{2} d_{KL}(\mathbb{P}_{1/2}, \mathbb{P}_{1/2+u}} = \sqrt{\frac{n}{2}\big[-\frac{1}{2}\log(1+2u) - \frac{1}{2}\log(1-2u)\big]}.$$

So

$$d_{TV}(\mathbb{P}_{1/2}, \mathbb{P}_{1/2+u}) \leq \frac{1}{2}\sqrt{-n\log(1-4u^2)} \leq \sqrt{u^2 n + nO(u^4)}.$$

So for $u \leq o(1/\sqrt{n})$, this concludes the estimation proof.

References

Bubeck, Sebastien, and Nicolo Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends in Machine Learning*, 5(1):1-122, 2013.

Cesa-Bianchi, Nicolo, and Gabor Lugosi. Prediction, learning, and games. *Cambridge University Press*, 2006.