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

1.3.2. Back to the two armed bandit problem

We consider the two-armed bandit problem from the last subsection. Let $\mathcal{S}$ be the set of all two-armed bandit problems with distributions that have support on $[0,1]$. Let $\bar{R}_n(S,\mathcal{A})$ be the pseudo-regret that algorithm $\mathcal{A}$ would suffer on problem $S \in \mathcal{S}$.

Theorem 1. *It holds that*

$$\inf_{\mathcal{A} \quad algorithm} \sup_{S\in\mathcal{S}} \bar{R}_n(S,\mathcal{A}) \geq \min\Big(\frac{\log(nu^2)}{640u}, nu/64\Big).$$

For $u \in (0,1/4]$, consider the bandit problems where the first distribution is a Dirac mass in $1/2+u/2$, and where the second distribution is a Bernoulli of parameter $1/2+u$ - let us write $\mathbb{P}_{1/2+u,\mathcal{A}}, \mathbb{E}_{1/2+u,\mathcal{A}}$ for the distribution (resp. expectation) of the data for this problem when algorithm $\mathcal{A}$ is used. Consider also the bandit problems where the first distribution is a Dirac mass in $1/2+u/2$, and where the second distribution is a Bernoulli of parameter $1/2$ - let us write $\mathbb{P}_{1/2,\mathcal{A}}, \mathbb{E}_{1+u,\mathcal{A}}$ for the distribution (resp. expectation) of the data for this problem when algorithm $\mathcal{A}$ is used. The previous theorem follows directly from the following lemma.

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

Lemma 1. *For $u \in [0, 1/4]$, it holds that*

$$\inf_{\mathcal{A} \quad algorithm} \Big[\mathbb{E}_{1/2+u,\mathcal{A}}[n - T_{2,n}] + \mathbb{E}_{1/2,\mathcal{A}} T_{2,n}\Big] \geq \min\Big(\frac{\log(nu^2)}{640u^2}, n/64\Big).$$

Proof Let $\mathcal{A}$ be an algorithm, we write for short $\mathbb{P}_{1/2+u}, \mathbb{E}_{1/2+u}$, for $\mathbb{P}_{1/2+u,\mathcal{A}}, \mathbb{E}_{1/2+u,\mathcal{A}}$. Let us write $(X_1, \ldots, X_{T_{2,n}})$ for the samples collected by sampling the second distribution.

Let for $T > 0$

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

Let $\Omega_T = \{T_{2,n} = T\}$. We have

$$\begin{aligned}\mathbb{P}_{1/2+u}(\Omega_T) &= \mathbb{E}_{1/2}\Big[\frac{L_{1/2+u}(X_1, \ldots, X_T)}{L_{1/2}(X_1, \ldots, X_T)}\mathbf{1}\{\Omega_T\}\Big] \\ &= \mathbb{E}_{1/2}\left[\exp\Big(\log(\frac{1+2u}{1-2u})\sum_i X_i + T\log(1-2u)\Big)\mathbf{1}\{\Omega_T\}\right].\end{aligned}$$

Consider now the event

$$\xi = \Big\{\forall T \leq n, |\sum_{i \leq T} X_i - T/2| \leq \sqrt{T\log(2t)}\Big\}.$$

Note that $\mathbb{P}_{1/2}(\xi) \geq 1 - 1/n^2$.

We have

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

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

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

So we have

$$\begin{aligned}\mathbb{P}_{1/2+u}(\Omega_T) &\geq \mathbb{E}_{1/2}\left[\exp\Big((4u - 8u^2)\big(T/2 - \sqrt{T\log(2T)}\big) + T(-2u - 2u^2)\Big)\mathbf{1}\{\Omega_T \cap \xi\}\right] \\ &\geq \mathbb{E}_{1/2}\left[\exp\Big(-6Tu^2 - 4u\sqrt{T\log(2T)}\Big)\mathbf{1}\{\Omega_T \cap \xi\}\right] \\ &\geq \exp\Big(-6Tu^2 - 4u\sqrt{T\log(2T)}\Big)\mathbb{E}_{1/2}\left[\mathbf{1}\{\Omega_T \cap \xi\}\right] \\ &\geq \exp\Big(-6Tu^2 - 4u\sqrt{T\log(2T)}\Big)[\mathbb{P}_{1/2}(\Omega_T) - 1/n^2] \\ &:= M^{-1}(T, u)[\mathbb{P}_{1/2}(T_{2,n} = T) - 1/n^2].\end{aligned}$$

i.e.

$$\mathbb{P}_{1/2+u}(T_{1,n} = n - T)(n - T) \geq M^{-1}(T, u)[\mathbb{P}_{1/2}(T_{2,n} = T) - 1/n^2](n - T).$$

This implies that

$$\begin{aligned}
\mathbb{E}_{1/2+u}T_{1,n} &\geq \sum_T M^{-1}(T, u)[\mathbb{P}_{1/2}(T_{2,n} = T) - 1/n^2](n - T) \\
&\geq \sum_T \exp(-6Tu^2 - 4u\sqrt{T\log(2T)})[\mathbb{P}_{1/2}(T_{2,n} = T) - 1/n^2](n - T) \\
&\geq \exp(-12\bar{T}u^2 - 8u\sqrt{\bar{T}\log(2\bar{T})})[\sum_{T\leq 2\bar{T}} \mathbb{P}_{1/2}(T_{2,n} = T) - 1/n](n - 2\bar{T}) \\
&\geq \exp(-2\bar{T}u^2 - 8u\sqrt{\bar{T}\log(2\bar{T})})[\mathbb{P}_{1/2}(T_{2,n} \leq 2\bar{T}) - 1/n](n - 2\bar{T}),
\end{aligned}$$

for any $\bar{T} \leq n$. Set $\bar{T} = \mathbb{E}_{1/2}T_{2,n}$. It holds that $\mathbb{P}_{1/2}(T_{2,n} \leq 2\bar{T}) \geq 1/2$, so that

$$\begin{aligned}
\mathbb{E}_{1/2+u}T_{1,n} &\geq \exp(-12\bar{T}u^2 - 8u\sqrt{\bar{T}\log(2\bar{T})})[1/2 - 1/n](n - 2\bar{T}) \\
&\geq \exp(-\max\Big(20\bar{T}u^2, \log(2\bar{T})\Big))(n - 2\bar{T})/4.
\end{aligned}$$

So this implies that

$$\begin{aligned}
\mathbb{E}_{1/2+u}T_{1,n} + \bar{T} &\geq \exp(-20u^2\bar{T} - \log(2\bar{T}))n/32 + \bar{T} \\
&\geq \min\Big(\frac{\log(nu^2)}{640u^2}, n/64\Big).
\end{aligned}$$

This concludes the 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.

1.4. Exercises : part 2

Lower bounds arguments. Consider the $2-armed$ stochastic bandit setting where the objective is to minimize the pseudi-regret.

For $u \in (0, 1/4]$, consider the bandit problems where the first distribution is a Dirac mass in $1/2 + u/2$, and where the second distribution is a Bernoulli of parameter $1/2 + u$ - let us write $\mathbb{P}_{1/2+u,\mathcal{A}}, \mathbb{E}_{1/2+u,\mathcal{A}}$ for the distribution (resp. expectation) of the data for this problem when algorithm $\mathcal{A}$ is used. Consider also the bandit problems where the first distribution is a Dirac mass in $1/2 + u/2$, and where the second distribution is a Bernoulli of parameter $1/2$ - let us write $\mathbb{P}_{1/2,\mathcal{A}}, \mathbb{E}_{1+u,\mathcal{A}}$ for the distribution (resp. expectation) of the data for this problem when algorithm $\mathcal{A}$ is used. The previous theorem follows directly from the following lemma.

1. Write the likelihood of T samples that are distributed according to a Bernoulli distribution of parameter μ.
2. Consider the event
$$\xi = \Big\{\forall T \leq n, |\sum_{i \leq T} X_i - T/2| \leq \sqrt{T\log(2t)}\Big\}.$$
Prove that $\mathbb{P}_{1/2}(\xi) \geq 1 - 1/n^2$.
3. Prove that for any $T \leq n$
$$\mathbb{P}_{1/2+u}(T_{2,n} = T) \geq \exp\Big(-6Tu^2 - 4u\sqrt{T\log(2T)}\Big)[\mathbb{P}_{1/2}(T_{2,n} = T) - 1/n^2].$$
4. Deduce from this that
$$\mathbb{E}_{1/2+u}T_{1,n} \geq \exp(-\max\Big(20\bar{T}u^2, \log(2\bar{T})\Big))(n - 2\bar{T})/4.$$
5. Conclude that
$$\inf_{\mathcal{A} \text{ algorithm}} \sup_{S \in \mathcal{S}} \bar{R}_n(S, \mathcal{A}) \geq \min\Big(\frac{\log(nu^2)}{640u}, nu/64\Big).$$
6. Recall Pinsker's inequality. Deduce the problem independent bound from it.