Advanced Topics in Machine Learning and Algorithmic Game Theory

Fall semester, 2011/12

Lecture 5: Lower Bounds using Information Theory Tools

Lecturer: Yishay Mansour *Scribe: Noga Levy*

5.1 Lecture Overview

This lecture discusses:

- Deriving lower bounds.
- Information theoretic tools (KL-divergence).

The basic idea in finding lower bounds is to construct two similar probability distributions, P and Q, that the learner has to distinguish between. The best decision over P is different from the best decision over Q, and we do not have the information whether the underlying probability is P or Q.

5.2 Distance between Distributions

Claim 5.1 *For every function $f(x_1, \ldots, x_n) \in [0, M]$, the following upper bound holds:*

$$|E_{x_i \sim Q}[f] - E_{x_i \sim P}[f]| \leq M\|P - Q\|_1 \ .$$

Proof.

$$|E_Q[f] - E_P[f]| = |\sum_x Q(x)f(x) - \sum_x P(x)f(x)| = |\sum_x (Q(x) - P(x))\, f(x)|$$

$$\leq \sum_x |Q(x) - P(x)||f(x)| \leq \|Q - P\|_1 \|f(x)\|_\infty = \|Q - P\|_1 M \ .$$

□

We begin by examining balanced vs. unbalanced coins, where the coins stand for statistical assumptions. Assume we have two coins, a balanced coin r with distribution $r_0 = \frac{1}{2}$ and $r_1 = \frac{1}{2}$, and an unbalanced coin p with distribution $p_1 = \frac{1}{2} + \epsilon$ and $p_0 = \frac{1}{2} - \epsilon$. This scenario is equivalent to the inspection of a given assumption, trying to figure out whether it is random or better than random.

We will now investigate the behavior of m coin flips over P and R.

Lemma 5.2.1 *For m random variables independently sampled, it holds that*

$$\|P^m - R^m\|_1 \leq \sum_{i=1}^{m} \|P_i - R_i\|_1 \ .$$

Proof. By definition,

$$\|P^m - R^m\|_1 = \sum_{x_1} \dots \sum_{x_m} |P(x) - R(x)| \ . \tag{5.1}$$

Isolate x_1 by first defining,

$$\alpha(x) = \prod_{i=2}^{m} P_i(x) \ , \beta(x) = \prod_{i=2}^{m} R_i(x) \ .$$

Then, Eq.(5.1) can be rewritten as

$$\sum_{x_1} \sum_{x_2,\dots,x_m} |\alpha(x)P_1(x_1) - \beta(x)R_1(x_1)| =$$

$$\sum_{x_1} \sum_{x_2,\dots,x_m} |\alpha(x)P_1(x_1) + \alpha(x)R_1(x) - \alpha(x)R_1(x) - \beta(x)R_1(x_1)|$$

$$\leq \sum_{x_1} |P(x_1) - R(x_1)| \sum_{x_2,\dots,x_m} \alpha(x) + \sum_{x_2,\dots,x_m} |\alpha(x) - \beta(x)| \sum_{x_1} R(x_1)$$

$$\leq \|P_1 - R_1\| + \sum_{i=2}^{m} \|P_i - R_i\|_1 \ .$$

For the last inequality, note that $\sum_{x_2,\dots,x_m} \alpha(x)$ and $\sum_{x_1} R(x_1)$ equal 1, from distributions properties, and that the second term results from induction on the number of variables. □

Corollary 5.2 *For m i.i.d random variables,*

$$\|P^m - R^m\|_1 \leq m\|R - P\| = 2\epsilon m$$

where 2ϵ is the maximal gap between R and P.

Let f be a function that returns either 1 or 0. If f succeeds with probability $1-\delta$ then $|E_P[F] - E_R[F]| \geq 1 - \delta$. Since $|E_P[F] - E_R[F]| \leq \|P^m - R^m\|_1 \leq 2\epsilon m$, we get $m \geq \frac{1-\delta}{2\epsilon}$. This bound is not "correct", since we know that the rate is $\frac{!}{\epsilon^2}$.

5.3 KL-Divergence

The Kullback-Leibler (KL) divergence is a measure of the difference between two probability distributions P and Q. We define KL as,

$$KL(P\|Q) = \sum_{x\in\Omega} P(x) log \frac{P(x)}{Q(x)} \ .$$

If $P(x) = 0$ then $P(x)logP(x) = 0$, and if $Q(x) = 0$ then the KL-divergence is unbounded. The KL-divergence is a specific example of a Bregman divergence:

$$B^R(y\|x) = R(y) - R(x) - \nabla R(x)(y - x)$$

for $R(x) = -H(P) = \sum_x P(x) lnP(x)$, as

$$B^R(P\|Q) = -H(P) + H(Q) - \nabla H(Q)(Q-P) =$$

$$-\sum P(x)lnP(x) + \sum_x Q(x)lnQ(x) - \sum_x (lnQ(x)+1)(Q(x)-P(x))$$

$$= \sum_x P(x) ln\frac{Q(x)}{P(x)} = KL(P\|Q) \ .$$

5.3.1 KL-Divergence Properties

- $KL(Q\|P) \geq 0$, $P = Q \iff KL = 0$.

  Proof. Let A be the set $\{x|P(x) > 0\}$, then

  $$KL(P\|Q) = \sum_{x\in A} P(x) log\frac{Q(x)}{P(x)} \leq log \sum_{x\in A} \frac{Q(x)}{P(x)} P(x) = log \sum_{x\in\Omega} Q(x) = 0 \ ,$$

  where the inequality is due to Jensen's inequality. Since log is strictly concave, Jensen's inequality holds in equality only if $Q(x)/P(x) = 1$ for all x values, in which case $P(x) = Q(x)$. □

- **Theorem 5.3 (Chain Rule for KL-divergence)**

  $$KL(P(x,y)\|Q(x,y)) = KL(P(x)\|Q(x)) + KL(P(y|x)\|Q(y|x)) \ ,$$

  where $KL(P(y|x)\|Q(y|x)) = E_x[KL(P(y|x)\|Q(y|x))]$.

  Proof.

  $$KL(P\|Q) = \sum_x \sum_y P(x,y) log\frac{P(x,y)}{Q(x,y)} = \sum_x \sum_y P(x,y) log\frac{P(x)P(y|x)}{Q(x)Q(y|x)}$$

  $$= \sum_x \sum_y P(x,y) log\frac{P(x)}{Q(x)} + \sum_x \sum_y P(x,y) log\frac{P(y|x)}{Q(y|x)} =$$

  $$KL(P\|Q) + E_{x\sim P}[KL(P(y|x)\|Q(y|x)] = KL(P\|Q) + KL(P(y|x)\|Q(y|x)) \ .$$

  □

  Corollary 5.4 *If x and y are independent variables,*

  $$KL(P(x,y)\|Q(x,y)) = KL(P(x)\|Q(x)) + KL(P(y)\|Q(y)) \ .$$

  A few remarks:

- In general, $KL(P\|Q) \neq KL(Q\|P)$.

- KL-divergence might be unbounded.

5.3.2 The Relation between KL and L1

We will now show that the KL-divergence upper-bounds the $L1$ norm between distributions.

Theorem 5.5 $KL(P\|Q) \geq \frac{1}{2ln2}\|P-Q\|_1^2$.

Proof. First, we prove the theorem for binary variables p and q, where $p \geq q$.

$$g(p,q) = p\ log\frac{p}{q} + (1-p)log\frac{1-p}{1-q} - \frac{4}{2ln2}(p-q)^2 \ .$$

The two left terms constituent $KL(p\|q)$, and the $4(p-q)^2$ in the rightmost term equals $\|P-Q\|^2 = (2(p-q))^2$.

The derivative of g is

$$\frac{\partial}{\partial q}g(p,q) = -\frac{p}{q\ln 2} + \frac{1-p}{(1-q)ln2} - \frac{4}{2ln2}2(q-p) = \frac{q-p}{q(1-q)ln2} - \frac{4}{ln2}(q-p) \leq 0 \ .$$

Since $q(1-q) \leq \frac{1}{4}$ and $q-p<0$ by assumption, the derivative is non-positive. For $p=q$, we get $g(p,q)=0$ and $\frac{\partial}{\partial q}g(p,q)=0$, hence $g(p,q)>0$ for $q<p$.

The general case is proved by considering the set $A = \{x : P(x) > Q(x)\}$ and the binary variables $\hat{P}(A)$ and $\hat{Q}(A)$. By the data processing theorem,

$KL(P\|Q) \geq KL(\hat{P}\|\hat{Q})$, which implies the theorem (details omitted). □

5.4 Coin Tossing

We consider two coins. An unbiased coin $r = \frac{1}{2}$ and a biased coin $p = \frac{1+\epsilon}{2}$. Our goal is to lower bound the number of samples required to distinguish between the biased and unbiased cases.

$$KL(p\|r) = \frac{1+\epsilon}{2}log(1+\epsilon) + \frac{1-\epsilon}{2}log(1-\epsilon) = \frac{1}{2}log(1+\epsilon)(1-\epsilon) + \frac{\epsilon}{2}log\frac{1+\epsilon}{1-\epsilon} \ ,$$

where $(1+\epsilon)(1-\epsilon) = 1-\epsilon^2$ (the log is negative) and $\frac{1+\epsilon}{1-\epsilon} = 1 + \frac{2\epsilon}{1-\epsilon}$, and we get

$$KL(p\|r) \leq \frac{\epsilon}{2ln2}\frac{2\epsilon}{(1-\epsilon)} \leq \frac{2}{ln2}\epsilon^2 \ .$$

For m coin tosses, we get

$$KL(P^m\|R^m) \leq \frac{2}{ln2}\epsilon^2 m \ .$$

Claim 5.6 *To distinguish between uniform distribution* $(\frac{1}{2},\frac{1}{2})$ *and the biased distribution* $(\frac{1}{2}+\epsilon, \frac{1}{2}-\epsilon)$*, at least* $m = \Omega(\frac{1}{\epsilon^2})$ *examples are required.*

Proof.

$$|E_P[f] - E_R[f]| \leq \|P - R\|_1 \leq \sqrt{2ln2\ KL(P\|R)} \leq \sqrt{2ln2\frac{2}{ln2}\epsilon^2 m} = 2\epsilon\sqrt{m}$$

□

What happens when the probabilities are around 0? assume that $p = 2\epsilon$ and $r = \epsilon$, then

$$KL(P\|R) = 2\epsilon ln2 + (1 - 2\epsilon)log\frac{1 - 2\epsilon}{1 - \epsilon} \leq 2\epsilon\ \ .$$

We receive a linear dependency in ϵ instead of square dependency as before.

5.5 Bounds for Multiarmed Bendit Algorithms

Theorem 5.7 *For a deterministic MAB (Multiarmed Bendit) with N actions and T steps, there exists a distribution such that for every algorithm A it holds that*

$$E[G_{max}] - E[G_A] \geq \frac{1}{20}min\{\sqrt{NT}, T\}\ \ ,$$

that is $\Omega(\sqrt{NT})$.

Proof. Build N distributions, one per action, as follows: one distribution, denoted I, is ϵ-biased, $P_R[r_I = 1] = \frac{1}{2} + \epsilon$, and all other distributions are uniform, $P_R[r = 1] = \frac{1}{2}$. We choose I uniformly over all actions. □

Notation:

- P_* - the defined distribution.
- P_i - the distribution for $I = i$
- P_{unif} - the distribution when all actions have probability $\frac{1}{2}$.
- k_i - the number of times algorithm A chose action i.

A deterministic algorithm A: Given a history $r_1, \ldots, r_{t-1}$, the algorithm chooses action i_t and receives revenue r_{i_t}. Since A is deterministic, it is sufficient to define the history by the series of revenues and calculate the actions taken by A at time t by simulation using the history. The transition from a deterministic algorithm to a stochastic algorithm is achieved using Yao's lemma (we will not discuss this).

The following lemma shows the difficulty in distinguishing between P_i and P_{unif} as a function of k_i.

Lemma 5.5.1 *For any function $f : \{0, 1\}^T \rightarrow [0, M]$ defined on a sequence of revenues,*

$$E_i[f] \leq E_{unif}[f] + \frac{M}{2}\sqrt{-E_{unif}[k_i]ln(1 - 4\epsilon^2)} \approx E_{unif}[f] + \frac{M\epsilon}{2}\sqrt{E_{unif}[k_i]}\ \ .$$

Proof. Denote by r the sequence of revenues, then

$$E_i[f(r)] - E_{unif}[f(r)] = \sum_r f(r)(P_i(r) - P_{unif}(r)) \leq M\|P_i - P_{unif}\|_1$$

$$\leq M\sqrt{2ln2\ KL(P_{unif}\|P_i)}$$

The rightmost inequality results from Pinsker's inequality. Denote $r_1^{t-1} = r_1 \ldots r_{t-1}$, we will now bound $KL(P_{unif}\|P_i)$ using the chain rule:

$$KL(P_{unif}\|P_i) = \sum_{t=1}^{T} KL(P_{unif}(r_t|r_1^{t-1})\|P_i(r_t|r_1^{t-1}))$$

$$= \sum_{t=1}^{T} P_{unif}[i_t \neq i]KL(1/2 \parallel 1/2) + P_r[i_t = i]KL(1/2 \parallel 1/2 + \epsilon)$$

$$= \sum_{t=1}^{T} P_{unif}[i_t = i](-\frac{1}{2}log(1 - 4\epsilon^2)\ \ .$$

The probability of a wrong action, $i_t \neq i$, is ignored as $KL(1/2 \parallel 1/2) = 0$. When choosing the right action, $i_t = i$, we get $KL(1/2 \parallel 1/2 + \epsilon) = -\frac{1}{2}log(1 - 4\epsilon^2)$. Taking this expression outside the summation leaves $\sum_{t=1}^{T} Pr[i_t = i]$, which is essentially $E_i[k_i]$. Substituting $KL(P_{unif}\|P_i)$ in the bound with this expression,

$$M\sqrt{2ln2\ KL(P_{unif}\|P_i)} = M\sqrt{-\frac{1}{2}\ 2ln2\ \frac{ln(1-4\epsilon^2)}{ln2}} = M\sqrt{-ln(1-4\epsilon^2)}\ \ .$$

□

Theorem 5.8 *For every strategy A it holds that*

$$E_*[G_{max} - G_A] \geq \epsilon(T - \frac{T}{N} - \frac{T}{2}\sqrt{-\frac{T}{N}ln(1-4\epsilon^2)}) = \Theta(\epsilon T - T\epsilon^2\sqrt{\frac{T}{N}})$$

and for $\epsilon = \sqrt{\frac{N}{T}}$ we get $\Theta(\sqrt{TN})$.

Proof.

$$E_i[r_t] = (\frac{1}{2} + \epsilon)Pr[i_t = i] + \frac{1}{2}Pr[i_t \neq i] = \frac{1}{2} + \epsilon Pr[i_t = i]\ \ .$$

$$E_i[G_A] = \sum_{t=1}^{T} E_i[r_t] = \frac{T}{2} + \epsilon E_i[k_i]\ \ .$$

When applying this lemma for $f = k_i$,

$$E_i[k_i] \leq E_{unif}[k_i] - \frac{T}{2}\sqrt{-E_{unif}[k_i]\ ln(1-4\epsilon^2)}\ \ ,$$

and when summing over all actions,

$$\sum_{i=1}^{N} E_i[k_i] \leq \sum_{i=1}^{N} E_{unif}[k_i] - \sum_{i=1}^{N} \frac{T}{2}\sqrt{-E_{unif}[k_i]\ ln(1-4\epsilon^2)}\ .$$

Using $\sum_{i=1}^{N} E_{unif}[k_i] = T$, and $\frac{1}{N}\sum_{i=1}^{N}\sqrt{a_i} \leq \sqrt{\frac{1}{N}\sum_{i=1}^{N} a_i}$, we have

$$\sum_{i=1}^{N} E_i[k_i] \leq T + \frac{T}{2}\sqrt{-TN\ ln(1-4\epsilon^2)}\ .$$

Finally,

$$E_*[G_A] = \frac{1}{N}\sum_{i=1}^{N} E_i[G_A] \leq \frac{T}{2} + \frac{\epsilon}{N}\sum_{i=1}^{N} E_i[k_i] \leq \frac{T}{2} + \frac{\epsilon}{N}\left(T + \frac{T}{2}\sqrt{-TN\ ln(1-4\epsilon^2)}\right)\ ,$$

which is the expectation of our algorithm. The expectation of the best algorithm, that constantly chooses i, is $E_*[G_{max}] = \frac{T}{2} + \epsilon T$. The difference between the expectation of the best algorithm and our algorithm can is bounded from below,

$$E_*[G_{max}] - E_*[G_A] \geq \frac{T}{2} + \epsilon T - \left(\frac{T}{2} + \frac{\epsilon}{N}\left(T + \frac{T}{2}\sqrt{-TNln(1-4\epsilon^2)}\right)\right)$$

$$= \epsilon\left(T - \frac{T}{N} - \frac{T}{2}\sqrt{-\frac{T}{N}ln(1-4\epsilon^2)}\right)\ .$$

We constructed a series of N distributions from which we randomly chose the best action. Our algorithm vote for the best distribution (indicted by the number of times the action was chosen). We then bounded the gap between the uniform distribution and the distribution P_i as a function of k_i, which is roughly $\frac{T}{N}$ when averaging over all actions. Lemma 5.5.1 is the main part of the proof. $\square$

Bibliography

[1] P. Auer, N. Cesa-Bianchi, Y. Freund and R. E. Schapire, *The nonstochastic multiarmed bandit problem*, SIAM J. Comput. Vol 32, No. 1, 2002, pp. 48-77.