Lecture 6: Minimax Lower Bounds Part I

16th September, 2025

Instructor: Shubhanshu Shekhar

We now begin our discussion of a collection of techniques for deriving fundamental lower bounds (i.e., converse results) on the minimax risk in statistical estimation problems. In this lecture, we present the first step which is common in most of the techniques, that involves reducing the problem of estimation to that of testing. Then, we introduce an approach based on binary hypothesis testing: a conceptually simple approach, that nevertheless, can be used to establish the (order) optimal minimax rates in several problems that we also discuss.

1   From Estimation to Testing

Recall our general formulation of *statistical decision theory*:

- Let $\{P_\theta : \theta \in \Theta\}$ denote a family of probability distributions on some observations space $(\mathcal{X}, \mathcal{F})$ indexed by a "parameter set" Θ endowed with a "pseudo-metric" $\rho : \Theta \times \Theta \to [0, \infty)$. Recall that by pseudo-metric, we mean that ρ is symmetric, satisfies triangle inequality, and $\rho(\theta, \theta) = 0$ for all $\theta \in \Theta$.

- Let $\boldsymbol{X} \sim P_\theta$ denote an $\mathcal{X}$-valued observation drawn from the distribution corresponding to the unknown parameter θ. In most cases of interest, we have $\mathcal{X} = \mathbb{X}^n$ for some space $\mathbb{X}$, and $\boldsymbol{X} = (X_1, \ldots, X_n)$ denotes n observations.

- Let $(\mathcal{W}, \mathcal{F}_\mathcal{W})$ denote the "decision space", and let $W \sim P_{W|\boldsymbol{X}}$ denote the $\mathcal{W}$-valued "decision" made by the statistician based on the observations $\boldsymbol{X} \equiv (X_1, \ldots, X_n)$. In most cases, we will consider $\mathcal{W} = \Theta$, in which case the decision rule $\widehat{\theta}$ is called an "estimator". In the sequel, we will overload the notation $\widehat{\theta}$ to represent both the estimate, $\widehat{\theta}(\boldsymbol{x})$ when $\boldsymbol{X} = \boldsymbol{x}$, and the stochastic kernel $P_{W|\boldsymbol{X}}$.

- To measure the quality of an estimator, we use a loss function $L(\theta, \theta') = \Phi \circ \rho(\theta, \theta')$ where $\Phi : [0, \infty) \to [0, \infty)$ is a non-decreasing function. This loss function can then be used to define the minimax risk, as

$$R_n^*(\Theta, \Phi \circ \rho) = \inf_{\widehat{\theta}} \sup_{\theta \in \Theta} \mathbb{E}_{\boldsymbol{X} \sim P_\theta} \left[L(\widehat{\theta}(X), \theta) \right] = \inf_{\widehat{\theta}} \sup_{\theta \in \Theta} \mathbb{E}_{\boldsymbol{X} \sim P_\theta} \left[\Phi \circ \rho(\widehat{\theta}(X), \theta) \right].$$

In practice, for most problems that we encounter, it will be difficult to exactly characterize the value of the minimax risk. Instead, the best we can hope to achieve is to establish the *minimax rate*; that is, how does the term $R_n^*(\Theta, \Phi \circ \rho)$ behave as a function of the "size of the problem", modulo constant terms.

In many problem instances, the observation $\boldsymbol{X}$ consists of n i.i.d. observations; that is, $\boldsymbol{X} = (Y_1, \ldots, Y_n) \overset{i.i.d.}{\sim} P_\theta$, with each $\mathcal{X} = \mathcal{Y}^n$. In such instances, we say that the minimax rate is ψ_n, if the following is true:

$$c \leq \liminf_{n \to \infty} \frac{\mathcal{R}(n, \Theta, \Phi \circ \rho)}{\psi_n} \leq \limsup_{n \to \infty} \frac{\mathcal{R}(n, \Theta, \Phi \circ \rho)}{\psi_n} \leq C,$$

for constants $0 < c \leq C < \infty$.

If we want to prove that an estimator $\widehat{\theta}$ is asymptotically rate-optimal, then we need to establish two statements:

- First, we need an *achievability* statement which says that there exists a constant C such that

$$\limsup_{n \to \infty} \frac{\sup_{\theta \in \Theta} \mathbb{E}_\theta \left[\Phi \circ \rho(\widehat{\theta}(\boldsymbol{X}), \theta) \right]}{\psi_n} \leq C < \infty.$$

- We also need a *converse* statement, which says that there exists a constant $c > 0$, such that

$$\liminf_{n \to \infty} \frac{\mathcal{R}(n, \Theta, \Phi \circ \rho)}{\psi_n} \geq c. \tag{1}$$

If $c = C$, then such an estimator $\widehat{\theta}$ is said to be asymptotically efficient. Information theoretic techniques play a key role in establishing the converse bounds such as (1). A common first step in such analysis is to reduce the problem of estimation to that of testing, as we formalize in the statement below.

Proposition 1.1. *For some integer $M \geq 2$, let $\Theta_M = \{\theta_1, \ldots, \theta_M\}$ denote a finite subset of Θ. Then, the minimax risk associated with (Θ, L) with $L = \Phi \circ \rho$ satisfies*

$$\mathcal{R}(\Theta, L) \geq \omega(\Theta_M) \inf_{\Psi : \mathcal{X} \to \Theta_M} \left(\frac{1}{M} \sum_{i=1}^{M} \mathbb{P}_{\theta_i} \left(\Psi(\boldsymbol{X}) \neq \theta_i \right) \right),$$

where $\omega(\Theta_M) = \min_{\theta_i \neq \theta_j} L(\theta_i, \theta_j)/2 = \min_{\theta_i \neq \theta_j} \Phi \circ \rho(\theta_i, \theta_j)/2$.

Remark 1.2. Observe that the first term $\omega(\Theta_M)$ depends on how "well-separated" the points in the set Θ_M are. In other words, for a fixed integer $M \geq 2$, the "best" choice of Θ_M should be one which is as widely-spaced as possible in the metric ρ.

The second term in the lower bound is simply the average probability of error in the M-ary hypothesis testing problem involving the parameters in the set Θ_M. To make this quantity large, we need the distributions $\{P_{\theta_i} : \theta_i \in \Theta_M\}$ to be "statistically indistinguishable" given $\boldsymbol{X}$.

These two requirements are contradictory to each other, and the best lower bounds are obtained by finding Θ_M that achieve the best trade-off between these terms.

Proof of Proposition 1.1. Since $\Theta_M \subset \Theta$, we immediately have

$$R_n^*(\Theta, L) = \inf_{\widehat{\theta}} \sup_{\theta \in \Theta} \mathbb{E}_\theta \left[L(\widehat{\theta}, \theta) \right] \geq \inf_{\widehat{\theta}} \sup_{\theta \in \Theta_M} \mathbb{E}_\theta \left[L(\widehat{\theta}, \theta) \right].$$

Let us define the minimum-distance estimator $\widehat{\Psi}(\boldsymbol{X}) = \operatorname{argmin}_{\theta' \in \Theta_M} \rho(\widehat{\theta}(X), \theta')$, which selects the element of Θ_M that is closest to the estimate $\widehat{\theta}(X)$ in ρ. Now consider the case of $\theta = \theta_i$, observe that

$$L(\widehat{\theta}(X), \theta_i) \geq \mathbf{1}_{\widehat{\Psi}(\boldsymbol{X}) \neq \theta_i} L(\widehat{\theta}(X), \theta_i) \geq \mathbf{1}_{\widehat{\Psi}(\boldsymbol{X}) \neq \theta_i} \omega(\Theta_M). \tag{2}$$

The last inequality follows from the fact that when $\widehat{\Psi}(\boldsymbol{X}) \neq \theta_i$, then there must exist a $\theta_j \in \Theta_M$, such that $\rho(\widehat{\theta}(\boldsymbol{X}), \theta_j) \leq \rho(\widehat{\theta}(\boldsymbol{X}), \theta_i)$, which in turn means that $\rho(\widehat{\theta}(\boldsymbol{X}), \theta_i) \geq (1/2)\rho(\theta_j, \theta_i)$. Or in other words, under this event, we have $L(\widehat{\theta}(\boldsymbol{X}), \theta_i) \geq L(\theta_j, \theta_i)/2$, and the inequality follows by the definition of $\omega(\Theta_M) = \min_{\theta', \theta'' \in \Theta_M} L(\theta', \theta'')/2$.

On taking expectation in (2), we get

$$\max_{\theta \in \Theta_M} \mathbb{E}_\theta \left[L(\widehat{\theta}(\boldsymbol{X}), \theta) \right] \geq \omega(\Theta_M) \max_{\theta \in \Theta_M} \mathbb{P}_\theta \left(\widehat{\Psi}(\boldsymbol{X}) \neq \theta \right) \geq \omega(\Theta_M) \frac{1}{M} \sum_{\theta \in \Theta_M} \mathbb{P}_\theta \left(\widehat{\Psi}(\boldsymbol{X}) \neq \theta \right),$$

where the last inequality simply lower-bounds the maximum with an average.

Let us denote by $p_M(\widehat{\Psi})$ the average probability of error of the estimator Ψ. Observe that p_M is simply the Bayesian probability of error of the M-ary hypothesis test $\widehat{\Psi}$. By design, each estimator $\widehat{\theta}$ is associated with such a minimum-distance hypothesis test. Hence, we have

$$\inf_{\widehat{\theta}} \max_{\theta \in \Theta_M} \mathbb{E}_\theta \left[L(\widehat{\theta}(\boldsymbol{X}), \theta) \right] \geq \omega(\Theta_M) \inf_{\widehat{\Psi} \equiv \widehat{\Psi}(\widehat{\theta})} p_M(\widehat{\Psi}) \geq \omega(\Theta_M) \inf_{\Psi : \mathcal{X} \to \Theta_M} p_M(\Psi).$$

There is a subtle fact involved in achieving the last inequality: if $\widehat{\theta}$ is a randomized estimator, then the corresponding minimum-distance hypothesis test $\widehat{\Psi}$ will be a randomized hypothesis test; that is $\widehat{\Psi} : \mathcal{X} \to \mathcal{P}(\Theta_M)$. However, in the final lower bound we have only restricted our attention to non-randomized tests $\Psi : \mathcal{X} \to \Theta_M$. The reason for this is that p_M is the Bayesian probability of error (with uniform prior), and it is known that randomization is not necessary for Bayesian hypothesis testing. In other words, there exists a non-randomized test Ψ^* that achieves the lowest average probability of error among all (randomized or non-randomized) tests. $\qquad\square$

2 Two-Point Method and Variants

Theorem 2.1 (Le-Cam's Two-Point Method). *Suppose $\Theta_2 \subset \Theta$ of size 2, and $\omega(\Theta_2) = L(\theta_1, \theta_2)/2$. We have the following lower bounds on the minimax risk, :*

$$\mathcal{R}(\Theta, L) \geq \omega(\Theta_2) \left(1 - TV(P_{\theta_1}, P_{\theta_2}) \right) \tag{3}$$

$$\mathcal{R}(\Theta, L) \geq \omega(\Theta_2) \max \left(1 - \sqrt{D_{KL}(P_{\theta_1} \| P_{\theta_2})/2}, (1/2) \exp\left(-D_{KL}(P_{\theta_1} \| P_{\theta_2}) \right) \right)$$

$$\mathcal{R}(\Theta, L) \geq \omega(\Theta_2) \left(1 - \sqrt{H^2(P_{\theta_1}, P_{\theta_2})(1 - H^2(P_{\theta_1}, P_{\theta_2})/4)} \right)$$

$$\mathcal{R}(\Theta, L) \geq \omega(\Theta_2) \max \left(1 - \sqrt{\chi^2(P_{\theta_1} \| P_{\theta_2})/2}, (1/2) \exp\left(-\chi^2(P_{\theta_1} \| P_{\theta_2}) \right) \right)$$

Proof of Lemma 2.1. We only need to prove the first inequality given in (3); the other follow directly by the inequalities between TV and other f-divergences that we had derived in a previous lecture.

The inequality (3) essentially is a direct consequence of Proposition 1.1, which implies that

$$R_n^*(\Theta, L) \geq \omega(\Theta_2) \inf_{\Psi:\mathcal{X}\to\Theta_2} \left(\mathbb{P}_{\theta_1}(\Psi(X) \neq \theta_1) + \mathbb{P}_{\theta_2}(\Psi(X) \neq \theta_2)\right).$$

Now, to each deterministic binary hypothesis test, we can associate a set $E = \{\Psi(\boldsymbol{X}) = \theta_1\}$, which implies that

$$R_n^*(\Theta, L) \geq \omega(\Theta_2) \inf_E \left(\mathbb{P}_{\theta_1}(E^c) + \mathbb{P}_{\theta_2}(E)\right) = \omega(\Theta_2) \left(1 - \sup_E \left(\mathbb{P}_{\theta_2}(E) - \mathbb{P}_{\theta_1}(E)\right)\right)$$

$$= \omega(\Theta_2) \left(1 - TV(P_{\theta_1}, P_{\theta_2})\right).$$

This completes the proof. $\square$

Note that the lower bound (3) is a product of two terms:

- The first term, $\omega(\Theta_2) = \Phi \circ \rho(\theta_1, \theta_2)$, is measure of separation between the two elements in Θ_2. To make this term larger, we need to select θ_1, θ_2, that are far from each other in ρ (since Φ is non-decreasing).

- The other term $1 - TV(P_{\theta_1}, P_{\theta_2})$ is a measure of closeness between P_{θ_1} and P_{θ_2}: to make this term larger, we need to bring the two distributions close to each other in TV distance.

Thus, to get the tightest lower bound we need to find the right balance between these two terms: often this is achieved by maximizing $\Phi \circ \rho(\theta_1, \theta_2)$, subject to the constraint that $TV(P_{\theta_1}, P_{\theta_2}) \leq 1/2$ (or any other constant less than 1); that is,

$$R_n^*(\Theta, L) = R_n^*(\Theta, \Phi \circ \rho) \geq \frac{1}{4} \sup_{\theta_1, \theta_2 : TV(P_{\theta_1}, P_{\theta_2}) \leq 1/2} \Phi \circ \rho(\theta_1, \theta_2).$$

This bound is often more convenient to handle.

3 Applications of the Two-Point Method

In this section, we illustrate the application of LeCam's method for deriving minimax lower bounds for proving the rate-optimality of the MLE for GLM parameter estimation in $d = 1$, and for the rate-optimality of the popular Upper Confidence Bound (UCB) algorithm for K-armed bandits.

3.1 Gaussian Location Model

Consider the following problem instance:

- $\boldsymbol{X} = (Y_1, \ldots, Y_n) \overset{i.i.d.}{\sim} N(\theta, \sigma^2)$, where $\theta \in \Theta = \mathbb{R}$ endowed with the metric $\rho(\theta, \theta') = |\theta - \theta'|$.

- The goal is to estimate the unknown mean (with known variance σ^2), under squared loss function; that is, $L(\theta, \theta') = |\theta - \theta'|^2$ (that is, $\Phi(t) = t^2$).

We know that the maximum likelihood estimate (MLE) in this case is defined as

$$\widehat{\theta} = \frac{1}{n} \sum_{i=1}^{n} Y_i = \frac{1}{n} \boldsymbol{X}^T \mathbf{1}_n, \quad \text{where} \quad \mathbf{1}_n = (1, \ldots, 1)^T \in \mathbb{R}^n.$$

For this estimator, the expected risk is equal to

$$\mathbb{E}_\theta \left[L(\widehat{\theta}, \theta) \right] = \mathbb{E}_\theta \left[\left(\frac{1}{n} \sum_{i=1}^{n} (Y_i - \theta) \right)^2 \right] = \frac{\sigma^2}{n}.$$

Now, let us prove that this estimator is minimax rate optimal. To do this, we consider $\Theta_2 = \{0, \Delta\}$ for some value Δ that we will specify later. Observe that $\omega(\Theta_2) = L(0, \Delta) = \Delta^2$. Since $P_\theta = p_\theta^n$ with $p_\theta \sim N(\theta, \sigma^2)$, we will use the lower bound that depends on relative entropy (since it is easy to evaluate for product measures, and especially for products of Gaussians). Observe that we have

$$D_{\mathrm{KL}}(P_{\theta_1} \parallel P_{\theta_2}) = \frac{n\Delta^2}{2\sigma^2}, \quad \text{which implies} \quad TV(P_{\theta_1}, P_{\theta_2}) \leq \sqrt{\frac{n\Delta^2}{4\sigma^2}},$$

by Pinsker's inequality. Thus, by choosing Δ to be the largest value that makes this upper bound on TV equal to $1/2$, we get

$$\mathcal{R}(\Theta, \Phi \circ \rho) \geq \frac{1}{4} \Delta^2 = \frac{1}{4} \left(\frac{\sigma^2}{2n} \right) = \frac{\sigma^2}{8n}.$$

This proves that the MLE is minimax rate-optimal.

3.2 Multi-armed Bandits

We now show that the two-point method can be used to derive the optimal minimax rates even in problems beyond the usual estimation. In particular, we look at the K-armed bandit problem, which is an instance of the general "interactive decision-making" problems introduced in an earlier lecture. In Homework 2, we will analyze the upper confidence bound (UCB) algorithm, which admits a regret bound of the order $\mathcal{O}(\sqrt{nK})$ (ignoring logarithmic factors), where n denotes the problem horizon. We will now use the the two-point method to establish the rate-optimality (modulo logarithmic factors) of this procedure. Recall the problem setup:

- A K-armed bandit problem is characterized by a parameter $\theta \in \Theta = \mathbb{R}^K$ representing the mean of the K "arms" (or distributions). Each arm $i \in [K]$ corresopnds to a distribution with mean $\theta[i]$, that is known to be σ^2-sub-Gaussian, with a known σ^2.

- For a given horizon n, the objective is to design a policy $\pi \equiv (\pi_1, \ldots, \pi_n)$, with $\pi_t : \mathbb{R}^{t-1} \times [K]^{t-1} \to \mathcal{P}([K])$, for drawing the arms in an adaptive manner.

- The measure of performance is the expected regret, defined as

$$\text{Reg}(n, \pi, \theta) = n \max_{i \in [K]} \theta[i] - \mathbb{E}_\theta \left[\sum_{t=1}^n X_{I_t, t} \right], \quad \text{where} \quad I_t \sim \pi_t \quad \text{for} \quad t \in [n].$$

Let $T_i(n, \pi)$ denote the number of times arm i has been pulled in the first n rounds by the policy π. Then, the regret can be re-written as

$$\text{Reg}(n, \pi, \theta) = n \max_{i \in [K]} \theta[i] - \sum_{i=1}^K \mathbb{E}_\theta[T_i(n, \pi)] \Delta_i, \quad \text{where} \quad \Delta_i := \max_{j \in [K]} \theta[j] - \theta[i]$$

is the suboptimality of the ith arm.

To apply the two-point method, we need to select two instances of the MAB problem. For some $\delta > 0$ to be decided later, define $\theta_1 = (\delta, 0, \ldots, 0) \in \Theta$. In this case, $i^*(\theta_1) := \text{argmax}_{i \in [K]} \theta_1[i] = 1$ is the optimal arm. Now, observe the following:

$$\sum_{j=2}^K \mathbb{E}_{\theta_1}[T_j(n, \pi)] \le n, \quad \Longrightarrow \quad \exists j \in \{2, \ldots, K\} \quad \text{s.t.} \quad \mathbb{E}_{\theta_1}[T_j(n, \pi)] \le \frac{n}{K-1}.$$

In other words, there must exist at least one arm that is (i) suboptimal for θ_1, and (ii) is sampled less than average by π under θ_1. Let us use this arm to define a new problem instance, whose mean is 2δ at the jth coordinate: $\theta_2 = (\delta, 0, \ldots, 2\delta, 0, \ldots, 0)$.[1]

In the next step, we will relate the regret of the policy π for the two problems in terms of the expected value of $T_1(n, \pi)$:

$$\text{Reg}(n, \pi, \theta_1) = \sum_{i=1}^K \mathbb{E}_{\theta_1}[T_i(n, \pi)](\delta - \theta_1[i]) = \delta \times (n - \mathbb{E}_{\theta_1}[T_1(n, \pi)]), \quad \text{and}$$

$$\text{Reg}(n, \pi, \theta_2) = \sum_{i=1}^K \mathbb{E}_{\theta_1}[T_i(n, \pi)](2\delta - \theta_1[i]) \ge \delta \times \mathbb{E}_{\theta_1}[T_2(n, \pi)].$$

Now, comes the cruicial part: we define the event $E_1 = \{T_1(n, \pi) < n/2\}$. Then, the above expression for regret bounds can we connected to the probability of E_1 as

$$\text{Reg}(n, \pi, \theta_1) \ge \frac{\delta n}{2} \times \mathbb{P}_{\theta_1}(E_1), \quad \text{and} \quad \text{Reg}(n, \pi, \theta_1) \ge \frac{\delta n}{2} \times \mathbb{P}_{\theta_2}(E_1^c)$$

This means that we can lower bound the minimax regret with the average of these two regrets:

$$\text{Reg}(n, \Theta) := \inf_\pi \sup_{\theta \in \Theta} \text{Reg}(n, \pi, \theta) \ge \frac{1}{2} \left(\text{Reg}(n, \pi, \theta_1) + \text{Reg}(n, \pi, \theta_2) \right) \ge \frac{n\delta}{4} \left(1 - (\mathbb{P}_{\theta_2}(E_1) - \mathbb{P}_{\theta_1}(E_1)) \right)$$

$$\ge \frac{n\delta}{4} \left(1 - TV(P_{\theta_1}, P_{\theta_2}) \right) \ge \frac{n\delta}{8} \exp\left(-D_{\text{KL}}(P_{\theta_1} \| P_{\theta_2}) \right). \quad (4)$$

---

[1]Note: this is a cruicial aspect of our construction for MABs, where we use the performance of π on the first problem instance θ_1, to guide the choice of the second problem instance θ_2. An "oblivious" application of the two-point method would not lead to the right lower bound.

To complete the proof, we will employ the divergence bound for interactive decision-making derived in Lemma **??**, according to which we have

$$D_{\mathrm{KL}}(P_{\theta_1} \parallel P_{\theta_2}) = \mathbb{E}_{\theta_1}[T_j(n, \pi)] D_{\mathrm{KL}}\left(N(0, \sigma^2) \parallel N(2\delta, \sigma^2)\right) = \mathbb{E}_{\theta_1}[T_j(n, \pi)]\frac{4\delta^2}{2\sigma^2} \leq \frac{2n\delta^2}{(K-1)\sigma^2}.$$

For the last inequality, we use the defining property of arm j. Plugging this back into (4), and selecting $\delta = \sigma\sqrt{(K-1)/2n}$, we get

$$\mathrm{Reg}(n, \Theta) := \inf_{\pi} \sup_{\theta \in \Theta} \mathrm{Reg}(n, \pi, \theta) \geq \frac{\sigma\sqrt{n(K-1)}}{8\sqrt{2}e} = \Omega(\sqrt{nK}).$$

3.2.1 Relative Entropy Decomposition for Bandits

> **Lemma 3.1.** *Consider two K-armed bandit instances with mean vectors θ_1 and θ_2. For some policy π and horizon $n \geq 1$, let $P_{\theta_1}^\pi$ and $P_{\theta_2}^\pi$ denote the joint distributions of the first n action-observation pairs induced by the policy π on the two problem instances. Then, we have*
>
> $$D_{KL}(P_{\theta_1}^\pi \parallel P_{\theta_2}^\pi) = \sum_{i=1}^{K} \mathbb{E}_{\theta_1}[T_i(n, \pi)] D_{KL}(f_{\theta_1,i} \parallel f_{\theta_2,i}).$$
>
> *Here we use $f_{\theta_j,i}$ to denote the distribution of the ith arm for problem instance j.*

The proof is essentially an application of the tower-rule of conditional expectation. In particular, let $H_n = (A_1, X_1, A_2, X_2, \ldots, A_n, X_n)$ denote the first n action-observation pairs. Then, for any realization $h_n = (a_1, x_1, \ldots, a_n, x_n)$, the likelihood ratio of $P_{\theta_1}^\pi$ and $P_{\theta_2}^\pi$ can be written as

$$\frac{p_{\theta_1}^\pi(h_n)}{p_{\theta_2}^\pi(h_n)} = \prod_{t=1}^{n} \frac{\pi_t(a_t|h_{t-1})f_{\theta_1,a_t}(x_t)}{\pi_t(a_t|h_{t-1})f_{\theta_2,a_t}(x_t)} = \prod_{t=1}^{n} \frac{f_{\theta_1,a_t}(x_t)}{f_{\theta_2,a_t}(x_t)}.$$

Hence, we can evaluate the relative entropy as

$$D_{\mathrm{KL}}(P_{\theta_1}^\pi \parallel P_{\theta_2}^\pi) = \mathbb{E}\left[\sum_{t=1}^{n} \log \frac{f_{\theta_1,A_t}(X_t)}{f_{\theta_2,A_t}(X_t)}\right] = \sum_{t=1}^{n} \mathbb{E}\left[\mathbb{E}\left[\log \frac{f_{\theta_1,A_t}(X_t)}{f_{\theta_2,A_t}(X_t)}\bigg|H_{t-1}\right]\right]. \qquad (5)$$

The second equality uses the tower property of conditional expectation. Now, observe that for any t, we have the Markov chain $H_{t-1} \longrightarrow A_t \longrightarrow X_t$; having selected the arm A_t, the observation X_t is conditionally independent of the past H_{t-1}. Using this, we can evaluate the inner expectation as follows, with $\ell_k(x) = f_{\theta_1,k}(x)/f_{\theta_2,k}(x)$:

$$\mathbb{E}\left[\log \ell_{A_t}(X_t)|H_{t-1}\right] = \mathbb{E}\left[\mathbb{E}\left[\log \ell_{A_t}(X_t)|A_t, H_{t-1}\right]|H_{t-1}\right]$$

$$= \mathbb{E}\left[\sum_{k=1}^{K} \mathbb{P}(A_t = k \mid H_{t-1})\mathbb{E}\left[\log \ell_k(X_t)|H_{t-1}, A_t = k\right]\bigg|H_{t-1}\right].$$

Now, the inner expectation is equal to $D_k := D_{\mathrm{KL}}(f_{\theta_1,k} \parallel f_{\theta_2,k})$, since conditioned on $A_t = k$, the observation X_t is independent of the past. Hence, we get

$$\mathbb{E}\left[\log \ell_{A_t}(X_t)|H_{t-1}\right] = \sum_{k=1}^{K} \mathbb{P}(A_t = k \mid H_{t-1})D_k.$$

Combining this with (5), we get

$$
\begin{aligned}
D_{\mathrm{KL}}(P_{\theta_1}^\pi \parallel P_{\theta_2}^\pi) &= \sum_{t=1}^{n} \mathbb{E}\left[\sum_{k=1}^{K} \mathbb{P}\left(A_t = k \mid H_{t-1}\right) D_k\right] \\
&= \sum_{t=1}^{n}\sum_{k=1}^{K} \mathbb{P}(A_t = k)D_k = \sum_{k=1}^{K} D_k \sum_{t=1}^{n} \mathbb{P}(A_t = k) \\
&= \sum_{k=1}^{K} D_k \mathbb{E}\left[\sum_{t=1}^{n} \mathbf{1}_{A_t=k}\right] = \sum_{k=1}^{K} D_k \mathbb{E}_{\theta_1}[T_k(n, \pi)]
\end{aligned}
$$

This completes the proof. $\qquad\square$

3.3   Examples where two-point methods are insufficient

Here are some problem instances in which the basic two-point method introduced in this lecture is insufficient:

Gaussian Location Model in $\mathbb{R}^d$. The simplest case in which the two-point method fails is when we wish to estimate the mean of Gaussian random vectors. In particular, consider the case of $\boldsymbol{X} = (Y_1, \ldots, Y_n) \overset{i.i.d.}{\sim} N(\theta, \sigma^2 I_d)$ for some $d \geq 2$. Then, the MLE $\widehat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$ satisfies

$$\mathbb{E}_\theta\left[\|\widehat{\theta}_n - \theta\|_2^2\right] = \frac{\sigma^2 d}{n}.$$

However, if we use a two-point method as in Section 3.1, the lower bound again turns out to be $\Omega(\sigma^2/n)$, and is unable to capture the dimension dependence.

Nonparametric Density Estimation. In a previous lecture, we showed that there exists a histogram estimator for (M, γ) Hölder continuous densities on $[0, 1]$ that achieves an integrated squared error bound $\mathcal{O}(n^{-2\gamma/(1+2\gamma)})$. Using the two-point method to establish the optimality of this procedure fails, as it only leads to $\Omega(n^{-1})$ lower bound. We will develop the appropriate techniques that obtain the correct lower bound in a future lecture.

Sparse Mean Estimation: Similar to the first example of this section, the two-point method cannot get the right sparsity dependent rate.

Minimax Hypothesis Testing and Estimation of Functionals. Another important application where the basic two-point method is insufficient is in establishing the optimal rates of functional estimation. We will discuss a natural generalization of the two-point method (referred to as the "method of two fuzzy hypotheses" by Tsybakov) to address such problems.

For these problems we need to develop more intricate techniques that will be our topic of discussion in the next few lectures.

Notes

The two-point method is based on a result of LeCam (1973), who established a connection between testing and the total-variation distance between the convex hull of two hypotheses classes. We have mostly followed the presentation approach of Tsybakov (2009, Chapter 2), with some minor modifications in the presentation. For the MAB lower bound, we have used the simplified proof given by Lattimore and Szepesvári (2020, Chapter 15).

References

T. Lattimore and C. Szepesvári. *Bandit algorithms*. Cambridge University Press, 2020.

L. LeCam. Convergence of estimates under dimensionality restrictions. *The Annals of Statistics*, pages 38–53, 1973.

A. B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer, 2009.