Home Assignment 1

Econometrics of High-Dimensional Models - ECON231C

Mauricio Vargas-Estrada
Master in Quantitative Economics
University of California - Los Angeles

Problem 1

Prove Hoeffding's lemma: for any mean-zero random variable Z satisfying $|Z|\leq a$ almost surely,

$$\mathbb{E}\left[\exp\left(\lambda Z_i\right)\right] \leq \exp\left(\frac{\lambda^2 a^2}{2}\right),$$

for all $\lambda > 0$.

Proof. Since $\exp\left(\lambda Z_i\right)$ is a convex function of Z_i, for all $z_i \in [-a, a]$,

$$\begin{aligned}
\exp\left(\lambda z_i\right) &\leq \frac{a - z_i}{a - (-a)} \exp\left(-a\lambda\right) + \frac{z_i - (-a)}{a - (-a)} \exp\left(a\lambda\right) \\
&\leq \frac{a - z_i}{2a} \exp\left(-a\lambda\right) + \frac{z_i + a}{2a} \exp\left(a\lambda\right)
\end{aligned}$$

taking expectation on both sides and knowing that Z is mean-zero, we have,

$$\begin{aligned}
\mathbb{E}\left[\exp\left(\lambda Z_i\right)\right] &\leq \mathbb{E}\left[\frac{a - Z_i}{2a} \exp\left(-a\lambda\right)\right] + \mathbb{E}\left[\frac{Z_i + a}{2a} \exp\left(a\lambda\right)\right] \\
&\leq \frac{a - \mathbb{E}\left[Z_i\right]}{2a} \exp\left(-a\lambda\right) + \frac{\mathbb{E}\left[Z_i\right] + a}{2a} \exp\left(a\lambda\right) \\
&\leq \frac{a}{2a} \exp\left(-a\lambda\right) + \frac{a}{2a} \exp\left(a\lambda\right) \\
&\leq \frac{1}{2} \exp\left(-a\lambda\right) + \frac{1}{2} \exp\left(a\lambda\right) \\
&\leq \exp\left(-a\lambda\right) + \exp\left(a\lambda\right)
\end{aligned}$$

Calculating the Taylor's expansion of the right-hand side of the inequality over λ and simplifying,

$$
\begin{aligned}
\exp(-a\lambda) + \exp(a\lambda) &= \left(1 - a\lambda + \frac{a^2\lambda^2}{2} - \frac{a^3\lambda^3}{6} + \dots\right) + \left(1 + a\lambda + \frac{a^2\lambda^2}{2} + \frac{a^3\lambda^3}{6} + \dots\right) \\
&= 1 + \frac{a^2\lambda^2}{2} + \frac{a^4\lambda^4}{24} \dots
\end{aligned}
$$

We recognize that the Taylor's expansion of $\exp\left(\frac{a^2\lambda^2}{2}\right)$ is:

$$
\exp\left(\frac{a^2\lambda^2}{2}\right) = 1 + \frac{a^2\lambda^2}{2} + \frac{a^4\lambda^4}{8} \dots
$$

meaning that:

$$
\begin{aligned}
1 + \frac{a^2\lambda^2}{2} + \frac{a^4\lambda^4}{24} \dots &\leq 1 + \frac{a^2\lambda^2}{2} + \frac{a^4\lambda^4}{8} \dots \\
\exp(-a\lambda) + \exp(a\lambda) &\leq \exp\left(\frac{a^2\lambda^2}{2}\right)
\end{aligned}
$$

Therefore

$$
\mathbb{E}\left[\exp(\lambda Z_i)\right] \leq \exp\left(\frac{\lambda^2 a^2}{2}\right)
$$

□

Problem 2

While proving the Hoeffding inequality, we said that two probabilities,

$$
\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu \geq t\right)
$$

and

$$
\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu \leq -t\right),
$$

can be bounded in the same way and did the derivation only for the former probability. Show that the latter probability is indeed bounded by the same quantity.

Proof. Lets define $Z_i = X_i - \mu, \forall i = 1, 2, \dots, n$. Then, we have:

$$
\bar{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i = \bar{X}_n - \mu
$$

and,

$$|Z_i| \leq a, \forall i = 1, 2, \ldots, n$$

then, whe can rewrite the former probability as:

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu \leq -t\right) = \mathbb{P}\left(\bar{Z}_n \geq -t\right)$$

Expanding the average $\bar{Z}_n$:

$$\mathbb{P}\left(\bar{Z}_n \geq -t\right) = \mathbb{P}\left(\sum_{i=1}^{n} Z_i \leq -nt\right)$$

for any $\lambda > 0$, we have [1]:

$$\mathbb{P}\left(\bar{Z}_n \geq -t\right) = \mathbb{P}\left(\lambda\sum_{i=1}^{n} Z_i \leq -\lambda nt\right)$$

and, by Markov's inequality:

$$\mathbb{P}\left(\lambda\sum_{i=1}^{n} Z_i \leq -\lambda nt\right) \leq \frac{\mathbb{E}\left[\exp(\lambda\sum_{i=1}^{n} Z_i)\right]}{\exp(-\lambda nt)}$$

since Z_i are independent and identically distributed, we can write:

$$\begin{aligned}\frac{\mathbb{E}\left[\exp(\lambda\sum_{i=1}^{n} Z_i)\right]}{\exp(-\lambda nt)} &= \frac{\prod_{i=1}^{n}\mathbb{E}\left[\exp(-\lambda Z_i)\right]}{\exp(-\lambda nt)} \\ &= \frac{\prod_{i=1}^{n}\exp(-1)\mathbb{E}\left[\exp(\lambda Z_i)\right]}{\exp(-\lambda nt)} \\ &= \frac{\prod_{i=1}^{n}\mathbb{E}\left[\exp(\lambda Z_i)\right]}{\exp(\lambda nt)}\end{aligned}$$

Applying the Hoeffding's lemma [2] to the above expression, we have:

1. Given that $f(x) = \lambda x$ is a monotonically increasing function when $\lambda > 0$
2. If X is a random variable such that $X \leq a$, then for any $\lambda > 0$, we have:

$$\mathbb{E}\left[\exp(\lambda X)\right] \leq \exp\left(\frac{\lambda^2 a^2}{2}\right)$$

$$
\begin{aligned}
\frac{\prod_{i=1}^{n} \mathbb{E}\left[\exp(\lambda Z_i)\right]}{\exp(\lambda n t)} &\leq \frac{\prod_{i=1}^{n} \exp\left(\frac{\lambda^2 a^2}{2}\right)}{\exp(\lambda n t)} \\
&\leq \frac{\exp\left(\frac{n\lambda^2 a^2}{2}\right)}{\exp(\lambda n t)} \\
&\leq \exp\left(\frac{n\lambda^2 a^2}{2} - \lambda n t\right)
\end{aligned}
$$

Because the above inequality holds for any $\lambda > 0$, we can optimize the right-hand side with respect to λ.

$$
\lambda^* = \underset{\lambda > 0}{\arg\min} \left\{ \frac{n\lambda^2 a^2}{2} - \lambda n t \right\}
$$

Calculating the F.O.C. with respect to λ, we get:

$$
n a^2 \lambda^* - n t = 0 \Rightarrow \lambda^* = \frac{t}{a^2}
$$

Substituting λ^* back into the inequality:

$$
\begin{aligned}
\mathbb{P}\left(\bar{Z}_n \geq -t\right) &\leq \exp\left(\frac{n\left(\frac{t}{a^2}\right)^2 a^2}{2} - \frac{t}{a^2} n t\right) \\
&\leq \exp\left(\frac{n t^2}{2a^2} - \frac{n t^2}{a^2}\right) \\
&\leq \exp\left(-\frac{n t^2}{2a^2}\right)
\end{aligned}
$$

replacing $\bar{Z}_n$ by $\bar{X}_n - \mu$:

$$
\mathbb{P}\left(|\bar{X}_n - \mu| \geq t\right) \leq \exp\left(-\frac{n t^2}{2a^2}\right)
$$

$\square$

Problem 3

(Tricky) While proving the Hoeffding inequality, we have used the following bound:

$$\mathbb{P}(X > \delta) \leq \frac{\mathbb{E}\left[e^{\lambda X}\right]}{e^{\lambda\delta}}, \quad \lambda > 0.$$

An alternative could be

$$\mathbb{P}(X > \delta) \leq \frac{\mathbb{E}\left[|X|^k\right]}{\delta^k}, \quad k \geq 0.$$

Show that if $X \geq 0$ a.s., then

$$\inf_{k=0,1,2,\ldots} \frac{\mathbb{E}\left[|X|^k\right]}{\delta^k} \leq \inf_{\lambda>0} \frac{\mathbb{E}\left[e^{\lambda X}\right]}{e^{\lambda\delta}}.$$

Proof. From the right hand side of the inequality, we note that the following expressions are equivalent:

$$\frac{\mathbb{E}\left[e^{\lambda X}\right]}{e^{\lambda\delta}} = \mathbb{E}\left[\frac{e^{\lambda X}}{e^{\lambda\delta}}\right] = \mathbb{E}\left[\left(\frac{e^X}{e^\delta}\right)^\lambda\right] = \mathbb{E}\left[\left(e^{X-\delta}\right)^\lambda\right]$$

Taking the Taylor's expansion of $\left(e^{X-\delta}\right)^\lambda$ around $\lambda = 0$:

$$\left(e^{X-\delta}\right)^\lambda = 1 + (X-\delta) + \lambda^2\frac{(X-\delta)^2}{2!} + \lambda^3\frac{(X-\delta)^3}{3!} + \cdots$$

Taking the expected value of that expansion:

$$\mathbb{E}\left[\left(e^{X-\delta}\right)^\lambda\right] = 1 + \lambda\mathbb{E}\left[(X-\delta)\right] + \lambda^2\frac{\mathbb{E}\left[(X-\delta)^2\right]}{2!} + \lambda^3\frac{\mathbb{E}\left[(X-\delta)^3\right]}{3!} + \cdots$$

We are interested in $\inf_{\lambda>0}\frac{\mathbb{E}\left[e^{\lambda X}\right]}{e^{\lambda\delta}}$. Given the above expansion, we can see that the infimum of the right hand side of the inequality can by approximated by takin the limit where $\lambda \to 0^+$:

$$\lim_{\lambda\to 0^+} 1 + \lambda\mathbb{E}\left[(X-\delta)\right] + \lambda^2\frac{\mathbb{E}\left[(X-\delta)^2\right]}{2!} + \lambda^3\frac{\mathbb{E}\left[(X-\delta)^3\right]}{3!} + \cdots = 1$$

Meaning that $\inf_{\lambda>0}\frac{\mathbb{E}\left[e^{\lambda X}\right]}{e^{\lambda\delta}} = 1$.

Now, we are interested in $\inf_{k=0,1,2,\ldots}\frac{\mathbb{E}\left[|X|^k\right]}{\delta^k}$. We can see that the left hand side of the inequality can be approximated by taking the limit where $k \to \infty$

$$\lim_{k\to\infty}\frac{\mathbb{E}\left[|X|^k\right]}{\delta^k} = \lim_{k\to\infty}\frac{\mathbb{E}\left[X^k\right]}{\delta^k}$$

Assuming that all the moments of X exist, we can see that the limit of the left hand side of the inequality depends on the value of δ.

If $0 < \delta \leq 1$, then:

$$\lim_{k\to\infty} \frac{\mathbb{E}\left[X^k\right]}{\delta^k} = \infty$$

Meaning that the infimum of the left hand side of the inequality is $\inf_{k=0,1,2,\dots} \frac{\mathbb{E}\left[|X|^k\right]}{\delta^k} = \frac{\mathbb{E}\left[|X|^0\right]}{\delta^0} = 1$.

If $\delta > 1$, then:

$$\lim_{k\to\infty} \frac{\mathbb{E}\left[X^k\right]}{\delta^k} = 0$$

Meaning that the infimum of the left hand side of the inequality is $\inf_{k=0,1,2,\dots} \frac{\mathbb{E}\left[|X|^k\right]}{\delta^k} = 0$.

Therefore, for any $\delta > 0$, and $\lambda > 0$, we have that:

$$\inf_{k=0,1,2,\dots} \frac{\mathbb{E}\left[|X|^k\right]}{\delta^k} \leq \inf_{\lambda>0} \frac{\mathbb{E}\left[e^{\lambda X}\right]}{e^{\lambda\delta}}.$$

□

Problem 4

> While proving the maximal inequality, i.e., a bound on
>
> $$\mathbb{P}\left(\max_{1\leq j\leq p}\left|\frac{1}{n}\sum_{i=1}^{n} X_{ij} - \mu_j\right| \geq t\right),$$
>
> we applied the union bound followed by the Hoeffding inequality. Show what happens if we replace the Hoeffding inequality by the Chebyshev inequality.

Theorem 1 (Chebyshev's inequality)**.** *Given a random variable X with mean μ and variance σ^2, then:*

$$\mathbb{P}\left(|X - \mu| \geq t\right) \leq \frac{\sigma^2}{t^2}, \quad \forall t > 0$$

The dimension of each random vector X_i is p, where $X_i = [X_{i,1}, \cdots, X_{i,p}]'$, and $X_{i,j}$ is the j-th component of the i-th random vector.

Also, recall that:

$$|X_{i,j} - \mu_j| \leq a, \quad ,\forall i = 1, \cdots, n, \quad \forall j = 1, \cdots, p, \quad \forall a > 0$$

By the union bound,

$$\mathbb{P}\left(\max_{1\leq j\leq p}\left|\frac{1}{n}\sum_{i=1}^{n} X_{i,j} - \mu_j\right| \geq t\right) \leq \sum_{j=1}^{p} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_{i,j} - \mu_j\right| \geq t\right)$$

By the Chebyshev's inequality,

$$\sum_{j=1}^{p} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_{i,j} - \mu_j\right| \geq t\right) \leq \sum_{j=1}^{p} \frac{\sigma_j^2}{nt^2}$$

We know that each component of the random vector X_i is bounded in mean by a, i.e., $|X_{i,j} - \mu_j| \leq a$. Therefore, $\sigma_j^2 \leq a^2$.

$$\sum_{j=1}^{p} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_{i,j} - \mu_j\right| \geq t\right) \leq \sum_{j=1}^{p} \frac{\sigma_j^2}{nt^2} \leq \sum_{j=1}^{p} \frac{a^2}{nt^2} = \frac{pa^2}{nt^2}$$

In other words, according to the Chebyshev's inequality:

$$\mathbb{P}\left(\max_{1\leq j\leq p}\left|\frac{1}{n}\sum_{i=1}^{n} X_{i,j} - \mu_j\right| \geq \frac{a\sqrt{p}}{\sqrt{\epsilon}\sqrt{n}}\right) \leq \epsilon$$

Or Using the big O notation,

$$\max_{1\leq j\leq p}\left|\frac{1}{n}\sum_{i=1}^{n} X_{i,j} - \mu_j\right| = O_p\left(\sqrt{\frac{p}{n}}\right)$$

Which indicates a rate of convergence slower than the one obtained by the Hoeffding's inequality.