Math 612: Homework 1

Instructor: Geoffrey Schiebinger

October 18, 2019

This homework is due at the beginning of class on Thursday October 24.

1 Warm up exercises on probability

Problem 1. *In this exercise you will prove Markov's inequality:*
If X is a nonnegative random variable, then for any $a > 0$ we have the following inequality:

$$\text{Prob}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Prove this statement.

Problem 2. *Prove Chebyshev's inequality:*
If X is a random variable with mean μ and variance σ^2, then for any $a > 0$ we have the following inequality:

$$\text{Prob}(|X - \mu| \geq a\sigma^2) \leq \frac{1}{a^2}.$$

Hint: This is an easy problem if you apply Markov's inequality!

These two inequalities are examples of *tail bounds*. They controls the probability that a random variable is bigger than some amount (related to the mean). More sophisticated tail bounds can be used to show that various estimators converge to the true value (which is usually the mean).

2 Programming exercise on spectral clustering

Problem 3. *Read the documentation for syzygy.*
https://intro.syzygy.ca/getting-started/

Please download the python notebook on spectral clustering from the course website and complete the exercises. Turn in an electronic copy of your notebook to me as well as a condensed summary of your findings in latex. Please polish your notebook by deleting any unnecessary cells you add for debugging purposes. A portion of the grade will be based on style.

3 Exercises on Markov chains

In this exercise you will examine the transition kernel of a discrete time Markov chain through simulation in a Jupyter notebook. Unlike section 2, you will need to create your own notebook for this.

Let $N_0, N_1, N_2, \dots, N_n$ denote a sequence of independent Gaussian random variables with distribution

$$N_i \sim \mathcal{N}(0, \epsilon),$$

for some variance parameter ϵ.

Consider the stochastic process X_t defined by $X_0 = 0$, and

$$X_{t+1} = \begin{cases} 1 & \text{if} \quad X_t + N_t \geq 1 \\ X_t + N_t & \text{if} \quad X_t + N_t \in (0,1) \\ 0 & \text{if} \quad X_t + N_t \leq 0 \end{cases}$$

Here $t = 0, 1, 2, 3, \ldots, n$ denotes discrete time. This is a random walk constrained to stay within $[0, 1]$.

Now define the process Z_t on the discrete space $\{0, \frac{1}{M}, \frac{2}{M}, \ldots, \frac{M-1}{M}, 1\}$ by rounding the process X_t to the nearest grid point:

$$Z_t = \frac{\text{Round}(MX_t)}{M}.$$

The function Round(x) returns the integer closest to the number x. For example Round$(3.75) = 4$. The transition matrix for Z_t is hard to compute by hand, but we will approximate it by running the process X_t many times.

Problem 4. *Set $M = 10$ and $\epsilon = 0.1$ and visualize the 10×10 transition matrix for Z_t with a heat-map. Denote the transition matrix by Π. The i, j entry of Π is $\Pi_{i,j} = Prob(Z_{t+1} = j | Z_t = i)$. Estimate this probability by simulating the process Z_t starting from i a total of L times. Denote this estimate by $\hat{\Pi}_L$. Make a heat map visualization of $\hat{\Pi}_L$ for $L \in \{3, 5, 10, 100\}$.*

Hint: To start the process from $Z_t = i$ exactly L times, pick a random X_t in the region corresponding to $Z_t = i$ and then simulate X_{t+1} a total of L times. For each of the L trials, compute Z_{t+1} by rounding X_{t+1}. Increment the ith row of $\hat{\Pi}$ a total of L times.

Problem 5. *For a matrix A, the Frobenius norm is $\|A\|_F$ is defined to be the square root of the sum of squared entries:*

$$\|A\|_F = \sqrt{\sum_{i,j} A_{i,j}^2}$$

Plot $\|\Pi_L - \Pi_{L+1}\|_F$ as a function of L and find the smallest value of L for which this quantity is less than 0.1. This illustrates the Law of Large Numbers.

4 Theoretical exercises on convex sets and functions

Problem 6. *Prove that the intersection of two convex sets is convex.*

Problem 7. *Prove that the maximum of two convex functions is convex.*