Homework 4

Jack Wilburn

Dec 1st, 2021

1 Q1: Trade-offs in sampling

1.1 Randomly pick samples of size (a) 20, (b) 100, (c) 400, and evaluate the probability that A is majority even in your sample (by running the experiment say 100 times and taking the count of times A is the majority in your sample). Write down the values you observe for these probabilities in the cases (a-c).

The probabilities for the three cases above were:

1. 0.51
2. 0.58
3. 0.76

1.2 What is the size of the sample you need for the probability above to become 0.9? (Your answer can be within 20% of the 'best' value.)

Running my simulation with a sample size of 1070 gives me a probability of 0.8994. That's the closest I could get the value to 0.9.

1.3 Suppose the population is more biased —55% of them vote for A and 45% of them vote for B— and re-solve part (b).

With a 55 percent bias, I found I only needed 180 people to have a 0.9 probability that A is in the majority. I ended up with 0.8989.

2 Q2: Satisfying ordering constraints

2.1 As a baseline, let us consider a *uniformly random* ordering. What is the expected number of constraints that are satisfied by this ordering? [*Hint:* define appropriate random variables whose sum is the quantity of interest, and apply the linearity of expectation.]

Let probability that each constraint is satisfied be represented by x_i, a binary random variable that is equal to 1 with probability p and 0 with probability $1-p$. Let X be the sum of these random variables. That is $X = x_1 + x_2 + ... + x_m$. By the linearity of expectation, $E(X) = E(x_1) + E(x_2) + ... + E(x_m) = mp$. Now we need to find p. By running through some examples on paper, I'm convinced that the value is $\frac{2}{3}$, which would lead us to a value of $E(X) = \frac{2m}{3}$

2.2 Use the definition of X above to conclude that $\Pr[X \geq E] \geq 1/m$.

3 Q3: Writing constraints

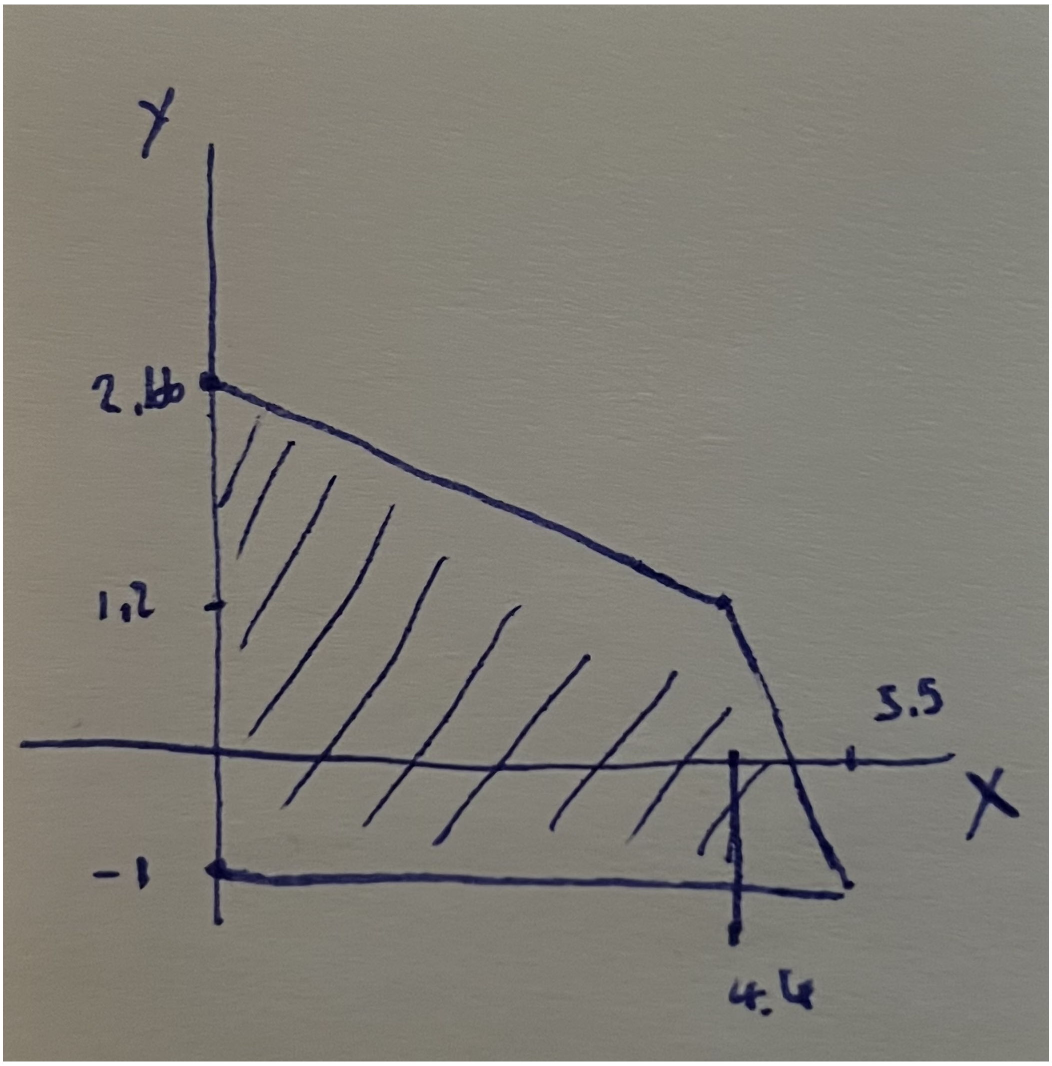
3.1 Does an optimal solution to this optimization problem *always yield* an optimal solution to TSP? Provide a proof or a counterexample with a full explanation.

These constraints do not always yield an optimal solution of the TSP. We're missing a constraint that says all the nodes must be connected in one cycle. The constraints we have just show that every node is visited, but it does not need to be on the same cycle. Here's a counter example:

Take node A, B, C, D and edges $AB = 1, BA = 1CD = 1, DC = 1AC = 10, DB = 10$. The optimal solution is AC, CD, DB, BA for a total of 22. The constraints we have would lead to 2 cycles, AB, BA and CD, DC, since this is the constraints don't specify that we're only allowed one cycle. In this example, every node is visited once (has 1 outgoing and 1 incoming edge chosen), but the solution is not a solution to the TSP.

4 Q4: LP Basics

4.1 Plot the feasible region for the linear program. (Diagram can be drawn approximately to scale + scanned).



4.2 Suppose the objective function is to *maximize* $x + y$. Find the maximum value and the point at which this is attained.

The maximum $x + y$ happens when $x = \frac{22}{5}, y = \frac{6}{5}, x + y = \frac{28}{5}$.

4.3 Say the maximum value found in part (b) is C. Then could you have concluded that $x + y \leq C$ by just "looking at" the equations? [*Hint:* does adding equations, possibly with constants, yield a bound?]

Yes, you can conclude by just looking at the equations. Take the equations $2x + y \leq 10, x + 3y \leq 8$, and multiply the first by 2 to yield $4x + 2y \leq 20, x + 3y \leq 8$. Now add the equations to get $5x + 5y \leq 28$ and divide by 5 for $x + y \leq \frac{28}{5}$, the value from above.

5 Q5: Identifying Corners

5.1 Prove that any such point is (a) a feasible point for the LP

To prove that the point is a feasible point, we just need to show that it satisfies the constraints. For any point $a_1, a_2, ..., a_n$ where $a_i \in (0, 1)$, that means that any a_i is either 0 or 1, and thus the first constraint is satisfied (i.e. $0 \leq 0 \leq 1$ and $0 \leq 1 \leq 1$). Since exactly k of the a_i are 1, the sum of vector is just k. That satisfies the second constraint (i.e. $k \leq k$). Thus, this point a is a feasible point.

5.2 and (b) a corner point of the polytope defining the feasible set.

To show that it's a corner point, I'll use the hint from the question and show that for any non-zero vector$z$, $x + z$ is an infeasible point. Let's consider one of the a_xs in a vector a as defined in the question (that is the sum of the a_is is k and each $a_i \in (0, 1)$). That a_x is either 0 or 1. If we add a none zero vector to a that has a 1 in xth position, then we'll be adding that to a_x to get either a 1 or a 2 back. If the new value is 2, we break constraint 1 because $0 \leq 2 \leq 1$ is false. If the new value is 1, then the sum of the vector a has gone from k to $k + 1$ breaking the second constraint. Thus for any none zero vector (i.e. it has at least a 1 at some position x), the constraints would be violated. Thus a is a corner point of the feasible set.

6 Q6: Checking feasibility vs optimization

6.1 Prove that using $O(\log(M/\epsilon))$ calls to the oracle, one can determine the optimum value of the LP above up to an error of $\pm\epsilon$, for any given accuracy $\epsilon > 0$.