ISyE 4133 – Advanced Optimization – Spring 2022

Problem Set 4, Due: April 11, 2022

Mathieu Dahan

Problem Set Rules:

1. Each student should hand in an individual problem set on Canvas.
2. You may interact with fellow students when preparing your homework solutions. However, at the end, you must write up solutions on your own. Duplicating a solution that someone else has written (verbatim or edited) is not permissible.
3. Late assignments will *not* be accepted. However, each student has one "flex day" during the semester (use it wisely) that grants a 24-hour extension for a problem set. It must be requested to one of the TAs before the original due date of the assignment. A flex day cannot be split into multiple fractional flex days!
4. The solutions and Python codes should be submitted electronically on Canvas prior to the deadline. Any handwritten solutions should be scanned so that they are easily read by the TAs, and then submitted electronically.
5. For questions involving computations, the Python code as well as the answers will be graded.
6. Show your work. Answers requiring a justification will not receive full grades if the justification is not provided. Similarly, answers requiring intermediate computations will not receive full grades if those computations are not shown in your solutions.

Problem 1

Gerrymandering New Mexico (35 points)

In the United States, each state is divided into small regions called districts for the purposes of representation in the House of Representatives, which is one of two chambers of the U.S. Congress. In every even-numbered year, the citizens who reside in each district can vote in an election to determine the U.S. Representative for that district. Representatives hold great power, as they can propose and vote on bills that later can become laws.

Each representative typically is affiliated with one of the two major political parties in the United States: the Democratic Party or the Republican Party. Each party naturally wants to ensure that they have as many representatives in Congress as possible. One way of achieving this is through gerrymandering.

Gerrymandering refers to the process of redrawing district boundaries so as to favor a particular political party. To illustrate this, suppose we have the hypothetical state shown in Figure 1, with three districts, indicated by the bold black lines. Each district is further subdivided along a grid into smaller subregions, where each subregion votes unanimously for either party. For simplicity, suppose that in this hypothetical example there is only one voter in each subregion.

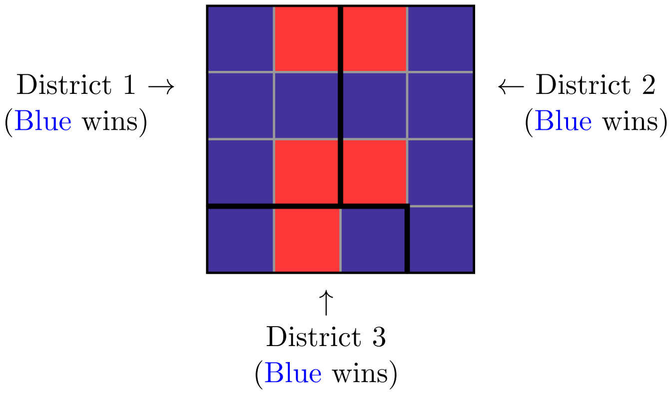

Figure 1: A hypothetical state, with three districts. With the district boundaries shown here, blue wins every district.

Based on the current district boundaries, the blue party has a majority in each district, so each district elects a blue representative. However, suppose we decide to redraw the boundaries as shown in Figure 7. The voting didn't change, but now the red party wins two of the three districts.

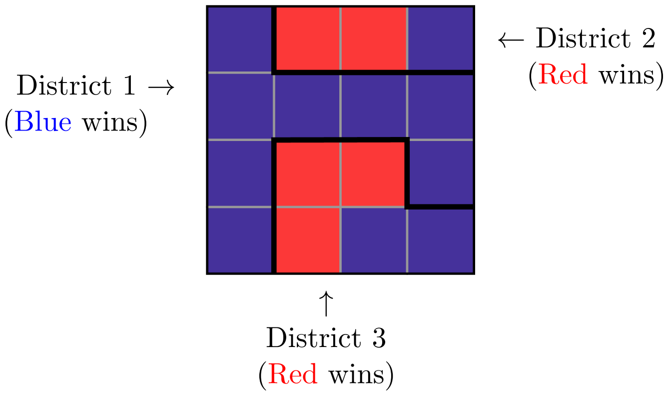

Figure 2: A hypothetical state, with three districts. With the district boundaries shown here, blue wins one district and red wins two districts.

In this problem, we will be exploring how to systematically manipulate these kinds of boundaries. We will be doing this specifically for the state of New Mexico in the United States, which currently has three districts, as shown in Figure 3. It is further subdivided into counties (shown by the faint lines in Figure 3). In many states, a district boundary can go through a county, but for this problem we will assume that a county must belong to one and only one district.

We have the voting record from the 2012 election for each county in New Mexico, which we will use as a proxy for how the county will vote in the next election (whether the majority will be Republican or Democratic). In the 2012 House of Representatives election, the Democratic party won in New Mexico's 1st and 3rd districts, while the Republican party won in the 2nd district. This data is provided in the file *Gerrymandering.xlsx*.

Suppose that we have the opportunity to gerrymander New Mexico. We would like to keep three districts, but redesign them so that we expect the Democratic party to take all three districts. Our decisions are which county should be assigned to which district. We need to make sure that each county is assigned to exactly one district, and each district consists of at least one county.

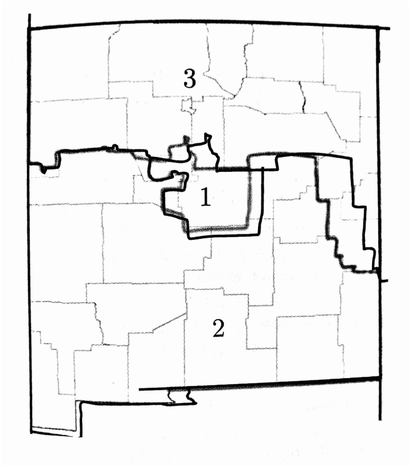

Figure 3: The current three district borders in the state of New Mexico.

The first scenario we will consider is selecting our objective to maximize the number of votes that the Democratic party wins district 2 by (remember that the Democratic party lost in the 2nd district). If we can get the Democratic party to win district 2 by a margin of at least 100 votes while still winning districts 1 and 3 by a margin of at least 100 votes in each district, the Democratic party will win all three districts. So we also need to make sure that the Democratic party still wins districts 1 and 3 by a margin of at least 100 votes.

Note: This problem is independent of the political opinion of the instructor.

a) (7 points) Formulate this problem as an integer optimization problem. Write out the decision variables, objective function, and constraints. You should use the data provided in the file *Gerrymandering.xlsx.* For the vote difference, you should use the numbers under Scenario l.

b) (5 points) Solve this optimization model. What is the optimal solution? By how many votes does the Democratic Party win each of the districts under this redistricting?

c) It would be nice to find a solution for which the Democratic party wins in all three districts, but for which the new assignments are not very different from the existing districts. In the spreadsheet, we have included information about the county assignments in 2012, that is, which counties belong to which districts. (Note that this is an approximate assessment because in New Mexico, districts are not exactly made up of counties.)

   i) (4 points) Let us change our objective so that it instead maximizes the number of districts that are not reassigned. What should the new objective be?

   ii) (3 points) Modify your problem to use the new objective. The old objective should become a constraint like the ones we have for districts 1 and 3 - we want to ensure that the Democratic party wins by a margin of at least 100 votes in district 2 as well. Then re-solve the problem. You should still be using the "Scenario 1" column for the vote differences. In the optimal solution, which counties have been re-assigned? How does this compare to the previous solution?

d) In addition to ensuring that the Democratic party wins in each district, we would also like to take into account the following considerations: (1) Exactly one of Santa Fe (county 27) or

Dona Ana (county 8) must be in district 2; and (2) Both Socorro (county 29) and Torrance (county 31) must be in the same district.

i) (4 points) What constraint(s) do we need to add to our model to make sure that these additional considerations are satisfied?

ii) (4 points) Add these constraints to your model and re-solve it. How does the optimal solution compare to the previous optimal solution?

e) (8 points) So far, we have been using only one voting scenario to design our districts. In this scenario, we have assumed that each county will vote in the representative election of its designated district according to the same example: For instance, Bernalillo county will vote for the Democratic candidate of its district, with a margin of 42,941 more voters (i.e., the number of Democratic votes from Bernalillo is 42,941 higher than the number of Republican votes).

   This is a problematic feature of the model, because voters will not vote in this exact way in future elections. In fact, if they vote sufficiently differently, the democratic party may not be able to win all of its representative elections.

   Let us change our formulation to make it more robust to changes in voter behavior. Suppose that in addition to the data we have been using so far (Scenario 1), we also wish to account for two other scenarios: Scenario 2 and Scenario 3. These scenarios are based on forecasts obtained from a separate prediction model. Furthermore, we want to make sure that the Democratic party wins by a large margin, so we will change the constraints to ensure that the Democratic party wins at least 12,000 more votes than the republicans in each district.

   To do this, we need to revisit our constraints that ensure that the Democratic party wins each district. In particular, the Democratic Party should win each district with a margin of at least 12,000 votes in every scenario.

   Add these constraints to your model, and re-solve it. What is the optimal solution now? How does it compare to the previous optimal solution?

Problem 2

Integer Programming Geometry (20 points)

Consider the following mixed-integer set.

$$S = \left\{ (x, y) : x + y \geq 1.5,\ x \in \mathbb{Z},\ y \geq 0 \right\},$$

where x can be any integer, and y is a continuous nonnegative variable.

a) (10 points) Draw the set S on a plane of (x, y), with x as the horizontal axis. On the same plot, also draw the convex hull of S.

b) (5 points) Is conv(S) a polyhedron? If it is, express conv(IP) as a polyhedron, using the minimum number of linear constraints. If it is not, give a justification.

c) (5 points) On a different plot, draw the feasible region of the linear programming relaxation of S, denoted as S'. Is S' the same as conv(S)?

Problem 3

Integer Programming Algorithms (45 points)

Consider the following integer program.

$$
\begin{aligned}
(IP) \quad \max \quad & 3x_1 - 2x_2 \\
\text{s.t.} \quad & 5x_1 - 2x_2 \le 7.5 \\
& -2x_1 + 5x_2 \le 7.5 \\
& x_1, x_2 \ge 0 \\
& x_1 \in \mathbb{Z},\ x_2 \in \mathbb{Z}.
\end{aligned}
$$

a) (6 points) Draw the feasible regions of (IP) and its linear programming relaxation. Graphically find the optimal solution and optimal objective value of the linear relaxation of (IP).

b) (6 points) Draw the integer hull of (IP), denoted as $\text{conv}(IP)$. Find the optimal solution and optimal objective value of (IP) by manual/graphical inspection.

c) (4 points) Write down the minimal set of linear constraints defining $\text{conv}(IP)$. In other words, express $\text{conv}(IP)$ as a polyhedron, using the minimum number of linear constraints.

d) For this question, we want to use the cutting plane algorithm to solve (IP). First, we derive an equivalent formulation of (IP).

   i) (5 points) Write down an equivalent formulation of (IP), where all the coefficients are integers. We denote the new formulation (IP_2). Mathematically argue that (IP) and (IP_2) are equivalent.

   ii) (4 points) Graph the feasible regions of (IP_2) and its linear programming relaxation. Is (IP_2) a stronger or weaker formulation than (IP)?

   iii) (10 points) Solve (IP_2) using the cutting-plane algorithm. At each iteration, solve the corresponding LP graphically or by manual inspection. When picking a fractional variable, be aware, one might be better than the other. *Note: Although you solve LPs graphically (with 2 variables), generating the cuts requires standard-form LPs.*

e) (10 points) Solve (IP) **(not** $(IP2)$**)** using the branch-and-bound algorithm. At each iteration, solve the corresponding LP relaxation graphically or by manual inspection. Detail all the steps of your work.