Homework on social choice and welfare

April 23, 2024

Electronic submission to Gradescope due **5:00pm Saturday 5/2**.

1 Solving optimization problems in python

Familiarise yourself with solving optimization problems in python. You may use some common softwares like (Diamond and Boyd 2016, CVXPY), (Google, OR-Tools). You may visit `cvxpy.org` for some examples on how to use cvxpy for the formulation of convex optimization problems. To gain more familiarity, formulate and solve the following problem on an LP solver. A linear program is formulated and solved in the colab notebook
`https://colab.research.google.com/drive/1dAQFO2yok3Y7SUUtoMq69_Hn7DYp8Xxf?usp=sharing`.

Please do not directly edit the colab notebook. Make a copy of the notebook and then solve the problems on it. While submitting please attach the relevant python scripts and/or excel files to enable us to give partial credits.

Problem 1

A pharmaceutical company has 10000mg of a certain drug that it wants to package in capsules of 4 sizes- a) extra-large b) large c) medium and d) small which can hold 100mg, 60 mg, 30 mg and 10 mg respectively. Past consumer research shows that the number of large capsules should be at least twice that of extra-large ones and the number of medium capsules should be at least thrice the number of large capsules. Further there must be at least 100 small capsules. Each extra-large capsule is sold for a profit of 5 dollars, large capsule for a profit of 4 dollars, medium capsule for a profit of 2 dollars and small capsule for a profit of 50 cents. Assuming that the company may manufacture fractional number of capsules, how many capsules of each type should it manufacture to maximise the profit?

Formulate the problem as a linear program and use python or excel to solve it. We recommend you to use python as it maybe useful later. Also report the maximum total profit earned by the company.

Check which primal constraints are tight and slack and verify whether the dual variables for the slack constraints are zero. Note that you can get the dual variables after solving the LP from either excel or python, as shown in class/on the python demo.

In the next problem we often refer to the set $\{1, 2, \ldots, m\}$ as $[m]$.

2 Scoring based rules for elections

Consider a generic social choice problem with n voters and $m < n$ candidates with each voter having a preferred ranking over all the candidates. We are interested in studying different scoring based aggregation rules to determine the winner in which a score of s_k is given to a candidate ranked at the k^{th} position and the candidate with highest score wins.

Scoring based rules: A scoring rule is determined by a sequence of real numbers $(s_1, s_2, \ldots, s_m)$ satisfying $s_1 \geq s_2 \geq \ldots \geq s_m$. A candidate who is ranked at k^{th} position by a voter i is assigned score of s_k by voter i and the winner is decided by calculating the cumulative score for each candidate by summing across all voters. Ties of scores maybe broken in an arbitrary way. An example is described below for two different scoring vectors.

Example 1. *Consider an election with three candidates A, B and C and three voters. The first voter ranks $A > C > B$, the second candidate ranks $B > A > C$ and the third candidate ranks $B > A > C$. We now discuss the winner under two different scoring vectors.*

- *Suppose a scoring vector of* $(1, 3/4, 0)$ *is chosen. In this case, candidate* A *gets a cumulative score of* $1 + 3/4 + 3/4 = 2.5$*, candidate* B *a cumulative score of* $2 + 0 = 2$ *and candidate* C *a cumulative score of* $3/4$ *and thus* A *is the winner.*
- *Suppose a scoring vector of* $(1, 1/4, 0)$ *is chosen. In this case, candidate* A *gets a cumulative score of* $1 + 1/4 + 1/4 = 1.5$*, candidate* B *a cumulative score of* $2 + 0 = 2$ *and candidate* C *a cumulative score of* $1/4$ *and thus* B *is the winner.*

We may observe that different scoring vectors may lead to different winners.

3 Different weights on votes for aggregation

However, in this problem we consider under-representations of minorities (in the election process) and over-weight/ underweight some votes. Let us denote the weight that we assign to voter i by w_i and denote the weight vector by $\mathbf{w} \in \mathbb{R}^n$ satisfying $\sum w_i = 1$.

Observe that in this problem the total score of a candidate is the weighted cumulative score by computing the weighted sum over all voters. We give an example below.

Example 2. *Consider an election with three candidates* A, B *and* C *and three voters. The first voter ranks* $A > C > B$*, the second candidate ranks* $B > A > C$ *and the third candidate ranks* $B > A > C$*. We now discuss the winner under the scoring vector* $(3, 2, 1)$ *under two different weight vectors. We assume a tie-breaking order with* $A > B > C$*.*

- *Suppose a weight vector* $(1/3, 1/3, 1/3)$ *is chosen. In this case, candidate* A *gets a cumulative score of* $1/3(3 + 2 + 2) = 7/3$*, candidate* B *a cumulative score of* $1/3(3 + 3 + 1) = 7/3$ *and candidate* C *a cumulative score of* $1/3(2 + 1 + 1) = 4/3$ *and thus* A *is the winner as* A *is preferred over* B *under the tie-breaking rule*
- *Suppose a a weight vector* $(0, 1/2, 1/2)$ *is chosen. In this case, candidate* A *gets a cumulative score of* $1/2(2 + 2) = 2$*, candidate* B *a cumulative score of* $1/2(3 + 3) = 3$ *and candidate* C *a cumulative score of* $1/2(1 + 1) = 1$ *and thus* B *is the winner.*

We may observe that different weighting vectors may lead to different winners.

Problem 2

Recall that in this problem, we over-weight and underweight some votes. We now restrict the weights to deviate from the uniform distribution by factor of α and β and thus restrict the weight vector $\mathbf{w}$ to lie in $S(\alpha, \beta) := \{\mathbf{w} : \sum_{i=1}^{N} w_i = 1; \frac{\alpha}{N} \leq w_i \leq \frac{\beta}{N}\}$. Suppose, we use the scoring vector with $s_i = m + 1 - i$ for every $i \in [m]$.

Now formulate a linear program to check whether there exists a weight vector in $S(\alpha, \beta)$ such that a candidate $c \in \mathcal{C}$ is the winner.

Also solve this linear program for attached file "vote ranking.csv" for different values of α and β in the table to check whether candidate A can be the winner. Solve the same linear program for candidate E as well.

α	β
0.9	2
0.8	2
0.7	2
0.8	3
0.7	5

Table 1: Various values of α and β

4 Knapsack voting problem

Recall the problem of knapsack voting with n voters and m projects and a total budget of B. Each voter $i \in [n]$ votes for a budget $v_i \in \mathbb{R}^m_{\geq 0}$ [1] satisfying the budget constraint $\sum_{i=1}^m v_{i,j} \leq B$.

Recall from class that we gave the formulation of a linear program to select a budget z that maximises the cumulative overlap utility by summing across all voters. The overlap utility between a budget $v \in \mathbb{R}^m_{\geq 0}$ and $z \in \mathbb{R}^m_{\geq 0}$ is given by $\sum_{j=1}^m \min(v_j, z_j)$.

Problem 3

The attached "austin data clean.csv" contains the preferred allocations of $n = 37,004$ voters to $m = 11$ different projects with a total budget of B= 54750 units. Solve the linear program (in python) to select a budget $z \in \mathbb{R}^m_{\geq 0}$ maximising the overlap cumulative utility.

Note that this problem is too big to solve using excel.

Problem 4*

This is much harder than the other problems and is completely optional. If you do this, you can skip all other problems in this HW. Or you can do it for an extra credit of 5% of the total class grade. You can also ask for one (and only one) hint.

Consider the same setup as problem 3. Consider the first vote in the csv file. Now, determine if there are voter weights (using the notation of problem 2) which satisfy $\alpha = 0.8, \beta = 2$ which can make this budget the optimum budget.

Note that this problem is too big to solve using excel.

Problem 5: Rent Division

This uses the same setup as the class. Four friends want to share the cost of a 4-room apartment. The total cost is \$4,000 and their valuations are given in the attached excel file. Use excel's LP solver to find the assignment that maximizes the total valuation. Without solving another LP, use the dual solution (on the sensitivities sheet) to find envy-free prices for the rooms under this assignment. The prices should sum to \$4,000 which might require adding or deleting a constant to the prices that you obtain from the dual solution.

It is not an acceptable solution for this problem to just guess the assignments and the prices. You have to actually use the LP solver and show how you used the dual solution to derive prices.

References

Steven Diamond and Stephen Boyd. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83):1–5, 2016.

Google. Or-tools - google optimization tools. `https://github.com/google/or-tools/`.

[1] $\mathbb{R}_{\geq 0}$ denotes the set of non-negative reals