ORIE 4120 Spring 2022
Dr. David A. Goldberg
Cornell

Homework 2, due Monday 3/14 at 5pm EST uploaded on canvas, 40 points

Q1 (12 points). In this problem, we will model a multi-stage inventory problem as a linear program. In particular, suppose you run a small business, selling a single item.

First, we will model the single-stage problem as a linear program.

Thus consider the following simple setup.

At the start of today (day 1), you have 0 inventory.

At the start of today, you select a (possibly fractional) amount of inventory o_1 to order.

You receive the o_1 units instantaneously.

Suppose it costs you c_1 dollars per unit of inventory you order.

You know the demand for your product that arrives today will be D_1.

Suppose all demand for the day arrives at a single moment in the middle of the day.

Suppose that (as in the EOQ model with backlogging), your net inventory level can go negative. In particular, if $D_1 > o_1$, then your net inventory level at the end of the day is negative. Of course, it may also be positive.

Let I_1 denote your net inventory level at the end of Day 1.

a. (1 point) Express I_1 explicitly in terms of o_1 and D_1

Suppose the cost structure with respect to the net inventory level is similar to that in the HW problem (in HW 1) on EOQ models, although here without any notion of infinitesimals or integrals. In particular, suppose that at the end of the day you incur a one-time inventory cost equal to $h_1 \times \max(0, I_1) + b_1 \times \max(0, -I_1)$.

b. (1 point) Express the total cost incurred on day 1, including ordering costs and inventory costs, explicitly in terms of o_1, c_1, D_1, b_1, h_1.

c. (4 points) Consider the following linear program

$\min c_1 \times o_1 + h_1 \times I_1^+ + b_1 \times I_1^-$

s.t.

$I_1 = o_1 - D_1$

$I_1^+ \geq I_1$

$I_1^+ \geq 0$

$I_1^- \geq -I_1$

$I_1^- \geq 0$

$o_1 \geq 0$

i. (2 points) Argue that in any feasible solution, I_1^+ must be at least $\max(0, I_1)$ and I_1^- must be at least $\max(0, -I_1)$

ii. (2 points) Argue that in any optimal solution, one would never set I_1^+ to be strictly greater than $\max(0, I_1)$; and similarly one would never set I_1^- to be strictly greater than $\max(0, -I_1)$. Conclude that in any optimal solution, it will hold that $I_1^+ = \max(0, I_1)$ and $I_1^- = \max(0, -I_1)$.

d. (1 point) Based on your answers to a.,b., and c., argue that the value for o_1 in the solution to the linear program from part c. gives the optimal amount to order in the stated inventory system.

Consider now a multi-period version of the problem considered in a. - d. In particular, suppose now you are optimizing an inventory system with similar dynamics over T days.

At the start of Day 1, you have 0 inventory.

At the start of each Day t (for $t \in [1, T]$), you select a (possibly fractional) amount of inventory o_t to order, and you receive this inventory amount instantaneously.

Suppose it costs you c_t dollars per unit of inventory you order at the start of Day t (note that this may be different for different days)

You know the demand for your product that arrives on day t will be D_t, and again suppose that all demand for a given day arrives at a single moment in the middle of that day.

Again suppose that, as before, you net inventory level can go negative, although as before it may also be positive.

Let I_t denote your inventory level at the end of Day t

e. (1 point) For $t \geq 2$, expresss I_t explicitly in terms of I_{t-1}, o_t, and D_t.

Suppose the cost structure with respect to the net inventory level is very similar each day to that given in the first part of the problem, but with the specific cost parameters changing each day. In particular, suppose that at the end of Day t you incur an inventory cost equal to $h_t \times \max(0, I_t) + b_t \times \max(0, -I_t)$.

f. (2 points) Write a linear program to optimize your ordering levels $o_1, \ldots, o_t$ over the time horizon to minimize your total cumulative ordering and inventory costs. This should in many ways resemble the formulation from c., but summed over all time periods. You may wish to introduce variables $I_1, \ldots, I_t$ and ${I^+}_1, \ldots, {I^+}_t, {I^-}_1, \ldots, {I^-}_t$ and constraints which enforce $I_t = I_{t-1} + o_t - D_t$ for all $t \in [2, T]$.

g. (2 points) Suppose in this part only that $c_t = 0$ for all t. In that case, how does the inventory problem simplify, and what is the optimal solution and cost?

Q2 (28 points). In this problem, we will model an airline which has to sell tickets for flights as a linear program, a problem sometimes called airline network revenue management. Of course, real models for such a problem are extremely complex, taking millions of constraints and variables into consideration. Here we consider a very simplified optimization model, but which still takes several of the core ideas into consideration. Also, we consider a formulation which is in some ways simple and intuitive, yet which would ultimately be very computationally inefficient in many cases. Indeed, airlines are generally very sophisticated users of modeling and optimization, and here we will not get into more sophisticated ways to model and optimize such problems.

- Suppose you are the airline's chief operations officer, and you are trying to optimize your ticket sales for the day. We will consider a very simple model where all demand for various flights at various prices is known in advance (for example from very sophisticated data-gathering), and the key problem is how to allocate seats and to whom. In a real model there are many unknowns and one must optimize using ideas from probability over the course of months, and we will see such models later in the semester. For now we ignore those complexities.

- Suppose there are n possible origin-destination pairings $T_1, \ldots, T_n$. Each such origin-destination pair consists of an ordered pair of cities, for example (Atlanta , Ithaca). I note that an origin-destination pair will NOT in general correspond to a direct flight, but instead a possible multi-hop trip which takes a flyer from a given origin city to a given destination city (for example there are no direct flights from Atlanta to Ithaca, yet one can buy a ticket from Atlanta to Ithaca). We consider only 1-way flights, so in this problem we do not worry about the notion of "round trip flights" etc.

- Suppose there are m possible individual flights $F_1, \ldots, F_m$. Each of these flights is a single segment / hop, each consisting of an ordered pair of cities and takeoff / landing times, for example (Syracuse at 10 am, Ithaca at 11 am). Each of these ordered pairs DOES correspond to a single direct flight (with its own particular plane and flight number) with no landings inbetween.

- Suppose there are N possible flight combinations $S_1, \ldots, S_N$. Each flight combination is a feasible ordered sequence of individual flights, for example $\big($ (Atlanta at 10 am , Syracuse at 1 pm), (Syracuse at 1pm , Ithaca at 2 pm) $\big)$. We will assume that each of these is a different feasible ordered sequence of connections, where a flight combination may also consist of only a single direct flight. Suppose that for each of the N flight combinations, the corresponding origin-destination pair appears as one of the n origin-destination pairs (although it is possible that the corresponding demand is 0).

- Let $A_{i,j} = 1$ if flight combination S_i is "consistent" with origin-destination pairing T_j, and 0 otherwise. Thus (for example) if $S_5 = \big($ (Atlanta at 10 am , JFK at 1 pm), (JFK at 1pm, Syracuse at 2pm), (Syracuse at 2pm , Ithaca at 3pm) $\big)$, and $T_4 = \big($ (Atlanta , Ithaca) $\big)$, then we would have $A_{5,4} = 1$. Note that a given flight combination is consistent with exactly one origin-destination pair (corresponding to the first city on its first leg and the last city on its last leg), since customers cannot "cheat the system" and "hop off" after the first leg of a multi-leg flight.

- Let $B_{i,j} = 1$ if flight F_i appears in flight combination S_j, and 0 otherwise. Thus (for example) if $S_5 = \big($ (Atlanta at 10 am , JFK at 1 pm), (JFK at 1pm, Syracuse at 2pm), (Syracuse at 2pm , Ithaca at 3pm) $\big)$, and $F_{11} =$ (Atlanta at 10 am , JFK at 1 pm), $F_{20} =$(JFK at 1pm, Syracuse at 2pm), $F_{1003} =$ (Syracuse at 2pm , Ithaca at 3pm), then $B_{11,5} = 1, B_{20,5} = 1, B_{1003,5} = 1$, and $B_{i,5} = 0$ for all other i.

- Let D_i denote the demand for origin-destination pair T_i. Again, this is all taking place for a given fixed day, and this is thus the customer demand for that origin-destination pairing on that given day. Namely, this is the number of potential flyers that want to go from that origin to that destination on this day. Note that this in no way considers the internal hops of flights (and we assume flyers cannot somehow try to cheat the system in that way), namely it represents the genuine demand for flights going from that origin to that destination on this day. We suppose there is no substitution of anything else, and that this demand is due to genuine travel needs (and we ignore effects arising from settings where one could choose from multiple local airports etc.). For simplicity, here we suppose flyers have no preference for the time, length, or price of their flight, just that on that day they get from the given origin to the given destination. This is of course not realistic, but we assume it here for simplicity. We will later address this unrealistic assumption in a certain way, again noting that in real airline operations the models used are very complex.

- It is natural to ask why not model the problem by having the demand be for given flight combinations, instead of given origin-destination pairs. The reason is that in many settings it will be more natural to estimate and model the demand for flying from a given origin to a given destination, as opposed to the demand for any given combination (which is in a sense competing with other specific bundles, as a flyer may have minimal preference for whether they e.g. fly through Philadelphia or Detroit all else equal). Of course, more sophisticated models would use data on all these aspects, along with associated probability distributions that somebody selects one flight over another, but here we consider a very simplified model.

- Let $p(S_i)$ denote the price the airline charges for flight combination S_i. Suppose this has been predetermined by the airline based on market research, as well as fuel costs, etc. Note that (for example) flights with more hops may in general cost less, as they are "less desirable" in some cases. Note that it makes sense to have the price on the combination instead of the origin-destination pair, as some combinations may be much more desirable for even the same origin-destination pair. Of course, in general the price and demand are dependent (higher price may lead to lower demand), and a more sophisticated model would factor all this in (we will later factor in a bit of this in this problem).

- Let M_i denote the maximum number of passengers flight F_i can hold.

- Let x_i denote the number of tickets you sell for flight combination S_i

- In this part of the problem, let us assume that a flyer demanding to take origin-destination pair T_i is indifferent to which flight combination is assigned, and also indifferent to the price (again not realistic, yet simplifying). Thus intuitively, the optimization model we build will attempt to route people in the way which works best for the airline monetarily, as in our model the flyers are indifferent to how they are routed, as long as they get from the given origin to the given destination. Note that it is not "obvious" what the output of such a model will be, since (for example) although there will be a preference to put people on more expensive flight combinations, it may be that more expensive flight combinations do not allow for flying the maximum number of people (as they may route people inefficiently), and thus the optimization model will have to implicitly balance these competing interests.

- Suppose there is no penalty for not meeting some amount of demand, beyond not having made any of the corresponding sales, and that one cannot make sales that exceed the corresponding demands.

- Suppose there is no future inventory of demand for future days, or any other such dynamics.

- Note that the only variables in this formulation are $x_1, x_2, \ldots, x_N$, namely your decisions about how much to sell of each flight combination.

- Suppose fractional sales are possible (even though in practice this would not be the case, and one would need to impose that certain variables were integers, leading to an integer program instead of a linear program).

With the above model in place, let us formulate a linear programming model.

a. (2 points) Express your total profit as a function of $x_1, \ldots, x_N$ and $\{p(S_i)\}$. Note that in this model there is no notion of "costs" (which are presumably factored into some larger model elsewhere), only a notion of profit from sales.

b. (4 points)

i. (2 points) For a given individual flight F_i, express the total number of people that you have assigned to take that flight as a function of $x_1, \ldots, x_N$ and the elements of the matrix B.

ii. (2 points) Write down a set of constraints enforcing that for your choice of $x_1, \ldots, x_N$, no individual flight has its capacity exceeded.

c. (4 points)

i. (2 points) For a given individual origin-destination pair T_i, express the total number of people you have assigned to leave from that origin and arrive at that destination as a function of $x_1, \ldots, x_N$ and the elements of the matrix A. Note that this should count the total number of people assigned to flight combinations whose first leg originates in the corresponding origin city of T_i and whose last leg ends in the corresponding destination city of T_i.

ii. (2 points) Write down a set of constraints enforcing that for your choice of $x_1, \ldots, x_N$, no individual origin-destination pair T_i has its demand exceeded.

d. (2 points) Combine your answers to a. - c. to write down a complete linear programming formulation to solve the airline's optimization problem.

e. (2 points) Intuitively, for the model your wrote down, will it be the case that every flight is filled, or that every origin-destination pair has its demand completely met? Why or why not?

f. (4 points) Suppose now (as a thought experiment, for this part of the question only) that the planes were super-sized, and that effectively $M_i = \infty$ for all i, namely that you can ignore the constraints requiring that no plane has its capacity exceeded.

i. (2 points) In this case, will it be true that all demand is met?

ii. (2 points) In this case, provide a simple and intuitive greedy-like algorithm / method that the airline could use to compute an optimal solution to their optimization problem, and explain in words why it would make sense that such an approach would be optimal.

Note that the above model does NOT capture the fact that a given customer will generally have preferences on the particular flight combinations, presumably due to their quality (e.g. 1 leg better than 3 legs) and price (e.g. cheaper is better). This would clearly be a first-order consideration in such a model. Let us modify the above as follows. In the remainder of the problem, we incorporate some such considerations into the problem, again in a very simplified manner.

- Suppose there are l different "price classes" for the ordered flight combinations. Namely, there are l prices $p_1 < p_2 < \ldots < p_l$ such that every flight combination has price belonging to one of these l values. Thus assume $p(S_i)$ always belongs to the set $\{p_1, \ldots, p_l\}$. Intuitively, suppose that higher price combinations correspond to "higher quality" combinations (e.g. better time, fewer hops, more in demand, etc.).
- Suppose that we segment the demand for a given origin-destination pair by which price/quality range the potential flyer is content with. For $i \in [1, n], j \in [1, l]$ and $k \in [j, l]$, let $D_{i,j,k}$ denote those potential flyers who demand origin-destination pair T_i and who are willing to buy any bundle whose price belongs

to exactly the set $\{p_j, \ldots, p_k\}$. Note that if all that mattered was how much a potential flyer was willing to pay then it would suffice to segment the demand by "those willing to pay at most p_j". However, it would be much more realistic to suppose that a given flyer also cares about the "quality" of the flight bundle, and thus some fliers may (for example) only be willing to take direct flights etc. This can be factored in by segmenting the demand not just by how much they are willing to pay, but instead by what "range" of prices they are comfortable with, thus capturing both an upper bound on how much they are willing to pay and a lower bound indicating they will not accept seats cheaper than some level (as that would be indicative of the flights being somehow undesirable, e.g. due to many hops). Of course, some population of demand will be fine with flights down to the cheapest price level. Suppose that these sets of demands are disjoint, i.e. each of these demand pools $D_{i,j,k}$ represents different potential flyers, and there is no notion of substitution etc. Note that when we model it this way, there is no longer a notion of D_i (althoug this could in principle be recovered by summing over the different price ranges). Also note that we are not considering any notion of first-class seats or anything like this on any given flight.

- Let $x_{i,j,k}$ denote the number of flight combination S_i you sell to a customer whose pay range is $\{p_j, \ldots, p_k\}$. Note that conceptually, this modeling approach is related to the demand segmentation we modeled in the brewery example in class.

g. (2 points) Express x_i, the total number of tickets you sell for flight combination S_i, in terms of the $x_{i,j,k}$ variables.

h. (2 points) Express the total number of tickets you sell to customers on origin-destination pair T_i whose pay range is $\{p_j, \ldots, p_k\}$ in terms of the $\{x_{r,j,k}, r = 1, \ldots, N\}$ variables and the entries of the matrix A.

i. (2 points) Write down a set of constraints enforcing that for your choice of $\{x_{i,j,k}\}$ variables, the number of tickets sold to potential flyers on each target-destination pair for each given price range does not exceed the number of customers that demand that given price-destination pair and are willing to pay exactly that price range.

j. (1 point) Interpret the following constraint (note that a similar kind of idea was discussed in the brewery example from class):

$$x_{i,j,k} = 0 \text{ for all i and all j,k} \in [1, l] \text{ for which } p(S_i) \notin [p_j, p_k].$$

k. (3 points) Combine all of the above insights with your original LP formulation from part d. to write down a new LP formulation modeling the airline's profit maximization problem in this more sophisticated model.