Notes on Linear Programming

Scribe: Jiajia Xie July 5, 2020

In this tutorial, I would like to introduce some basic knowledge about linear programming (LP). This tutorial will introduces a fundamental problem formulation in LP.

Mixing Problem [1]

An operation research problem often appears in oil industry is to determine the best way of processing crude oil so that the cost of production is minimized. To answer this problem, we have to formulate a problem but in purely mathematical expression which allows us to analyze the problem nature, to derive the dual problem (the dual problem of minimizing cost can be maximizing the profit), or to implement or derive efficient algorithm to solve the problem. Formulation can be a difficult and many people get stumped at the beginning. We should start the formulation step by step to eventually present the problem of minimizing the cost of oil processing subject to different requirements. To begin, please keep the following three questions in mind:

1. What choices do I have? What can I control. These tells you what your decision variables are.

2. What are the data? Remember, data is not something I can control with. Data will involve in the function of modeling but they will just serve as some constants.

3. What is my objective? This will give you the objective function.

4. What limits my choices? What is impossible for me to do? This will give you the constraints.

Now, let us begin. An oil and gas company has M (this is data) oil fields and N processing plants. We use indices $i \in \{1, 2, ...M\}$ to denote each oil field and $j \in \{1, 2...N\}$ to denote each processing plant. Crude oil is extracted from oil field at a cost of c_i dollars per barrel and its percentage of sulfur (a type of pollutant) is a_i. All extracted oil is transported to the processing plant (so there will be a transportation cost). At each processing plant, crude from various oil plants are blended together. The cost of transporting from oil field i to processing plant j is $f_{i,j}$ dollars per barrel.

Consider the first question: what choice do I have? An oil company needs to extract and allocate crude oil from each oil field i to each plant j, so we can use $x_{i,j}$ as the decision variable. The second question is to determine the objective of the problem. In this case, the cost of producing and transporting oil is the sum of extraction and transportation $\sum_{i=1}^{M}\{\sum_{j=1}^{N} x_{i,j} f_{i,j} + c_i \sum_{j=1}^{N} x_{i,j}\}$. What are the limits? Things get tricky here. Can each oil field produce a negative number of barrels? Absolutely no! So there should be a constraint on $x_{i,j} \geq 0$ for every i and j.

Now, let us start to write an incomplete version of the formulation. In general, the formulation has the objective function with $\min_x$ or $\max_x$ indicating the direction of optimization and the decision variable x, and s.t represents *subject to* constraints:

$$\begin{array}{rl} \min\limits_{x} & \sum_{i=1}^{M} \sum_{j=1}^{N} x_{i,j} \left(f_{i,j} + c_i \right) \\ \text{s.t.} & x_{i,j} \;\; \geq \;\; 0, \; \forall i \in \{1, ...M\}, \forall j \in \{1, ...N\} \end{array} \tag{1}$$

Carefully review this formulation. Does it make sense? Doesn't all $x_{i,j} = 0$ yield the optimal solution? That is, the solution of this formulation tells you that the best way of minimizing the cost of production is not producing, which is obviously not true in the real-world. So what are we missing? Review you economic class, a market consists of supply and demand. An producer like the oil company must meet the demand from the market, denoted by D number of barrels. There is another market constraint that the decision needs to satisfy: $\sum_{i=1}^{M} \sum_{j=1}^{N} x_{i,j} \geq D$, which indicating the total produced oil should be greater than or equal to the market demand.

Is that good enough? Note that an oil company in the real-world can cause severe pollution, such as sulfur. Local governments typically assign quota to the plants to limit sulfur emission. Let us say, an quota of u_j is assigned to the processing plant j. There is another pollution constraint $\sum_{i=1}^{M} x_{i,j} a_i \leq u_j$ for every plant j.

Is that good enough? We must think about the capacity of oil fields and processing plants. Let us say, the processing plant j has an inventory limit that cannot exceed p_j barrels. The production limit of the oil field i cannot exceed o_i. We add another two constraints: $\sum_{i=1}^{M} x_{i,j} \leq p_j$ for every j, $\sum_{j=1}^{N} x_{i,j} \leq o_i$ for every i.

Putting the above constraints together, let us see the formulation again:

$$\begin{array}{rr} \min\limits_{x} & \sum_{i=1}^{M} \sum_{j=1}^{N} x_{i,j} \left(f_{i,j} + c_i \right) \\ \text{s.t.} & x_{i,j} \geq 0, \; \forall i \in \{1, ...M\}, \; \forall j \in \{1, ...N\} \\ & \displaystyle\sum_{i=1}^{M} (1 - a_i) \sum_{j=1}^{N} x_{i,j} \geq D \\ & \displaystyle\sum_{i=1}^{M} x_{i,j} a_i \leq u_j, \; \forall j \in \{1, ...N\} \\ & \displaystyle\sum_{i=1}^{M} x_{i,j} \leq p_j, \; \forall j \in \{1, ...N\} \\ & \displaystyle\sum_{j=1}^{N} x_{i,j} \leq o_i, \; \forall i \in \{1, ...M\} \end{array} \tag{2}$$

Now you have the mathematical formulation of the oil company problem to minimize the cost. Note that the above objective and constraint functions are all linear in the control variable x. Any optimization problem with such property is called Linear Programming. An LP problem is guaranteed to be solved efficiently i.e. existing many polynomial time algorithms.

Exercise: Product Manufacturing Problem [1]

The company ZinCo. manufactures a dairy-based food additive. They need to plan their production for the next 12 months. Initially they have 100 kg of the product in stock. The demand for the product in month t is D_t kg. They can produce more than the demand, and hold the excess in inventory. Because of spoilage, only 95% of the inventory in month t is available in month $t+1$. To have a smooth production schedule, this company requires that, starting form the first month, the absolute difference between production in month t and $t+1$ should not exceed L_t kg. The cost of production in month t is C_t dollars per kg and the cost of holding the product is H_t dollars per kg.

You may use the simplified assumptions that production finishes at the beginning of the month, and demand is fulfilled at the beginning of the month, and the spoilage only occurs at the end of the month.

Formulate an mathematical expression of the above manufacturing problem in LP.

References

[1] C. A. Tovey, *Linear Optimization and Duality: A Modern Exposition.* Chapman and Hall/CRC, 2020.