School of Mathematics
Dr Kit Searle

Risk and Logistics

Homework Assignment 1

January 27, 2025

Instructions

1. For this assignment, you will have to write a report and produce some code, preferably using Xpress, but python is also ok.
2. You should attempt all questions.
3. The total marks for this assignment are: 15.
4. Explain in your report what you have been doing in a concise and reproducible way. Write full sentences, not just the final results. The style in which you present your answers is part of the marking.
5. Structure your code, include comments, and avoid unnecessary loops, conditions, and calculations. The style in which you write your code is also part of the marking.
6. You don't have to do everything by the book. **Use your imagination!**

   If you have an idea for a new method or one that helps you to improve the efficiency or effectiveness (or both!) of an existing method, without deteriorating the other too much, then please go ahead. The questions are kept deliberately vague to allow you that freedom (it would be boring – both for you and me – if I'd tell you exactly what to do and what not).

   I am happy with any method that is reasonable and does the job, but I expect you to properly explain and motivate why you think this is a good idea - and, ideally, back it up with some empirical results. Marks will be awarded for clever ideas (to some degree, independent of whether they work or not), and the overall efficiency, i.e. runtime, and effectiveness, i.e. solution quality, of your heuristics.
7. The strict deadline for handing-in your assignment is **16:00 on Friday, 14 February 2024**.

   Please upload exactly two files on Gradescope: a PDF with your report and a zip file containing all your codes and, if applicable, your data files.

Preliminaries

I have uploaded on Learn

- *CaseStudyData.txt*: a file containing the data of the case study in Xpress format. The file contains two additional fields: *CustomerDemandPeriods* and *CustomerDemandPeriodScenarios*. The former contains the demand per customer, product group, and year, and the latter the demand per customer, product group, year, and scenario (see below for more information on "years" and "scenarios".
- *CaseStudy.mos*: a Mosel template that already includes all variable declarations and routines for reading data which you can use as a starting point.

Introduction

While *Tartan Trade* is aiming at winning over 5% of all people living in Scotland as customers, it is very unlikely that they will reach this target within the first year, or even within the first five years. Instead, they anticipate that their customer base will be fairly small in the first year, but then – hopefully – keep on growing over about 8-10 years, reaching the 5% target at the end of Year 10, at the latest. However, all of this is only an estimate based on forecasts and the actual demand may very from customer to customer and, moreover, deviate quite significantly from the predicted demand.

All data given and all decisions taken in the basic Warehouse Location Problem refer to one generic time period, e.g. 5 years or 10 years. Hence, constructing all warehouses at the start will result in many warehouses being (heavily) underutilized in the first few years and others being, possibly, overutilized in later years. Instead, it would be more appropriate to build up the supply chain step-by-step over the years as customer demand increases. That is, we would not only decide where to build warehouses, but also in which year to build them.

The goal of this exercise to extend the model for the multi-echelon capacitated WLP (MECWLP) by allowing for a gradual build-up of the supply chain over time and to account for uncertainty in customer demands.

Aggregation (4 marks)

To obtain optimal or even just near-optimal solutions for the case study data you will inevitably have to aggregate the set of customers, or the set of candidate locations, or even both. Once you do so, please clearly explain what you are doing and why. I am not expecting you to use or develop any sophisticated aggregation method; a simple one will do as long as it is properly motivated.

I leave the decision to you when to start aggregating, and whether to aggregate only customers or candidate locations, or both.

The MECWLP with multiple time periods (5 marks)

To start with, we divide the planning horizon into ten *time periods*, one for each year. In the following, denote the set of time periods, i.e. years, by $T = \{1, \ldots, |T|\}$. To account for the periods, all data must now be given with respect to time, at least in principle. To simplify matters for this assignment, only the customer demands will be time-dependent; all other parameters remain time-invariant. In the following, the demand of customer $i \in I$ in period $t \in T$ is denoted as b_i^t. Moreover, having more than one time period, f_j now only accounts for the cost of actually constructing a warehouse at j. The annual cost of keeping a warehouse at j running is now given as a separate operating cost g_j (the cost is proportional to the warehouse capacity, but independent of the amount of goods handled by the warehouse).

Once a warehouse has been built, it remains in operation until the end of the planning horizon and, accordingly, we have to pay the annual operating costs until the end. Moreover, each warehouse can serve a different set of customer in each time period and each customer can be served by a different warehouse in each period.

For simplicity, we assume that a warehouse can be constructed instantaneously, i.e. if we decide to build a warehouse in period t, then the warehouse is ready to serve demands already in period t. While this is obviously unrealistic, it is straight-forward to adapt the model to starting construction one or more periods earlier, and this adaptation doesn't change the nature of the formulation, it only messes up the indexing of the variables.

The task is now to determine where to build warehouses and in which time period, and which warehouse is to supply which customer in each period such that all customer demand is served at minimal cost (= setup cost and operating cost for warehouses plus transportation cost).

Questions

Answer the following questions about the problem described above. For each question, briefly explain in the report what you are doing.

(1) Derive a mixed-integer linear programming formulation for the MECWLP with multiple time periods. Include the formulation in your report and explain the meaning of your constraints and variables.

Try to find a computationally efficient formulation, i.e. a formulation that finds an optimal solution in as little time as possible.

(2) Implement your formulation in Mosel and test it with the case study data.

Include in your report where and when warehouses should be built, the different costs of the solution (construction, operation, transportation), the run time to find the (optimal) solution, and the MIP gap for non-optimal solutions (see more at the end).

Note: The data file includes a field *CustomerDemandPeriods* that contains the customer demands per product group and period.

The MECWLP with multiple time periods and uncertain demands (6 marks)

Having to accurately predict the demand in ten years time is almost impossible. Thus, you will now have incorporate uncertainty into your formulation. We can interpret the given customer demand per period, b_i^t, as the predicted demand of customer i in period t. As forecasts are less accurate the further they are in the future, we assume that the actual demand of a customer in a period can deviate from the predicted demand by up to 10% more for each time period. That is, if b_i^t is the predicted demand in period t, then the actual demand $\hat{b}_i^t$ can deviate by up to $t \cdot 10\%$ from b_i^t, i.e. $\hat{b}_i^t \in [(1 - 0.1 \cdot t)b_i^t, (1 + 0.1 \cdot t)b_i^t]$. For simplicity we assume that all scenarios are equally likely.

Questions

Answer the following questions about the problem described above. For each question, briefly explain in the report what you are doing.

(a) Using scenarios, derive a mixed-integer linear programming formulation for the deterministic equivalent of the MECWLP with multiple time periods and uncertain demands so that all subproblems are feasible, i.e. we have sufficient capacity to serve the demand of any scenario. Include the formulation in your report and explain the meaning of your constraints and variables.

(b) Implement your formulation in Mosel.

Include in your report where and when warehouses should be built, the different expected costs of the solution (construction, operation, transportation), the run time to find the (optimal) solution, and the MIP gap for non-optimal solutions (see more at the end).

Moreover, compare the solution to the one from the previous question and interpret your results.

Note: The data file includes a field *CustomerDemandPeriodScenarios* that contains the customer demands per product group, period, and scenario for a total of 100 scenarios.

You do not have to use all scenarios that are provided. If you use fewer, e.g. 20, then please use the first twenty scenarios.

If you think it helpful, then please let me know and I can create more scenarios. I kept it to 100, since the data file was starting to get very big.

(c) In (a) you were asked derive a formulation that ensures that all subproblems are feasible. While this seems sensible at first glance, it is a very conservative approach that is typically not used in practice. The main issue is that a few high-demand scenarios will force the formulation to open a lot of warehouses in order to have enough capacity in place to cope with these high-demand scenarios – scenarios which maybe very unlikely to happen but still exert an unduly large influence on the final solution.

Can you come up with an idea for a less conservative formulation? Maybe one that does not have to serve all demand or allows to violate warehouse capacities? If so, then try derive a mixed-integer programming formulation for it. You do not have to implement the formulation. Please properly outline your reasoning.

Note: Just deleting high-demand scenarios from the data set is not a viable option.

Some helpful Xpress commands

It is possible to limit the run time of the optimization using the Xpress command

```
setparam("XPRS_MAXTIME", s)
```

where `s` is the run time in seconds. For example, `setparam("XPRS_MAXTIME", 1800)` stops the optimization after half an hour and returns the best found feasible solution. The command needs to be included before the `minimize` or `maximize` command.

If the minimization is stopped before reaching the optimal solution, then it is interesting to get a quality guarantee for the best found feasible solution by reporting the "MIP Gap". The MIP Gap is the relative percentage deviation of the objective value of the best found feasible solution from the best lower bound found by Xpress (and constitutes an upper bound on how far away the best found solution is at most from the optimum). In IVE, the MIP gap is given on the "Stats" tab. For Workbench, it can be computed using

```
100 * (getobjval - getparam("XPRS_BESTBOUND")) / getobjval
```

If you are using IVE, then the "Stats" tab will inform you about the progress of the optimization, especially concerning the MIP gap. For Workbench, there is no such tab. But including the command

```
setparam("XPRS_VERBOSE", true)
```

will print a status log. Again, the command needs to be included before `minimize` or `maximize` (it also works for IVE).