15.094 Homework 4 Solutions

Was due: 04/21/2021

1 Polyhedral uncertainty and extreme points

$$\begin{aligned} \text{minimize} \quad & c^T x + \max_{b \in \mathcal{U}} q^T y(b) \\ \text{subject to} \quad & Tx + Wy(b) \geq b, \quad \forall b \in \mathcal{U}. \end{aligned} \tag{1}$$

Let $\mathcal{U} \subseteq \mathbb{R}^n$ be a polyhedron and denote its extreme points by $ext(\mathcal{U})$.

Show that the objective value of Problem (1) is equal to that of

$$\begin{aligned} \text{minimize} \quad & c^T x + \max_{b \in ext(\mathcal{U})} q^T y(b) \\ \text{subject to} \quad & Tx + Wy(b) \geq b, \quad \forall b \in ext(\mathcal{U}) \end{aligned} \tag{2}$$

Hint 1: Problem (1) is equivalent to

$$\text{minimize} \quad c^T x + \max_{b \in \mathcal{U}} Q(x, b)$$

where the second-stage cost function is given by

$$Q(x, b) = \begin{array}{ll} \text{minimize} & q^T y \\ \text{subject to} & Tx + Wy \geq b. \end{array}$$

Hint 2: You may use the fact that maximizing a convex function over a polyhedron $\mathcal{P}$ is equivalent to maximizing over $ext(\mathcal{P})$.

Solution:

Problem (1) can be rewritten as

$$\underset{x}{\text{minimize}} \quad c^T x + \max_{b \in \mathcal{U}} Q(x, b)$$

where the second-stage cost function is defined as

$$Q(x, b) = \min_y \left\{ q^T y : Wy \geq b - Tx \right\}.$$

Applying duality to the second-stage cost function, we obtain

$$Q(x, b) = \max_\lambda \left\{ (b - Tx)^T \lambda : W^T \lambda = q, \ \lambda \geq 0 \right\}.$$

From the dual representation, we observe that $Q(x, b)$ is the pointwise maximum of linear functions which are linear in b and x, and so $Q(x, b)$ is convex in b for every fixed x. Therefore, the optimum of $\max_{b \in \mathcal{U}} Q(x, b)$ is obtained when b is an extreme point of $\mathcal{U}$.

2 Feasibility of LDRs

Consider the two-stage adaptive optimization problem

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & y_1(b) + y_2(b) \leq 1 \\ & y_1(b) \geq b_1 \\ & y_1(b) \geq -b_1 \\ & y_2(b) \geq b_2 \\ & y_2(b) \geq -b_2 \\ & \quad \forall b \in \mathcal{U} \end{array} \tag{3}$$

where the uncertainty set is

$$\mathcal{U} = \left\{ b \in \mathbb{R}^2 : \|b\|_1 \leq 1 \right\}.$$

First, show that Problem (3) has a feasible solution. Then, show that Problem (3) becomes infeasible if we restrict to linear decision rules.

Solution:

The decision rules $y_1(b) = |b_1|$ and $y_2(b) = |b_2|$ are feasible for the unrestricted problem. For the linear decision rule approach, it is clear that

$$\begin{aligned} y_1(b) &= y_1^0 + y_1^1 b_1 + y_1^2 b_2 \\ y_2(b) &= y_2^0 + y_2^1 b_1 + y_2^2 b_2 \end{aligned}$$

will satisfy the bottom four constraints if and only if

$$y_1^0, y_2^0 \geq 1.$$

However, when $b = (0, 0)$, the first constraint becomes

$$1 + 1 \leq 1,$$

and thus we have infeasibility.

3 Adaptive Facility Location Problem

In this problem, we will explore a canonical example of adaptive robust optimization, which is the facility location problem. The nominal problem is given by the following, and is provided in `HW4_facilityLocation.jl`, along with the `plot_solution` method which will allow you to compare your results.

$$\begin{aligned}
\min_{y(\cdot),x} \quad & \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}y_{ij} + \sum_{i=1}^{n} f_i x_i \\
s.t. \quad & \sum_{i=1}^{n} y_{ij} \geq d_j, \quad \forall d \in \mathcal{D},\ \forall j \in [m], \\
& \sum_{j=1}^{m} y_{ij} \leq s_i x_i, \quad \forall i \in [n], \\
& y_{ij} \geq 0, \quad \forall i \in [n],\ \forall j \in [m], \\
& x \in \{0,1\}^n.
\end{aligned}$$

$i \in [n]$ is a possible location for a facility, and $j \in [m]$ is a demand node. $c_{i,j}$ is the cost of transportation of goods from facility i to destination j, and s_i is the capacity of facility i. It costs f_i to construct facility i, and the demand at location j is d_j and uncertain, lying within uncertainty set $\mathcal{D}$. The binary construction decisions $\mathbf{x}$ precede the delivery decisions $\mathbf{y}$, and thus the problem may be formulated as an adaptive robust optimization problem.

3.1

Please solve the nominal optimization problem using the provided code in `HW4_facilityLocation.jl`. Provide the plot using the `plot_solution` function provided, as well as total, facility and transportation costs.

Solution:

The nominal solution is in Figure 1.

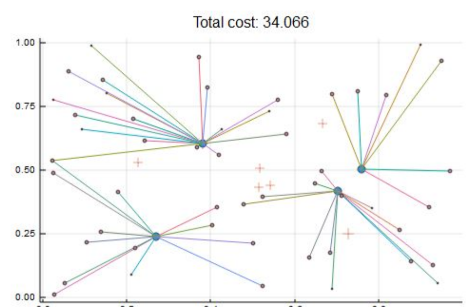


Figure 1: The nominal solution, with total, facility and transportation costs of 34.07, 21.51 and 12.56 respectively.

3.2

We want to account for uncertainties in demand. Assume we know that our demand uncertainty is correlated among nodes based on pairwise distances between nodes i, j with the following rule:

$$P_{i,j} = \begin{cases} 0.2 \exp\left(-\frac{1}{R_D} D_{i,j}\right), & \text{if } D_{i,j} \leq R_D \\ 0, & \text{otherwise} \end{cases}$$

where $D_{i,j}$ is the Euclidian distance between nodes i and j, and where $(Pz)_j$ is the uncertain perturbation on demand d_j at node j, and R_D is the radius of influence. We will assume that z belongs to a budget uncertainty with $||z||_\infty \leq 1$ and $||z||_1 \leq 5$.

Using this set, please formulate in JuMP/JuMPeR the robust facility location problem with fixed recourse, where $\mathbf{y}(\mathbf{z}) = \mathbf{y}$. Solve it with $R_D = 0.25$, providing the plot, the facility cost and transportation cost.

Please provide intuition about the choice of 1 and 5 as safety factors to the uncertainty set, especially considering the shape of $P_{i,j}$.

Solution:
The robust (fixed recourse) problem is the following:

$$\begin{aligned}
\min_{y(\cdot),x} \quad & \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}y_{ij} + \sum_{i=1}^{n} f_i x_i \\
s.t. \quad & \sum_{i=1}^{n} u_{ij} \geq d_j + (Pz)_j, \ \forall z \in \mathcal{Z}, \ j \in [m], \\
& \sum_{j=1}^{m} u_{ij} \leq s_i x_i, \ \forall z \in \mathcal{Z}, \ i \in [n], \\
& \mathcal{Z} = \{z : ||z||_\infty \leq \rho, \ ||z||_1 \leq \Gamma\}, \\
& y_{ij} \geq 0, \quad \forall i \in [n], \ \forall j \in [m], \\
& x \in \{0,1\}^n.
\end{aligned}$$

We have formulated and solved the problem using both JuMPeR's RobustModel, as well as using a direct reformulation in JuMP. You can find both codes in `HW4_facilityLocation_Solution.jl`. The result is 6 facilities instead of 4 as in the nominal case.

The intuition behind the budget uncertainty set with $\rho = 1$ and $\Gamma = 5$ is that we are allowing up to 5 maximal magnitude (1) perturbations in z, or a convex combination of them not going above a value of 5. These perturbations are inverse exponentially shaped by the kernel that describes P, almost like ripples in a lake.

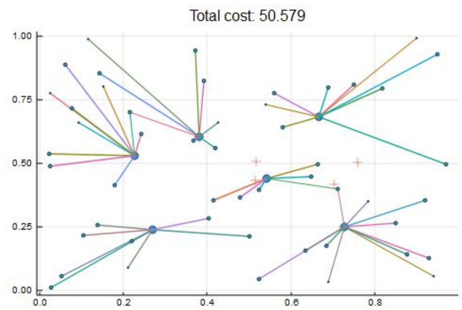


Figure 2: The robust (fixed recourse) solution to the facility location problem results in 6 facilities. The total, facility and transportation costs are 50.58, 34.36 and 16.22 respectively.

3.3

We are exploring the network effects of the radius of influence R_D on our facility locations and transportation costs. Please solve the fixed recourse problem for $R_D \in \{\epsilon : 0.05 : 0.6\}$, providing the following tabulated data:

- Value of R_D
- Sum of all values in matrix P
- Total cost
- Facility cost
- Transportation cost
- Number of nonzero elements of P
- Optimal number of facilities

(Note the ϵ to avoid the singularity at 0. $\epsilon = 10^{-5}$ should be sufficient.) Provide network plots of all unique resulting facility configurations.

Assess the results and explain your intuition about the influence of R_D. Do the results make sense? How is R_D related to the uncertainty set/scenarios?

Solution:

To do this, you have to regenerate the P matrix for each value of R_D, and reoptimize. The resulting solutions can be generated using `HW4_facilityLocation_Solution.jl`, and are presented in Table 3.

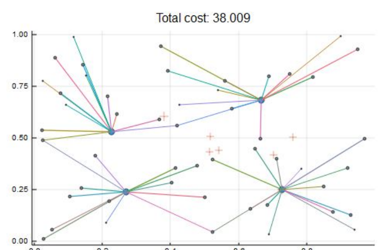
The unique configurations are shown in Figure 4.

The intuition behind R_D is that there are many ways to change the size of the uncertainty set in a problem, and that it is not just a function of ρ and Γ for the budget uncertainty set. In different circumstances, it may be more appropriate to use one or the other. The intuition behind the budget set

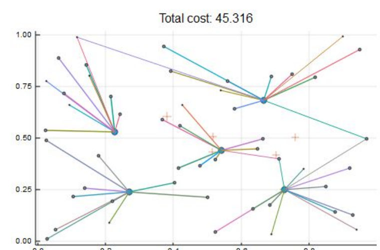
16x8 DataFrame

Row	R_D	sumP	Gamma	f+c	f	c	nP	nx
	Any	Any	Any	Any	Any	Any	Any	Any
1	1.0e-5	10.0	5	36.4976	22.6604	13.8372	50	4.0
2	0.05	10.3	5	36.56	22.6604	13.8997	54	4.0
3	0.1	14.9422	5	38.009	23.4904	14.5186	98	4.0
4	0.15	22.4184	5	43.8096	28.5491	15.2605	170	5.0
5	0.2	32.7188	5	45.3161	28.5491	16.767	268	5.0
6	0.25	43.9281	5	50.579	34.3567	16.2223	368	6.0
7	0.3	61.2547	5	51.2101	34.3567	16.8534	538	6.0
8	0.35	77.8581	5	51.5536	34.3567	17.1969	688	6.0
9	0.4	94.319	5	51.8302	34.3567	17.4735	830	6.0
10	0.45	113.116	5	52.0571	34.3567	17.7004	998	6.0
11	0.5	133.544	5	52.2463	34.3567	17.8896	1182	6.0
12	0.55	151.77	5	52.4063	34.3567	18.0496	1332	6.0
13	0.6	171.48	5	52.5433	34.3567	18.1866	1500	6.0
14	0.65	191.27	5	52.662	34.3567	18.3053	1666	6.0
15	0.7	208.173	5	52.7662	34.3567	18.4095	1792	6.0
16	0.75	226.811	5	52.8585	34.4381	18.4204	1942	6.0

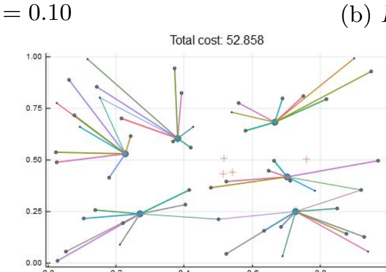
Figure 3: Increased radius of influence results in worse worst-case outcomes, increasing the size of the uncertainty sets and changing the set of facilities required.



(a) $R_D = 0.10$



(b) $R_D = 0.20$



(c) $R_D = 0.75$

Figure 4: As the radius of influence grows, more facilities are required to address the worst-case demand outcomes.

remains; we still allow up to $\Gamma = 5$ maximal bounded perturbations. However, these perturbations have a greater area of effect as R_D increases. As a surrogate to see this, if we look at the sum of all values in matrix P, we see that it grows with R_D.

3.4

Please formulate in JuMP/JuMPeR the adaptive robust problem with linear policies, where $\mathbf{y}(\mathbf{z}) = \mathbf{Y}\mathbf{z} + \mathbf{Y}_0$. Solve (or try to solve) the problem for the nominal $R_D = 0.25$.

(Note: This problem will take a lot longer to solve. Give it time; depending on your machine it might take 24 hours or more. Solving the problem within your terminal instead of a jupyter notebook will allow you to track its convergence through the Gurobi log. This will give some value insight into what is happening; we encourage you to optimize at least until the first feasible integer solution, given by an asterisk in the log. No worries if you don't have the time or patience for this; we would like to see that you can formulate the problem.)

How do our facility allocations change? Why might affine decision rules confer a big cost reduction compared to fixed recourse in this case? (Hint: think about the server optimization problem from HW2.)

Solution:
The adaptive problem with affine policies is given below:

$$
\begin{aligned}
\min_{u,V,x} \quad & \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}y_{ij} + \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{m} c_{ij}V_{ijk}z_j + \sum_{i=1}^{n} f_i x_i \\
s.t. \quad & \sum_{i=1}^{n} u_{ij} + \sum_{i=1}^{n}\sum_{k=1}^{m} V_{ijk}z_k \geq d_j + (Pz)_j, \ \forall z \in \mathcal{Z}, \ j \in [m], \\
& \sum_{j=1}^{m} u_{ij} + \sum_{j=1}^{m}\sum_{k=1}^{m} V_{ijk}z_k \leq s_i x_i, \ \forall z \in \mathcal{Z}, \ i \in [n], \\
& \mathcal{Z} = \{z : ||z||_\infty \leq \rho, \ ||z||_1 \leq \Gamma\}, \\
& u_{ij} + V_{ijk}z_k \geq 0, \quad \forall i \in [n], \ \forall j \in [m], \ \forall k \in [m] \\
& x \in \{0,1\}^n.
\end{aligned}
$$

In practice, this is a large scale mixed integer optimization problem with more than 100,000 constraints and 80,000 variables, due to the nature of the robust counterpart of the budget uncertainty set, as well as the scale of the adaptive variables.

As part of the homework, we provided the network plot in Figure 5 with associated costs.

Using this solution, we see clearly that our facility locations change dramatically compared to the robust case, most notably requiring only 4 facilities instead of 6. The intuition behind is similar to the intuition in the collaborative server case in HW2. Since affine policies allow facilities to act jointly to address demand perturbations, the capacity constraints of each facility are less

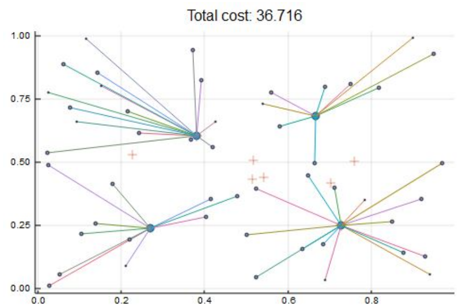


Figure 5: The adaptive facility location result, using greedy heuristic for binary variables.The total, facility and transportation costs of 36.72, 23.23 and 13.39 respectively.

constraining, thus requiring fewer facilities. In the fixed recourse case, such a collaboration cannot happen.

In `HW4_facilityLocation_Solution.jl`, we generate the solution in Figure 5 by greedily picking the binary facility location decisions, since the original ARO formulation was not tractable in a reasonable time. A closer approximation is in `HW4_facilityLocation_Adaptive.jl`, where we restrict the space of possible linear policies by only allowing nonnegative u's and V's. The solution we get is in Figure 6.

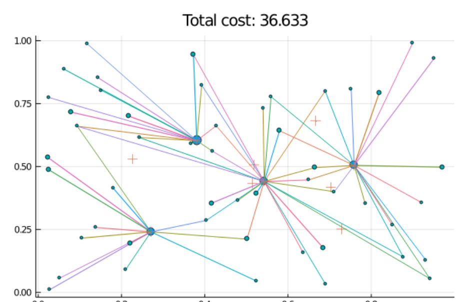


Figure 6: The adaptive facility location result, using robust counterpart with nonnegative u and V. The total, facility and transportation costs of 36.63, 21.43 and 13.29 respectively.

Note that these solutions are generated using potential answers to Part 3.6!

3.5

Please provide your source code.

3.6 Extra credit (10 points)

In practice, we didn't solve the binary ARO problem directly, since the solution did not converge after over 24 hours on an 8-core i7 machine. We used an inexact but good heuristic to dramatically speed up the solution process, which reduced solution time to less than a minute. Come up with a method to speed up the solution of the binary ARO problem, and present the results.

Solution:

We came up with 2 methods to dramatically speed up the solution process. Possible solutions are not limited to these two.

The first relies on reducing the binary optimization problem into a linear problem by taking a greedy approach to the facility location variables x. We relax the binary constraints so that $0 \leq x_i \leq 1, \ \forall \ i \in [n]$. At every iteration, we rank the values of x, and constrain the maximum non-integer value to 1. We repeat until convergence. Note that this is suboptimal, but ends up resulting in a high-performing result nonetheless. We provide the first method in `HW4_facilityLocation_Solution.jl`, and the results are in Figure 5.

The second is a tighter solution to the full ARO problem, and imposes a nonnegativity constraint on the V-component of the LDRs, with a simple constraint $V_{i,j,k} \geq 0, \ \forall \ i, j, k$. Note that this restricts the LDR, so it is not necessarily optimal to the original problem which allows $V \in \mathbb{R}^{i \times j \times k}$. It is not obvious why this should help dramatically, but it is likely because of the nonnegativity constraints being difficult to solve in general due to there being so many of them. Since worst case outcomes are likely going to be positive, this makes these set of constraints a lot easier to satisfy. This second method is provided in in `HW4_facilityLocation_Adaptive.jl`, and uses the robust counterpart directly in JuMP. The result is provided in Figure 6.

3.7 Extra credit (15 points)

Thinking about the relationship between $\sum_{\forall i,j} P_{i,j}$, Γ and total cost, it is clear that the possible uncertain scenarios grow as the radius of influence increases. This means that we are not doing an apples-to-apples comparison of our locations as we change R_D (and thus P) at the same safety factors ρ and Γ. Devise and present a method to scale P (and optionally ρ and Γ) so that (1) your scaled $\{\bar{P}, \ \bar{\rho}, \ \bar{\Gamma} : R_D = 0.25\}$ is the same as the nominal case, (2) the inverse exponential properties in P with pairwise distances are preserved, and (3) the comparison is more fair.

Assess your new results by solving the fixed recourse problem for $R_D \in \{\epsilon : 0.05 : 0.6\}$ and providing the same tabulated data as in 3.3. Provide the unique facility configurations as well. Does your facility allocation change in this new comparison? Does the behavior of allocations with respect to R_D fit your intuition?

Solution:

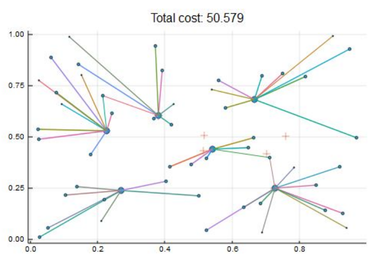
There are many approaches to this problem. The following is one such approach. In `HW4_facilityLocation_Solution.jl`, we make sure that the column-wise sum of values of P (i.e. total demand perturbation over all nodes due to each z_i) is constant. This essentially diffuses the influence of z_i across multiple nodes, giving it an area of effect whose radius is defined by R_D, without increasing the actual magnitude of the adverse effect. Doing this results in the following results:

Row	R_D	sumP	Gamma	f+c	f	c	nP	nx
	Any	Any	Any	Any	Any	Any	Any	Any
1	1.0e-5	43.9281	5	52.525	34.3567	18.1683	50	6.0
2	0.05	43.9281	5	52.525	34.3567	18.1683	54	6.0
3	0.1	43.9281	5	52.525	34.3567	18.1683	98	6.0
4	0.15	43.9281	5	52.423	34.3567	18.0663	170	6.0
5	0.2	43.9281	5	51.6044	34.3567	17.2476	268	6.0
6	0.25	43.9281	5	50.579	34.3567	16.2223	368	6.0
7	0.3	43.9281	5	44.8358	28.5491	16.2867	538	5.0
8	0.35	43.9281	5	43.8225	28.5491	15.2735	688	5.0
9	0.4	43.9281	5	43.1721	28.5491	14.623	830	5.0
10	0.45	43.9281	5	38.5213	23.3234	15.1979	998	4.0
11	0.5	43.9281	5	37.6706	22.6604	15.0103	1182	4.0
12	0.55	43.9281	5	37.2643	22.6604	14.6039	1332	4.0
13	0.6	43.9281	5	36.9052	22.6604	14.2449	1500	4.0
14	0.65	43.9281	5	36.6089	22.6604	13.9486	1666	4.0
15	0.7	43.9281	5	36.3782	22.6604	13.7178	1792	4.0
16	0.75	43.9281	5	36.1622	22.6604	13.5018	1942	4.0

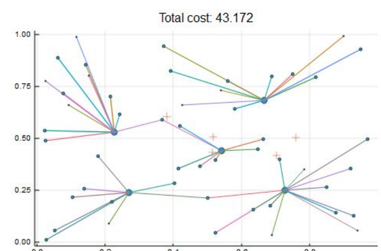
Figure 7: Results for a "fairly-scaled" P-matrix.

Interestingly enough, now the trend is reversed, meaning that a larger area of influence actually results in fewer facilities required for the robust case. The unique facility configurations are in Figure 8.

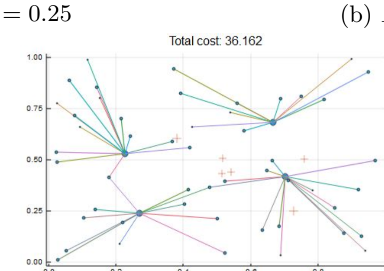
These results also make sense intuitively, since if we keep the 1-norm magnitude of the perturbations the same but instead diffuse them across many nodes, it is possible that fewer of the capacity constraints are hit when there is a demand perturbation.



(a) $R_D = 0.25$



(b) $R_D = 0.40$



(c) $R_D = 0.75$

Figure 8: As the radius of influence grows, fewer facilities can handle the diffused, but equal magnitude worst-case demand spikes.