15.094, Problem Set 3 solutions

Due: 18 March 2015 at 9am

Problem 1 - The Price of Robustness (40 points)

We consider an instance of the facility location problem which consists of F candidate facilities (potential sites where a facility can be opened) and C demand points that must be serviced (possibly from a combination of facilities). Opening facility $f \in \mathcal{F} := \{1, \ldots, F\}$ incurs a cost c_f, while servicing all the demand of customer $c \in \mathcal{C} := \{1, \ldots, C\}$ from facility f incurs a cost d_{fc} (delivery cost). We assume that the cost of opening a facility is precisely known (perfect information). The servicing costs are uncertain, but it is known that d_{fc}, $f \in \mathcal{F}$, $c \in \mathcal{C}$ are independent, symmetric and bounded random variables with support $\mathcal{S}_{fc} := \left[\overline{d}_{fc}(1-\rho), \overline{d}_{fc}(1+\rho)\right]$, where $\rho > 0$ and $\overline{d}_{fc}$ denotes the nominal value of d_{fc}.

We wish to be immunized against variations in the servicing costs when at most $\Gamma \in \{0, \ldots, FC\}$ of these costs can deviate from their nominal values.

(a) (10 points) Formulate the robust facility location problem that minimizes costs in the worst-case realization of the uncertain parameters.

(b) (10 points) Reformulate the robust facility location problem as a deterministic optimization problem.

(c) (20 points) Let $\rho = 5\%$. Using the data in the companion Excel spreadsheet, investigate the price of robustness in this problem. That is, calculate the worst-case cost incurred and the associated bound on the violation probability in dependence of Γ. Plot the trade-off curves.

Solution:

(a) The robust facility location problem is

$$
\begin{aligned}
\min_{\mathbf{x},\mathbf{y}} \quad & \left(\sum_{f=1}^{F} c_f y_f + \sum_{f=1}^{F}\sum_{c=1}^{C} \bar{d}_{fc} x_{fc} + \max_{|S| \le \Gamma} \sum_{fc \in S} \rho \bar{d}_{fc} x_{fc} \right) \\
\text{s.t.} \quad & \sum_{f=1}^{F} x_{fc} = 1 \quad c = 1, \ldots, C \\
& x_{fc} \le y_f \quad f = 1, \ldots, F, \quad c = 1, \ldots, C \\
& 0 \le x_{fc} \le 1, \quad y_f \in \{0, 1\} \qquad f = 1, \ldots, F, \quad c = 1, \ldots, C.
\end{aligned}
$$

(b) The robust counterpart of the facility location problem is

$$
\begin{aligned}
\min_{\mathbf{x},\mathbf{y}} \quad & \sum_{f=1}^{F} c_f y_f + \sum_{f=1}^{F}\sum_{c=1}^{C} \bar{d}_{fc} x_{fc} + z\Gamma + \sum_{f=1}^{F}\sum_{c=1}^{C} p_{fc} \\
\text{s.t.} \quad & \sum_{f=1}^{F} x_{fc} = 1 \quad c = 1,\ldots,C \\
& x_{fc} \le y_f \quad f = 1,\ldots,F, \quad c = 1,\ldots,C \\
& z + p_{fc} \ge \rho \bar{d}_{fc} x_{fc} \quad f = 1,\ldots,F, \quad c = 1,\ldots,C \\
& z \ge 0, \\
& 0 \le x_{fc} \le 1, \quad y_f \in \{0,1\}, \quad p_{fc} \ge 0, \qquad f = 1,\ldots,F, \quad c = 1,\ldots,C.
\end{aligned}
$$

(c) We use the combinatorial probability bound shown in Lecture 5. Example code for solving the facility location problem in Julia is provided below:

```
using JuMP, Gurobi

function FacLoc(Gamma)

  m = Model(solver = GurobiSolver())

  ## Parameters
  rho = 0.05
  cost_fac = [20 10 10 15]
  cost_del = [30 35 30 35;
  40 40 40 30;
  35 40 35 40;
  30 35 35 30;
  40 45 40 30;
  30 35 35 40;
  40 25 30 30;
  30 35 35 30;
  35 25 35 30;
  35 35 50 35;
  30 35 40 40;
  35 40 45 40]'

  ## Build model
  F = size(cost_del, 1)
  C = size(cost_del, 2)

  @defVar(m, 0 <= x[1:F,1:C] <= 1)
  @defVar(m, y[1:F], Bin)
  @defVar(m, z >= 0)
  @defVar(m, p[1:F,1:C] >= 0)

  @setObjective(m, Min, sum{ cost_fac[i]*y[i], i=1:F} + sum{
      cost_del[i,j]*x[i,j] + p[i,j], i=1:F, j=1:C} + Gamma*z)

  for j=1:C
    @addConstraint(m, sum{ x[i,j], i=1:F} == 1) # Customers must be serviced
    for i=1:F
      @addConstraint(m, x[i,j] <= y[i]) # Customer can only be serviced by
          selected facilities
```

```
    @addConstraint(m, z + p[i,j] >= rho*cost_del[i,j]*x[i,j]) # Robustness
        constraint
  end
end

solve(m)

## Output variables
#  xvals = zeros(F,C)
#  yvals = zeros(F,1)
#  for i=1:F
#    yvals[i] = getValue(y[i])
#    for j=1:C
#      xvals[i,j] = getValue(x[i,j])
#    end
#  end
#  println(yvals[:])
#  println(xvals)
return getObjectiveValue(m)

end
```

The trade-off plots are provided in Figure 1:

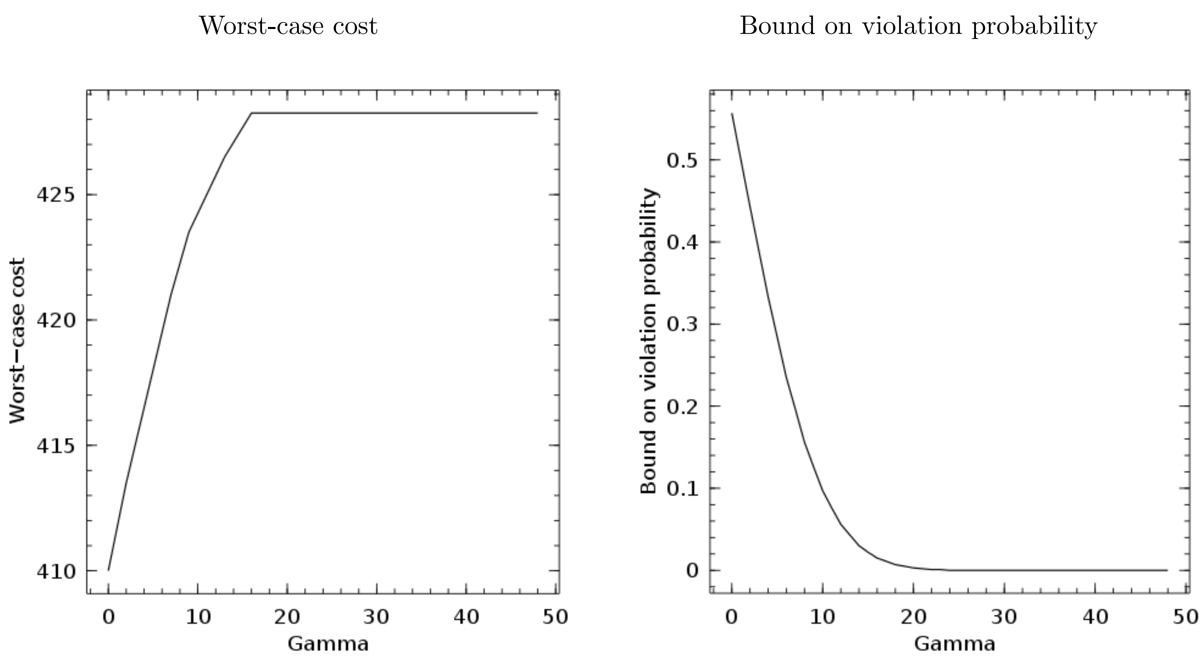

Figure 1: Companion figures for the solution to Problem 1

Problem 2 - Probabilistic Guarantees (T/F) (35 points)

For each of the following statements, indicate if the statement is true or false. Provide also a brief justification, sketch of a proof, or counterexample.

(a) Assume throughout that $\tilde{\mathbf{u}} \sim \mathbb{P}^*$ is a random variable coming from some distribution $\mathbb{P}^*$ which we may not know.

i. (5 points) Suppose $\mathcal{U}_1$ implies a probabilistic guarantee at level ϵ_1 and $\mathcal{U}_2$ implies a probabilistic guarantee at level ϵ_2. Furthermore, suppose $\epsilon_1 < \epsilon_2$. Then for any set $\mathcal{X}$, we have that

$$\min_{\mathbf{x}\in\mathcal{X}} \max_{\mathbf{u}\in\mathcal{U}_1} \mathbf{u}^T\mathbf{x} \geq \min_{\mathbf{x}\in\mathcal{X}} \max_{\mathbf{u}\in\mathcal{U}_2} \mathbf{u}^T\mathbf{x}$$

ii. (5 points) Recall the data-driven uncertainty set described in Lecture 7 using a hypothesis test for the mean:

$$\mathcal{U} = \left\{ \mathbf{u} \in \mathbb{R}^d : (\mathbf{u} - \hat{\boldsymbol{\mu}})^T \boldsymbol{\Sigma}^{-1} (\mathbf{u} - \hat{\boldsymbol{\mu}}) \leq \left(\Gamma + \sqrt{1/\epsilon - 1}\right)^2 \right\}$$

Here $\hat{\boldsymbol{\mu}}$ is the sample mean and $\Gamma \equiv \frac{R^2}{N}\left(2 + \sqrt{2\log(1/\delta)}\right)$. Let $\mathcal{U}_1$ be the result of applying this construction to this data with parameter $\delta = \delta_1$, and let $\mathcal{U}_2$ be the result from applying this construction with parameter $\delta = \delta_2$. Assume $\delta_1 < \delta_2$. Then for any set $\mathcal{X}$, we have that

$$\min_{\mathbf{x}\in\mathcal{X}} \max_{\mathbf{u}\in\mathcal{U}_1} \mathbf{u}^T\mathbf{x} \geq \min_{\mathbf{x}\in\mathcal{X}} \max_{\mathbf{u}\in\mathcal{U}_2} \mathbf{u}^T\mathbf{x}$$

(b) For the remaining parts, consider the following two stage adaptive linear optimization problem. Again assume that $\tilde{\mathbf{b}} \sim \mathbb{P}^*$ is a random variable coming from some distribution we may not know.

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y}(\cdot)} \quad & \mathbf{c}^T\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y}(\mathbf{b}) \geq \mathbf{b} \quad \forall \mathbf{b} \in \mathcal{U} \end{aligned} \tag{1}$$

Assume that $\mathcal{U} \subseteq \mathbb{R}^m$.

i. (5 points) Consider the robust approximation to this problem, i.e. where we impose that $\mathbf{y}(\cdot) \equiv \mathbf{y}_0 \in \mathbb{R}^{n_2}$. Suppose that $\mathcal{U}$ implies a probabilistic guarantee at level ϵ. Then, if $(\mathbf{x}, \mathbf{y}_0)$ are robust feasible, they are feasible with probability at least $1 - m\epsilon$ to problem (1).

ii. (5 points) Consider the affine approximation to this problem, i.e. where we impose that $\mathbf{y}(\cdot) \equiv \mathbf{F}\mathbf{b} + \mathbf{y}_0$ for some matrix $\mathbf{F} \in \mathbb{R}^{n_2 \times m}$ and vector $\mathbf{y}_0 \in \mathbb{R}^m$. Then, if $(\mathbf{x}, \mathbf{y}(\cdot))$ are robust feasible, they are feasible with probability at least $1 - m\epsilon$ to problem (1).

iii. (5 points) Now consider the fully adaptive version of this problem where $\mathbf{y}(\cdot)$ is permitted to be any function of the data. Then, if $(\mathbf{x}, \mathbf{y}(\cdot))$ are robust feasible, they may not be feasible with probability at least $1 - m\epsilon$ to problem (1).

iv. (5 points) Finally, suppose $\mathbb{P}(\tilde{\mathbf{b}} \in \mathcal{U}) \geq 1 - \epsilon$. (Recall from the lecture this is a *stronger* requirement than implying a probabilistic guarantee). Then, if $(\mathbf{x}, \mathbf{y}(\cdot))$ are robust feasible, they will be feasible with probability at least $1 - \epsilon$ to problem (1).

v. (5 points) **Important:** What does this problem tell you about constructing uncertainty sets for multistage optimization problems?

Solution:

(a) i. FALSE. The fact that $\mathcal{U}_1$ and $\mathcal{U}_2$ imply a probabilistic guarantee at levels ϵ_1 and ϵ_2, respectively with $\epsilon_1 < \epsilon_2$ is not sufficient to infer the inequality.
As an example, suppose the objective function involves a single random parameter $\tilde{u}$ which is known to be distributed according to a standard normal. Define $\mathcal{U}_1 := [-1, 1]$ and $\mathcal{U}_2 := [0, +\infty]$. Then, $\mathcal{U}_1$ and $\mathcal{U}_2$ imply a probabilistic guarantee at levels $\epsilon_1 = \frac{1}{3}$ and $\epsilon_2 = \frac{1}{2}$, respectively. Suppose now that $\mathcal{X} := \{1\}$ (a singleton). Then,

$$\min_{x\in\mathcal{X}} \max_{u\in\mathcal{U}_1} ux = \max_{u\in[-1,1]} u = 1 \quad \text{and} \quad \min_{x\in\mathcal{X}} \sup_{u\in\mathcal{U}_2} ux = \sup_{u\in[0,\infty]} ux = \infty.$$

Thus,

$$\min_{x\in\mathcal{X}} \max_{u\in\mathcal{U}_1} ux < \min_{x\in\mathcal{X}} \sup_{u\in\mathcal{U}_2} ux,$$

despite the fact that $\epsilon_1 < \epsilon_2$.

ii. TRUE. It can be shown that the derivative of Γ with respect to δ is $-\frac{R^2}{\sqrt{2N}\delta\sqrt{\log(1/\delta)}} < 0$ thus Γ is a strictly decreasing function of δ. Since $\delta_1 < \delta_2$, it holds that $\Gamma_1 > \Gamma_2$ which in turn implies that $\mathcal{U}_2 \subset \mathcal{U}_1$. This naturally implies that the optimal objective value of the optimization problem on the right is not greater that the optimal objective value of the problem on the left.

(b) i. TRUE. If we impose $\mathbf{y}(\cdot) \equiv \mathbf{y}$, problem (1) becomes

$$\begin{array}{ll} \min\limits_{\mathbf{x},\mathbf{y}} & \mathbf{c}'\mathbf{x} \\ \text{s.t.} & \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y} \geq \mathbf{b} \quad \forall \mathbf{b} \in \mathcal{U}. \end{array}$$

The ith constraint of this problem is affine in $\mathbf{b}$. Since $\mathcal{U}$ implies a probabilistic guarantee at level ϵ, by definition, each constraint is violated with probability no greater than ϵ. Using the Bonferroni inequality, this implies that the probability that any one of the m constraints is violated is no greater than $m\epsilon$. Thus, the probability that none of the constraints are violated is $1 - m\epsilon$.

ii. TRUE. The same argument as above applies in this case also since the problem has fixed recourse.

iii. TRUE. As an example, suppose that problem (1) involves a single uncertain parameter $\tilde{b}$ ($m = 1$) known to be uniformly distributed ($\mathbb{P}^* = U[0, 1]$) and presents a single uncertain inequality constraint given by

$$x + y(b) \geq b \quad \forall b \in \mathcal{U}.$$

Define $\mathcal{U} := [\frac{1}{4}, \frac{3}{4}]$ and let $\epsilon = \frac{1}{4}$.

A. We first show that $\mathcal{U}$ implies a probabilistic guarantee with level ϵ.
Fix $x,\ d \in \mathbb{R}$. If $ux \leq d \quad \forall u \in \mathcal{U}$, then:
- If $x = 0$: We must have $d \geq 0$, whence $\mathbb{P}(\tilde{u}x > d) = 0 \leq \epsilon$.
- If $x > 0$: Taking $u = \frac{3}{4}$, we must have $\frac{d}{x} \geq \frac{3}{4}$, whence $\mathbb{P}^*(\tilde{u}x > d) = 1 - \frac{d}{x} \leq \epsilon$.
- If $x < 0$: Taking $u = \frac{1}{4}$, we must have $\frac{d}{x} \leq \frac{1}{4}$, whence $\mathbb{P}^*(\tilde{u}x > d) = \frac{d}{x} \leq \epsilon$.

So for any x and d, the implication

$$ux \leq d \quad \forall u \in \mathcal{U} \quad \Rightarrow \quad \mathbb{P}^*(\tilde{u}x > d) \leq \epsilon$$

holds, i.e., $\mathcal{U}$ implies a probabilistic guarantee at level ϵ.

B. We construct a robust feasible solution to this problem. Let $x^\star = 0$ and

$$y^\star(b) = \begin{cases} b & \text{if } b \in \mathcal{U} \\ b - 1 & \text{if } b \notin \mathcal{U}. \end{cases}$$

Then, the pair $(x^\star, y^\star(\cdot))$ is a robust feasible solution to our problem. Indeed, $x^\star + y^\star(b) = b \quad \forall b \in \mathcal{U}$.

C. We now show that despite the fact that $(x^\star, y^\star(\cdot))$ is robustly feasible, it is feasible with probability less than $1-\epsilon$ (infeasible with probability greater than ϵ). First, note that $x^\star + y^\star(b) = b - 1 < b$ for all $b \notin \mathcal{U}$. Thus, $(x^\star, y^\star(\cdot))$ satisfies $x^\star + y^\star(b) \geq b$ if and only if $b \in \mathcal{U}$. Therefore,

$$\mathbb{P}^*((x^\star, y^\star(\cdot)) \text{ is feasible}) = \mathbb{P}^*(\tilde{b} \in \mathcal{U}) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2} < \frac{3}{4} = 1 - \epsilon.$$

iv. TRUE. $(\mathbf{x}, \mathbf{y}(\cdot))$ are feasible to problem (1) with probability at least $\mathbb{P}(\tilde{\mathbf{b}} \in \mathcal{U})$, since they are robust feasible. (Note that $1 - \epsilon > 1 - m\epsilon$: this holds independent of the number of constraints.)

v. The bound $1 - m\epsilon$ in parts (a) and (b) is typically very loose as the union bound is loose (note that it is possible to construct examples where it is tight). (Side note 1: This is not hard to do... You should try it!) (Side note 2: A possible research question is: "Are there simple conditions that can be validated with data to ensure that a tighter bound than $1 - m\epsilon$ holds in parts (a), (b)?")

Consequently, if it is very important to guarantee theoretically that your solution be feasible with a specified probability, say 90%, you have two choices:

A. Pick an uncertainty set $\mathcal{U}_1$ which implies a guarantee at level $1 - 10\%/m$ and then use a static (robust) or affine policy.

B. Pick an uncertainty set $\mathcal{U}_2$ such that $\mathbb{P}^\star(\tilde{\mathbf{b}} \in \mathcal{U}_2) \geq 90\%$. Then use any policy you want.

In general, if m is large, $\mathcal{U}_1$ will probably be very big, and the solution obtained from option i. might be overly conservative and perform poorly in practice. In this case, option ii. might be a better choice. However, for smaller m, $\mathcal{U}_2$ might be much larger than $\mathcal{U}_1$ (we saw this in lecture for the case $m = 1$) and affine policies are frequently close to optimal for many kinds of problems (we also saw this in lecture). Thus, in this case, option ii. might be a better choice.

Of course, if a provable guarantee isn't as important to you in a particular application, out-of-sample testing, tuning, and using application specific knowledge can often do better than either options i. or ii.

Problem 3 - A two stage adaptive RO problem (25 points)

Consider the two-stage adaptive robust problem

$$\begin{aligned} \min_{u_1 \in \mathbb{R}} \quad & c u_1 + \max_{\substack{w_1 \in \mathbb{R}: \\ -1 \le w_1 \le 1 \\ L \le u_1 + w_1 \le U}} h(x_1 + u_1 + w_1) \\ \text{s.t.} \quad & L \le u_1 \le U, \end{aligned} \tag{2}$$

where $c, x_1 \in \mathbb{R}$ are fixed.

(a) (10 points) Suppose $c > 0$, $0 < L < 1 < L + 1 < U$, and $h(y) = y$. Derive an expression for the optimal cost of (2). *Hint:* it should an affine function of c.

(b) (5 points) What would change if you repeated (a) for $c \le 0$? (No need to do the full analysis, mention which cases (if any) you would consider.)

(c) (10 points) Outline your approach for an arbitrary convex function $h(.)$. (Write down the optimal w_1 for the inner problem, the final minimization problem, and the optimal u_1 for the outer problem. Is this outer problem still convex?)

Solution:

(a) The inner problem becomes

$$\begin{aligned} \max_{-1 \le w \le 1} \quad & u_1 + w_1 + x_1 \\ \text{s.t.} \quad & L - u_1 \le w_1 \le U - u_1 \end{aligned}$$

It is easy to see that the optimal w_1^* in this case is given by

$$w_1^*(u_1) = \max\left\{\min\left\{1, U - u_1\right\}, -1\right\} \tag{3}$$

As $U - u_1 \ge 0 > -1$, we have

$$w_1^*(u_1) = \min\left\{1, U - u_1\right\} \tag{4}$$

The outer problem now becomes

$$\begin{aligned} \min_{u_1} \quad & c u_1 + x_1 + u_1 + \min\left\{1, U - u_1\right\} \\ \text{s.t.} \quad & L \le u_1 \le U \end{aligned}$$

We split this into 2 cases : $1 \le U - u_1$ and $U - u_1 \le 1$. Now, we write this as the minimum of these 2 problems:

$$\begin{aligned} \min_{u_1} \quad & (c+1) u_1 + x_1 + 1 \\ \text{s.t.} \quad & L \le u_1 \le U - 1 \end{aligned}$$

and

$$\begin{aligned} \min_{u_1} \quad & c u_1 + x_1 + U \\ \text{s.t.} \quad & U - 1 \le u_1 \le U \end{aligned}$$

As $c > 0$, the optimal objective is:

$$z^* = x_1 + \min\{(c+1)L + 1, c(U-1) + U\} \tag{5}$$

Comparing the two terms in the second term above, after some algebra, and using $L < U + 1$, we see that the optimal cost is

$$z^* = L(c+1) + 1 + x_1$$

(b) Need to split this into 2 further cases : $c \leq -1$, and $-1 < c \leq 0$.

(c) For any arbitrary convex h, the optimal objective to the inner problem can be written as

$$z^*_{\text{Inner}}(u_1) = \arg\max\left\{h^1(u_1), h^2(u_1)\right\} \tag{6}$$

where

$$\begin{aligned} h^1(u_1) &= h(u_1 + x_1 + \min\{1, U - u_1\}) \\ h^2(u_1) &= h(u_1 + x_1 + \max\{-1, L - u_1\}) \end{aligned}$$

and $w_1^*(u_1)$ is

$$w_1^*(u_1) = \begin{cases} \min\{1, U - u_1\} & \text{if } h^1 > h^2 \\ \max\{-1, L - u_1\} & \text{else .} \end{cases} \tag{7}$$

This is due to the fact that the optimum will be at an extreme point, as we are maximizing a convex function in this case.
The outer problem now becomes

$$\begin{aligned} \min_{u_1} \quad & cu_1 + z^*_{\text{Inner}}(u_1) \\ \text{s.t.} \quad & L \leq u_1 \leq U \end{aligned} \tag{8}$$

where

$$z^*_{\text{Inner}}(u_1) = \max\{h(x_1 + u_1 + \min\{1, U - u_1\}), h(x_1 + u_1 + \max\{-1, L - u_1\})\}$$

The outer problem is not convex. For instance, when $h(y) = y$, we can see that the problem is concave.

Problem 4 (20 points, **OPTIONAL EXTRA CREDIT**)

As we have seen, the primary constraint encountered in two-stage adaptive optimization is something of the form

$$\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y}(\boldsymbol{\xi}) \le \mathbf{d} \; \forall \boldsymbol{\xi} \in \mathcal{U},$$

where $\mathbf{A}, \mathbf{B}$, and $\mathbf{d}$ are all known. In reality, $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{d}$ are often estimated. If there are uncertainty in $\mathbf{A}$ and $\mathbf{d}$, we can handle this in the usual robust optimization framework. The focus of this question is what if $\mathbf{B}$ is uncertain? For simplicity we will only consider a basic model of uncertainty and only look at a single constraint. Fix $\mathbf{a}, d$ of appropriate dimensions, and consider the constraint

$$\mathbf{a}'\mathbf{x} + \mathbf{b}'\mathbf{y}(\boldsymbol{\xi}) \le d \; \forall \boldsymbol{\xi} \in \mathcal{U}, \mathbf{b} \in \mathcal{V}, \tag{9}$$

where $\mathbf{y}(\boldsymbol{\xi}) = \bar{\mathbf{y}} + \mathbf{E}\boldsymbol{\xi}$, and $\bar{\mathbf{y}}, \mathbf{E}$ are decision variables in the outer problem.

(a) Propose a general solution technique for *tractably* solving such a problem with a constraint as given in (9) under general (convex) sets $\mathcal{U}$ and $\mathcal{V}$. You should formally prove any tractability claims you make. Can you even make such a claim for well-structured (e.g. polyhedral) sets $\mathcal{U}$ and $\mathcal{V}$?

(b) The type of uncertainty assumed aboved requires (essentially) independence in uncertainty between $\mathbf{b}$ and $\mathbf{y}$. This is likely unrealistic. Let us consider instead the constraint

$$\mathbf{a}'\mathbf{x} + \mathbf{b}(\boldsymbol{\xi})'\mathbf{y}(\boldsymbol{\xi}) \le d \; \forall \boldsymbol{\xi} \in \mathcal{U},$$

where now we take for example $\mathbf{b}(\boldsymbol{\xi}) = \bar{\mathbf{b}} + \mathbf{D}\boldsymbol{\xi}$, where $\bar{\mathbf{b}}$ and $\mathbf{D}$ are fixed. How do you solve such a problem?

Solution: Left to your own imagination.