**ORI 397 Optimization Under Uncertainty Spring 2020**

# Assignment 3

Due: Thursday, February 27

**Problem 1.** Consider the following linear optimization problem with an expectation constraint:

$$\begin{array}{ll} \inf\limits_{\boldsymbol{x}\in\mathcal{X}} & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X} \\ & \mathbb{E}[f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})] \le \delta. \end{array} \tag{1}$$

Here, $\mathcal{X}$ is a polytope and $\delta$ is a prescribed constant. How would you recommend solving the following instances of (1) *efficiently*? Be specific.

(a) Let $f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) = \sum_{n\in[N]} \tilde{\xi}_n x_n$. Assume that $\tilde{\xi}_n$, $n \in [N]$, are independent exponential random variables with mean values $\mu_1, \ldots, \mu_N$ and variances $\sigma_1^2, \ldots, \sigma_N^2$, respectively.

(b) Let $f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) = \|\tilde{\boldsymbol{\xi}} - \boldsymbol{x}\|_2^2$. Assume that $\tilde{\xi}_n$, $n \in [N]$, are independent normal random variables with mean values $\mu_1, \ldots, \mu_N$ and variances $\sigma_1^2, \ldots, \sigma_N^2$, respectively.

(c) Let $\delta \in (0, \frac{1}{2})$, and

$$f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) = \mathbb{I}\left\{ \sum_{n\in[N]} \tilde{\xi}_n x_n > t \right\}.$$

Here, $t$ is a fixed target cost we do not want to exceed, and $\mathbb{I}\{\cdot\}$ is an indicator function that takes the value one if its argument is true and zero otherwise. Assume that $\tilde{\xi}_n$, $n \in [N]$, are independent normal random variables with mean values $\mu_1, \ldots, \mu_N$ and variances $\sigma_1^2, \ldots, \sigma_N^2$, respectively.

**Problem 2.** A risk measure $\mathcal{R}$ is said to be coherent if it satisfies the following properties:

- **Monotonicity:** If for some decisions $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ we have $\mathbb{P}\left(\ell(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \le \ell(\boldsymbol{y}, \tilde{\boldsymbol{\xi}})\right) = 1$ then $\mathcal{R}[\ell(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})] \le \mathcal{R}[\ell(\boldsymbol{y}, \tilde{\boldsymbol{\xi}})]$.
- **Positive homogeneity:** For any $\lambda \in \mathbb{R}_+$, we have $\mathcal{R}[\lambda \ell(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})] = \lambda \mathcal{R}[\ell(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})]$ for all $\boldsymbol{x} \in \mathcal{X}$.
- **Subadditivity:** We have $\mathcal{R}[\ell(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) + \ell(\boldsymbol{y}, \tilde{\boldsymbol{\xi}})] \le \mathcal{R}[\ell(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})] + \mathcal{R}[\ell(\boldsymbol{y}, \tilde{\boldsymbol{\xi}})]$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$.
- **Translation invariance:** For any $c \in \mathbb{R}$, we have $\mathcal{R}[\ell(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) + c] = \mathcal{R}[\ell(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})] + c$ for all $\boldsymbol{x} \in \mathcal{X}$.

The class of coherent risk measures has become very attractive since the financial crisis of 2008 as it was observed that the value-at-risk (VaR) fails to satisfy the subadditivity property (thus, the use of VaR as a risk measure may discourage a diversified portfolio). Show via a counterexample that VaR is indeed not coherent as it fails to satisfy the subadditivity property above.

An alternative way to quantify risks is by using the conditional value-at-risk (CVaR). Show that CVaR is a coherent risk measure as it satisfies all of the above properties.

**Problem 3** (Digital Communication)**.** A signal $\boldsymbol{s} = [0\ 0]^\top \in \mathbb{R}^2$ is transmitted over a noisy communication channel and is perturbed by an additive random noise $\tilde{\boldsymbol{\xi}} \in \mathbb{R}^2$ governed by an unknown probability distribution $\mathbb{P}$. The noise $\tilde{\boldsymbol{\xi}}$ is known to obey the mean values

$$\mathbb{E}_{\mathbb{P}}[\tilde{\xi}_1] = 0,\ \mathbb{E}_{\mathbb{P}}[\tilde{\xi}_2] = 0$$

and the absolute deviations

$$\mathbb{E}_{\mathbb{P}}[|\tilde{\xi}_1|] \le 0.1,\ \mathbb{E}_{\mathbb{P}}[|\tilde{\xi}_2|] \le 0.1.$$

The signal $\boldsymbol{s}$ can be recovered exactly if its perturbed version $\boldsymbol{s} + \tilde{\boldsymbol{\xi}}$ resides within the 1-norm ball $\{\boldsymbol{z} \in \mathbb{R}^2 : \|\boldsymbol{z}\|_1 \le 1\}$. Formulate a linear program that computes the worst-case probability of correct detection

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left(\|\underbrace{\boldsymbol{s}}_{\binom{0}{0}} + \tilde{\boldsymbol{\xi}}\|_1 \le 1\right), \quad (2)$$

$$= \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}(\|\tilde{\xi}\| \le 1) = \inf \int \mathbb{I}_{[|\xi_1|+|\xi_2|\le 1]}\, \nu(d\xi)$$
$$\text{s.t. } \nu(\cdot) \ge 0,\quad \int \nu(d\xi) = 1,\quad \int \xi\, \nu(d\xi) = 0,\quad \int |\xi|\, \nu(d\xi) \le 0.1e$$

where

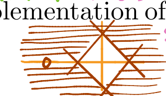
$$\mathcal{P} = \left\{ \mathbb{P} \in \mathcal{P}_0(\mathbb{R}^2) : \begin{array}{l} \mathbb{E}_{\mathbb{P}}[\tilde{\xi}_1] = 0,\ \mathbb{E}_{\mathbb{P}}[\tilde{\xi}_2] = 0, \\ \mathbb{E}_{\mathbb{P}}[|\tilde{\xi}_1|] \le 0.1,\ \mathbb{E}_{\mathbb{P}}[|\tilde{\xi}_2|] \le 0.1 \end{array} \right\}.$$

show the deviations

$$= \sup\ \alpha - 0.1\gamma^\top e$$
$$\text{s.t. } \alpha \in \mathbb{R},\ \beta \in \mathbb{R}^2,\ \gamma \in \mathbb{R}^2_+$$
$$\mathbb{I}_{[|\xi_1|+|\xi_2|\le 1]} \ge \alpha + \beta^\top\xi - \gamma^\top|\xi| \quad \forall \xi \in \mathbb{R}^2$$

$$1 \ge \alpha + \beta^\top\xi - \gamma^\top|\xi|,\ \forall \xi \in \mathbb{R}^k \Leftrightarrow 1 \ge \sup_{\xi\in\mathbb{R}^2} \alpha + \beta^\top\xi - \gamma^\top|\xi|$$ ← can be formulated as LP

$$0 \ge \alpha + \beta^\top\xi - \gamma^\top|\xi|,\ \forall \xi \in \mathbb{R}^2 : |\xi_1|+|\xi_2| \ge 1 \Leftrightarrow 0 \ge \sup\ \alpha + \beta^\top\xi - \gamma^\top|\xi| \ \text{ s.t. } \xi\in\mathbb{R}^2,\ |\xi_1|+|\xi_2|\ge 1$$ ← 4 constraints where RHS involves LP




$$\begin{cases} 0 \ge \sup \cdots \\ 0 \ge \sup \cdots \\ 0 \ge \sup \cdots \\ 0 \ge \sup \cdots \end{cases}$$
…为LP

Implement the linear program in a MATLAB function `wcprob.m` defined as follows.

```
[prob]=wcprob()
```

Here, the function `wcprob.m` outputs the worst-case probability (2). Submit your implementation of `wcprob.m` to **Canvas**.

**Problem 4.** You are given a data set containing measurements for 40 flowers from two different species. The data set is available in the file `data.mat` on **Canvas**. Here, the vectors $\{\boldsymbol{y}_i\}_i$ are the data points. The components of $\boldsymbol{y}_i$ describe the following measurements for the $i$th flower: the sepal length $y_{i1}$ and the petal length $y_{i2}$. Flowers in the data set have been labeled as belonging to either one of the two species. The components of vector $\boldsymbol{z} \in \{-1, 1\}^{40}$ are the labels. Flower $i$ belongs to species 1 if $z_i = 1$ and to species 2 if $z_i = -1$.

In this problem, we aim to construct a linear classifier that will be able to classify a new data point accurately into one of the two species. Specifically, we seek for parameters $\boldsymbol{w} = (w_1, w_2)$ and $b$ so that a data point corresponding to one species is contained in the half-space $\{(x_1, x_2) \in \mathbb{R}^2 : w_1x_1 + w_2x_2 - b \ge 1\}$, while that corresponding to the other species is contained in the half-space $\{(x_1, x_2) \in \mathbb{R}^2 : w_1x_1 + w_2x_2 - b \le -1\}$.

(a) **(Chance Constrained Programming)** We assume that the pair $(\tilde{\boldsymbol{y}}, \tilde{z})$ of data point and its label is random and governed by the empirical distribution

$$\mathbb{P}\Big((\tilde{\boldsymbol{y}}, \tilde{z}) = (\boldsymbol{y}_i, z_i)\Big) = \frac{1}{40} \qquad \forall i = 1, \ldots, 40.$$

Formulate a *chance constrained program* whose optimal solution is the parameters $\boldsymbol{w} = (w_1, w_2)$ and $b$ of the linear classifier that satisfy $\|(\boldsymbol{w}, b)\|_\infty \le 100$ and that with probability at least $1 - \epsilon$ correctly classifies the random data point $\tilde{\boldsymbol{y}}$. Formulate a mixed-integer linear program for the problem, and implement it in a MATLAB function `ccp.m` defined as follows.

```
[w,b]=ccp(epsilon)
```

Here, the function `ccp.m` takes as input the tolerance level $\epsilon$, and outputs the coefficients $\boldsymbol{w}$ and $b$ of the linear classifier. Submit your implementation of `ccp.m` to **Canvas**. Report and describe the optimal solution for $\epsilon = 0.1$. Provide a plot of the data points and the optimal classifier (hyperplane).

*Hint: This is a feasibility problem; there is no objective function.*

(b) **(CVaR Approximation)** The chance constrained program is generically intractable and does not scale well with large input data sizes. Formulate a tractable conservative approximation using the conditional value-at-risk. Implement this formulation in MATLAB (you don't have to submit the code). Report and describe the optimal solution for $\epsilon = 0.1$.

$\inf 0$

s.t. $w \in \mathbb{R}^2,\ b \in \mathbb{R}$

$\|(w, b)\|_\infty \leq 100$

$\mathbb{P}(\tilde{z}(w^T \tilde{y} - b) \geq 1) \geq 1 - \epsilon$ ← (b) use definition we discussed in class reformula as a VaR constraint. replace VaR with CVaR and get the answer.

$\Leftrightarrow \frac{1}{N} \sum_i \mathbb{I}[z_i(w^T y_i - b) \geq 1] \geq 1 - \epsilon$

$g \in \{0, 1\}^N$

$g_i = 1$ iff $z_i(w^T y_i - b) \geq 1$ ← big-M formulation

$\frac{1}{N} e^T g \geq 1 - \epsilon$