15.094, Problem Set 2

Due: 4 March 2015 at 9am EST

Problem 1 - Data-driven RO (30 points)

You are given the set $\mathcal{S} := \{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_T\}$ which consists of T historical samples for $\mathbf{r} \in \mathbb{R}^k$, the (random) vector of returns on k stocks.

(a) (15 points) Propose two data-driven uncertainty sets for $\mathbf{r}$ that imply a probabilistic guarantee for the true distribution of the returns at level ϵ with confidence δ. Clearly state any assumptions made on the support of the true distribution, the number of samples available, etc.

(b) (15 points) You wish to invest in a portfolio of these stocks, i.e., you wish to allocate your wealth among these assets. Formulate (for both uncertainty sets) the robust counterpart of the portfolio problem that maximizes the worst-case return on your investment.

Hint: See paper [22] from the syllabus.

Problem 2 - Robust 0-1 Optimization (30 points)

Consider the robust combinatorial optimization problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}'\mathbf{x} + \max_{\substack{S,T: \\ S\subseteq N, |S|\leq \Gamma_1 \\ T \subseteq M, |T| \leq \Gamma_2}} \left(\sum_{j\in S} d_j x_j + \sum_{k \in T} f_k x_k \right) \\ \text{subject to} \quad & \mathbf{x} \in X \subseteq \{0,1\}^{2n} \end{aligned} \tag{1}$$

where $N = \{1, \ldots, n\}$ and $M = \{n+1, \ldots, 2n\}$. Assume that $d_1 \geq d_2 \geq \ldots \geq d_n \geq 0$ and $f_{n+1} \geq f_{n+2} \geq \ldots \geq f_{2n} \geq 0$, Γ_1, Γ_2 both positive integers, and X is a subset of $\{0,1\}^{2n}$.

Essentially, what we are modeling here is that at most Γ_1 of $\{c_1, \ldots, c_n\}$ and Γ_2 of $\{c_{n+1}, \ldots, c_{2n}\}$ can vary from their nominal values.

(a) (10 points) Using ideas from Lecture 5, write down the resulting robust counterpart of (1).

(b) (20 points) Suppose we have a specialized fast subroutine for solving problems of the form

$$\begin{aligned} \min_{\mathbf{x}} \quad & \bar{\mathbf{c}}'\mathbf{x} \\ \text{subject to} \quad & \mathbf{x} \in X \subseteq \{0,1\}^{2n} \end{aligned} \tag{2}$$

Propose an algorithm which solves problem (1) using the above subroutine.

Problem 3 - Convex duality (30 points)

(a) (10 points) Consider the RO problem

$$\begin{aligned} \max \ & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \ & \mathbf{a}'\mathbf{x} \leq \mathbf{b}, \ \forall \mathbf{a} \in \mathcal{U}, \end{aligned} \tag{3}$$

where

$$\mathcal{U} = \{\mathbf{a} \,|\, \mathbf{a} = \bar{\mathbf{a}} + \mathbf{\Delta u}, \|\mathbf{u}\| \leq 1\} \tag{4}$$

for a given matrix (of appropriate dimensions) $\mathbf{\Delta}$ and norm $\|\cdot\|$.

Write down (with proof) the robust counterpart of problem (3) when the norm used to define (4) is $\ell_1 \cap \ell_\infty$, defined by

$$\|\mathbf{u}\|_{1\cap\infty} = \max\left\{\frac{1}{\Gamma}\|\mathbf{u}\|_1, \|\mathbf{u}\|_\infty\right\}$$

for a fixed positive constant Γ.

(b) (20 points) Consider

$$\min_{\mathbf{x}\in\mathcal{X}} \ \max_{\xi\in\Xi} \ \xi'\mathbf{x}, \tag{5}$$

where $\Xi \subseteq \mathbb{R}^k$ denotes the uncertainty set, and $\mathcal{X} := \{\mathbf{x} \in \mathbb{R}^n \ : \ \mathbf{Ax} \leq \mathbf{b}\}$ with $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$. Reformulate (5) as a deterministic optimization problem with a *finite* number of constraints in the case when

$$\Xi := \left\{\xi \in \mathbb{R}^k \ : \ \exists \zeta \in \mathbb{R}^l \text{ with } \|\mathbf{P}\xi + \mathbf{Q}\zeta + \mathbf{h}\|_2 \leq \mathbf{p}'\xi + \mathbf{q}'\zeta + h\right\}.$$

where $\mathbf{P} \in \mathbb{R}^{m\times k}$, $\mathbf{Q} \in \mathbb{R}^{m\times l}$, $\mathbf{h} \in \mathbb{R}^m$, $\mathbf{p} \in \mathbb{R}^k$, $\mathbf{q} \in \mathbb{R}^l$ and $h \in \mathbb{R}$ are fixed.

Also, discuss the conditions required for this new problem you wrote down derived to be *equivalent* to (5).

Problem 4 - Using JuMPeR (10 points)

Consider the following robust optimization problem

$$\begin{aligned} \min_{x_1,x_2} \ & x_1 + x_2 \\ \text{s. t.} \ & x_1, x_2 \in \mathbb{R} \\ & x_1 \geq 0, x_2 \geq 0 \\ & a_1x_1 + a_2x_2 \geq 1 \quad \forall\, (a_1, a_2) \in \Xi, \end{aligned} \tag{6}$$

where $\Xi = \{(a_1, a_2) \in \mathbb{R}^2 : 0 \leq a_1, a_2 \leq 1, a_1 + a_2 \leq 1\}$. Solve (6) using JuMPeR. Provide us with your commented code and the optimal solution obtained.

N.B. See JuMPeR and Iain's talk in Recitation 2 (on Friday, February 20) for references.

Problem 5 (20 points, <u>**OPTIONAL EXTRA CREDIT**</u>)

Consider the robust optimization problem

$$\begin{aligned} \max_{\mathbf{x} \in \mathbb{R}^n} \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & \mathbf{a}'\mathbf{x} \leq b \quad \forall \mathbf{a} \in \mathcal{U}, \end{aligned} \tag{7}$$

where the uncertainty set is

$$\mathcal{U} = \left\{ \mathbf{a} \in \mathbb{R}^n \ : \ \exists \xi \in \mathbb{R}^k \text{ with } \mathbf{a} = \overline{\mathbf{a}} + \mathbf{A}\xi, \|\xi\|_p \leq \rho \right\}.$$

In class (Lecture 4) it was shown that if the ξ_i are independent random variables with support $[-1, 1]$, the optimal solution to (7) will satisfy $\mathbb{P}(\mathbf{a}'\mathbf{x} > b) \leq e^{\frac{-\rho^2}{2}}$ if $p = 2$. In other words, a high level of robustness can be achieved with (astonishingly) low values of ρ, when one considers an ℓ_2-norm based uncertainty set. In this problem, we would like to explore whether similar results can be obtained in the case of general norm uncertainty sets.

(a) (10 points) For $p = 2$, we have seen that ρ can be chosen independent of k (the dimension of ξ) to ensure that $\mathbb{P}(\mathbf{a}'\mathbf{x} > b) \leq \epsilon$. For general ℓ_p-norm, can ρ be chosen independent of k?

(b) (5 points) Does the bound improve if the distributions are all identical? Answer the question even for $p = 2$.

(c) (5 points) Does the bound improve if the distributions are symmetric around the mean? Answer the question even for $p = 2$.