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.

Solution:

There are a variety of different solutions. Any proposal with clearly stated assumptions that makes sense for modeling stock returns will receive full-credit. The following solutions are indicative.

(a) Possible sets:

i. $\mathcal{U}^M$: If we do not believe that we can learn the correlation structure among the $\mathbf{r}_t$ – e.g. if T is very small –then $\mathcal{U}^M$ is an appropriate choice. Assume only that there exist lower and upper bounds $\mathbf{r}^{(0)}, \mathbf{r}^{(T+1)}$ s.t. $\mathbf{r}^{(0)} \leq \mathbf{r}^{(t)} \leq \mathbf{r}^{(T+1)}$ for all $t = 1, \ldots, T$. Define s by

$$s = \min \left\{ l \in \mathbb{T} : \sum_{j=l}^{T} \binom{T}{j} (\epsilon/k)^{T-j} (1 - \epsilon/k)^j \leq \frac{\delta}{2k} \right\},$$

and let $s = T + 1$ if this set is empty. Notice we have adapted notation from the paper. Then let

$$\mathcal{U}^M = \left\{ \mathbf{r} \in \mathbb{R}^k : r_i^{(N-s+1)} \leq r_i \leq r_i^{(s)}, \quad i = 1, \ldots, k \right\},$$

where $r_i^{(j)}$ is the j-th largest value among components r_i. In practice, we would tune the value of δ, perhaps using cross-validation.

ii. $\mathcal{U}^{\chi^2}$: We assume that the distribution of future returns is discrete, and takes values on each of the $\mathbf{r}_t$.[1] Suppose there are J unique values among the $\mathbf{r}_t$, indexed by j. Let $\hat{p}_j$ be the empirical proportion of observations that have value $\mathbf{r}_j$. Take $\chi^2_{J-1,1-\delta}$ to be the $1-\delta$ quantile of a χ^2 distribution with $J-1$ degrees of freedom. Then

$$\mathcal{U}^{\chi^2} = \Bigg\{ \mathbf{r} \in \mathbb{R}^k : \mathbf{r} = \sum_{j=1}^{J} q_j \mathbf{r}_j \quad \mathbf{q} \geq \mathbf{0}, \quad \mathbf{e}'\mathbf{q} = 1, \quad \mathbf{q} \leq \frac{1}{\epsilon}\mathbf{p}, \mathbf{q} \geq \mathbf{0}, \mathbf{e}'\mathbf{p} = 1,$$
$$\sum_{j=1}^{J} t_j \leq \frac{\chi^2_{J-1,1-\delta}}{2T}, \left\| \begin{pmatrix} 2(p_j - \hat{p}_j) \\ p_j - 2t_j \end{pmatrix} \right\|_2 \leq p_j + 2t_j, \quad j = 1, \ldots, J \Bigg\}.$$

iii. $\mathcal{U}^{CS}$ Let $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ be the sample mean and sample covariance, respectively. We can take:

$$\mathcal{U}^{CS} = \left\{ \mathbf{r} \in \mathbb{R}^k : \|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\| \leq \Gamma_1, \|\boldsymbol{\Sigma} - \hat{\boldsymbol{\Sigma}}\| \leq \Gamma_2, \begin{pmatrix} \frac{1-\epsilon}{\epsilon} & (\mathbf{r} - \boldsymbol{\mu})' \\ \mathbf{r} - \boldsymbol{\mu} & \boldsymbol{\Sigma} \end{pmatrix} \succcurlyeq \mathbf{0} \right\}.$$

The parameters Γ_1 and Γ_2 should be chosen by bootstrapping.

(b) Reformulations: We consider the reformulations in context of the following problem:

$$\begin{aligned} \max \quad & z && \text{(1a)} \\ \text{s.t.} \quad & \mathbf{e}'\mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}, && \text{(1b)} \\ & \mathbf{r}'\mathbf{x} \geq z \quad \forall \mathbf{r} \in \mathcal{U} && \text{(1c)} \end{aligned}$$

Observe that for any set $\mathcal{U}$, the robust counterpart of equation (1c) is simply $\phi_{\mathcal{U}}(\mathbf{x}) \leq z$ for $\phi_{\mathcal{U}}$ the support function of $\mathcal{U}$.

i. $\mathcal{U}^M$: A simple way to do this is to use LP duality to form the robust counterpart directly. A more clever way is to use the formula for $\phi_{\mathcal{U}^M}$ from the paper and the observation above. In either case, we replace equation (1c) with

$$\begin{aligned} & \textstyle\sum_{i=1}^{k} t_i \leq z \\ & t_i \geq x_i r_i^{(s)} \\ & t_i \geq x_i r_i^{(N-s+1)} \end{aligned}$$

ii. $\mathcal{U}^{\chi^2}$: An explicit formula for $\phi_{\mathcal{U}^{\chi^2}}$ was not given in this revision of the paper, so we have to form the robust counterpart directly using Lagrange duality, see Helper Material 1.

iii. $\mathcal{U}^{CS}$: We can either form the robust counterpart directly using Lagrange duality, or, if you study the proof of Theorem 9, you will notice that $\phi_{\mathcal{U}^{CS}}$ is given by the right hand side of equation (32). In either case, we replace equation (1c) with:

$$\begin{aligned} & \hat{\boldsymbol{\mu}}^T \mathbf{x} + \Gamma_1 t_1 + \sqrt{\frac{1-\epsilon}{\epsilon}} t_2 \leq z \\ & \|\mathbf{x}\| \leq t_1 \\ & \mathbf{C}\mathbf{x} = \mathbf{y} \\ & \|\mathbf{y}\| \leq t_1 \end{aligned}$$

[1] This is a reasonably common assumption in financial engineering. A similar approach is to first cluster the returns into a smaller number of groups representing "market modes", like "bull-market", "bear-market", etc, and then work with the cluster centers.

where $\mathbf{C}'\mathbf{C} = \hat{\mathbf{\Sigma}} + \Gamma_2 \mathbf{I}$ is a Cholesky decomposition.

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

Consider the robust combinatorial optimization problem

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

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 (2).

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

$$\begin{aligned} \min_{\mathbf{x}} \quad & \bar{\mathbf{c}}'\mathbf{x} \\ \text{subject to} \quad & \mathbf{x} \in X \subseteq \{0,1\}^{2n} \end{aligned} \tag{3}$$

Propose an algorithm which solves problem (2) using the above subroutine.

Solution:

(a) Using the fact that the dual of the problem (for a given $\mathbf{y} \in \mathbb{R}^n$)

$$\begin{aligned} \max_{\mathbf{u}} \quad & \mathbf{u}'\mathbf{y} \\ \text{s.t.} \quad & \sum_{i=1}^{n} u_i \leq \Gamma \\ & \mathbf{0} \leq \mathbf{u} \leq \mathbf{1} \end{aligned} \tag{4}$$

is given by

$$\begin{aligned} \min_{\theta, \mathbf{p}} \quad & \Gamma\theta + \sum_{i=1}^{n} p_i \\ \text{s.t.} \quad & p_i + \theta_i \geq y_i \\ & \mathbf{p} \geq \mathbf{0}, \theta \geq 0 \end{aligned} \tag{5}$$

we can write the robust counterpart as:

$$\begin{aligned} \min_{\mathbf{x}, \theta_1, \theta_2} \quad & \mathbf{c}'\mathbf{x} + \sum_{j=1}^{n} \max(d_j x_j - \theta_1, 0) + \sum_{k=n+1}^{2n} \max(f_k x_k - \theta_2, 0) + \Gamma_1 \theta_1 + \Gamma_2 \theta_2 \\ \text{s.t.} \quad & \theta_1 \geq 0, \theta_2 \geq 0 \\ & \mathbf{x} \in X \subseteq \{0,1\}^{2n} \end{aligned} \tag{6}$$

(b) As $\forall i, x_i = 0$ or 1, the robust counterpart reduces to

$$\begin{aligned}
\min_{\mathbf{x},\theta_1,\theta_2} \quad & \mathbf{c}'\mathbf{x} + \sum_{j=1}^{n} \max(d_j - \theta_1, 0)x_j + \sum_{k=n+1}^{2n} \max(f_k - \theta_2, 0)x_k + \Gamma_1\theta_1 + \Gamma_2\theta_2 \\
\text{s.t.} \quad & \theta_1 \geq 0, \theta_2 \geq 0 \\
& \mathbf{x} \in X \subseteq \{0,1\}^{2n}
\end{aligned} \tag{7}$$

It is easy to see that θ_1 can only take one of the $n+1$ values d_i, and similarly, θ_2 can only take one of the $n+1$ values f_i. Thus, the robust problem reduces to solving $(n+1)^2$ linear binary problems (for which we have a fast sub routine), where in each problem, θ_1 and θ_2 are fixed to one of d_i and f_k respectively. The optimal robust solution x^* will be the solution of one of the $(n+1)^2$ problems with the least objective.

Problem 3 - Convex duality (30 points)

(a) (10 points) Consider the RO problem

$$\begin{aligned}
\max \quad & \mathbf{c}'\mathbf{x} \\
\text{s.t.} \quad & \mathbf{a}'\mathbf{x} \leq \mathbf{b}, \ \forall \mathbf{a} \in \mathcal{U},
\end{aligned} \tag{8}$$

where

$$\mathcal{U} = \{\mathbf{a} \,|\, \mathbf{a} = \bar{\mathbf{a}} + \mathbf{\Delta u}, \|\mathbf{u}\| \leq 1\} \tag{9}$$

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

Write down (with proof) the robust counterpart of problem (8) when the norm used to define (9) 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

$$\min_{\mathbf{x}\in\mathcal{X}} \ \max_{\xi\in\Xi} \ \xi'\mathbf{x}, \tag{10}$$

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 (10) 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 (10).

Solution:

(a) The problem we are considered involves rewriting

$$\max_{\|\mathbf{u}\|\leq 1} \mathbf{x}'\boldsymbol{\Delta}\mathbf{u}.$$

Let $\mathbf{y} = \boldsymbol{\Delta}'\mathbf{x}$. Then observe that

$$\begin{aligned}
\max_{\|\mathbf{u}\|\leq 1} \mathbf{y}'\mathbf{u} = \;& \begin{array}{ll} \max & \mathbf{y}'\mathbf{u} \\ \text{s.t.} & \|\mathbf{u}\|_1 \leq \Gamma \\ & \|\mathbf{u}\|_\infty \leq 1 \end{array} \\
= \;& \begin{array}{ll} \max & \sum_i |y_i| u_i \\ \text{s.t.} & \sum_i u_i \leq \Gamma \\ & \mathbf{u} \leq \mathbf{1} \\ & \mathbf{u} \geq \mathbf{0} \end{array} \\
= \;& \begin{array}{ll} \min & \Gamma p + \mathbf{1}'\mathbf{q} \\ \text{s.t.} & p\mathbf{1} + \mathbf{q} \geq |\mathbf{y}| \\ & p, \mathbf{q} \geq 0 \end{array} .
\end{aligned}$$

Rewriting the constraint $p\mathbf{1} + \mathbf{q} \geq |\mathbf{y}|$ as

$$p\mathbf{1} + \mathbf{q} \geq \boldsymbol{\Delta}'\mathbf{x} \text{ and } p\mathbf{1} + \mathbf{q} \geq -\boldsymbol{\Delta}'\mathbf{x},$$

we arrive at the reformulation

$$\begin{array}{ll} \max & \mathbf{c}'\mathbf{x} \\ \text{s.t.} & \bar{\mathbf{a}}'\mathbf{x} + \Gamma p + \mathbf{1}'\mathbf{q} \leq b \\ & p\mathbf{1} + \mathbf{q} \geq \boldsymbol{\Delta}'\mathbf{x} \\ & p\mathbf{1} + \mathbf{q} \geq -\boldsymbol{\Delta}'\mathbf{x} \\ & p, \mathbf{q} \geq 0. \end{array} .$$

There are of course other ways of deriving this, and if your answer looks different you should see if a change of variables gives you this form of the answer.

(b) This is SOCP duality and can be found in a variety of references, but we'll reproduce everything here. For simplicity, we ignore all bold letterings and replace the unbolded h with α. One method we will look at is as follows: we need to rewrite the inner problem

$$\max_{\xi \in \Xi} \xi' x.$$

Observe that

$$\max_{\xi \in \Xi} x'\xi = \max_{\xi,\zeta} \min_{\mu \geq 0} x'\xi - \mu(\|P\xi + Q\zeta + h\|_2 - (p'\xi + q'\zeta + \alpha)).$$

This is still hard to use. Recall that $\|y\|_2 = \max_{\|u\|_2 \leq 1} u'y$, and so

$$\begin{aligned}
\max_{\xi \in \Xi} x'\xi &= \max_{\xi,\zeta} \min_{\mu \geq 0} x'\xi - \mu(\|P\xi + Q\zeta + h\|_2 - (p'\xi + q'\zeta + \alpha)) \\
&= \max_{\xi,\zeta} \min_{\mu \geq 0} x'\xi - \mu\left(\max_{\|u\|_2 \leq 1} u'(P\xi + Q\zeta + h) - (p'\xi + q'\zeta + \alpha\right)) \\
&= \max_{\xi,\zeta} \min_{\substack{\mu \geq 0 \\ \|u\|_2 \leq 1}} x'\xi - \mu\left(u'(P\xi + Q\zeta + h) - (p'\xi + q'\zeta + \alpha)\right) \\
&= \max_{\xi,\zeta} \min_{\|v\|_2 \leq \mu} x'\xi - \left(v'(P\xi + Q\zeta + h) - \mu(p'\xi + q'\zeta + \alpha)\right),
\end{aligned}$$

where we have made the change of variables $v = \mu u$. Now

$$
\begin{aligned}
\max_{\xi,\zeta} \min_{\|v\|_2 \le \mu} x'\xi - &\left(v'(P\xi + Q\zeta + h) - \mu(p'\xi + q'\zeta + \alpha)\right) \\
&\le \min_{\|v\|_2 \le \mu} \max_{\xi,\zeta} x'\xi - \left(v'(P\xi + Q\zeta + h) - \mu(p'\xi + q'\zeta + \alpha)\right) \\
&= \min_{\|v\|_2 \le \mu} \mu\alpha - v'h + \max_{\xi,\zeta} (x' - v'P - \mu p')\xi + (-v'Q + \mu q')\zeta \\
&= \begin{array}{ll} \min & \mu\alpha - v'h \\ \text{s.t.} & x - P'v - \mu p = 0 \\ & Q'v - \mu q = 0 \\ & \|v\|_2 \le \mu \end{array}.
\end{aligned}
$$

Hence, an upper bound for the original problem be written as

$$
\begin{array}{ll}
\min\limits_{v,\mu,x} & \mu\alpha - v'h \\
\text{s.t.} & x - P'v - \mu p = 0 \\
& Q'v - \mu q = 0 \\
& \|v\|_2 \le \mu \\
& x \in \mathcal{X}
\end{array}.
$$

A condition which would make the two problems equivalent is something such as Slater's condition (strictly feasible points). More generally, see conditions for strong duality.

Another way to solve this problem is to introduce auxiliary variables before doing anything. In this case, we rewrite our inner problem as

$$
\max_{\xi \in \Xi} x'\xi = \begin{array}{ll}
\max\limits_{\xi,\zeta,\mu,v} & x'\xi \\
\text{s.t.} & v = P\xi + Q\zeta + h \\
& \mu = p'\xi + q'\zeta + \alpha \\
& \mu \ge 0 \\
& \|v\|_2 \le \mu
\end{array}.
$$

You can then work from here similarly to the way given above (Nishanth discussed this approach in recitation).

Problem 4 - Using JuMPeR (10 points)

Consider the following robust optimization problem

$$
\begin{array}{ll}
\min\limits_{x_1,x_2} & x_1 + x_2 \\
\text{s.t.} & x_1, x_2 \in \mathbb{R} \\
& x_1 \ge 0, x_2 \ge 0 \\
& a_1 x_1 + a_2 x_2 \ge 1 \quad \forall\, (a_1, a_2) \in \Xi,
\end{array} \tag{11}
$$

where $\Xi = \{(a_1, a_2) \in \mathbb{R}^2 : 0 \le a_1, a_2 \le 1, a_1 + a_2 \le 1\}$. Solve (11) 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.

Solution:

This problem is infeasible. Note that the constraint $a_1x_1 + a_2x_2 \geq 1$ for all $\mathbf{a} \in \Xi$ can never be satisfied because $\mathbf{0} \in \Xi$. The constraint $0 \geq 1$ can of course never be satisfied.

Problem 5 (20 points, **OPTIONAL EXTRA CREDIT**)

Consider the robust optimization problem

$$\begin{array}{ll} \max\limits_{\mathbf{x}\in\mathbb{R}^n} & \mathbf{c}'\mathbf{x} \\ \text{s. t.} & \mathbf{a}'\mathbf{x} \leq b \quad \forall \mathbf{a} \in \mathcal{U}, \end{array} \tag{12}$$

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 (12) 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 $p = 2$.

Solution: Left to your own imagination.