15.094, Problem Set 1 Solutions

Martin S. Copenhaver and Nishanth Mundru

Due: 18 February 2014 at 5pm EST

Problem 1 (20 points)

Let d_t be the demand for a product at time $t \in \mathcal{T} := \{1, \ldots, T\}$. From historical data, we have estimated the mean and covariance matrix for $\mathbf{d} = (d_t)_{t \in \mathcal{T}} \in \mathbb{R}^T$, which we denote by $\boldsymbol{\mu} \in \mathbb{R}^T$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{T \times T}$, respectively.

(a) (10 points) Formulate an uncertainty set for $\mathbf{d}$ that takes into account this historical information.

(b) (10 points) Suppose that the rank of $\boldsymbol{\Sigma}$ is $\kappa \ll T$. Formulate an uncertainty set for $\mathbf{d}$ that incorporates this new information.

Solution: There are many possible solutions to this problem. Here, we only propose some of them.

(a) Suppose that $\boldsymbol{\Sigma}$ is positive definite and define $\mathbf{C} = \boldsymbol{\Sigma}^{-1/2}$. We let Γ denote the budget of uncertainty and propose the following uncertainty set for $\mathbf{d}$:

$$\mathcal{R} := \left\{ \mathbf{d} \in \mathbb{R}^T \ : \ \|\mathbf{C}(\mathbf{d} - \mu)\|_p \leq \Gamma \right\},$$

where $p \in [1, \infty]$. Indeed, in the case when the demand enters the problem in a linear fashion, this set will guarantee that each constraint will be feasible with high probability (dependent on the choice of Γ).

We note that when assets are uncorrelated (i.e., $\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_t^2)_{t \in \mathcal{T}}$) and $p = 1$, this uncertainty set reduces to

$$\mathcal{R}_1 := \left\{ (d_1, \ldots, d_t) \in \mathbb{R}^T \ : \ \sum_{t \in \mathcal{T}} \left| \frac{d_t - \mu_t}{\sigma_t} \right| \leq \Gamma \right\}.$$

In the case when assets are uncorrelated and $p = \infty$, it becomes

$$\mathcal{R}_\infty := \left\{ (d_1, \ldots, d_t) \in \mathbb{R}^T \ : \ \left| \frac{d_t - \mu_t}{\sigma_t} \right| \leq \Gamma \quad \forall t \in \mathcal{T} \right\}.$$

(b) When the matrix $\boldsymbol{\Sigma}$ is not full rank, it is not invertible and neither is its square root. The approach from (a) is thus not applicable directly.

One way of solving this is to just use the "pseudo-inverse" of Σ, and replicate the set from part (a). For any matrix B, its pseudo-inverse is given by

$$\tilde{B}^{-1} = U \operatorname{diag}(\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_r}, 0, \ldots, 0) V' \tag{1}$$

, where

$$B = U\text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0)V' \tag{2}$$

and $\sigma_1 \geq \dots \geq \sigma_r > 0$ are the singular values of B.

An alternative way of arriving at the **same conclusion** is by considering a factor model. We thus model the demand vector as a linear function of a small number K ($K \ll T$) of i.i.d. factors corrupted by zero mean i.i.d. additive noise, as follows

$$\mathbf{d} = \mathbf{A}\mathbf{f} + \epsilon.$$

The mean and standard of the factors are 0 and I, respectively. The matrix A is basically the square root of the pseduo-inverse of Σ. We let σ_ϵ denote the standard deviation of the error. Assume all these constants are known (or computed from the data). A suitable uncertainty set is then given by:

$$\mathcal{R}_{\text{CLT}} := \left\{ \begin{array}{l} \mathbf{d} \in \mathbb{R}^T \ : \exists \epsilon \in \mathbb{R}^T, \ \mathbf{f} \in \mathbb{R}^F \text{ with} \\ \mathbf{d} = \mathbf{A}\mathbf{f} + \epsilon \\ |\mathbf{e}'\mathbf{f} - F\mu_f| \leq \Gamma\sigma_f\sqrt{F} \\ |\mathbf{e}'\epsilon| \leq \Gamma_\epsilon\sigma_\epsilon\sqrt{T} \end{array} \right\},$$

where $\mathbf{e}$ denotes a vector of all ones of appropriate length, while Γ and Γ_ϵ are suitably chosen parameters that control the degree of immunization desired by the decision-maker.

We remark that the sets for $\boldsymbol{\epsilon}$ and $\mathbf{f}$ in the formulation above could have been constructed by using our answer to part (a) of this problem rather than the CLT formulation, yielding the set

$$\mathcal{R}_p := \left\{ \begin{array}{l} \mathbf{d} \in \mathbb{R}^T \ : \exists \epsilon \in \mathbb{R}^T, \ \mathbf{f} \in \mathbb{R}^F \text{ with} \\ \mathbf{d} = \mathbf{A}\mathbf{f} + \epsilon \\ \left\| \frac{\mathbf{f} - \mu_f \mathbf{e}}{\sigma_f} \right\|_p \leq \Gamma \\ \left\| \frac{\boldsymbol{\epsilon}}{\sigma_\epsilon} \right\|_p \leq \Gamma_\epsilon \end{array} \right\}.$$

Numerous variants of these sets would also be acceptable.

Problem 2 (25 points)

Every instance of robustification we have focused on thus far has considered constraints of the form

$$\mathbf{a}'\mathbf{x} \leq b \ \forall \mathbf{a} \in \mathcal{U},$$

where $\mathcal{U} \subseteq \mathbb{R}^n$. In this exercise we will focus on what happens with equality constraints.

(a) (10 points) Consider the robust equality constraint

$$\mathbf{a}'\mathbf{x} = b \ \forall \mathbf{a} \in \mathcal{U},$$

where $\mathcal{U} = \{\mathbf{a} : \mathbf{A}\mathbf{a} \leq \mathbf{d}\}$. Using the fact that

$$\mathbf{a}'\mathbf{x} = b \ \forall \mathbf{a} \in \mathcal{U} \iff \mathbf{a}'\mathbf{x} \leq b \ \forall \mathbf{a} \in \mathcal{U} \quad \text{and} \quad \mathbf{a}'\mathbf{x} \geq b \ \forall \mathbf{a} \in \mathcal{U}$$

rewrite the uncertain constraint $\mathbf{a}'\mathbf{x} = b \ \forall \mathbf{a} \in \mathcal{U}$ as a deterministic, finite number of linear inequality constraints (using the usual duality methods).

(b) (10 points) Despite the fact that part (a) suggests that we can in fact robustify equality constraints, this is essentially never done. The primary reason is that robust equality constraints often lead to infeasibility. In this part we consider the homogenous case when $b = 0$. Prove that $\{\mathbf{x} : \mathbf{a}'\mathbf{x} = b \ \forall \mathbf{a} \in \mathcal{U}\} = \{\mathbf{0}\}$ if and only if $\mathcal{U}$ contains a basis for $\mathbb{R}^n$ (here, basis has the usual linear algebra definition: a linear independent spanning set).

(c) (5 points) Now consider the non-homogenous case where $b \neq 0$. State and prove an analogous claim as to that given in part (b).

Hint: affine independence.

N.B. Note that in particular, this implies that if $\mathcal{U}$ is full-dimensional, then the robustified equality constraint is infeasible (if $b \neq 0$).

N.B. Is there hope for equality constraints? Yes. It is often the case that often one really cares about only one of the constraints $\mathbf{a}'\mathbf{x} \leq b$ or $\mathbf{a}'\mathbf{x} \geq b$, and so you can robustify the one that matters. Further, based on problem structure, it is also the case that you may only need to include one of these, and you can guarantee that it is binding (at equality) in all optimal solutions. This is a very standard technique throughout optimization theory.

Solution:

(a) We first consider $\mathbf{a}'\mathbf{x} \leq b \ \forall \mathbf{a} \in \mathcal{U}$. This can be written equivalently as $\max_{\mathbf{a} \in \mathcal{U}} \mathbf{a}'\mathbf{x}$ which via strong LP duality is equivalent to

$$\begin{array}{ll} \min & \mathbf{p}'\mathbf{d} \\ \text{s.t.} & \mathbf{A}'\mathbf{p} = \mathbf{x} \\ & \mathbf{p} \geq \mathbf{0}. \end{array}$$

This gives rise to the constraints

$$\mathbf{A}'\mathbf{p} = \mathbf{x}, \ \mathbf{p} \geq \mathbf{0}, \ \text{and } \mathbf{d}'\mathbf{p} \leq b.$$

Likewise, the constraint $\mathbf{a}'\mathbf{x} \geq b \ \forall \mathbf{a} \in \mathcal{U}$ can be written as

$$\mathbf{A}'\mathbf{q} = \mathbf{x}, \ \mathbf{q} \leq \mathbf{0}, \ \text{and } \mathbf{d}'\mathbf{p} \geq b.$$

(Based on changing signs, etc., this is one of several different ways of writing the duals.) Putting things together, the robust equality constraints can be written as

$$\mathbf{A}'\mathbf{p} = \mathbf{x}, \ \mathbf{p} \geq \mathbf{0}, \ \mathbf{d}'\mathbf{p} \leq b, \ \mathbf{A}'\mathbf{q} = \mathbf{x}, \ \mathbf{q} \leq \mathbf{0}, \ \text{and } \mathbf{d}'\mathbf{p} \geq b.$$

(b) Let us first prove the right-to-left direction. Suppose that $\mathcal{U}$ contains a basis, say, $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$, where $\mathcal{U} \subseteq \mathbb{R}^n$. Let $\mathbf{A}$ denote the matrix with ith row $\mathbf{a}_i$. Then by assumption $\mathbf{A}$ is invertible, and so the only solution to $\mathbf{A}\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$, which concludes this direction.

Conversely, if $\mathcal{U}$ does not contain a basis, then $\mathcal{U}^{\perp}$ (the *orthogonal complement* of $\mathcal{U}$) must be positive-dimensional (by any myriad of theorems, e.g., "rank-nullity" theorem). Pick any nonzero $\mathbf{x} \in \mathcal{U}^{\perp}$, and this satisfies the desired properties, concluding this direction and hence the proof.

(c) The correct statement is that $S := \{\mathbf{x} : \mathbf{a}'\mathbf{x} = b \ \forall \mathbf{a} \in \mathcal{U}\} = \emptyset$ if and only if $\mathbf{0} \in \operatorname{aff}(\mathcal{U})$, where $\operatorname{aff}(\mathcal{U})$ is the affine hull of $\mathcal{U}$, namely, the set of all affine combinations of vectors in $\mathcal{U}$. This is essentially Fredholm's alternative.

The right-to-left direction is simple: if $\mathbf{0} \in \operatorname{aff}(\mathcal{U})$, then there exist $\mathbf{a}_1, \ldots, \mathbf{a}_k \in \mathcal{U}$ and $\lambda_i \in \mathbb{R}$ so that

$$\sum_i \lambda_i \mathbf{a}_i = \mathbf{0} \quad \text{and} \quad \sum_i \lambda_i = 1.$$

The existence of such $\mathbf{a}_i$ clearly demonstrate that the set S must be empty, since $0 \neq b$. The reverse direction can be adapted from part (b) or using a direction construction.

Problem 3 (30 points)

The typical mantra of robust optimization is that "the robust analogue of (Optimization Problem of Type X) is an (Optimization Problem of Type X) of comparable size." We will explore the validity of such a statement for robust LPs. Modern LP solvers can solve massive scale LPs (with millions of variables and constraints). It is possible to solve even larger problems when the systems of inequalities are well-structured or sparse. Sparse LPs are ones for which many of the coefficients in the constraints are zeros. The focus of the question is, *what happens to sparsity properties of an LP under robustification?*

(a) (10 points) Consider the constraint $\mathbf{a}'\mathbf{x} \leq b \ \forall \mathbf{a} \in \mathcal{U}$, where $\mathcal{U} = \{\mathbf{a} : \|\mathbf{a} - \widehat{\mathbf{a}}\|_\infty \leq \epsilon\}$ (here $\widehat{\mathbf{a}} \in \mathbb{R}^n$ and $\epsilon > 0$ are fixed; $\|\cdot\|_\infty$ denotes the ℓ_∞ norm, with $\|\mathbf{a}\|_\infty = \max_i |a_i|$). Rewrite the semi-infinite constraint as a finite number of linear inequality constraints. *Hint: you will likely need to use auxiliary variables.*

(b) (5 points) Comment on the sparsity of the new representation in part (a). If $\widehat{\mathbf{a}}$ has mostly zero entries (in which case we would call the constraint $\widehat{\mathbf{a}}'\mathbf{x} \leq b$ "sparse"), are the constraints in the new representation of the uncertain constraint still sparse? Please be as specific as possible.

(c) (10 points) Repeat parts (a) and (b) with a different uncertainty set:

$$\mathcal{U} = \{\mathbf{a} : \|\mathbf{a} - \widehat{\mathbf{a}}\|_1 \leq \epsilon\}.$$

(Here $\|\mathbf{a}\|_1 = \sum_i |a_i|$.)

(d) (5 points) What differences, if any, did you observe between parts (b) and (c)? Offer an explanation for why there are (or are not) differences. Does this say anything about how to choose uncertainty sets?

Solution: This problem can be solved in multiple ways, but we will just do the simplest way here.

(a) Because the ℓ_1 norm is the dual of the ℓ_∞ norm, we can apply the theory developed in lecture to note that

$$\max_{\|\mathbf{a}-\widehat{\mathbf{a}}\|_\infty \leq \epsilon} \mathbf{a}'\mathbf{x} = \widehat{\mathbf{a}}'\mathbf{x} + \epsilon\|\mathbf{x}\|_1.$$

Therefore, we require that $\widehat{\mathbf{a}}'\mathbf{x} + \epsilon\|\mathbf{x}\|_1 \leq b$. Note that this is not a linear formulation. To create a linear reformulation, we add auxiliary variables $\mathbf{z} \in \mathbb{R}^n$ with

$$z_i \geq x_i \text{ and } z_i \geq -x_i.$$

Then the constraint $\widehat{\mathbf{a}} + \epsilon\|\mathbf{x}\|_1 \leq$ becomes

$$\widehat{\mathbf{a}}'\mathbf{x} + \epsilon \sum_i z_i \leq b,\ z_i \geq x_i\ \forall i,\ \text{and } z_i \geq -x_i\ \forall i.$$

(b) While the constraints $z_i \geq x_i$ and $z_i \geq -x_i$ are sparse, we now have the one dense constraint $\widehat{\mathbf{a}}'\mathbf{x} + \epsilon \sum_i z_i \leq b$, which could make things difficult (remember from studying Dantzig-Wolfe in an introductory optimization class that sometimes adding even a small number of constraints which are not similar to an existing well-structured problem can make solving the problem substantially more difficult).

(c) Likewise, here ℓ_1 leads to the dual norm ℓ_∞, and so we have the constraint

$$\widehat{\mathbf{a}}'\mathbf{x} + \epsilon\|\mathbf{x}\|_\infty \leq b.$$

As before we add an auxiliary variable z, and now we have

$$\widehat{\mathbf{a}}'\mathbf{x} + \epsilon z \leq b, z \geq x_i\ \forall i,\ \text{and } z \geq -x_i\ \forall i.$$

Now we essentially preserve the sparsity of the original constraint, and note that we only had to add a single auxiliary variable.

(d) The first choice of uncertainty set led to an introduction of a dense constraint, in constrast to the second set. While this might not give us general wisdom about how to choose an uncertainty set, it certainly points out that if we wish to preserve some certain problem structure in a robust counterpart that the original problem has (which in our case for this exercise is sparsity), then we should pay careful attention to how we ultimately choose our sets.

Problem 4 (25 points)

Consider the following robust optimization problem

$$\begin{aligned} \min_{x_1,x_2} \quad & x_1 + x_2 \\ \text{s.t.} \quad & x_1, x_2 \in \mathbb{R} \\ & \left.\begin{aligned} x_1 \geq \xi_1 \\ x_2 \geq \xi_2 \end{aligned}\right\} \quad \forall\ (\xi_1, \xi_2) \in \Xi, \end{aligned} \tag{3}$$

where $\Xi = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_1| + |\xi_2| \leq 1\}$.

(a) (20 points) Formulate and solve the robust counterpart of (3). Provide us with your commented code and the optimal solution obtained.

(b) (5 points) Do you find the solution intuitive? Is it what you expected?

N.B. For this exercise, you may use the programming language and solver of your choice. For LPs and SOCPs, the solvers most commonly used by academics are CPLEX and Gurobi (both freely available for students).

Solution: In this problem, we use Matlab and Yalmip (to formulate the optimization problems) and CPLEX (to solve the optimization problems).

(a) In order to solve this problem, we may use the robust counterpart provided in the lectures. For this purpose, we introduce an auxiliary variable x_3 that we constrain to take on the value of 1 and define $\mathbf{x} = (x_1, x_2, x_3)$. Problem 4 is then expressible as

$$\begin{array}{ll} \min\limits_{\mathbf{x} \in \mathbb{R}^3} & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} & \mathbf{a}_1^\top \mathbf{x} \le 0 \quad \forall \mathbf{a}_1 \in \mathcal{U}_1 \\ & \mathbf{a}_2^\top \mathbf{x} \le 0 \quad \forall \mathbf{a}_2 \in \mathcal{U}_2 \\ & \mathbf{e}_3^\top \mathbf{x} = 1, \end{array}$$

where $\mathbf{e}_3$ denotes the 3rd basis vector in $\mathbb{R}^3$, the cost vector is given by $\mathbf{c} = (1, 1, 0)$ while the uncertainty sets are defined through

$$\mathcal{U}_1 := \left\{ \mathbf{a}_1 \in \mathbb{R}^3 \ : \ \mathbf{a}_1 = \mathbf{A}_1 \xi + \bar{\mathbf{a}}_1, \ \|\xi\|_1 \le 1 \right\} \text{ with } \mathbf{A}_1 := \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \text{ and } \bar{\mathbf{a}}_1 = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix},$$

and

$$\mathcal{U}_2 := \left\{ \mathbf{a}_2 \in \mathbb{R}^3 \ : \ \mathbf{a}_2 = \mathbf{A}_2 \xi + \bar{\mathbf{a}}_2, \ \|\xi\|_1 \le 1 \right\} \text{ with } \mathbf{A}_2 := \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \text{ and } \bar{\mathbf{a}}_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}.$$

From the lectures, the robust counterpart is expressible as

$$\begin{array}{ll} \min\limits_{\mathbf{x} \in \mathbb{R}^3} & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} & \bar{\mathbf{a}}_1^\top \mathbf{x} + \|\mathbf{A}_1^\top \mathbf{x}\|_\infty \le 0 \\ & \bar{\mathbf{a}}_2^\top \mathbf{x} + \|\mathbf{A}_2^\top \mathbf{x}\|_\infty \le 0 \\ & \mathbf{e}_3^\top \mathbf{x} = 1. \end{array} \tag{4}$$

Here, we have used the fact that the dual norm of the ℓ_1 norm is the ℓ_∞ norm. We now proceed to implement the robust counterpart. The code is provided below.

```
% define the decision variables
x = sdpvar(3,1);
% define the data matrices
c = [1;1;0];
a1_bar = [-1;0;0];
a2_bar = [0;-1;0];
A1 = zeros(3,2);
A1(3,1) = 1;
A2 = zeros(3,2);
A2(3,2) = 1;
% define the objective function
objective = c'*x;
% define the three constraints
C = [a1_bar' * x + max( abs( A1' * x ) ) <= 0];
C = C + [a2_bar' * x + max( abs( A2' * x ) ) <= 0];
C = C + [ [0 0 1] * x == 1 ];
% select the solver
ops = sdpsettings('solver','cplex')
```

```
% solve the problem with chosen solver
out = solvesdp(C,objective,ops)
% display the solution and the optimal objective
display( double(x) )
display( double(objective) )
```

The optimal solution obtained was $\mathbf{x}^\star = (1,1,1)$ while the optimal objective was 2.

(b) Before answering this question, we provide a slightly different approach to formulate the robust counterpart which provides further intuition. We note:

$$x_1 \geq \xi_1 \quad \forall(\xi_1,\xi_2) \in \Xi \;\Leftrightarrow\; x_1 \geq \xi_1 \quad \forall \xi_1 \in \Xi_1 \;\text{ and }\; x_2 \geq \xi_2 \quad \forall(\xi_1,\xi_2) \in \Xi \;\Leftrightarrow\; x_2 \geq \xi_2 \quad \forall \xi_2 \in \Xi_2,$$

where Ξ_1 (Ξ_2) denotes the projection of Ξ onto the ξ_1 (ξ_2) axis, i.e.

$$\Xi_1 := \{\xi_1 \in \mathbb{R} \,:\, \exists \xi_2 \text{ with } (\xi_1,\xi_2) \in \Xi\} = [-1,1]$$

$$\Xi_2 := \{\xi_2 \in \mathbb{R} \,:\, \exists \xi_1 \text{ with } (\xi_1,\xi_2) \in \Xi\} = [-1,1].$$

The robust counterpart is then given by

$$\begin{array}{ll} \min\limits_{x_1,x_2} & x_1 + x_2 \\ \text{s.t.} & x_1, x_2 \in \mathbb{R} \\ & x_1 \geq 1 \\ & x_2 \geq 1. \end{array}$$

We remark that this problem is precisely equivalent to (4) (after elimination of x_3).

From this formulation, it becomes apparent that it is not important whether the uncertainties appearing in different constraints are or not linked to each other. All that matters is the projection of the uncertainty set onto the space of the data of each individual constraint. This result may seem counter-intuitive: in fact, we could have replaced the uncertainty set by $\Xi_1 \times \Xi_2 = [-1,1]^2$ (the Euclidean product of the projections) and obtained the same robust counterpart.

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

Fix $\mathbf{a}_i \in \mathbb{R}^n$ and $\epsilon_i > 0$ for $i = 1, \ldots, m$. Fix $p \in [1,\infty]$ and let $\|\cdot\|_p$ denote the usual ℓ_p norm. Let $\mathcal{U}_i = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{a}_i - \mathbf{x}\|_p \leq \epsilon_i\}$. Write the matrix

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1' \\ \vdots \\ \mathbf{a}_m' \end{pmatrix} \in \mathbb{R}^{m \times n},$$

and suppose that $\mathbf{b} \in \mathbb{R}^m$ so that $P = \{\mathbf{x} : \mathbf{A}\mathbf{x} \leq b\}$ is a full-dimensional, bounded polyhedron. Let $\text{vol}(S)$ denote the n-dimensional volume for a set $S \subseteq \mathbb{R}^n$. Compute explicitly (or provide estimates of) the ratio

$$\frac{\text{vol}\left(\{\mathbf{x} : \mathbf{a}'\mathbf{x} \leq b_i \;\forall \mathbf{a} \in \mathcal{U}_i \;\forall i\}\right)}{\text{vol}(P)}.$$

Can you provide an estimate in the case when only one ϵ_i is nonzero? How does your answer depend on p?

Hint: it is instructive to focus on different examples, such as hypercubes, simplices, etc.

Solution: Left to your own imagination.