Semidefinite and Second Order Cone Programming Seminar
Fall 2012
Project: Robust Optimization and its Application of Robust Portfolio Optimization

Instructor: Farid Alizadeh
Author: Ai Kagawa

12/12/2012

1 General Concept of the Robust Optimization

There are several different methods to treat uncertainty in literacy. First, sensitivity analysis which is a post-optimization method analyzing stability of generated solutions. Second, stochastic programming is a modeling approach, and this method is limited to that the uncertainty is stochastic in nature. Third, "Robust Mathematical Programming" is that a candidate solutions allowing to violate the "scenario realizations" of the constraints instead of considering fixed "nominal" data for several scenarios considered, and add penalty terms in the objective function. [2]

The robust optimization is a method to handle deterministic uncertainty in optimization problems [4]. Majority of uncertain data are due to prediction errors, measurement errors, and implementation errors. Moreover, small errors in data values can make optimal solutions highly infeasible, and the robust optimization prevent the infeasible solutions. This paper first summarizes the basic concept of robust optimization [1] and shows its application of robust portfolio optimization.

2 Uncertain Linear Optimization

2.1 Uncertain Linear Problems and their Robust Counterparts

The ordinary linear programming (LP) form is

$$\begin{array}{ll} \min_x & c^T x + d \\ \text{s.t.} & Ax \le b, \end{array}$$

where $x \in \mathbb{R}^n$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}$, $A \in \mathbb{R}^{n \times m}$, and $b \in \mathbb{R}^m$.

Now, consider the case that input data (c, d, A, b) are uncertain. Then, we represent the uncertain LP as the following form.

Definition 1 *The uncertain LP form is*

$$\begin{array}{ll} \min_x & c^T x + d \\ \textit{s.t.} & Ax \le b \\ & (c, d, A, b) \in \mathcal{U}, \end{array}$$

where an uncertainty set $\mathcal{U} \in \mathbb{R}^{(m+1)\times(n+1)}$.

In the robust optimization problem, we assume that the uncertainty set $\mathcal{U}$ is affinity parameterized by a perturbation vector $\mathcal{C}$ associated with a given perturbation set $\mathcal{Z}$.

$$\mathcal{U} = \{D = D_0 + \sum_{l=1}^{L} \mathcal{C}_l D_l : \mathcal{C} \in \mathcal{Z} \subset \mathbb{R}^L\}$$

where the data of the problem, the nominal data, and the basic shift, respectively, as

$$D = \begin{pmatrix} c^T & d \\ A & b \end{pmatrix}, D_0 = \begin{pmatrix} c_0^T & d_0 \\ A_0 & b_0 \end{pmatrix}, \text{ and } D_l = \begin{pmatrix} c_l^T & d_l \\ A_l & b_l \end{pmatrix}.$$

The robust feasible solutions are guaranteed to be feasible for all admissible data.

Definition 2 *A vector* $x \in \mathbb{R}^n$ *is a robust feasible solution to the uncertain robust linear programming if it satisfies all realization of the constraints with the uncertainty set such that*

$$a^T x \le b \quad \forall (c, d, A, b) \in \mathcal{U}.$$

Definition 3 *For a given* x*, the robust objective value,* $\hat{c}(x)$*, in the uncertain linear programming is the largest value of the true objective value* $c^T x + d$ *over all realizations of the data from the uncertain set such that*

$$\hat{c}(x) = \sup_{(c,d,A,b)\in\mathcal{U}} (c^T x + d).$$

Definition 4 *The Robust Counterpart of the uncertain LP is*

$$\begin{array}{ll} \min_x & \hat{c}(x) = \sup_{(c,d,A,b)\in\mathcal{U}}(c^T x + d) \\ s.t. & Ax \le b \quad \forall (c,d,A,b) \in \mathcal{U} \end{array}$$

It can also rewritten as

$$\begin{array}{ll} \min_{x,t} & t \\ s.t. & c^T x + d \le t \\ & Ax \le b \\ & \forall (c,d,A,b) \in \mathcal{U}. \end{array}$$

An uncertain LP can be transformed as an uncertain LP with certain objective, and it is the same as the robust counterpart. Furthermore, the robust optimal solution to the uncertain LP is the optimal solution to the Robust Counterpart.

2.2 Tractability of Robust Counterparts

From the computational point of view, the RC of uncertain LP with the uncertainty set $\mathcal{U}$ is computationally tractable when the convex uncertainty set $\mathcal{U}$ is computationally tractable.

3 Robust Approximations of Scalar Chance Constraints

3.1 Specification of an Uncertainty Set: Probabilistic vs Uncertain-But-Bounded Perturbations

Two different ways to specify a given uncertainty set are using probabilistic perturbations and uncertain-but-bounded perturbations. The method using probabilistic perturbation is a natural way using the prior knowledge. On the other hand, the uncertain-but-bounded method is a implicit way to specify the model by choosing the proper worst-case-oriented decisions, but it gives clear prediction.

3.2 Chance Constraints and their Safe Tractable Approximations

Let's consider the uncertain linear constraint:

$$a^T x \le b \quad [a;b] = [a^0;b^0] + \sum_{l=1}^{L} \mathcal{C}_l [a^l;b^l]$$

where $\mathcal{C}$ is a perturbation vector, which is random with completely known probability distribution P. As in Definition 2, a vector x is a robust feasible solution with respect to the given perturbation set $\mathcal{Z}$ that is closed convex hull to support P if only if the given x satisfies the uncertain constraint. Then, it is natural to handle deterministic uncertainty in the context of uncertain linear programming using the formulation of chance constraint as

$$p(x) \equiv \mathsf{Prob}_{\mathcal{C}\sim P}\{\mathcal{C} : [a^0]^T x + \sum_{l=1}^{L} \mathcal{C}_l [a^l]^T x > b^0 + \sum_{l=1}^{L} \mathcal{C}_l b^l\} \leq \epsilon$$

where $\mathsf{Prob}_{\mathcal{C}\sim P}$ is the probability associated with the distribution P and $\epsilon \in (0,1)$ is a small tolerance.

Since the chance constraints often require a computational difficulties, we replace a chance constraint with its computationally tractable safe approximation.

Definition 5 *Let $\mathcal{S}$ be a system of convex constraints on x and variable v. Then, $\mathcal{S}$ is a safe convex approximation of chance constraint if x is component of every feasible solution (x, v) of $\mathcal{S}$ is feasible for the chance constraint.*

If convex constraints forming $\mathcal{S}$ are efficiently tractable, then the safe convex approximation $\mathcal{S}$ of the chance constraint is computationally tractable.

3.3 Safe Tractable Approximations of Scalar chance Constraints

Chance constraint is associated with randomly perturbed constraint, which is an uncertain linear inequality, and a given distribution P of random perturbations. It is rare to have full information of the distribution P in reality, and the only information we have is that P belongs to a given distribution family, $\mathcal{P}$. Then, we can represent the chance constraint as the ambiguous chance constraint.

$$\forall (P \in \mathcal{P}) : \mathsf{Prob}_{\mathcal{C}\sim P}\{\mathcal{C} : [a^0]^T x + \sum_{l=1}^{L} \mathcal{C}_l [a^l]^T x > b^0 + \sum_{l=1}^{L} \mathcal{C}_l b^l\} \leq \epsilon$$

Assume that $\mathcal{C}_l$ are the random variables satisfying:

1. $E[\mathcal{C}_l] = 0$
2. $|\mathcal{C}_l| \leq 1, l = 1, ..., L$
3. $\{\mathcal{C}_l\}_{l=1}^{L}$ are independent.

Moreover, observe that the body of the ambiguous chance constraint can be represented as

$$\eta \equiv \sum_{l=1}^{L} ([a^l]^T x - b^l) \mathcal{C}_l \leq b^0 - [a^0]^T x.$$

With the assumption of $\mathcal{C}_l$ and fixed x, η is a random variable with

1. Mean: $E[\eta] = 0$
2. Standard Deviation:

$$\mathrm{StD}[\eta] = \sqrt{\sum_{l=1}^{L}([a^l]^T x - b^l)^2 E[\mathcal{C}_l^2]} \leq \sqrt{\sum_{l=1}^{L}([a^l]^T x - b^l)^2}.$$

For the general conclusion of the safe tractable approximation, η is "nearly never" greater than $\Omega\sqrt{\sum_{l=1}^{L}([a^l]^T x - b^l)^2}$, where Ω is a safety parameter. Hence, the large Ω, the less chances for η to be larger than the outlined quantity. Therefore, we call

$$\Omega\sqrt{\sum_{l=1}^{L}([a^l]^T x - b^l)^2} \leq b^0 - [a^0]^T x$$

as a parametric safe tractable approximation of the randomly perturbed ambiguous chance constraint.

4 Globalized Robust Counterparts of Uncertain LP

The robust counterpart assumed that no violations of uncertainty constraints with a given perturbation set. On the other hand, globalized robust counterparts is the method considering the violation of the given perturbation set. This situation often arises in real data such as extreme values of data. Such "soft" constraints against the data uncertainty can be handle by ensuring controlled deterioration of the constraint when the data violate the uncertainty set.

Let $\mathcal{Z}_+$ be the set of "physically possible" perturbations such that $\mathcal{Z} \subset \mathcal{Z}_+$ where $\mathcal{Z}$ is the "normal range" of the perturbations.

Then the violation of constrain where $\mathcal{C} \in \mathcal{Z}_+ \backslash \mathcal{Z}$ ("physically possible" perturbation set is outside the "normal range") is bounded by a constraint times the distance from $\mathcal{C}$ to $\mathcal{Z}$.

$$[a^0 + \sum_{l=1}^{L} \mathcal{C}_l a^l]^T x - [b^0 + \sum_{l=1}^{L} \mathcal{C}_l b^l] \leq \alpha \cdot \mathrm{dist}(\mathcal{C}, \mathcal{Z}) \quad \forall\ \mathcal{C} \in \mathbb{Z}_+$$

where $\alpha \geq 0$ is a given "global sensitivity".

Definition 6 *Given $\alpha \geq 0$ and a perturbation structure $(\mathcal{Z}, \mathcal{L}, \|\cdot\|)$, a x is a globally robust feasible solution to uncertain LP with a given global sensitivity α if x satisfies the semi-ininfinite constraint which is also called globalized robust counterpart (GRC) of the uncertain constraint,*

$$[a^0 + \sum_{l=1}^{L} \mathcal{C}_l a^l]^T x \leq [b^0 + \sum_{l=1}^{L} \mathcal{C}_l b^l] + \alpha \cdot \mathrm{dist}(\mathcal{C}, \mathcal{Z}|\mathcal{L}) \quad \forall\ \mathcal{C} \in \mathcal{Z}_+,$$

where

$$\mathcal{Z}_+ = \mathcal{Z} + \mathcal{L} = \{\mathcal{C} = \mathcal{C}' + \mathcal{C}'' : \mathcal{C}' \in \mathcal{Z},\ \mathcal{C}'' \in \mathcal{L}\}, \quad \mathcal{L} \textit{ is a closed convex cone,}$$

and

$$\mathtt{dist}(\mathcal{C}, \mathcal{Z}|\mathcal{L}) = \inf_{\mathbb{C}'}\{\|\mathcal{C} - \mathcal{C}'\| : \mathcal{C}' \in \mathcal{Z},\ \mathcal{C} - \mathcal{C}' \in \mathcal{L}\}.$$

5 Uncertain Conic Optimization

The formulation of a conic optimization problem is

$$\begin{array}{ll} \min_x & c^T x \\ \text{s.t.} & Ax - b \in \mathbf{K}, \end{array}$$

where $x \in \mathbb{R}^m$ and $K \subset \mathbb{R}^m$ is a closed pointed convex cone with a non-empty interior. Also, $x \mapsto Ax - b$ is a given affine mapping from $\mathbb{R}^n$ to $\mathbb{R}^m_+$.

There are major three types of cones for convex programs.

1. Direct products of nonnegative rays, $\mathbf{K}$ is a non-negative orthant $\mathbb{R}^m$. These cones give the following linear optimization.

$$\begin{array}{ll} \min_x & c^T x \\ \text{s.t.} & a_i^T x - b_i \geq 0, 1 \leq i \leq m \end{array}$$

2. Direct products of Lorentz or Second-order cones.

$$L^k = \left\{ x \in \mathbb{R}^k : x_k \geq \sqrt{\sum_{j=1}^{k-1} x_j^2} \right\}$$

These cones form the second order conic optimization.

$$\begin{array}{ll} \min_x & c^T x \\ \text{s.t.} & \|A_i x - b_i\|_2 \leq c_i^T x - d_i, 1 \leq i \leq m \end{array}$$

3. Direct Products of semidefinite cones S^k_+, which are the cones of positive semidefinite $k \times k$ matrices.

   S^k_+ is the Euclidean space of the inner product $\langle A, B \rangle = \mathrm{Tr}(AB) = \sum_{i,j=1}^k A_{ij} B_{ij}$.

   Semidefinite cones lead to the Semidefinite Programming (SDP) problem,

$$\begin{array}{ll} \min_x & c^T x + d \\ \text{s.t.} & \mathcal{A}^i x - B_i \succeq 0 \quad i = 1, \dots m \end{array}$$

   where

$$x \mapsto \mathcal{A}_i x - B_i \equiv \sum_{j=1}^{n} x_j A^{ij} - B_i$$

   is an affine mapping from $\mathbb{R}^n$ to S^{k_i}, A^{ij}, B_i are symmetric matrices of $k_i \times k_i$, and $A \succeq 0$.

5.1 Uncertain Conic Problems and Robust Counterparts

Assume that uncertain data are affinely parameterized by a perturbation vector $\mathcal{C} \in \mathbb{R}^L$ where a given perturbation set $\mathcal{Z} \subset \mathbb{R}^L$.

$$(c, d, \{A_i, b_i\}_{i=1}^m) = (c^0, d^0, \{A_i^0, b_i^0\}_{i=1}^m) + \sum_{l=1}^{L} (c^l, d^l, \{A_i^l, b_i^l\}_{i=1}^m)$$

Definition 7 *A* $x \in \mathbb{R}^n$ *is a robust feasible solution to the uncertain conic problems if it satisfies all realization of the perturbation vector in the perturbation set such that*

$$[A_i^0 + \sum_{l=1}^{L} \mathcal{C}_l A_i^l]x - [b_i^0 + \sum_{l=1}^{L} \mathcal{C}_l b_i^l] \in Q_i \quad \forall i, 1 \leq i \leq m, \mathcal{C} \in \mathcal{Z}$$

where $Q_i \subset \mathbb{R}^{k_i}$ *is nonempty closed convex set such that*

$$Q_i = \{u \in \mathbb{R}^{k_i} : Q^{il}u - Q^{il} \in K_{il}, \quad l = 1 \ldots, L_i\}$$

Definition 8 *The Robust Counterpart of an uncertain conic problem is*

$$\begin{array}{ll} \min_{x,t} & t \\ s.t. & [c^0 + \sum_{l=1}^{L} \mathcal{C}_l c^l]^T x + [d^0 + \sum_{l=1}^{L} \mathcal{C}_l d^l] - t \in Q_0 = \mathbb{R}_- \\ & [A_i^0 + \sum_{l=1}^{L} \mathcal{C}_l A_i^l]x - [b_i^0 + \sum_{l=1}^{L} \mathcal{C}_l b_i^l] \in Q_i \quad , 1 \leq i \leq m \\ & \forall \mathcal{C} \in \mathcal{Z}. \end{array}$$

5.2 Uncertain SDP and their Robust Counterparts

Again, the formulation of the uncertain SDP is

$$\begin{array}{ll} \min_x & c^T x + d \\ \text{s.t.} & \mathcal{A}_i(x) \equiv \sum_{j=1}^{n} x_j A^{ij} - B_i \in \mathcal{S}_+^{k_i} \quad i = 1, \ldots, m \end{array}$$

$$\Updownarrow$$

$$\begin{array}{ll} \min_x & c^T x + d \\ \text{s.t.} & \mathcal{A}_i(x) \equiv \sum_{j=1}^{n} x_j A^{ij} - B_i \succeq 0 \quad i = 1, \ldots m \end{array}$$

where A^{ij}, B_i are symmetric matrices of $k_i \times k_i$. The constraints of SDP are Linear Matrix Inequality (LMI) constraints. Another formulation of SDP is

$$\begin{array}{ll} \min_x & c^T x + d \\ \text{s.t.} & Aix - b_i \in Q_i \quad i = 1 \ldots m \end{array}$$

where nonempty sets

$$Q_i = \{u \in \mathbb{R}^{p_i} : \mathcal{Q}_{il}(u) \equiv \sum_{s=1}^{p_i} u_s Q^{sil} - Q^{il} \succeq 0, \quad l = 1, \ldots, L_i\}$$

and $Q_i = \mathcal{S}_+^{k_i} \quad i = 1, \dots, m$.

The robust counterpart of uncertain SDP at a given perturbation level $\rho > 0$

$$\begin{array}{ll} \min_{x,t} & t \\ \text{s.t.} & [[c^n]^T + d^n] + \sum_{l=1}^{L} \mathcal{C}_l[[c^l]^T x + d^l] \le t \\ & [A_i^n + b_i^n] + \sum_{l=1}^{L} \mathcal{C}_l[A_i^l + b_i^l] \in Q_i \quad , 1 \le i \le m \\ & \forall \mathcal{C} \in \rho\mathcal{Z} \end{array}$$

RC of an uncertain LMI is generally computationally intractable. Therefore, we have three generic cases such that RCs to be computationally tractable.

1. Scenario perturbation set
2. Unstructured Norm-Bounded Uncertainty
3. Ellipsoidal Uncertainty

6 Application: Robust Portfolio Optimization

6.1 Notations

Let $w \in W$ are the weights of a portfolio. Let $x \in \mathbb{R}^n$ be one-period return of n stocks.For a probability measure μ, we have:

1. First Moment of Return (Mean): $\int_S x \, d\mu = \hat{x} = [\hat{x}_1 \dots \hat{x}_n]$
2. Second Moment of Return: $\int_S xx^T \, d\mu = \Sigma$
3. Covariance Matrix: $\Gamma = \Sigma - \hat{x}\hat{x}^T$

6.1.1 Markoviz Mean Variance Model Formulation

The basic Makoviz Mean Variance (MMV) model is

$$\begin{array}{ll} \max_{w \in \mathbb{R}^n} & \hat{x}^T w \\ \text{s.t.} & w^T \Gamma w \le \gamma^2 \\ & w \in \bar{W}, \end{array}$$

where γ is a standard deviation target, and $\bar{W} = \{w \in \mathbb{R}^n : \sum_{i=1}^n w_i = 1, w \ge 0\}$. By the Lagrangian of this optimization is

$$\begin{array}{ll} \max_{w \in \mathbb{R}^n} & \hat{x} w - \lambda w^T \Gamma w \\ \text{s.t.} & w \in \bar{W}, \end{array}$$

where $\lambda \ge 0$ is the Lagrangian multiplier associated with the portfolio risk constraint, and it is also interpreted as the relative importance associated with the risk.

6.2 Robustness to first estimation errors

Now, consider the model of the worst case scenario on the expected returns by assuming that the true expected return $\bar{x}_t$ is normally distributed and lies in confidence region represented by the ellipsoid such that

$$(\bar{x}_e - \bar{x}_t)^T \Gamma_e^{-1} (\bar{x}_e - \bar{x}_t) \leq \kappa^2$$

where $\bar{x}_e$ is the estimated expected returns and their covariance matrix Γ_e with probability η, where $\kappa^2 = \chi_n^2(1-\eta)$ and χ_n^2 is the inverse cumulative distribution function of the chi-square distribution function with n degree of freedom.

For a fixed portfolio $\hat{w}$, the worst-case return is

$$\begin{array}{ll} \max_{\bar{x}_e - \bar{x}_t} & (\bar{x}_e - \bar{x}_t)^T \hat{w} \\ \text{s.t.} & (\bar{x}_e - \bar{x}_t)^T \Gamma_e^{-1} (\bar{x}_e - \bar{x}_t) \leq \kappa^2. \end{array}$$

By its Lagrangian, $\bar{x}_t^T \hat{w} = \bar{x}_e^T \hat{w} - \kappa \|\Gamma_e^{1/2} \hat{w}\|$ where $\|\Gamma_e^{1/2} \hat{w}\|$ is the standard deviation of the estimated total return. Therefore, the robust portfolio selection problem is

$$\begin{array}{ll} \max_{w \in \mathbb{R}^n} & \hat{x}_e^T w - \kappa \|\Gamma_e^{1/2} w\| \\ \text{s.t.} & w^T \Gamma w \leq \gamma^2 \\ & w \in \bar{W}. \end{array}$$

This model is interpreted as the trade off between the expected return and the standard deviation of the expected return, and the term $\|\Gamma_e^{1/2} w\|$ has the effect of reducing estimation error in the portfolio optimization problem. Its second-order cone programming problem formulation is

$$\begin{array}{ll} \max_{w \in \mathbb{R}^n} & \hat{x}_e^T w - \kappa t \\ \text{s.t.} & w \in \bar{W} \\ & \|\Gamma^{1/2} w\| \leq \gamma \\ & \|\Gamma_e^{1/2} w\| \leq t. \end{array}$$

6.3 Robustness to first and second moment estimation errors

The minimum variance formulation of the portfolio selection problem is

$$\begin{array}{ll} \min_{w \in \mathbb{R}^n} & w^T \Gamma w \\ \text{s.t.} & w \in \bar{W} \\ & w^T \bar{x} \geq R, \end{array}$$

where R is the lower limit of the target expected return. Consider the portfolio optimization problem by the point of view of the worst-case return and covariance matrix. Then, we can formulate the following min-max portfolio optimization.

$$\begin{array}{ll} \min_{w\in\mathbb{R}^n} \max_{(\Sigma,\bar{x})\in\mathcal{U}} & w^T \Gamma w \\ \text{s.t.} & w \in \bar{W} \\ & \min_{\bar{x}\in\mathcal{U}} w^T\bar{x} \geq R, \end{array}$$

This model is assuming that the distributions of the first moment and the second moment is independent, and it is not satisfied in practice.

6.4 Robustness to second moment estimation errors

Assume that the expected returns $\bar{x}$ are knows and fixed. Then consider the worst-case variance portfolio optimization in SDP formulation. Therefore, the worst-case variance is depended on the second moment.

$$\begin{array}{ll} \sup_{\mu} & \langle ww^T, \int_S xx^T \, d\mu \rangle - \langle w, \int_S x \, d\mu \rangle^2 \\ \text{s.t.} & \int_S x \, d\mu = \hat{x} \\ & \int_S d\mu = 1 \\ & \mu(x) \geq 0. \end{array}$$

This model is an infinite dimensional optimization, and it cannot be solved directly. However, applying equivalent lemma, it can be transformed as the standard min-max SDP. Furthremore, applying the strong duality for the max part of the standard min-max SDP, it can be rewritten as min SDP. Therefore, the problem can be solved by a standard SDP solver. [3]

References

[1] A. Ben-Tal, L. El Ghaoui, A. Nemirovski, *Robust Optimization*, 2009

[2] A. Ben-Tal, A. Nemirovski, *Robust Convex Optimization*, 1998

[3] K. Ye, P. Parpas, B. Rustem, *Robust portfolio optimization: a conic programming approach*, Comput Optim Appl (2012) 52:463481

[4] L. El Ghaoui, *Robust Optimization and Application*, power-point presentation on robust optimization, IMA Tutorial, March 11, 2003

[5] S. Ceria, R. Stubbs, *Incorporating Estimation Errors into Portfolio Selection: Robust Portfolio Construction*, Axioma Research Paper No. 003, May, 2006