Selected Topics in Robust Convex Optimization

Arkadi Nemirovski
School of Industrial and Systems Engineering
Georgia Institute of Technology

- Optimization programs with uncertain data and their Robust Counterparts
- Tractability of Robust Counterparts
- Robust Optimization and Chance Constraints

- Optimization programs with uncertain data and their Robust Counterparts
- Tractability of Robust Counterparts
- Robust Optimization and Chance Constraints

♣ Robust Optimization is a methodology for processing *uncertain optimization problems*

$$\min_x \{f(x,\zeta) : F(x,\zeta) \in \mathbf{K}\}$$

- $x \in \mathbb{R}^n$ is the *decision vector*
- $\zeta \in \mathbb{R}^d$ is the *data* (or data perturbation)
- $f(x,\zeta) : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$ and $F(x,\zeta) : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ are given functions, and $\mathbf{K} \subset \mathbb{R}^m$ is a given set.

$f(\cdot,\cdot), F(\cdot,\cdot), \mathbf{K}$ form the *structure* of the uncertain problem.

♣ In contrast to Stochastic Programming, RO does not assume stochastic nature of data ζ and uses instead *uncertain-but-bounded* uncertainty model: ζ runs through a given (typically, compact) *uncertainty set* $\mathcal{Z} \subset \mathbb{R}^d$.

♣ **RO, people:**

- **1973:** A.L. Soyster (LP)
- **1997:** P. Kouvelis & G. Yu (IP)
- **1997 –:** L. El Ghaoui & H. Lebret & F. Oustry; A. Ben-Tal & A. Nemirovski } (CP)
- **2000 –:** E. Adida, A. Atamturk, A. Beck, D. Bertsimas, C. Bhattacharyya, H.-G. Bock, S. Boyd, G. Calafiore, M. Diehl, Y. Eldar, E. Erdogan, L. Grate, E. Guslitzer, B. Golany, D. Goldfarb, C. Hol, G. Iyengar, M. Jordan, E. Kostina, O. Kostyukova, G. Lanckriet, A. Nilim, M. Sim, D. Pachamanova, C. Roos, C. Schrerer, A. Sood, A. Thiele, J.-Ph. Vial, M. Zhang,...

♠ **RO, applications:**

• Structural/Circuit/Network Design • Control • Signal Processing • Machine Learning • Portfolio Optimization • Inventory...

Example of uncertain LP: Multi-product inventory with backlogged demand and shared warehouse capacity

$$
\begin{array}{lll}
\min\limits_{C,x_t,y_t,w_t,v_t} & C & \text{[inventory management cost]} \\
\text{s.t.} & \sum_{t=1}^{T}[c_{\rm h}'y_t + c_{\rm b}'w_t + c_{\rm p}'v_t] \leq C & \text{[cost description]} \\
& x_{t+1} = x_t + v_t - \zeta_t & \text{[balance equations]} \\
& x_t \leq y_t,\ 0 \leq y_t & \text{[bounds on } [x_t]_+ \text{]} \\
& -x_t \leq w_t,\ 0 \leq w_t & \text{[bounds on backlogged demand]} \\
& \underline{v}_t \leq v_t \leq \overline{v}_t & \text{[bounds on orders]} \\
& q'y_t \leq Q & \text{[warehouse capacity bound]}
\end{array}
$$

- $x_t \in \mathbb{R}^d$: inventory state at time t
- $v_t \in \mathbb{R}^d$: replenishment orders at time t
- $y_t \in \mathbb{R}^d$: stored items at time t
- $w_t \in \mathbb{R}^d$: backlogged demand at time t
- $q \in \mathbb{R}^d_+$: storage requirements
- $c_{\rm o} \in \mathbb{R}^d_+$: ordering costs
- $c_{\rm h} \in \mathbb{R}^d_+$: holding costs
- $c_{\rm b} \in \mathbb{R}^d_+$: backlog penalty

- Uncertain data: demands $\zeta = [\zeta_1; ...; \zeta_T]$

$$\boxed{\left\{\min_x \{f(x,\zeta) : F(x,\zeta) \in \mathbf{K}\} : \zeta \in \mathcal{Z}\right\}} \qquad \text{(Unc)}$$

Assume that our "decision environment" is such that

- All decisions x_j should be made before ζ "reveals itself" and thus should be independent of ζ
- The constraints are "hard": their violations cannot be tolerated
- We intend to take full care of all data $\zeta \in \mathcal{Z}$ and do not care what happens when $\zeta \notin \mathcal{Z}$.

Under these assumptions, seemingly the only meaningful way to process (Unc) is to solve the *Robust Counterpart*

$$\min_{t,x} \{t : \forall \zeta \in \mathcal{Z} : f(x,\zeta) \leq t, F(x,\zeta) \in \mathbf{K}\} \qquad \text{(RC)}$$

of the uncertain problem.

Example: To build the RC of the Inventory problem, we use balance equations to eliminate the states and pass to the RC of the resulting inequality constrained problem, thus arriving at

$$\begin{array}{ll} \min\limits_{C,y_t,w_t,v_t} & C \\ \text{s.t.} & \left.\begin{array}{l} \sum_{t=1}^{T}[c_{\mathrm{h}}' y_t + c_{\mathrm{b}}' w_t + c_{\mathrm{p}}' v_t] \leq C \\ x_1 + \sum_{\tau=1}^{t-1}[v_\tau - \zeta_\tau] \leq y_t,\ 0 \leq y_t \\ -x_1 - \sum_{\tau=1}^{t-1}[x_\tau - \zeta_\tau] \leq w_t,\ 0 \leq w_t \\ \underline{v}_t \leq v_t \leq \overline{v}_t,\ q' y_t \leq Q \end{array}\right\} \forall \zeta \in \mathcal{Z} \end{array} \qquad \text{(RC)}$$

Note: When $\mathcal{Z}$ is a computationally tractable convex set, the semi-infinite problem (RC) is computationally tractable. E.g., when $\mathcal{Z}$ is polyhedral: $\mathcal{Z} = \{\zeta : \exists u : P\zeta + Qu + r \geq 0\}$, RC can be converted into an explicit LP program of sizes polynomial in T, d and the sizes of the representation of $\mathcal{Z}$.

♣ **Extending the notion of RC:** *Adjustable/Affinely Adjustable Robust Counterpart*.

♠ Assumption "All decisions x_j should be independent of ζ" is too restrictive in many applications:

• Some of x_j are "analysis variables" which do not represent decisions at all and can therefore depend on the entire data.

> **Examples:** • Converting the constraint $\sum_i |a_i^T x - b_i| \leq t$ with uncertain a_i, b_i into $-y_i \leq a_i^T x - b_i \leq y_i$, $\sum_i y_i \leq t$, it is natural to allow the analysis variables y_i to "adjust themselves" to the actual data.
>
> • In the Inventory problem, the actual decisions are the replenishment orders v_t and the inventory management cost C; the remaining variables x_t, y_t, w_t are analysis ones, and we can allow these variables to "adjust themselves" to the actual data.

- In dynamical decision-making, some of the decisions x_j should be made when the actual data becomes partially known and thus can depend on the corresponding portions of the data

  Example: In the Inventory problem with uncertain demand, replenishment orders v_t of day t usually can depend on the actual demands at days $1, ..., t-1$.

♣ To account for adjustability, we allow for every x_j to depend on a prescribed portion $P_j\zeta$ of ζ: $x_j = X_j(P_j\zeta)$, thus arriving at Adjustable Robust Counterpart

$$\min_{t,\{X_j(\cdot)\}_{j=1}^n} \{t : \forall \zeta \in \mathcal{Z} : f(X(\zeta),\zeta) \le t, F(X(\zeta),\zeta) \in \mathbf{K}\} \quad \text{(ARC)}$$
$$[X(\zeta) = \{X_j(P_j\zeta)\}]$$

Note: ARC is infinite-dimensional and thus is typically heavily computationally intractable. Seemingly the only applicable technique is Dynamic Programming ⇒ *"curse of dimensionality"*

♣ To overcome, to some extent, intractability of ARC, we restrict the decision rules to be *affine*: $X_j(P_j\zeta) = \xi_j^0 + \xi_j^T P_j\zeta$, thus arriving at the Affinely Adjustable Robust Counterpart

$$\min_{t,\{\xi_j^0,\xi_j\}_{j=1}^n} \{t : \forall \zeta \in \mathcal{Z} : f(X(\zeta),\zeta) \le t, F(X(\zeta),\zeta) \in \mathbf{K}\} \quad \text{(AARC)}$$
$$[X(\zeta) = \{\xi_j^0 + \xi_j^T P_j\zeta\}_{j=1}^n]$$

Example: The only "actual decisions" in the Inventory problem are orders v_t. Assume that v_t can depend on the preceding demands $\zeta^{t-1} = [\zeta_1; ...; \zeta_{t-1}]$. To build the AARC, we

- introduce linear decision rules for the orders $v_t = v_t^0 + V_t\zeta^{t-1}$
- make x_t, y_t, w_t affine functions of ζ: $x_t = x_t^0 + X_t\zeta$, $y_t = y_t^0 + Y_t\zeta$, $w_t = w_t^0 + W_t\zeta$, thus ending up with

$$\begin{array}{ll}\min\limits_{C,v_t^0,V_t,\ldots,w_t^0,W_t} & C \\ \text{s.t.} & \left.\begin{array}{l}\sum_{t\le T}[c_{\mathrm{h}}'[y_t^0+Y_t\zeta]+c_{\mathrm{b}}'[w_t+W_t\zeta]+c_{\mathrm{p}}'[v_t^0+V_t\zeta^{t-1}]]\le C\\ x_{t+1}^0+X_{t+1}\zeta = x_t^0+X_t\zeta+v_t^0+V_t\zeta^{t-1}-\zeta_t\\ x_t^0+X_t\zeta\le y_t^0+Y_t\zeta,\ 0\le y_t^0+Y_t\zeta\\ -[x_t^0+X_t\zeta]\le w_t^0+W_t\zeta,\ 0\le w_t^0+W_t\zeta\\ \underline{v}_t\le v_t^0+V_t\zeta^{t-1}\le\overline{v}_t,\ q'[y_t^0+Y_t\zeta]\le Q\end{array}\right\}\forall\zeta\in\mathcal{Z}\end{array} \tag{AARC}$$

Note: The AARC of the Inventory problem is computationally tractable provided that $\mathcal{Z}$ is so. E.g., when $\mathcal{Z}$ is a polyhedral set, (AARC) is equivalent to an explicit LP program.

Example (continued): Consider *single-product* Inventory with $N = 10$ and a box uncertainty set: $(1-\rho)\zeta^{\mathrm{n}} \leq \zeta \leq (1+\rho)\zeta^{\mathrm{n}}$. Here the ARC is well within the grasp of Dynamic Programming.

♣ *How large are the gaps in the chain* Opt(ARC) $\leq$ Opt(AARC) $\leq$ Opt(RC) *?*

• We built a sample of 768 Inventory problems with uncertainty of 10% − 50% by picking at random cost coefficients, storage capacity and nominal demand trajectory ζ^{n} and subsequent filtering out problems with infeasible ARC's.

♠ It turns out that Opt(ARC) = Opt(AARC) in *every one* of these 768 problems!

Note: This phenomenon disappears when passing from minimizing the worst-case inventory management cost to minimizing the average of this cost.

♠ Opt(RC) was typically essentially worse than Opt(ARC) = Opt(AARC):

Range of $\frac{\text{Opt(RC)}}{\text{Opt(ARC)}}$	1	$(1, 2]$	$(2, 10]$	$(10, 1000]$	∞
Frequency in the sample	38%	23%	14%	11%	15%

♣ *In the RO context, affine decision rules not necessarily are bad!*

- Optimization programs with uncertain data and their Robust Counterparts
- Tractability of Robust Counterparts
- Robust Optimization and Chance Constraints

♣ Robust Counterparts of uncertain problem are *semi-infinite* programs and thus can be intractable even when all instances of the uncertain problem are easy to solve.

⇒ *When Robust Counterparts are computationally tractable?*
What to do if it is not the case?

♠ We focus on *uncertain affinely perturbed LP/CQP/SDP problems*

$$\left\{\min_x \left\{c_\zeta^T x + d_\zeta : A_\zeta^i x + b_\zeta^i \in \mathbf{K}_i,\, i = 1, ..., m\right\} : \zeta \in \mathcal{Z}\right\}$$

with fixed recourse:

- $c_\zeta, d_\zeta, A_\zeta^i, b_\zeta^i$: affine in ζ
- Fixed recourse [automatically valid for the RC]: All coefficients of the *adjustable* variables x_j (those with $P_j \neq 0$) are certain (i.e., independent of ζ).
- $\mathbf{K}_i$: nonnegative rays/Lorentz cones/semidefinite cones (uncertain LP, CQP, SDP, respectively).

♠ We always assume that $\mathcal{Z}$ *is given by a strictly feasible semidefinite representation*

$$\mathcal{Z} = \{\zeta : \exists u : \mathcal{P}(\zeta, u) \succ 0\}$$

($\mathcal{P}(\cdot)$: affine in (ζ, u)).

♣ Investigating tractability of Robust Counterparts of uncertain affinely perturbed LP/CQP/SDP problems with fixed recourse reduces to investigating tractability of *semi-infinite affinely perturbed conic inequalities*

$$\forall \zeta \in \mathcal{Z} : A_\zeta x + b_\zeta \in \mathbf{K} = \left\{ \begin{array}{ll} \text{nonnegative ray} & \text{[Uncertain LP]} \\ \text{Lorentz cone} & \text{[Uncertain CQP]} \\ \text{semidefinite cone} & \text{[Uncertain SDP]} \end{array} \right.$$

$$\left[A_\zeta, b_\zeta\text{: affine in } \zeta \right]$$

♣ Tractability of a semi-infinite affinely perturbed conic inequality

$$\forall \zeta \in \mathcal{Z} : A_\zeta x + b_\zeta \in \mathbf{K}$$

depends on the tradeoff between the geometries of $\mathbf{K}$ and $\mathcal{Z}$ – the more complicated is $\mathcal{Z}$, the simpler should be $\mathbf{K}$.

♣ **"Trivial case": Scenario-generated uncertainty set**

Theorem. The RC/AARC of an uncertain affinely perturbed LP/CQP/SDP problem

$$\left\{\min_x \left\{c_\zeta^T x + d_\zeta : A_\zeta^i x + b_\zeta^i \in \mathbf{K}_i,\ i = 1, ..., m\right\} : \zeta \in \mathcal{Z}\right\}$$

with fixed recourse and with *scenario-generated* uncertainty set $\mathcal{Z} = \text{Conv}\{\zeta^1, ..., \zeta^N\}$ is computationally tractable.

♣ "Solvable case": Uncertain LP

Theorem. The RC/AARC of uncertain affinely perturbed LP problem

$$\left\{\min_x \left\{c_\zeta^T x + d_\zeta : A_\zeta^i x + b_\zeta^i \in \mathbf{K}_i,\, i = 1,...,m\right\} : \zeta \in \mathcal{Z}\right\}$$
$$[\mathbf{K}_i : \text{given by explicit lists of linear inequalities}]$$

with fixed recourse is computationally tractable. With $\mathcal{Z}$ given by a strictly feasible LP/CQP/SDP representation, the RC/AARC is an explicit LP/CQP/SDP program of sizes polynomial in the size of instances and the size of the representation of the uncertainty set.

♣ Aside of a number of highly specific particular cases, semi-infinite conic quadratic/linear matrix inequalities are computationally intractable. Whenever it is the case, a natural course of actions in the RO context is to replace an intractable semi-infinite conic inequality with its *safe tractable approximation*.

Definition. Consider a semi-infinite conic inequality

$$\forall \zeta \in \mathcal{Z} : A_\zeta x + b_\zeta \in \mathbf{K} \qquad \text{(C)}$$

and let $0 \in \mathcal{Z}$. We embed (C) into the parametric family of semi-infinite conic inequalities

$$\forall (\zeta \in \rho\mathcal{Z}) : A_\zeta x + b_\zeta \in \mathbf{K} \qquad (\mathrm{C}_\rho)$$

($\rho \geq 0$: *uncertainty level*).

A system of convex constraints (S_ρ) in variables x and additional variables u is called a *safe tractable approximation* of (C_ρ) tight within factor $\vartheta \geq 1$, if

- [tractability] The constraints in (S_ρ) are efficiently computable
- [safety] Whenever x can be extended to a feasible solution of (S_ρ), x is feasible for (C_ρ)
- [tightness] Whenever x can*not* be extended to a feasible solution of (S_ρ), x is *not* feasible for ($\mathrm{C}_{\vartheta\rho}$).

♣ **Safe tractable approximation of semi-infinite Conic Quadratic Inequality**

$$\forall \zeta = [\zeta^\ell; \zeta^r] \in \rho\mathcal{Z} : \|A_{\zeta^\ell} x - b_{\zeta^\ell}\|_2 \leq c_{\zeta^r}^T x - d_{\zeta^r} \qquad (\mathrm{C}_\rho)$$
$$\left[A_{\zeta^\ell}, b_{\zeta^\ell}, c_{\zeta^r}, d_{\zeta^r} \text{ are affine in } \zeta\right]$$

♠ (C_ρ) **is computationally tractable** when:

• [simple ellipsoidal uncertainty] $\mathcal{Z}$ is an ellipsoid centered at the origin

• [side-wise uncertainty with unstructured norm-bounded lhs perturbations] $\mathcal{Z} = \mathcal{Z}^\ell \times \mathcal{Z}^r$, $\mathcal{Z}^\ell = \{\zeta^\ell \in \mathbb{R}^{p\times q} : \|\zeta^\ell\| \leq 1\}$ and

$$A_{\zeta^\ell} x - b_{\zeta^\ell} \equiv P(x) + L^T(x)\zeta^\ell R(x),$$

where $P(x)$, $L(x)$, $R(x)$ are affine in x and either $L(\cdot)$, or $R(\cdot)$ are constant.

Example: The semi-infinite Least Squares inequality

$$\forall(\zeta \in \mathbb{R}^{p\times q}, \|\zeta\| \le \rho) : \left\| \left[[A^{\mathrm{n}}, b^{\mathrm{n}}] + L^T \zeta R \right] [x; 1] \right\|_2 \le t \qquad (\mathrm{C}_\rho)$$

admits *exact* SDP representation

$$\left[\begin{array}{c|c|c} tI - \lambda L^T L & & A^{\mathrm{n}}x + b^{\mathrm{n}} \\ \hline & \lambda I & \rho R[x;1] \\ \hline [A^{\mathrm{n}}x + b^{\mathrm{n}}]^T & \rho[R[x;1]]^T & t \end{array}\right] \succeq 0 \qquad (\mathrm{S}_\rho)$$

in variables t, x, λ: (t, x) is feasible for (C_ρ) iff it can be extended to a feasible solution to (S_ρ).

$$\forall \zeta = [\zeta^\ell; \zeta^r] \in \rho\mathcal{Z} : \|A_{\zeta^\ell} x - b_{\zeta^\ell}\|_2 \leq c_{\zeta^r}^T x - d_{\zeta^r} \qquad (\mathrm{C}_\rho)$$

♠ (C_ρ) **admits tight tractable safe approximation** when the uncertainty is side-wise: $\mathcal{Z} = \mathcal{Z}^\ell \times \mathcal{Z}^r$ and

- [$\cap$-ellipsoidal lhs perturbation set]

$$\mathcal{Z}^\ell = \{\zeta^\ell : [\zeta^\ell]^T Q_j \zeta^\ell \leq 1,\ 1 \leq j \leq M\} \text{ with } Q_j \succeq 0, \sum_j Q_j \succ 0$$

$\Rightarrow$$(\mathrm{C}_\rho)$ *admits an explicit safe SDP approximation tight within the factor* $\vartheta = O(1)\sqrt{\ln M}$.

- [structured norm-bounded lhs perturbations]

$$\mathcal{Z}^\ell = \{\zeta^\ell = \{\zeta_j^\ell\}_{j=1}^J : \zeta_j^\ell \in \mathbb{R}^{p_j \times q_j} : \|\zeta_j^\ell\| \leq 1\}$$
$$A_{\zeta^\ell} x - b_{\zeta^\ell} \equiv P(x) + \sum_j L_j^T(x) \zeta_j^\ell R_j(x)$$

where $P(x)$, $L_j(x)$, $R_j(x)$ are affine in x and for every j either $L_j(\cdot)$, or $R_j(\cdot)$ are constant.

$\Rightarrow$$(\mathrm{C}_\rho)$ *admits an explicit safe SDP approximation tight within the factor* $\vartheta = \frac{\pi}{2}$.

Example: ["Robust Least Squares Antenna/Filter Design"] The semi-infinite Least Squares inequality

$$\forall(\zeta, \|\zeta\|_\infty \leq \rho) : \|A^{\mathrm{n}}(I + \mathrm{Diag}\{\zeta\})x - b^{\mathrm{n}}\|_2 \leq t \qquad (\mathrm{C}_\rho)$$

with "implementation errors $x_j \mapsto (1 + \zeta_j)x_j$" admits safe SDP approximation

$$\left[\begin{array}{c|c|c} t - \sum_j \lambda_j & [A^{\mathrm{n}}x - b^{\mathrm{n}}]^T & \\ \hline A^{\mathrm{n}}x - b^{\mathrm{n}} & tI & \rho A^{\mathrm{n}}\mathrm{Diag}\{x\} \\ \hline & \rho[A^{\mathrm{n}}\mathrm{Diag}\{x\}]^T & \mathrm{Diag}\{\lambda_1, ..., \lambda_n\} \end{array}\right] \succeq 0 \qquad (\mathrm{S}_\rho)$$

which is tight within the factor $\frac{\pi}{2}$.

♣ **Safe tractable approximation of semi-infinite Linear Matrix Inequality**

$$\forall(\zeta \in \rho\mathcal{Z}) : \mathcal{A}_\zeta(x) \succeq 0 \qquad (\mathrm{C}_\rho)$$
$$\left[\mathcal{A}_\zeta(x) : \text{ bi-affine in } x,\ \zeta\right]$$

♠ Aside of scenario-generated uncertainty set, the only known case when (C_ρ) admits a tight tractable approximation is the case of *structured norm-bounded uncertainty*:

$$\begin{array}{rcl} \mathcal{Z} & = & \{\zeta = \{\zeta_j\}_{j=1}^M : \zeta_j \in \mathbb{R}^{p_j \times p_j}, \|\zeta_j\| \leq 1, \zeta_j = \xi_j I,\ j \in \mathcal{J}\} \\ \mathcal{A}_\zeta(x) & = & \mathcal{A}^{\mathrm{n}}(x) + \sum_j \left[\mathcal{L}_j^T \zeta_j \mathcal{R}_j(x) + \mathcal{R}_j^T(x)\zeta_j^T \mathcal{L}_j\right] \\ & & \left[\mathcal{A}^{\mathrm{n}}(x), \mathcal{R}_j(x) \text{ are affine in } x\right] \end{array}$$

Theorem. The semi-infinite LMI

$$\forall \left(\{\zeta_j\}_{j=1}^{M} \|\zeta_j\| \leq \rho \, \forall j, \zeta_j = \xi_j I_{p_j},\, j \in \mathcal{J} \right) : \\ \mathcal{A}^{\mathrm{n}}(x) + \sum_j \left[\mathcal{L}_j^T \zeta_j \mathcal{R}_j(x) + \mathcal{R}_j^T(x) \zeta_j^T \mathcal{L}_j \right] \succeq 0 \qquad (\mathrm{C}_\rho)$$

with structured norm-bounded uncertainty admits a safe tractable SDP approximation. The tightness factor ϑ of this approximation depends solely on the largest size of the scalar perturbation blocks

$$\mu = \begin{cases} 1, & \mathcal{J} = \emptyset \\ \max_{j \in \mathcal{J}} p_j, & \mathcal{J} \neq \emptyset \end{cases}$$

and does not exceed $\vartheta(\mu)$, where $\vartheta(\cdot)$ is a universal function such that $\vartheta(1) = \frac{\pi}{2}, \vartheta(2) = 2,\, \vartheta(k) \leq \pi\sqrt{k}$. In the case of $M = 1$ (single perturbation block), the approximation is exact: $\vartheta = 1$.

Applications in: Robust Structural Design, Lyapunov Stability Analysis/Synthesis under interval uncertainty, etc.

- Optimization programs with uncertain data and their Robust Counterparts
- Tractability of Robust Counterparts
- Robust Optimization and Chance Constraints

♣ RO does *not* assume stochastic nature of uncertain data and uses instead uncertain-but-bounded model of data perturbations.

However: *Stochastic nature of uncertainty, if any, can be utilized in the RO framework.*

♣ Consider Uncertain LP with stochastic data. The entity of primary interest here is a randomly perturbed linear constraint

$$w_0(x) + \sum_{\ell=1}^{d} \zeta_\ell w_\ell(x) \leq 0 \qquad (C)$$

$$\left[\begin{array}{l} \bullet\ x\text{: decision vector} \quad \bullet\ w_0(x), ..., w_d(x)\text{: affine} \\ \bullet\ \zeta_1, ..., \zeta_d \in \mathbb{R}\text{: ``primitive'' random perturbations} \end{array}\right]$$

♠ A natural way to process (C) is to pass to a chance version

$$\text{Prob}\left\{w_0(x) + \sum_{\ell=1}^{d} \zeta_\ell w_\ell(x) > 0\right\} \leq \epsilon \qquad (C_\epsilon)$$

of the constraint (A. Charnes, W. Cooper, G. Symonds, 1958).

$$\pi(w) \equiv \text{Prob}\{\zeta^w \equiv w_0 + \sum_{\ell=1}^{d} \zeta_\ell w_\ell > 0\} \leq \epsilon \qquad (C_\epsilon)$$

♠ There exists significant literature on chance constraints (T. Badics, D. Dentcheva, A. Dupačová, L. Miller, A. Prékopa, A. Ruszczynski, B. Vizvari, H. Wagner,...)

However: *In general, (C_ϵ) is difficult to process:*

- *In many cases, the feasible set of a chance constraint is non-convex;*
- *Even when convex, the feasible set of (C_ϵ) can be "computationally intractable":*

> When $\zeta \sim \text{Uniform}([0,1]^d)$, $\pi(w)$ is quasi-convex (C. Lagoa et al., 2005). However, unless P=NP, it is impossible to compute $\pi(w)$ with accuracy $\delta > 0$ in time polynomial in d, total binary length of (rational) w and $\ln(1/\delta)$ [L. Khachiyan, 1989].

$$\pi(w) \equiv \text{Prob}\{\zeta^w \equiv w_0 + \sum_{\ell=1}^d \zeta_\ell w_\ell > 0\} \leq \epsilon \qquad (C_\epsilon)$$

♣ When (C_ϵ) "as it is" is difficult to process, one can look for a *safe tractable approximation of (C_ϵ) – a computationally tractable convex set W_ϵ such that*

$$W_\epsilon \subset \{w : \pi(w) \leq \epsilon\} \qquad (*)$$

♠ A way to build a safe tractable approximation of (C_ϵ):

Take a convex function $\gamma : \mathbb{R} \to \mathbb{R}$ with $\lim_{s \to -\infty} \gamma(s) = 0$, $\gamma(0) = 1$ and set $\Gamma(w) = \mathbf{E}\{\gamma(\zeta^w)\}$.

Note: $\pi(w) \leq \Gamma(\alpha w)\ \forall \alpha > 0$ and $\Gamma(\cdot)$ is convex, whence the set $W_\epsilon = \text{cl}\{w : \exists \alpha > 0 : \Gamma(\alpha w) \leq \epsilon\}$ is convex and satisfies $(*)$.

$\Rightarrow$ *Whenever $\Gamma(\cdot)$ is efficiently computable, W_ϵ is a safe tractable approximation of (C_ϵ).*

$$\gamma(\cdot):\mathbb{R}\to\mathbb{R}_+, \gamma(0)=1: \text{ convex, } \lim_{s\to-\infty}\gamma(s)=0$$

$$\Rightarrow \quad \Gamma(w)\geq \mathbf{E}\left\{\gamma\left(w_0+\sum_{\ell=1}^d\zeta_\ell w_\ell\right)\right\}: \text{ convex}$$

$$\Rightarrow \quad W_\epsilon \equiv \mathrm{cl}\,\{w:\exists\alpha>0:\Gamma(\alpha w)\leq\epsilon\}\subset\{w:\mathrm{Prob}\{\zeta^w>0\}\leq\epsilon\}$$

Note: W_ϵ is a closed convex cone such that

$$W_\epsilon=\{w\in\mathbb{R}^{d+1}:w_0+\sum_{\ell=1}^d\zeta_\ell w_\ell\leq 0\ \forall\zeta\in\mathcal{Z}_\epsilon\}$$

for an appropriate convex compact set $\mathcal{Z}_\epsilon$.

$\Rightarrow$ *The safe approximation $w(x)\in W_\epsilon$ of the chance constraint*

$$\mathrm{Prob}\{w_0(x)+\sum_{\ell=1}^d\zeta_\ell w_\ell(x)>0\}\leq\epsilon$$

is nothing but the Robust Counterpart

$$\forall\zeta\in\mathcal{Z}_\epsilon: w_0(x)+\zeta_1 w_1(x)+...+\zeta_d w_d(x)\leq 0$$

of the uncertain affinely perturbed linear constraint

$$w_0(x)+\zeta_1 w_1(x)+...+\zeta_d w_d(x)\leq 0$$

with properly defined perturbation set $\mathcal{Z}_\epsilon$.

♣ **How to choose $\gamma(\cdot)$?**

♠ As far as the conservatism of the upper bound

$$\text{Prob}\{\zeta^w > 0\} \leq \inf_{\alpha>0} \mathbf{E}\{\gamma(\zeta^{\alpha w})\}$$

$$[\gamma(\cdot) \geq 0: \text{ convex nondecreasing, } \gamma(0) = 1]$$

is concerned, the best choice of $\gamma(\cdot)$ is $\gamma(s) = \max[1 + s, 0]$. This choice leads to the famous *Conditional Value at Risk* safe approximation

$$\underbrace{\min_{\beta\in\mathbb{R}} \left[\beta + \frac{1}{\epsilon}\mathbf{E}\{[\zeta^w - \beta]_+\}\right]}_{\text{CVaR}_\epsilon(w)} \leq 0$$

of the chance constraint

$$\text{Prob}\{\zeta^w > 0\} \leq \epsilon.$$

However: Typically, $\text{CVaR}_\epsilon(\cdot)$ is difficult to compute...

♣ **Ensuring computability: Bernstein approximation [Pinter, 1989; Nem.&Shapiro, 2005]**

Observation: Consider a chance constraint

$$\text{Prob}\{w_0 + \sum_{\ell=1}^{d} \zeta_\ell w_\ell > 0\} \leq \epsilon \qquad (C_\epsilon)$$

and assume that

- $\zeta_1, ..., \zeta_d$ are independent
- We are smart enough to build efficiently computable convex bounds $\Phi_\ell(r) \geq \ln\left(\mathbb{E}\{e^{r\zeta_\ell}\}\right)$ on the logarithmic moment-generating functions of ζ_ℓ, $\ell = 1, ..., d$.

Choosing $\gamma(s) = e^s$, one can set

$$\Gamma(w) = \exp\{w_0 + \sum_{\ell=1}^{d} \Phi_\ell(w_\ell)\}$$

thus arriving at a *tractable* safe approximation of (C_ϵ).

Example (one of many): Range and mean a priori information on ζ_ℓ. Consider a chance constraint

$$\begin{array}{c}\text{Prob}\{w_0(x)+\sum_{\ell=1}^d \zeta_\ell w_\ell(x)>0\}\leq\epsilon\\ \bullet\ \zeta_\ell\in[-1,1]:\ \text{independent}\ \bullet\ \mathbf{E}\{\zeta_\ell\}\in[\mu_\ell^-,\mu_\ell^+]\end{array}\qquad (C_\epsilon)$$

The associated Bernstein approximation of (C_ϵ) is

$$\begin{array}{c}\forall\zeta\in\mathcal{Z}_\epsilon: w_0(x)+\sum_{\ell=1}^d\zeta_\ell w_\ell(x)\leq 0,\\ \mathcal{Z}_\epsilon=\{\zeta:\sum_{\ell=1}^d\phi_\ell(\zeta_\ell)\leq 2\ln(1/\epsilon)\}\\ \phi_\ell(s)=\left\{\begin{array}{ll}(1+s)\ln\left(\frac{1+s}{1+\mu_\ell^-}\right)+(1-s)\ln\left(\frac{1-s}{1-\mu_\ell^-}\right) & ,-1\leq s\leq\mu_\ell^-\\ 0 & ,\mu_\ell^-\leq s\leq\mu_\ell^+\\ (1+s)\ln\left(\frac{1+s}{1+\mu_\ell^+}\right)+(1-s)\ln\left(\frac{1-s}{1-\mu_\ell^+}\right) & ,\mu_\ell^+\leq s\leq 1\end{array}\right.\end{array}\qquad \text{(Br)}$$

$$\underbrace{\begin{array}{c}\forall \zeta \in \mathcal{Z}_\epsilon : w_0(x) + \sum_{\ell=1}^d \zeta_\ell w_\ell(x) \leq 0, \\ \mathcal{Z}_\epsilon = \{\zeta : \sum_{\ell=1}^d \phi_\ell(\zeta_\ell) \leq 2\ln(1/\epsilon)\},\ \phi_\ell(s) = ...\end{array}}_{\text{“Entropy uncertainty”, Nem.\&Shapiro, 2006}} \qquad (Br) \quad \Leftrightarrow \quad \text{Explicit CP}$$

$$\Downarrow$$

$$\text{Prob}\{w_0(x) + \sum_{\ell=1}^d \zeta_\ell w_\ell(x) > 0\} \leq \epsilon \qquad (C_\epsilon)$$

• $\zeta_\ell \in [-1, 1]$: independent • $\mathbf{E}\{\zeta_\ell\} \in [\mu_\ell^-, \mu_\ell^+]$

♠ (Br) can be further safely approximated by

$$\underbrace{\begin{array}{c}\forall \zeta \in \mathcal{Z}_\epsilon^+ : w_0(x) + \sum_\ell \zeta_\ell w_\ell(x) \leq 0 \\ \mathcal{Z}_\epsilon^+ = \{\|\zeta\|_\infty \leq 1\} \bigcap \left[\begin{array}{c}\{\|\zeta\|_2 \leq \sqrt{2\ln(1/\epsilon)}\} \\ +\{\mu^- \leq \zeta \leq \mu^+\}\end{array}\right]\end{array}}_{\text{“Ball-Box uncertainty”, Ben-Tal \& Nem., 2000}} \qquad (BB) \quad \Leftrightarrow \quad \text{Explicit CQP}$$

$$\forall \zeta \in \mathcal{Z}_\epsilon^+ : w_0(x) + \sum_\ell \zeta_\ell w_\ell(x) \leq 0$$
$$\mathcal{Z}_\epsilon^+ = \{\|\zeta\|_\infty \leq 1\} \bigcap \left[\begin{array}{c} \{\|\zeta\|_2 \leq \sqrt{2\ln(1/\epsilon)}\} \\ +\{\mu^- \leq \zeta \leq \mu^+\} \end{array} \right] \quad \text{(BB)}$$

$\Leftrightarrow$ Explicit CQP

$\Downarrow$

$$\forall \zeta \in \mathcal{Z}_\epsilon : w_0(x) + \sum_{\ell=1}^d \zeta_\ell w_\ell(x) \leq 0,$$
$$\mathcal{Z}_\epsilon = \{\zeta : \sum_{\ell=1}^d \phi_\ell(\zeta_\ell) \leq 2\ln(1/\epsilon)\}, \ \phi_\ell(s) = \ldots \quad \text{(Br)}$$

$\Leftrightarrow$ Explicit CP

$\Downarrow$

$$\mathrm{Prob}\{w_0(x) + \sum_{\ell=1}^d \zeta_\ell w_\ell(x) > 0\} \leq \epsilon \quad (C_\epsilon)$$

• $\zeta_\ell \in [-1, 1]$: independent • $\mathbf{E}\{\zeta_\ell\} \in [\mu_\ell^-, \mu_\ell^+]$

♠ (BB) is further safely approximated by

$$\forall \zeta \in \mathcal{Z}_\epsilon^{++} : w_0(x) + \sum_\ell \zeta_\ell w_\ell(x) \leq 0$$
$$\underbrace{\mathcal{Z}_\epsilon^{++} = \{\|\zeta\|_\infty \leq 1\} \bigcap \left[\begin{array}{c} \{\|\zeta\|_1 \leq \sqrt{2d\ln(1/\epsilon)}\} \\ +\{\mu^- \leq \zeta \leq \mu^+\} \end{array} \right]}_{\text{"Budgeted uncertainty", Bertsimas \& Sim, 2003}} \quad \text{(Bd)}$$

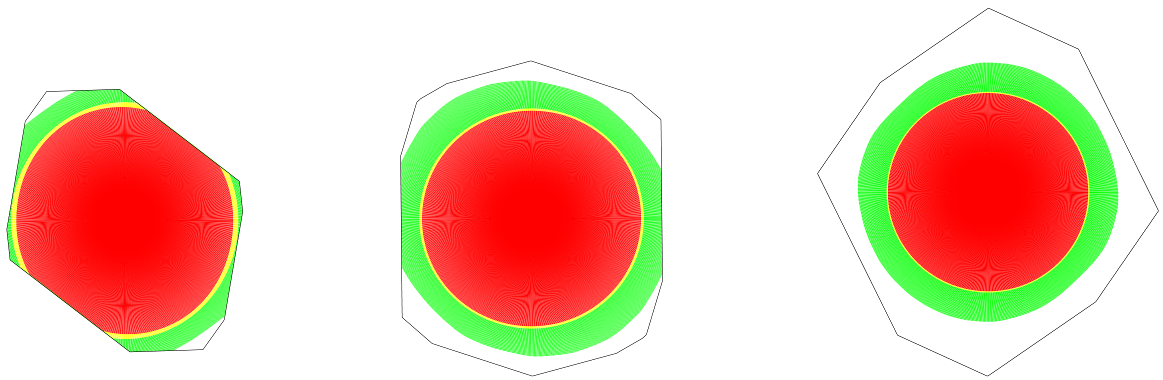
$\Leftrightarrow$ Explicit LP

$$\forall \zeta \in \mathcal{Z} : w_0(x) + \sum_{\ell=1}^{d} \zeta_\ell w_\ell(x) \leq 0 \qquad (\text{RC}[\mathcal{Z}])$$

$$\Downarrow$$

$$\text{Prob}\{w_0(x) + \sum_{\ell=1}^{d} \zeta_\ell w_\ell(x) > 0\} \leq \epsilon \qquad (C_\epsilon)$$

• $\zeta_\ell \in [-1, 1]$: independent • $E\{\zeta_\ell\} \in [\mu_\ell^-, \mu_\ell^+]$



$d = 64$ $d = 128$ $d = 256$

Random 2D cross-sections of Entropy (red), Ball-Box (yellow) and Budgeted (green) uncertainty sets, $\epsilon = 0.005$, $\mu_\ell^\pm = 0$.
Black: cross-section of the support $\{\|\zeta\|_\infty \leq 1\}$ of ζ.