C20 Robust Optimization
Example Paper

Kostas Margellos

kostas.margellos@eng.ox.ac.uk

Michaelmas Term 2024

Problems

1. Let $(\delta_1, \ldots, \delta_m)$ be a given m multi-sample used to construct a hypothesis H_m to approximate a target set T. Assume that $T \supseteq H_m$ for any multi-sample. The Probably Approximately Correct (PAC) learning paradigm is summarized in the following statement

$$\mathbb{P}^m \Big\{ \delta_1, \ldots, \delta_m : \; \mathbb{P}\Big(\delta \in T \setminus H_m \Big) \leq \epsilon \Big\} \geq 1 - q(m, \epsilon),$$

with $\lim_{m \to \infty} q(m, \epsilon) = 0$.

   (a) What are the random variables and events, and with respect to which probabilities are they measured? Justify your answers in all cases.

   (b) Comment on the roles of ϵ and $q(m, \epsilon)$? Why do we require that $\lim_{m \to \infty} q(m, \epsilon) = 0$?

   (c) A decision maker informs you that for this learning problem there exists a compression set with cardinality $d < m$. Fix any $\beta \in (0, 1)$. Determine an upper-bound $\epsilon(m, \beta)$ on $\mathbb{P}\Big(\delta \in T \setminus H_m \Big)$ as a function of m and β, such that, with confidence at least $1 - \beta$, the probability that H_m does not agree with T on a new realization δ of the uncertainty is at most $\epsilon(m, \beta)$.
   Note that your answer will also depend on d.

2. Consider the following $\min - \max$ uncertain optimization problem

$$\min_{x \in \mathbb{R}^{n_x}} \; \max_{\delta \in \Delta} \; f(x, \delta),$$

where x is a vector with n_x decision variables, δ is an uncertain parameter taking values in the set Δ, and f is convex in x. Let $\delta, \ldots, \delta_m$ be m

independent samples of δ according to a possibly unknown distribution $\mathbb{P}$, and consider the following $\min - \max$ scenario program

$$\min_{x \in \mathbb{R}^{n_x}} \max_{i=1,\ldots,m} f(x, \delta_i).$$

(a) Perform an epigraphic reformulation of the scenario program, introducing the scalar epigraphic variable $\gamma \in \mathbb{R}$. Provide a mathematical expression for the probability that the constraints of the reformulated problem are violated when a new sample δ is extracted.

(b) For a given $\epsilon \in (0, 1)$, provide an expression for the confidence $1 - q(m, \epsilon)$ with which the probability of constraint violation of part (a) is at most equal to ϵ.

(c) What is the implication of your answer in part (b) on the original $\min - \max$ scenario program.

3. Consider m samples in $\mathbb{R}$, i.e., $\delta_1, \ldots, \delta_m$, $i = 1, \ldots, m$, extracted according to an unknown probability measure $\mathbb{P}$. Consider the problem of determining the minimum length interval that contains all samples. See Figure 1 for a pictorial representation.

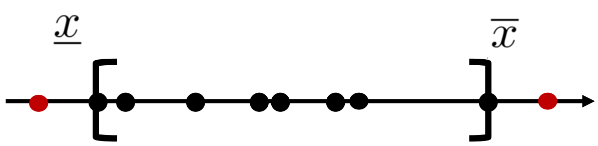

Figure 1: Minimum width interval.

(a) Parameterize the interval using either its start- and end-point (denoted by $\underline{x}$ and $\overline{x}$, respectively in Figure 1), or equivalently, by its center and semi-width length. Show that the problem of determining the minimum length interval enclosing all samples can be written as a convex scenario program.

(b) *Probabilistic robustness:* Determine the number of samples m that would be required to guarantee that, with confidence at least $1 - 10^{-6}$, the probability that a new sample extracted according to $\mathbb{P}$ lies outside the minimum length interval is at most 5%.

4. Consider $m = 1650$ points with coordinates $(u_i, y_i) \in \mathbb{R}^2$, $i = 1, \ldots, m$. The points are extracted independently according to an unknown, contunuous probability measure $\mathbb{P}$. Consider the problem of determining the minimum vertical width strip that contains all points. See Figure 2 for a pictorial representation[1].

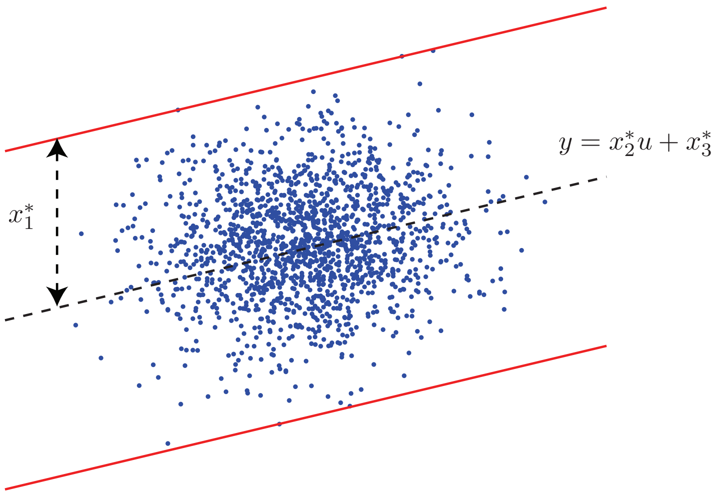

Figure 2: Minimum vertical width strip.

   (a) Consider the parameterization shown in Figure 2, where x_1 denotes the semi-width length, and x_2, x_3 encode the median line of the strip. Formulate the problem of determining the minimum vertical width strip that contains all points as a convex scenario program, and denote by (x_1^*, x_2^*, x_3^*) its optimal solution.

   (b) *Probabilistic robustness:* Provide a probabilistic certificate (confidence), with which the probability that a new sample extracted according to $\mathbb{P}$ lies outside the minimum vertical width strip is at most 1%.

   (c) How does the computed confidence compare with the one computed in the minimum radius disk program worked out in the lecture notes? Which result is more conservative and why?

5. Let $\delta_1, \ldots, \delta_m$ be m independent samples of an uncertain parameter δ, distributed according to a possibly unknown distribution $\mathbb{P}$. Consider the

[1]Figure taken from "The exact feasibility of randomized solutions of uncertain convex programs", by M. Campi and S. Garatti, SIAM Journal on Optimization, 19(3), 1211-1230, 2008.

following family of autonomous, linear dynamical systems/plants

$$\dot{x}(t) = A(\delta_i)x(t), \text{ for } i = 1, \ldots, m, \tag{1}$$

where $x \in \mathbb{R}^n$ is the system state, and for each realization of the uncertain parameter δ, $A(\delta) \in \mathbb{R}^{n\times n}$ governs its evolution.

Let $\gamma \in \mathbb{R}$, and consider the following minimization program

$$\begin{aligned} &\min_{P,\gamma} \ \gamma \\ &\text{subject to} \quad P = P^\top \succ 0 \\ &\qquad -I \preceq A^\top(\delta_i)P + PA(\delta_i) \preceq \gamma I, \ \text{for all } i = 1, \ldots, m, \end{aligned}$$

where I is an identity matrix with appropriate dimension.

(a) *Quadratic stability:* Provide a range of values for γ for the family of plants in (1) to be asymptotically stable.

(b) What type of probabilistic statement can you offer for the optimal solutions P^*, γ^* of the aforementioned program?
Comment on the probability that a new new plant (i.e., a new δ giving rise to $A(\delta)$), is asymptotically stable.

Note that the lower bound $-I$ in the constraint is introduced to ensure boundedness of the solutions, as when the matrix $A^\top(\delta_i)P + PA(\delta_i)$ becomes negative for some $i = 1, \ldots, m$ (and since it depends linearly on P), γ could drift to negative infinity.

6. Figure 3 shows three different scenario programs, each of them with two decision variables x_1 and x_2. In all cases, each constraint is V-shaped (e.g., 1-norm constraint). The feasibility region for each constraint is outside the shaded part, and the arrow indicates the optimization direction, corresponding to minimizing x_2.

   (a) For each case indicate which constraints are of support.

   (b) Which of those cases correspond to convex minimization programs? How is your answer on the number of support constraints related to convexity?

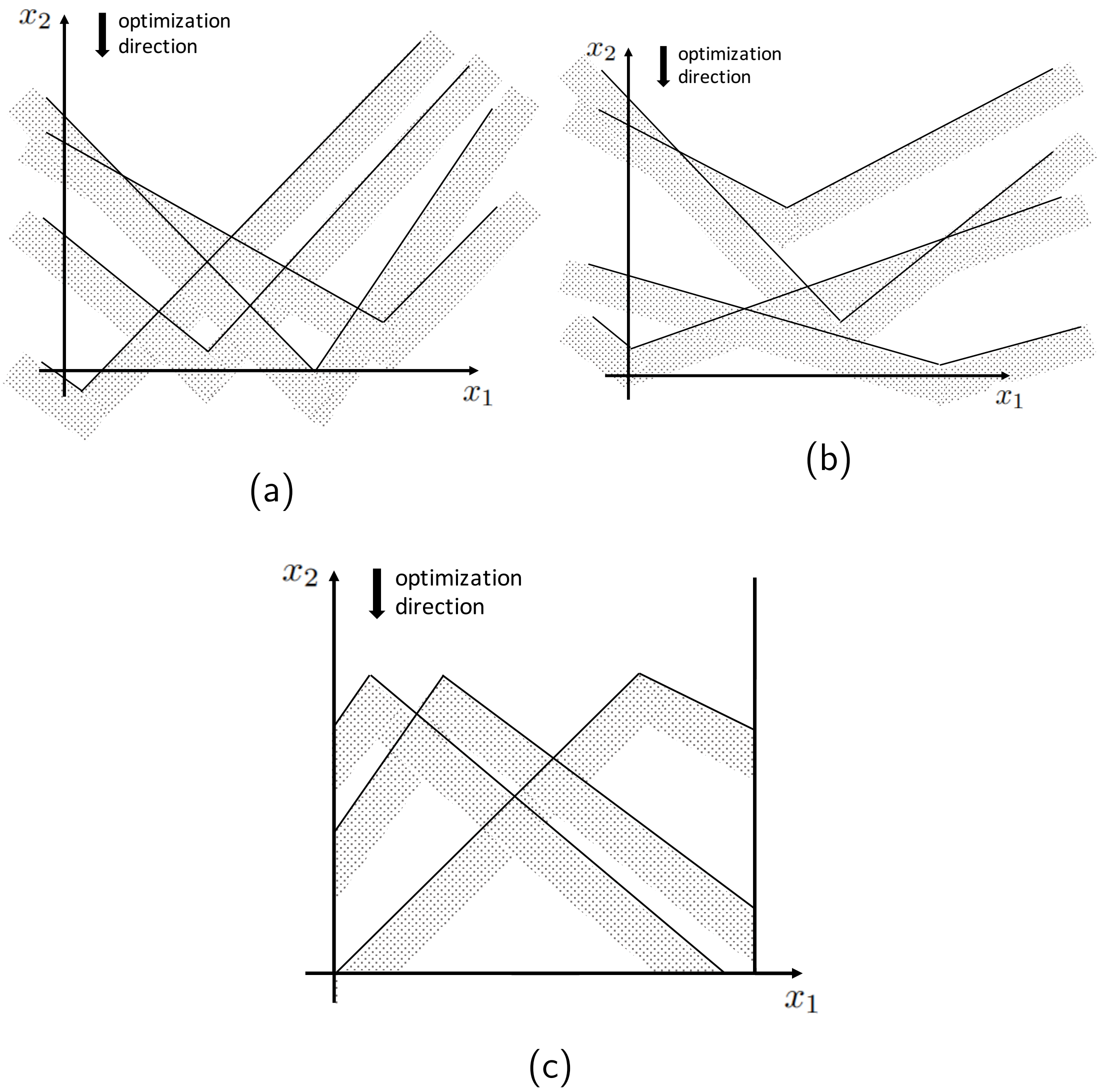

Figure 3: Different scenario programs with V-shaped constraints.

Justify your answers in all cases.

7. *Bounds on the expected value of the probability of constraint violation.*

   (a) Consider the minimum width interval of Problem 3. Determine the number of samples that would be required for the expected value of the probability that a new sample lies outside the minimum width interval to be less than or equal to 0.05.

   (b) Consider now the minimum vertical width strip of Problem 4. Determine the number of samples that would be required for the expected value of the probability that a new sample lies outside the minimum vertical width strip to be less than or equal to 0.01.