Convex Robust Optimization

Boris Houska

Numerical Optimization

Typical objectives:

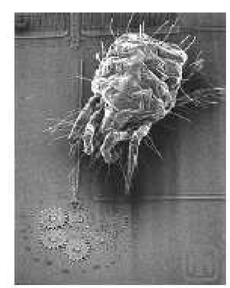
- Minimize traveling time,
- Minimze costs,
- Reduce emissions,
- Save energy, ...

Mathematical Formulation:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & F_0(x) \\ \text{subject to} \quad & F_i(x) \leq 0 \end{aligned}$$

Many engineering applications:






What is Robust Optimization?

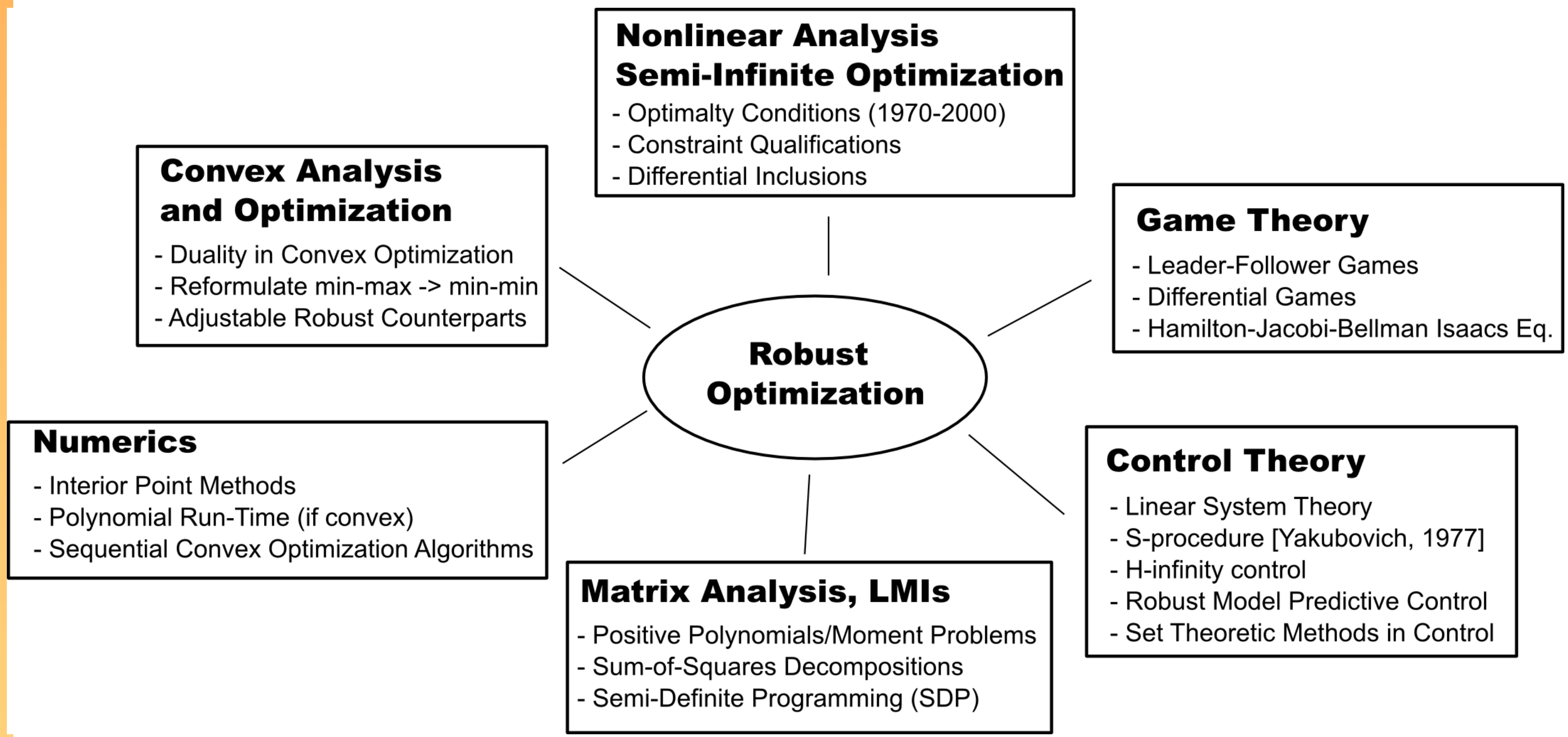
In practice:

- Mismatch between mathematical model and real world
- External disturbances
- How to ensure safe operation?

Robust formulation:

$$\begin{array}{ll} \min\limits_{x} & \max\limits_{w \in W} F_0(x,w) \\ \text{s.t.} & \max\limits_{w \in W} F_i(x,w) \leq 0 \end{array}$$

Overview: Robust Optimization



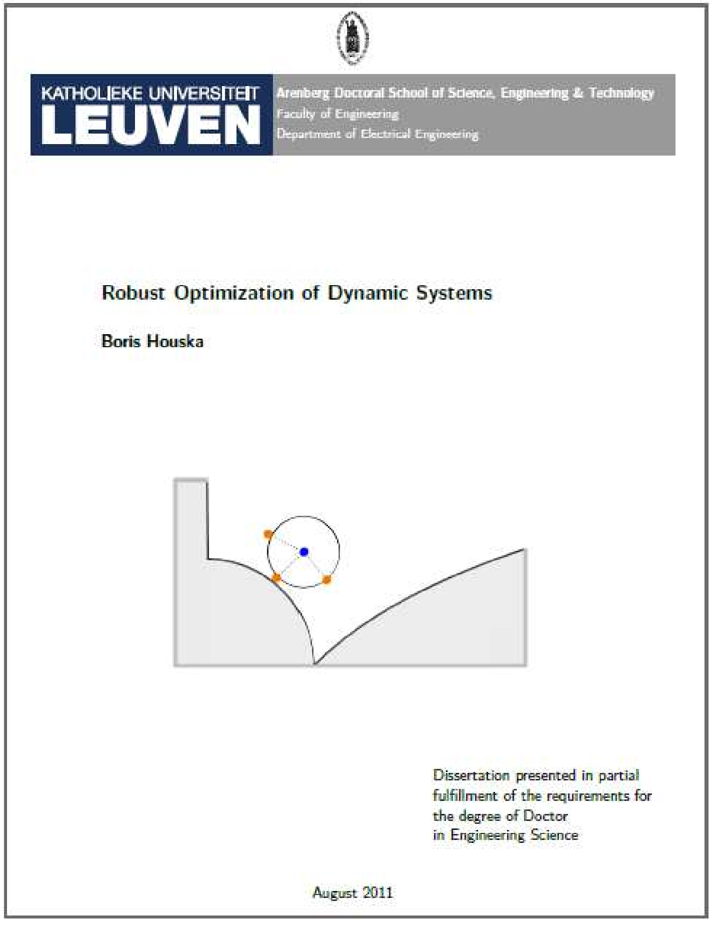
PhD Thesis

Part I:

- Finite Dimensional Robust Optimization

Part II:

- Robust Optimization of Dynamic Systems



KATHOLIEKE UNIVERSITEIT LEUVEN

Arenberg Doctoral School of Science, Engineering & Technology
Faculty of Engineering
Department of Electrical Engineering

Robust Optimization of Dynamic Systems

Boris Houska

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor in Engineering Science

August 2011

Overview

- **The convex optimization perspective on robust optimization**
- The S-procedure for Quadratic Forms
- Inner- and Outer Ellipsoidal Approximations

Semi-Infinite Optimization Problems

Notation: x denotes optimization variable, w denotes uncertainty.

Semi-Infinite Optimization Problems

Notation: x denotes optimization variable, w denotes uncertainty.

Robust Feasibility Problem:

$$\mathcal{F} \quad := \quad \left\{ x \in \mathbb{R}^{n_x} \;\middle|\; \begin{array}{rrcl} \forall w \in W : & F_1(x,w) & \leq & 0 \\ & & \vdots & \\ \forall w \in W : & F_m(x,w) & \leq & 0 \end{array} \right\} .$$

Semi-Infinite Optimization Problems

Notation: x denotes optimization variable, w denotes uncertainty.

Robust Feasibility Problem:

$$\mathcal{F} \quad := \quad \left\{ x \in \mathbb{R}^{n_x} \;\middle|\; \begin{array}{rrcl} \forall w \in W : & F_1(x,w) & \leq & 0 \\ & \vdots & & \\ \forall w \in W : & F_m(x,w) & \leq & 0 \end{array} \right\} .$$

Semi-Infinite Optimization Problem (SIP):

$$\min_{x} \max_{w \in W} \; F_0(x,w) \quad \text{s.t.} \quad x \in \mathcal{F} \; ,$$

Equivalent Min-Max Formulation

Lower-level robust counterpart functions:

$$\forall x \in \mathbb{R}^n : \qquad V_i(x) \quad = \quad \max_{w \in W} \; F_i(x,w) \qquad \text{with} \;\; i \in \{0, \ldots, m\} \; .$$

Equivalent Min-Max Formulation

Lower-level robust counterpart functions:

$$\forall x \in \mathbb{R}^n : \qquad V_i(x) \;=\; \max_{w \in W} \; F_i(x, w) \qquad \text{with} \;\; i \in \{0, \ldots, m\} \; .$$

Equivalent bi-level formulation:

$$\min_x \;\; V_0(x) \quad \text{s.t.} \quad V_i(x) \, \leq \, 0 \quad \text{for all} \;\; i \in \{1, \ldots, m\} \; .$$

Equivalent Min-Max Formulation

Lower-level robust counterpart functions:

$$\forall x \in \mathbb{R}^n : \qquad V_i(x) \;=\; \max_{w \in W} \; F_i(x,w) \qquad \text{with} \;\; i \in \{0, \ldots, m\} \; .$$

Equivalent bi-level formulation:

$$\min_x \; V_0(x) \quad \text{s.t.} \quad V_i(x) \,\leq\, 0 \quad \text{for all} \;\; i \in \{1, \ldots, m\} \; .$$

Problem: we have a bi-level problem: parametric lower-level maximization and upper level minimization.

Special Cases

Observation: If we can find $V_i(x)$ explicitly, we obtain a standard NLP.

Special Cases

Observation: If we can find $V_i(x)$ explicitly, we obtain a standard NLP.

Notation for Ellipsoids:

$$\mathcal{E}(Q,q) \;\; = \;\; \left\{ q + Q^{\frac{1}{2}} v \;\middle|\; \exists v \in \mathbb{R}^n : \;\; v^T v \;\leq\; 1 \right\} .$$

Special Cases

Observation: If we can find $V_i(x)$ explicitly, we obtain a standard NLP.

Notation for Ellipsoids:

$$\mathcal{E}(Q,q) \;=\; \left\{ q + Q^{\frac{1}{2}} v \;\middle|\; \exists v \in \mathbb{R}^n : \;\; v^T v \;\leq\; 1 \right\} .$$

Example: Functions F_i uncertainty affine:

$$F_i(x,w) \;=\; c_i(x)^T w + d_i(x)$$

for some functions $c_i : \mathbb{R}^{n_x} \to \mathbb{R}^{n_w}$ and $d_i : \mathbb{R}^{n_x} \to \mathbb{R}$, while the set $W \;:=\; \mathcal{E}(Q,q)$ is an ellipsoid. Then:

$$V_i(x) \;=\; \max_{w \in \mathcal{E}(Q,q)} \; c_i(x)^T w + d_i(x) \;\;=\;\; \sqrt{c_i(x)^T Q c_i(x)} + c_i(x)^T q + d_i(x) \, .$$

Other Special Cases

Example: robust least squares [El-Ghaoui and Lebret, 1997]

$$F_i(x, w) \;:=\; \|\,(A + \Delta)x\,\|_2 \;-\; d$$

Uncertainty vector can be written as $w := \text{vec}(\Delta)$. For ellipsoidal uncertainty we may assume suitable scaling:

$$W \quad := \quad \{\; \Delta \mid \; \|\Delta\|_{\mathrm{F}} \;\leq\; 1 \,\} \;.$$

Other Special Cases

Example: robust least squares [El-Ghaoui and Lebret, 1997]

$$F_i(x, w) \;:=\; \| \, (A + \Delta)x \, \|_2 \; - \; d$$

Uncertainty vector can be written as $w := \text{vec}(\Delta)$. For ellipsoidal uncertainty we may assume suitable scaling:

$$W \quad := \quad \{ \; \Delta \; | \; \|\Delta\|_{\mathrm{F}} \; \leq \; 1 \, \} \; .$$

Use the triangle inequality:

$$\| \, (A + \Delta)x \, \|_2 \; \leq \; \| \, Ax \, \|_2 + \| \, \Delta x \, \|_2 \; \leq \; \| \, Ax \, \|_2 + \| \, x \, \|_2 \; .$$

This inequality is tight for $\Delta^* \; := \; \frac{A \, x x^T}{\|Ax\| \; \|x\|}$.

Other Special Cases

Example: robust least squares [El-Ghaoui and Lebret, 1997]

$$F_i(x, w) \;:=\; \|\,(A + \Delta)x\,\|_2 \;-\; d$$

Uncertainty vector can be written as $w :=$ vec(Δ). For ellipsoidal uncertainty we may assume suitable scaling:

$$W \;\;:=\;\; \{\;\Delta \mid \;\|\Delta\|_{\mathrm{F}} \;\leq\; 1\,\}\;.$$

We have found that

$$V_i(x) \;=\; \max_{\Delta \in W} \;\|\,(A + \Delta)x\,\|_2 \;-\; d \;\;=\;\; \|\,Ax\,\|_2 + \|\,x\,\|_2 \;-\; d\;.$$

Other Special Cases

Example: robust least squares [El-Ghaoui and Lebret, 1997]

$$F_i(x, w) \;:=\; \|\,(A + \Delta)x\,\|_2 \;-\; d$$

Uncertainty vector can be written as $w :=$ vec(Δ). For ellipsoidal uncertainty we may assume suitable scaling:

$$W \quad := \quad \{\;\Delta \mid \;\|\Delta\|_{\mathrm{F}} \;\leq\; 1\,\}\;\;.$$

For $A := \left(\widehat{A}, b\right)$, $\Delta \;:=\; \left(\widehat{\Delta},\, \delta\right)$, and $x \;:=\; \left(y^T,\, 1\,\right)^T$:

$$\begin{aligned} &\min_y \max_{\|\Delta\|_{\mathrm{F}}^2 + \|\delta\|_2^2 \,\leq\, 1} \left\|(\widehat{A} + \widehat{\Delta})y + (b + \delta)\right\|_2 \\ &= \min_y \left\|\widehat{A}y + b\right\|_2 + \sqrt{\|y\|_2^2 + 1}\;, \end{aligned}$$

Special Cases (cont.)

Robust SOCP (Ben-Tal and Nemirovski, 1998)

$$F_i(x, w) \ := \ \| \, (A + \Delta)x \, \|_2 - (c + \delta)^T x \ ,$$

$\Delta \in \mathbb{R}^{m \times n}$ and $\delta \in \mathbb{R}^n$ are unknown:

$$W \ = \ \{ \, (\Delta, \delta) \ | \ \|\Delta\|_{\mathrm{F}} \ \leq \ 1 \ \text{ and } \ \|\delta\|_2 \ \leq \ 1 \ \}$$

Special Cases (cont.)

Robust SOCP (Ben-Tal and Nemirovski, 1998)

$$F_i(x, w) \;:=\; \|\,(A + \Delta)x\,\|_2 - (c + \delta)^T x\;,$$

$\Delta \in \mathbb{R}^{m\times n}$ and $\delta \in \mathbb{R}^n$ are unknown:

$$W \;=\; \{\,(\Delta, \delta)\;|\;\|\Delta\|_{\mathrm{F}} \,\le\, 1 \;\text{ and }\; \|\delta\|_2 \,\le\, 1\,\}$$

Combine the results from the previous two examples:

$$V_i(x) \;:=\; \max_{(\Delta,\delta)\in W}\;\|\,(A + \Delta)x\,\|_2 - (c+\delta)^T x \;\;=\;\|\,Ax\,\|_2 - c^T x + 2\,\|\,x\,\|_2\;.$$

With the same triangle-inequality trick: LPs, QPs, or QCQPs with uncertain data can all be written as SOCPs.

The Convex Optimization Perspective

Recall:

$$\min_x \; V_0(x) \quad \text{s.t.} \quad V_i(x) \leq 0 \quad \text{for all} \;\; i \in \{1, \ldots, m\} \; .$$

Definition of Lower Level Convexity: We say that a robust optimization problem is lower level convex if the uncertainty set W is convex, while the functions $F_i(x, \cdot) : W \to \mathbb{R}$ are for all indices $i \in \{1, \ldots, m\}$ and for all $x \in \mathcal{F}$ concave functions in w.

Duality: from Min-Max to Min-Min

Assume lower-level convexity and

$$W \;=\; \{\, w \in \mathbb{R}^{n_w} \;\mid\; B(w) \,\leq\, 0 \,\} \;.$$

If W has a non-empty interior (Slater's constraint qualification):

$$V_i(x) \;=\; \inf_{\lambda_i > 0} D_i(x, \lambda_i) \;.$$

with $$D_i(x,\lambda_i) \;:=\; \max_{w} F_i(x,w) - \lambda_i^T B(w) \;.$$

Duality: from Min-Max to Min-Min

Assume lower-level convexity and

$$W \;=\; \{\, w \in \mathbb{R}^{n_w} \;\mid\; B(w) \;\leq\; 0 \,\} \;.$$

If W has a non-empty interior (Slater's constraint qualification):

$$V_i(x) \;=\; \inf_{\lambda_i > 0} D_i(x, \lambda_i) \;.$$

$$\text{with} \qquad D_i(x, \lambda_i) \;:=\; \max_w F_i(x, w) - \lambda_i^T B(w) \;.$$

Main Idea: augment the upper level optimization variable x by the dual optimization variables $\lambda \;:=\; (\, \lambda_0 \; \ldots, \lambda_m \,)$:

$$\inf_{x, \lambda > 0} \;\; D_0(x, \lambda_0) \quad \text{s.t.} \quad D_i(x, \lambda_i) \;\leq\; 0 \;.$$

Special Case: Polytopic Uncertainty

Example:

$$F_i(x, w) \;=\; c_i(x)^T w + d_i(x) \quad \text{and} \quad W := \{\, w \mid Aw \;\leq\; b \,\}$$

Special Case: Polytopic Uncertainty

Example:

$$F_i(x,w) \;=\; c_i(x)^T w + d_i(x) \quad \text{and} \quad W := \{\, w \mid Aw \;\leq\; b \,\}$$

We can use dual linear programming:

$$\begin{aligned} V_i(x) \quad &= \quad \max_{w} \;\; c_i(x)^T w + d_i(x) \qquad \text{s.t.} \qquad Aw \;\leq\; b \\ &= \quad \min_{\lambda_i \geq 0} \;\; b^T \lambda_i + d_i(x) \qquad\quad \text{s.t.} \qquad A^T \lambda_i \;=\; c_i(x)\;. \end{aligned}$$

Special Case: Polytopic Uncertainty

Example:

$$F_i(x, w) \ = \ c_i(x)^T w + d_i(x) \quad \text{and} \quad W := \{ \, w \mid Aw \ \leq \ b \, \}$$

Robust counterpart problem reduces to a standard NLP:

$$\begin{aligned} \min_{x, \lambda_0, \ldots, \lambda_m} \quad & b^T \lambda_0 + d_0(x) \\ \text{s.t.} \quad & 0 \ \geq \ b^T \lambda_i + d_i(x) \\ & 0 \ \leq \ \lambda_i \\ & 0 \ = \ A^T \lambda_i - c_i(x) \qquad \text{for all} \quad i \in \{1, \ldots, m\} \; . \end{aligned}$$

Remark: if c_i and d_i are affine in x, we obtain an LP.

Special Case: Semi-Definite Uncertainty Set Models

Remark: The above example generalizes one-to-one for

$$W \;:=\; \left\{ w \;\middle|\; \sum_{j=1}^{n_w} A_j w_j \;\preceq\; B \right\} \;,$$

in this case the robust counterpart functions are of the form

$$\begin{aligned} V_i(x) \;&=\; \max_{w \in W} \;\; c_i(x)^T w + d_i(x) \\ &=\; \min_{\Lambda_i \succeq 0} \;\; \mathrm{Tr}\left(B^T \Lambda_i\right) + d_i(x) \quad \text{s.t.} \quad \mathrm{Tr}\left(A_j^T \Lambda_i\right) \;=\; c_{i,j}(x) \;. \end{aligned}$$

Upper Level Convexity

Simple but important observation:

We always have upper-level convexity if the functions F_i are convex in x. This result is independent of how the uncertainty w enters.

Upper Level Convexity

Simple but important observation:

We always have upper-level convexity if the functions F_i are convex in x. This result is independent of how the uncertainty w enters.

Proof: The maximum over convex functions is convex!

Upper Level Convexity

Simple but important observation:

We always have upper-level convexity if the functions F_i are convex in x. This result is independent of how the uncertainty w enters.

Proof: The maximum over convex functions is convex!

Remark: The reverse statement is not true.

Example: Upper Level Convexity

Example: Consider the unconstrained scalar min-max problem

$$\min_x \max_w F_0(x,w) \qquad \text{with} \quad F_0(x,w) := -x^2 + bxw - w^2$$

for some constant $b \geq 2$. The function F_0 is for no fixed w convex in x, but

$$V_0(x) = -x^2 + \frac{1}{4}(bx)^2$$

is convex for $b \geq 2$.

Example: Upper Level Convexity

Example: Consider the unconstrained scalar min-max problem

$$\min_x \max_w F_0(x,w) \qquad \text{with} \quad F_0(x,w) \;:=\; -x^2 + bxw - w^2$$

for some constant $b \geq 2$. The function F_0 is for no fixed w convex in x, but

$$V_0(x) \;=\; -x^2 + \frac{1}{4}(bx)^2$$

is convex for $b \geq 2$.

The robust counterpart problem can be "easier" to solve than the original optimization problem;

"Robustificiation" can lead to "Convexification".

Overview

- The convex optimization perspective on robust optimization
- **The S-procedure for Quadratic Forms**
- Inner- and Outer Ellipsoidal Approximations

The S-Procedure for Quadratic Forms

Basic Idea: Consider possibly non-convex QCQPs

$$V \quad := \quad \max_x \; x^T H_0 x + g_0^T x + s_0 \qquad \text{s.t.} \quad x^T H_i x + g_i^T x + s_i \; \leq \; 0$$

The S-Procedure for Quadratic Forms

Basic Idea: Consider possibly non-convex QCQPs

$$V \quad := \quad \max_x \ x^T H_0 x + g_0^T x + s_0 \qquad \text{s.t.} \quad x^T H_i x + g_i^T x + s_i \ \leq \ 0$$

Notation:

$$H(\lambda) \quad := \quad H_0 - \sum_{i=1}^{m} \lambda_i H_i \ , \quad g(\lambda) \ := \ g_0 - \sum_{i=1}^{m} \lambda_i g_i \ , \qquad (1)$$

and $s(\lambda) \quad := \quad s_0 - \sum_{i=1}^{m} \lambda_i s_i \ .$

The S-Procedure for Quadratic Forms

Basic Idea: Consider possibly non-convex QCQPs

$$V \quad := \quad \max_x \; x^T H_0 x + g_0^T x + s_0 \qquad \text{s.t.} \quad x^T H_i x + g_i^T x + s_i \;\leq\; 0$$

Dual Problem:

$$\widehat{V} \quad := \quad \inf_{\lambda > 0} \; \max_x \;\; x^T H(\lambda) x \; + \; g(\lambda)^T x \; + \; s(\lambda)$$

$$= \quad \inf_{\lambda > 0} \;\; \frac{1}{4} g(\lambda)^T H(\lambda)^{-1} g(\lambda) \; + \; s(\lambda) \quad \text{s.t.} \quad H(\lambda) \; \prec \; 0 \, .$$

The S-Procedure for Quadratic Forms

Basic Idea: Consider possibly non-convex QCQPs

$$V \quad := \quad \max_x \ x^T H_0 x + g_0^T x + s_0 \qquad \text{s.t.} \quad x^T H_i x + g_i^T x + s_i \ \le \ 0$$

Dual Problem:

$$\begin{aligned}
\widehat{V} \quad &:= \quad \inf_{\lambda > 0} \ \max_x \ x^T H(\lambda) x \ + \ g(\lambda)^T x \ + \ s(\lambda) \\
&= \quad \inf_{\lambda > 0} \ \frac{1}{4} g(\lambda)^T H(\lambda)^{-1} g(\lambda) \ + \ s(\lambda) \quad \text{s.t.} \quad H(\lambda) \ \prec \ 0 \, . \\
&= \quad \min_{\lambda \ge 0 \, , \, \gamma} \ \gamma \quad \text{s.t.} \quad \begin{pmatrix} s(\lambda) \ - \ \gamma & \frac{1}{2} g(\lambda)^T \\ \frac{1}{2} g(\lambda) & H(\lambda) \end{pmatrix} \ \preceq \ 0
\end{aligned}$$

S-Lemma

Standard duality: $V \leq \widehat{V}$.

S-Lemma

Standard duality: $V \leq \widehat{V}$.

Remark on Suboptimality Estimates: For special classes of QCQPs explicit bounds on the sub-optimality of the approximation $\widehat{V}$ are known. For example, in the context of the Maximum Cut problem (Goemans). More general sub-optimality estimates have been developed by Henrion, Nemirovski, and Nesterov.

- M.X. Goemans and D.P. Williamson. Improved approximation algorithms for Maximum Cut and satisfiability problems using semidefinite programming. *Journal of ACM*, 42:1115–1145, 1995.
- Y. Nesterov. Semidefinite relaxation and non-convex quadratic optimization. *Optimization Methods and Software*, 12:1–20, 1997.
- D. Henrion, S. Tarbouriech, and D. Arzelier. LMI Approximations for the Radius of the Intersection of Ellipsoids: A Survey. *Journal of Optimization Theory and Applications*, 108(1):1–28, 2001.

S-Lemma

Standard duality: $V \leq \widehat{V}$.

Remark on Suboptimality Estimates: For special classes of QCQPs explicit bounds on the sub-optimality of the approximation $\widehat{V}$ are known. For example, in the context of the Maximum Cut problem (Goemans). More general sub-optimality estimates have been developed by Henrion, Nemirovski, and Nesterov.

- M.X. Goemans and D.P. Williamson. Improved approximation algorithms for Maximum Cut and satisfiability problems using semidefinite programming. *Journal of ACM*, 42:1115–1145, 1995.
- Y. Nesterov. Semidefinite relaxation and non-convex quadratic optimization. *Optimization Methods and Software*, 12:1–20, 1997.
- D. Henrion, S. Tarbouriech, and D. Arzelier. LMI Approximations for the Radius of the Intersection of Ellipsoids: A Survey. *Journal of Optimization Theory and Applications*, 108(1):1–28, 2001.

S-Procedure in Robust Optimization

Example:

$$F_i(x,w) \quad = \quad w^T H_i(x) w + g_i(x)^T w \; .$$

Assume that the uncertainty set is an intersection of ellipsoids,

$$W \quad := \quad \bigcap_{j \in \{1,\ldots,N\}} \mathcal{E}(Q_j, q_j) \; .$$

S-Procedure in Robust Optimization

Example:

$$F_i(x,w) \quad = \quad w^T H_i(x) w + g_i(x)^T w \ .$$

Assume that the uncertainty set is an intersection of ellipsoids,

$$W \quad := \quad \bigcap_{j \in \{1,\dots,N\}} \mathcal{E}(Q_j, q_j) \ .$$

$$\widehat{V}_i(x) \ := \ \min_{\lambda_i \geq 0 \, , \, \gamma_i} \quad \gamma_i \quad \text{s.t.} \quad \begin{pmatrix} s_i(x,\lambda_i) \ - \ \gamma_i & \frac{1}{2} g_i(x,\lambda_i)^T \\ \frac{1}{2} g_i(x,\lambda_i) & H_i(x,\lambda_i) \end{pmatrix} \preceq \ 0$$

are upper bounds on the functions V_i, $\widehat{V}_i(x) \geq V_i(x)$.

S-Procedure in Robust Optimization

Example:

$$F_i(x,w) \quad = \quad w^T H_i(x) w + g_i(x)^T w \ .$$

Assume that the uncertainty set is an intersection of ellipsoids,

$$W \quad := \quad \bigcap_{j \in \{1,\ldots,N\}} \mathcal{E}(Q_j, q_j) \ .$$

Conservative reformulation given by

$$\min_{x,\gamma,\lambda_0,\ldots,\lambda_m} \quad \gamma_0 \quad \text{s.t.} \quad \left\{ \begin{array}{l} \forall i \in \{1,\ldots,m\}: \quad 0 \ \geq \ \gamma_i \quad , \quad 0 \ \leq \ \lambda_i \ , \\ 0 \ \succeq \ \begin{pmatrix} s_i(x,\lambda_i) \ - \ \gamma_i & \frac{1}{2} g_i(x,\lambda_i)^T \\ \frac{1}{2} g_i(x,\lambda_i) & H_i(x,\lambda_i) \end{pmatrix} \ . \end{array} \right.$$

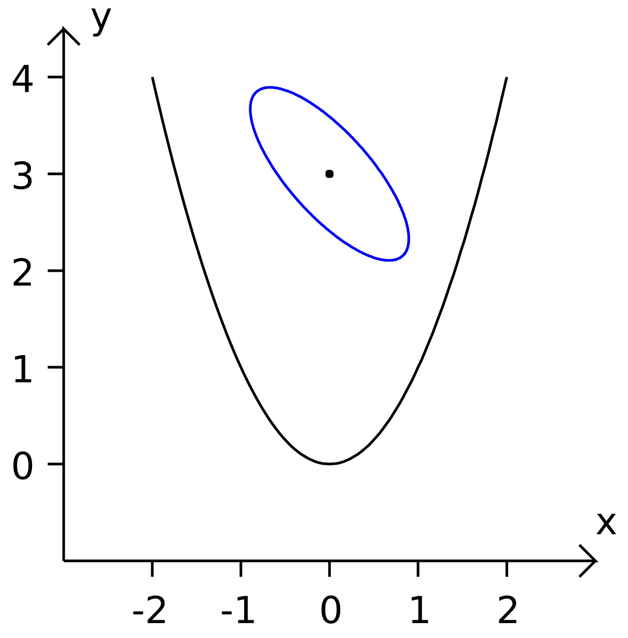
Tight Version of the S-Procedure

Theorem [Yakubovich, 1977] If we have a QCQP with only one constraint, the S-procedure yields a tight bound.

- The proof is not so trivial.
- Basis for LMI formulations of H_∞ control and almost all LMI-based robust control results from 1980-2000.

- V.A. Yakubovich. S-procedure in nonlinear control theory. *Vestnik Leningrad University*, 4:73–93, 1977.

Robust Optimization: An Example



Robust optimization problem:

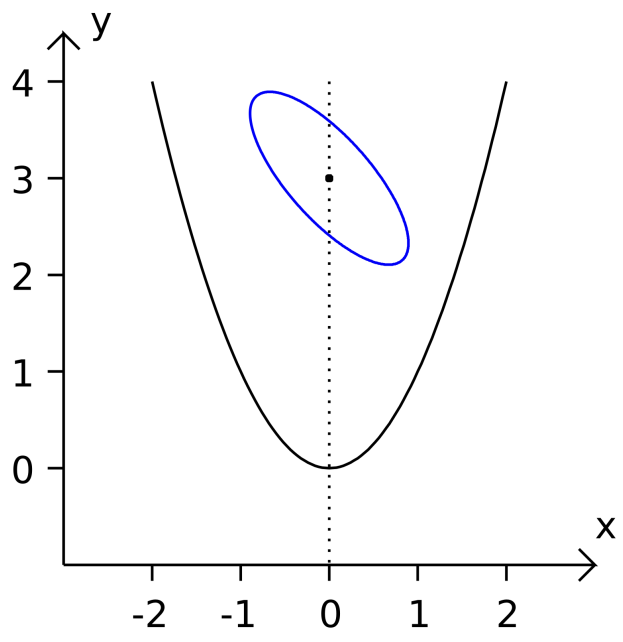
$$\min_{x,y} \quad y$$

$$\text{s.t.} \quad (x+v)^2 - (y+w) \leq 0$$

for all $(v, w) \in \mathcal{E}$.

Assumption: $\mathcal{E}$ is a given ellipsoidal uncertainty set.

Robust Optimization: An Example



Robust optimization problem:

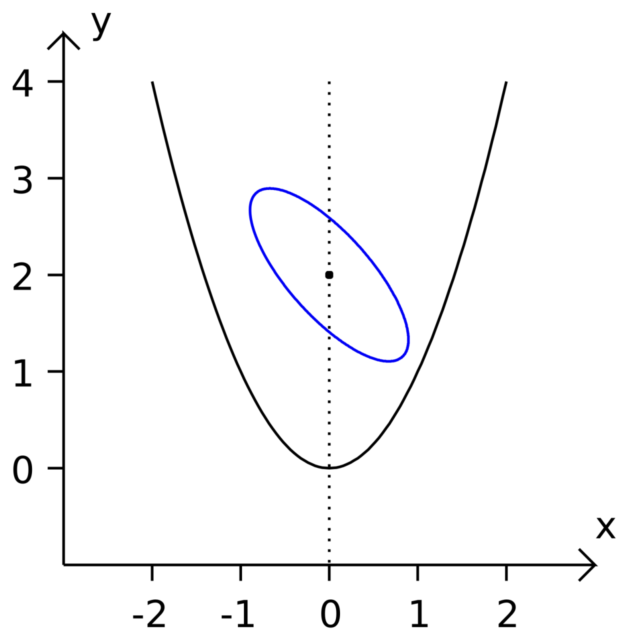
$$\min_{x,y} \quad y$$

$$\text{s.t.} \quad (x+v)^2 - (y+w) \leq 0$$

for all $(v, w) \in \mathcal{E}$.

Assumption: $\mathcal{E}$ is a given ellipsoidal uncertainty set.

Robust Optimization: An Example



Robust optimization problem:

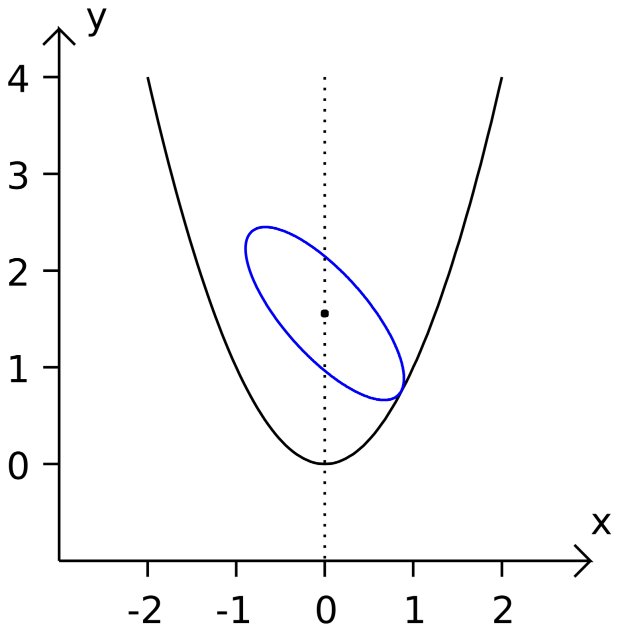
$$\min_{x,y} \quad y$$

$$\text{s.t.} \quad (x+v)^2 - (y+w) \leq 0$$

for all $(v, w) \in \mathcal{E}$.

Assumption: $\mathcal{E}$ is a given ellipsoidal uncertainty set.

Robust Optimization: An Example



Robust optimization problem:

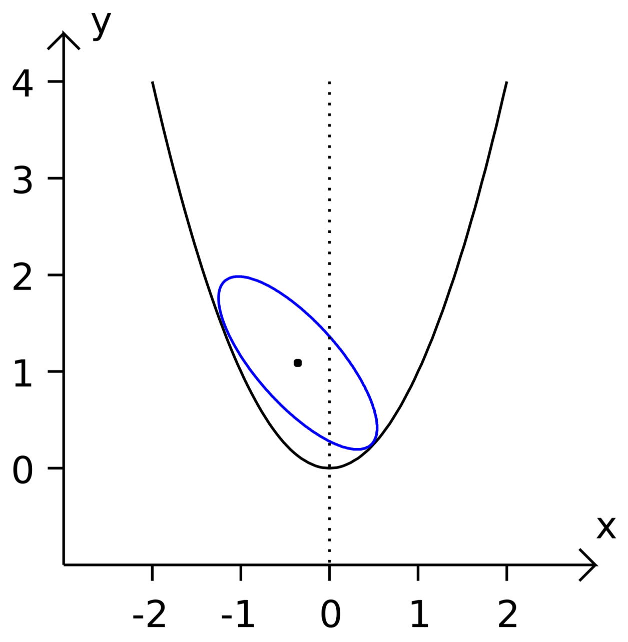
$$\min_{x,y} \quad y$$

$$\text{s.t.} \quad (x+v)^2 - (y+w) \leq 0$$

for all $(v, w) \in \mathcal{E}$.

Assumption: $\mathcal{E}$ is a given ellipsoidal uncertainty set.

Robust Optimization: An Example



Robust optimization problem:

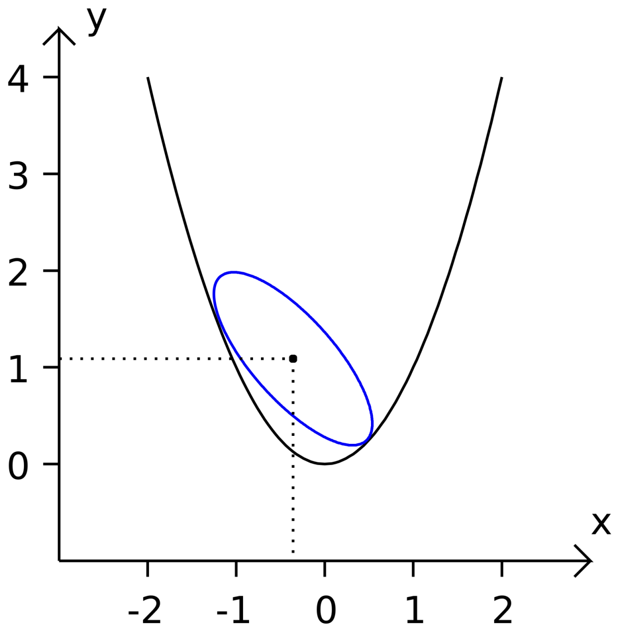
$$\min_{x,y} \quad y$$

$$\text{s.t.} \quad (x+v)^2 - (y+w) \le 0$$

for all $(v, w) \in \mathcal{E}$.

Assumption: $\mathcal{E}$ is a given ellipsoidal uncertainty set.

Robust Optimization: An Example



Robust optimization problem:

$$\min_{x,y} \quad y$$

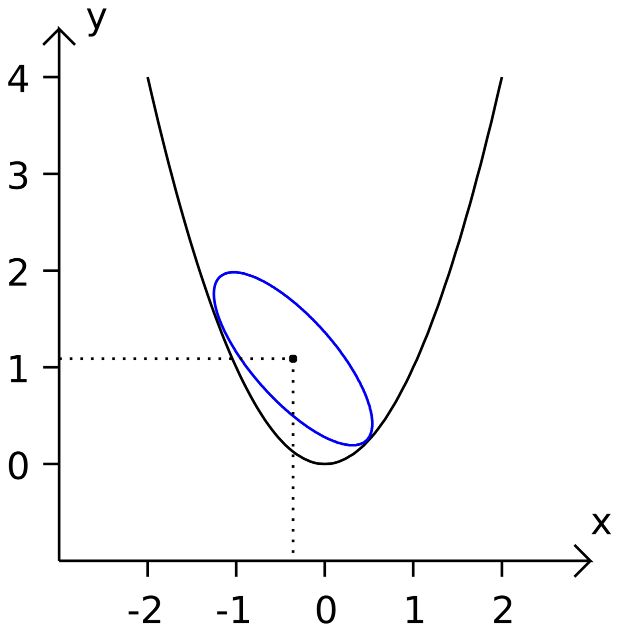
$$\text{s.t.} \quad (x+v)^2 - (y+w) \leq 0$$

for all $(v,w) \in \mathcal{E}$.

Assumption: $\mathcal{E}$ is a given ellipsoidal uncertainty set.

Question: How to find optimal solution numerically?

Robust Optimization: An Example

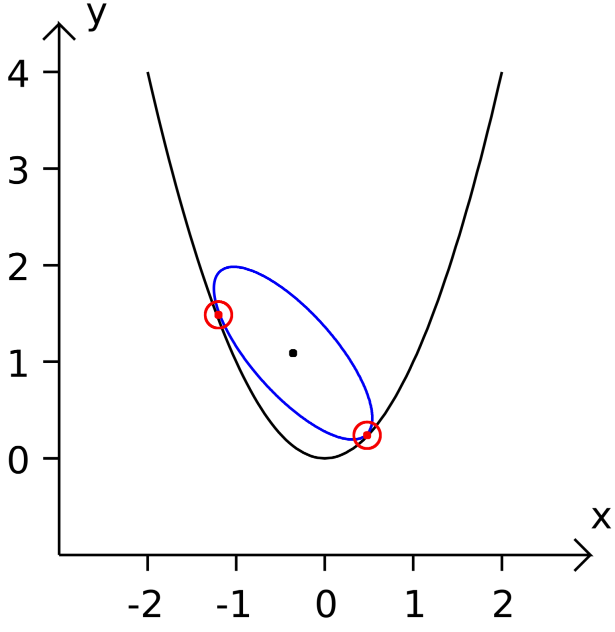


Regard as a min-max problem:

$$\min_{x,y} \quad y$$

$$\text{s.t.} \quad \max_{(v,w)\in\mathcal{E}} (x+v)^2 - (y+w) \le 0$$

Robust Optimization: An Example



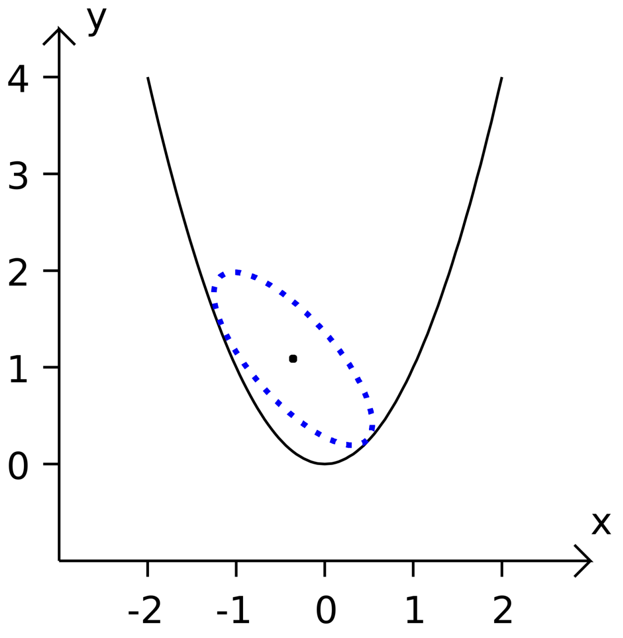
Regard as a min-max problem:

$$\min_{x,y} \quad y$$

$$\text{s.t.} \quad \max_{(v,w)\in\mathcal{E}} (x+v)^2 - (y+w) \leq 0$$

Problem: There are two local maxima in the optimal solution.

Robust Optimization: An Example



Regard as a min-max problem:

$$\min_{x,y} \quad y$$

$$\text{s.t.} \quad \max_{(v,w)\in\mathcal{E}} (x+v)^2 - (y+w) \leq 0$$

Problem: There are two local maxima in the optimal solution.

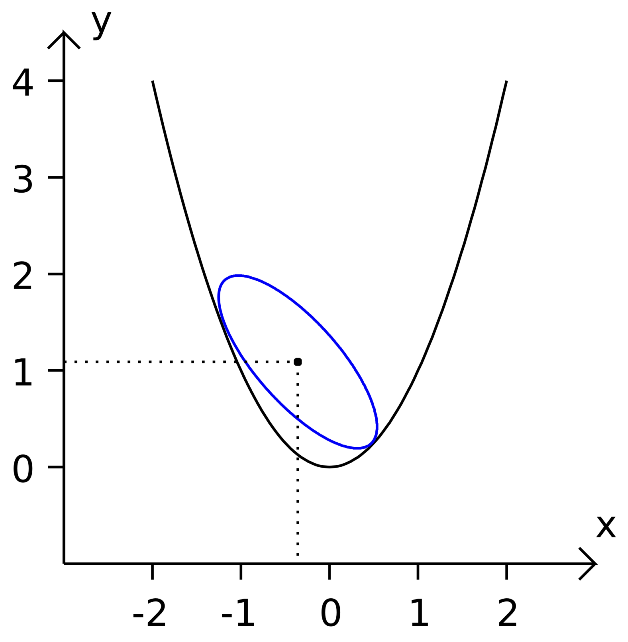
One Possibility: Check the inequality for all points in the ellipsoid.

Selection of Literature on Semi-Infinite Optimization

- R. Hettich and H.T. Jongen. Semi-infinite programming: Conditions of optimality and applications. Optimization Techniques, Lecture Notes in Control and Inform. Sci. 7, J. Stoer, Springer, 1978.
- R. Hettich and K. Kortanek. Semi infinite programming:Theory, Methods, and Application, volume 35. SIAM Review, 1993.
- H.T. Jongen, J.J. Rückmann, and O. Stein. Generalized semi-infinite optimization: A first order optimality condition and examples. Mathematical Programming, pages 145-158, 1998.
- C.A. Floudas and O. Stein. The Adaptative Convexification Algorithm: a Feasible Point Method for Semi-Infinite Programming. SIAM Journal on Optimization, 18(4):1187-1208, 2007.
- ...

Solution using the S-Procedure

Find the solution $(x^*, y^*) = (-0.35..., 1.08...)$ by convex optimization:



Define:

$$Q := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad q(x) := \begin{pmatrix} 2x \\ -1 \end{pmatrix}$$

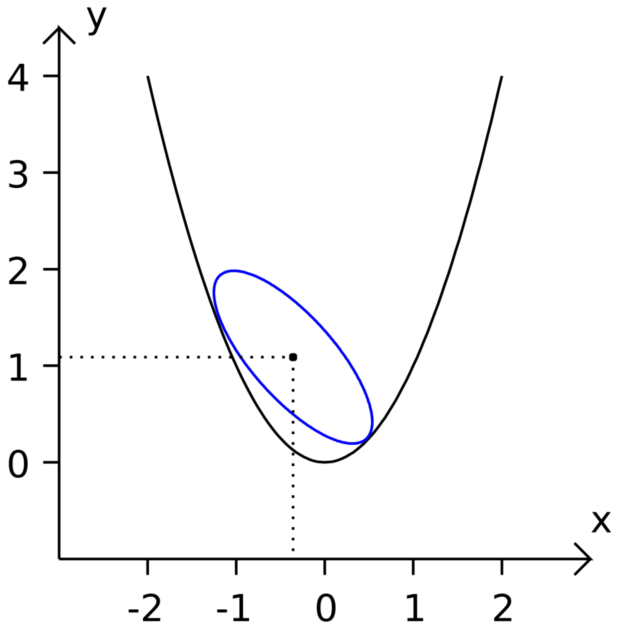
and

$$\Sigma := \begin{pmatrix} 0.8 & -0.6 \\ -0.6 & 0.8 \end{pmatrix}^{-1}$$

$$\boxed{\min_{x,y,\lambda>0.8} \quad y \quad \text{s.t.} \quad x^2 - y + \frac{1}{4} q(x)^T \left(\lambda\Sigma - Q\right)^{-1} q(x) + \lambda \;\leq\; 0 \,.}$$

Solution using the S-Procedure

Find the solution $(x^*, y^*) \;=\; (\,-0.35...\,,\,1.08...\,)$ by convex optimization:



Alternative formulation as LMI:

$$\min_{x,y,\lambda} \quad y$$

$$\text{s.t.} \quad \begin{pmatrix} y-\lambda & q(x)^T & x \\ q(x) & \lambda\Sigma - Q & 0 \\ x & 0 & 1 \end{pmatrix} \succeq 0$$

$$\min_{x,y,\lambda>0.8} \quad y \quad \text{s.t.} \quad x^2 - y + \frac{1}{4} q(x)^T \left(\lambda\Sigma - Q\right)^{-1} q(x) + \lambda \;\leq\; 0\,.$$

S-Procedure in Robust Stability Analysis

Question: Under which conditions is the system

$$\dot{x}(t) = Ax(t) + Bw(t) \ , \quad z(t) = Cx(t)$$

quadratically stable for all w with $w(t)^2 \leq \gamma^2 z(t)^2$?

S-Procedure in Robust Stability Analysis

Question: Under which conditions is the system

$$\dot{x}(t) = Ax(t) + Bw(t)\ , \quad z(t) = Cx(t)$$

quadratically stable for all w with $w(t)^2 \leq \gamma^2 z(t)^2$?

- Lyapunov stability condition: $\exists P \succeq 0, \quad 2x^T P(Ax + Bw) < 0$.
- Non-convex quadratic constraint: $w(t)^2 \leq \gamma^2 z(t)^2$

S-Procedure in Robust Stability Analysis

Question: Under which conditions is the system

$$\dot{x}(t) = Ax(t) + Bw(t)\ , \quad z(t) = Cx(t)$$

quadratically stable for all w with $w(t)^2 \leq \gamma^2 z(t)^2$?

- Lyapunov stability condition: $\exists P \succeq 0, \quad 2x^T P(Ax + Bw) < 0$.
- Non-convex quadratic constraint: $w(t)^2 \leq \gamma^2 z(t)^2$

S-Procedure yields "Circle Criterion":

$$\begin{pmatrix} AP + PA^T + \gamma^2 C^T C & PB \\ B^T P & -2 \end{pmatrix} \prec 0\ , \quad P \succeq 0$$

Overview

- The convex optimization perspective on robust optimization
- The S-procedure for Quadratic Forms
- **Inner- and Outer Ellipsoidal Approximations**

Support Functions

Definition of support function:

$$V(c) \quad := \quad \max_{x} \; c^T x \quad \text{s.t.} \quad x \in \mathcal{F} \; .$$

Support Functions

Definition of support function:

$$V(c) \quad := \quad \max_{x} \; c^T x \quad \text{s.t.} \quad x \in \mathcal{F} \; .$$

Support of an ellipsoid:

$$V(c) \; = \; \max_{x \in \mathcal{E}(Q,q)} c^T x \; = \; \sqrt{c^T Q c} + c^T q \; ,$$

Support Functions

Definition of support function:

$$V(c) \quad := \quad \max_x \; c^T x \quad \text{s.t.} \quad x \in \mathcal{F} \; .$$

Support of an ellipsoid:

$$V(c) \;=\; \max_{x \in \mathcal{E}(Q,q)} c^T x \;=\; \sqrt{c^T Q c} + c^T q \; ,$$

If $\mathcal{F}$ is compact and convex:

$$\mathcal{F} \quad = \quad \bigcap_{c \in \mathbb{R}^n \setminus \{0\}} \mathcal{H}(c) \; ,$$

where $\mathcal{H}(c) := \left\{ x \in \mathbb{R}^n \;\middle|\; c^T x \;\leq\; V(c) \right\}$.

Minkowski Sum of Ellipsoids

- The sum of these ellipsoids is defined as the standard Minkowski sum:

$$\sum_{i=1}^{N} \mathcal{E}(Q_i, q_i) \quad := \quad \left\{ \sum_{i=0}^{N} x_i \ \in \mathbb{R}^n \ \middle| \ x_i \in \mathcal{E}(Q_i, q_i) \right\} \ .$$

Minkowski Sum of Ellipsoids

- The sum of these ellipsoids is defined as the standard Minkowski sum:

$$\sum_{i=1}^{N} \mathcal{E}(Q_i, q_i) \quad := \quad \left\{ \sum_{i=0}^{N} x_i \ \in \mathbb{R}^n \ \middle| \ x_i \in \mathcal{E}(Q_i, q_i) \right\} \ .$$

- Examples: intervals and zonotopes:

$$\sum_{i=1}^{m} \mathcal{E}\left(a_i a_i^T \right) \quad = \quad \left\{ \sum_{i=1}^{m} \lambda_i a_i \ \in \mathbb{R}^n \ \middle| \ -1 \ \leq \ \lambda_i \ \leq \ 1 \right\} \ .$$

Minkowski Sum of Ellipsoids

- The sum of these ellipsoids is defined as the standard Minkowski sum:

$$\sum_{i=1}^{N} \mathcal{E}(Q_i, q_i) \quad := \quad \left\{ \sum_{i=0}^{N} x_i \in \mathbb{R}^n \;\middle|\; x_i \in \mathcal{E}(Q_i, q_i) \right\} .$$

- Examples: intervals and zonotopes:

$$\sum_{i=1}^{m} \mathcal{E}\left(a_i a_i^T \right) \quad = \quad \left\{ \sum_{i=1}^{m} \lambda_i a_i \in \mathbb{R}^n \;\middle|\; -1 \leq \lambda_i \leq 1 \right\} .$$

- Application: discrete-time systems

$$x^+ = Ax + Bw , \quad x \in \mathcal{E}(Q_x); \;\; w \in \mathcal{E}(Q_w)$$

then $\qquad x^+ \in \mathcal{E}\left(AQ_x A^T \right) + \mathcal{E}\left(BQ_w B^T \right)$

Support Function of the Sum of Ellipsoids

Let's compute the support function

$$V(c) \;=\; \max_{x_1,\ldots,x_N} \; c^T \left(\sum_{i=1}^{N} x_i \right) \quad \text{s.t.} \quad x_i^T Q_i^{-1} x_i \;\leq\; 1\;.$$

- Convex maximization problem; $x_1 = \ldots = x_N = 0$ is feasible.

Support Function of the Sum of Ellipsoids

Let's compute the support function

$$V(c) \;=\; \max_{x_1,\ldots,x_N} \; c^T\left(\sum_{i=1}^N x_i\right) \quad \text{s.t.} \quad x_i^T Q_i^{-1} x_i \;\le\; 1\;.$$

- Convex maximization problem; $x_1 = \ldots = x_N = 0$ is feasible.
- We can use duality to find

$$\begin{aligned} V(c) \;&=\; \inf_{\lambda > 0} \max_{x_1,\ldots,x_N} \; \sum_{i=1}^N \left(c^T x_i \,-\, \lambda_i x_i^T Q_i^{-1} x_i + \lambda_i \right) \\ &=\; \inf_{\lambda > 0} \; \sum_{i=1}^N \frac{c^T Q_i c}{4\lambda_i} + \sum_{i=1}^N \lambda_i \;. \end{aligned}$$

Support Function of the Sum of Ellipsoids

Let's compute the support function

$$V(c) \;=\; \max_{x_1,\ldots,x_N} \; c^T\left(\sum_{i=1}^N x_i\right) \quad \text{s.t.} \quad x_i^T Q_i^{-1} x_i \;\le\; 1\;.$$

- Convex maximization problem; $x_1 = \ldots = x_N = 0$ is feasible.
- We can use duality to find

$$\begin{aligned} V(c) \;&=\; \inf_{\lambda>0}\;\max_{x_1,\ldots,x_N}\;\sum_{i=1}^N \left(c^T x_i \,-\, \lambda_i x_i^T Q_i^{-1} x_i + \lambda_i\right) \\ &=\; \inf_{\lambda>0}\;\sum_{i=1}^N \frac{c^T Q_i c}{4\lambda_i} + \sum_{i=1}^N \lambda_i\;. \end{aligned}$$

Support Function: Sum of Ellipsoids

- We can use duality to find

$$
\begin{aligned}
V(c) &= \inf_{\lambda>0} \max_{x_1,\ldots,x_N} \sum_{i=1}^{N} \left(c^T x_i - \lambda_i x_i^T Q_i^{-1} x_i + \lambda_i \right) \\
&= \inf_{\lambda>0} \sum_{i=1}^{N} \frac{c^T Q_i c}{4\lambda_i} + \sum_{i=1}^{N} \lambda_i \ .
\end{aligned}
$$

Support Function: Sum of Ellipsoids

- We can use duality to find

$$V(c) = \inf_{\lambda > 0} \max_{x_1, \ldots, x_N} \sum_{i=1}^{N} \left(c^T x_i - \lambda_i x_i^T Q_i^{-1} x_i + \lambda_i \right)$$

$$= \inf_{\lambda > 0} \sum_{i=1}^{N} \frac{c^T Q_i c}{4 \lambda_i} + \sum_{i=1}^{N} \lambda_i \; .$$

- Idea: use the tight version of the AM-GM inequality:

$$\inf_{\kappa > 0} \frac{a}{4\kappa} + \kappa \, b = \sqrt{ab} \, , \tag{2}$$

which holds for all $a, b \in \mathbb{R}_+$.

Support Function: Sum of Ellipsoids

- We can use duality to find

$$V(c) = \inf_{\lambda > 0} \max_{x_1, \ldots, x_N} \sum_{i=1}^{N} \left(c^T x_i - \lambda_i x_i^T Q_i^{-1} x_i + \lambda_i \right)$$

$$= \inf_{\lambda > 0} \sum_{i=1}^{N} \frac{c^T Q_i c}{4\lambda_i} + \sum_{i=1}^{N} \lambda_i \; .$$

$$V(c) = \inf_{\lambda > 0} \inf_{\kappa > 0} \sum_{i=1}^{N} \frac{c^T Q_i c}{4\kappa\lambda_i} + \sum_{i=1}^{N} \kappa\lambda_i = \inf_{\lambda > 0} \sqrt{c^T Q(\lambda) c} \; .$$

with $\forall \lambda \in \mathbb{R}^N_{++}$: $\quad Q(\lambda) := \left(\sum_{i=1}^{N} \frac{1}{\lambda_i} Q_i \right) \left(\sum_{i=1}^{N} \lambda_i \right) \; .$

Ellipsoidal Calculus (Outer Approx.)

Theorem [Kurzhanski (and earlier Russian literature)]: Define

$$\mathbb{D}^+ \;:=\; \left\{ \lambda \in \mathbb{R}^N_{++} \;\middle|\; \sum_{i=1}^N \lambda_i \;\leq\; 1 \right\} .$$

For every $\lambda \in \mathbb{D}^+$ we have

$$\forall \lambda \in \mathbb{D}^+ : \qquad \sum_{i=1}^N \mathcal{E}(Q_i) \;\subseteq\; \mathcal{E}\left(\sum_{i=1}^N \frac{1}{\lambda_i} Q_i \right) .$$

Ellipsoidal Calculus (Outer Approx.)

Theorem [Kurzhanski (and earlier Russian literature)]: Define

$$\mathbb{D}^+ := \left\{ \lambda \in \mathbb{R}^N_{++} \;\middle|\; \sum_{i=1}^{N} \lambda_i \leq 1 \right\} .$$

For every $\lambda \in \mathbb{D}^+$ we have

$$\forall \lambda \in \mathbb{D}^+ : \qquad \sum_{i=1}^{N} \mathcal{E}(Q_i) \subseteq \mathcal{E}\left(\sum_{i=1}^{N} \frac{1}{\lambda_i} Q_i \right) .$$

The parameterized outer approximation is tight, i.e.,

$$\sum_{i=1}^{N} \mathcal{E}(Q_i) = \bigcap_{\lambda \in \mathbb{D}^+} \mathcal{E}\left(\sum_{i=1}^{N} \frac{1}{\lambda_i} Q_i \right) .$$

Ellipsoidal Calculus (Inner Approx.)

Theorem [Kurzhanski (and earlier Russian literature)]: Define

$$\mathbb{D}^- \; := \; \left\{ S \in \left(\mathbb{R}^{n\times n}\right)^N \;\middle|\; S_i \, S_i^T \; \preceq I \quad \text{for all} \;\; i \in \{1,\ldots,N\} \right\} .$$

For every set of matrices $S \in \mathbb{D}^-$ we have

$$\forall S \in \mathbb{D}^- : \quad \sum_{i=1}^N \mathcal{E}(Q_i) \quad \supseteq \quad \mathcal{E}\left(\left(\sum_{i=1}^N Q_i^{\frac{1}{2}} S_i \right) \left(\sum_{i=1}^N Q_i^{\frac{1}{2}} S_i \right)^T \right) .$$

Ellipsoidal Calculus (Inner Approx.)

Theorem [Kurzhanski (and earlier Russian literature)]: Define

$$\mathbb{D}^- \ := \ \left\{ S \in \left(\mathbb{R}^{n\times n}\right)^N \ \middle| \ S_i \, S_i^T \ \preceq I \quad \text{for all} \ \ i \in \{1,\ldots,N\} \right\} .$$

For every set of matrices $S \in \mathbb{D}^-$ we have

$$\forall S \in \mathbb{D}^- : \quad \sum_{i=1}^N \mathcal{E}(Q_i) \ \supseteq \ \mathcal{E}\left(\left(\sum_{i=1}^N Q_i^{\frac{1}{2}} S_i \right) \left(\sum_{i=1}^N Q_i^{\frac{1}{2}} S_i \right)^T \right) .$$

The inner approximation is tight, i.e.,

$$\sum_{i=1}^N \mathcal{E}(Q_i) \ = \ \bigcup_{S \in \mathbb{D}^-} \mathcal{E}\left(\left(\sum_{i=1}^N Q_i^{\frac{1}{2}} S_i \right) \left(\sum_{i=1}^N Q_i^{\frac{1}{2}} S_i \right)^T \right) .$$

Summary

- Semi-infinite programs $\leftrightarrow$ min-max NLP
- Upper- and lower level convexity (max over convex fcns is convex)
- Robust counterpart functions + explicit examples
- Using duality: min-max $\leftrightarrow$ min-min.
- S-procedure (approximations and tight version)
- Support functions of convex sets
- Inner- and outer ellipsoidal approximations