Some useful facts about ellipsoids and their application to robust optimal control

Florian Messerer

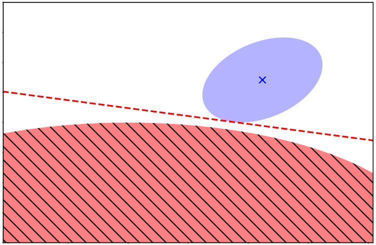
Systems Control and Optimization Laboratory, University of Freiburg

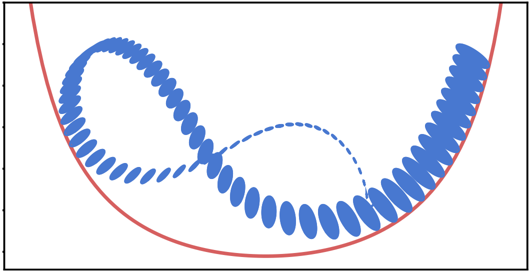
syscop talk

May 11, 2021

Introduction

1. Some basic knowledge about ellipsoids
2. Some stochastics
3. A little bit of robust optimization
4. Uncertainty description of dynamical systems





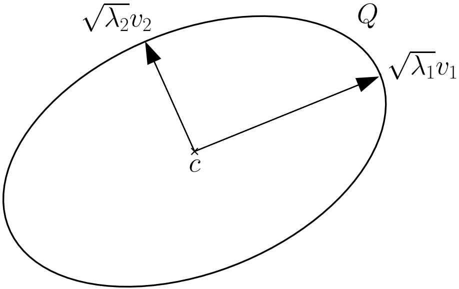
Definition and construction of ellipsoids

- Define ellipsoid by center $c \in \mathbb{R}^n$ and shape matrix $Q \in \mathbb{S}^n_{++}$ $(Q \succ 0)$

$$\begin{aligned}\mathcal{E}(Q,c) &:= \{x \in \mathbb{R}^n \mid \|x-c\|_{Q^{-1}} \leq 1\} \\ &= \{x \in \mathbb{R}^n \mid (x-c)^\top Q^{-1}(x-c) \leq 1\}\end{aligned}$$

- Denote by λ_i, v_i, $i = 1, \dots, n$, the eigenvalues / -vectors (normalized) of Q.
- Eigendecomposition $Q = V \Lambda V^\top$ with $\Lambda = \mathrm{diag}(\lambda)$, $VV^\top = I$

$$\mathcal{E}(Q,c) = \{V\Lambda^{\frac{1}{2}}w + c \mid w \in \mathbb{R}^n,\ w^\top w \leq 1\}$$



Definition and construction of ellipsoids (cont.)

- Relax to $Q \in \mathbb{S}^n_+$ $(Q \succeq 0)$
- $\lambda_i = 0$ for some $i \to$ degenerate ellipsoid
- For non-invertible Q, we can still use

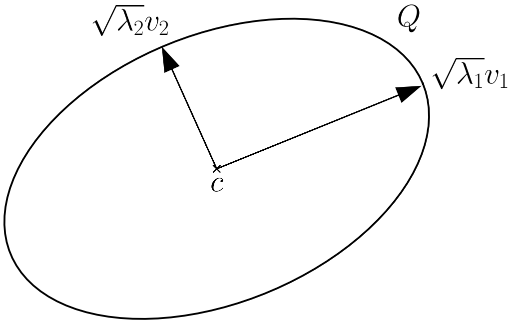
$$\mathcal{E}(Q, c) := \{Q^{\frac{1}{2}} w + c \mid w \in \mathbb{R}^n,\ w^\top w \leq 1\}$$

  with $Q = Q^{\frac{1}{2}} Q^{\frac{1}{2}}$ (unique, $Q^{\frac{1}{2}} = V \Lambda^{\frac{1}{2}} V^\top$)
- or just any $W \in \mathbb{R}^{n \times n}$ with $Q = W W^\top$

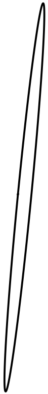
$$\mathcal{E}(Q, c) = \{W w + c \mid w \in \mathbb{R}^n,\ w^\top w \leq 1\}$$

- or non-square $W \in \mathbb{R}^{n \times m}$ with $Q = W W^\top$

$$\mathcal{E}(W W^\top, c) = \{W w + c \mid w \in \mathbb{R}^m,\ w^\top w \leq 1\}$$



Size of ellipsoids



- Most obvious measure of size is volume (with $\bar{V}_n$ volume of unit ball $\mathcal{E}(I_n)$)

$$\mathrm{Vol}(\mathcal{E}(Q)) = \bar{V}_n \prod_{i=1}^{n} \sqrt{\lambda_i} = \bar{V}_n \sqrt{\det Q}$$

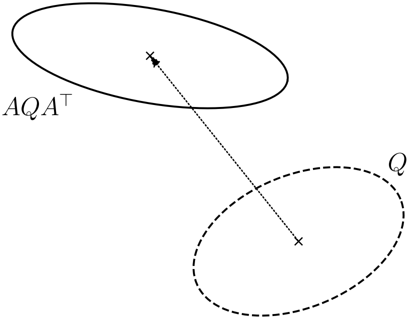
$\det Q = \det V \Lambda V^\top = \det V V^\top \Lambda = \det \Lambda = \prod_{i=1}^{n} \lambda_i$

- Often more intuitive measure

$$\sum_{i=1}^{n} \sqrt{\lambda_i}^2 = \sum_{i=1}^{n} \lambda_i = \operatorname{Tr} Q$$

$\operatorname{Tr} Q = \operatorname{Tr} V \Lambda V^\top = \operatorname{Tr} V^\top V \Lambda = \operatorname{Tr} \Lambda = \sum_{i=1}^{n} \lambda_i$

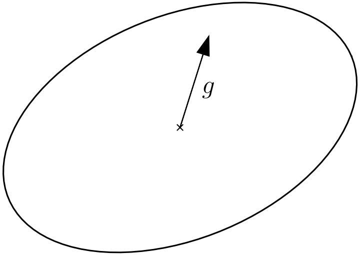
Affine transformation



$$A \in \mathbb{R}^{m\times n},\ b \in \mathbb{R}^m,$$

$$\begin{aligned}
& A\mathcal{E}(Q,c) + b, \\
:= & \{Ax + b \mid x \in \mathcal{E}(Q,c)\} \\
= & \{A(Q^{\frac{1}{2}}w + c) + b \mid w \in \mathcal{E}(I_n)\} \\
= & \{AQ^{\frac{1}{2}}w + Ac + b \mid w \in \mathcal{E}(I_n)\} \\
= & \mathcal{E}(AQ^{\frac{1}{2}}{Q^{\frac{1}{2}}}^\top A^\top, Ac + b) \\
= & \mathcal{E}(AQA^\top, Ac + b)
\end{aligned}$$

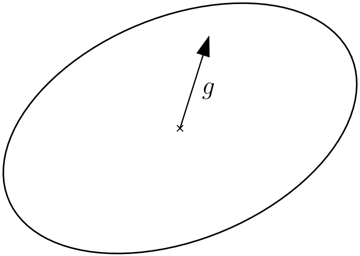
Maximum over linear function



$$\max_{x \in \mathbb{R}^n} \; g^\top x \quad \text{s.t.} \quad x \in \mathcal{E}(Q)$$
$$= \max_{w \in \mathbb{R}^n} \; g^\top Q^{\frac{1}{2}} w \quad \text{s.t.} \quad w \in \mathcal{E}(I)$$
$$= \max_{w \in \mathbb{R}^n} \; \bar{g}^\top w \quad \text{s.t.} \quad w \in \mathcal{E}(I)$$

$$w^* = \frac{\bar{g}}{\|\bar{g}\|} = \frac{Q^{\frac{1}{2}} g}{\sqrt{g^\top Q g}}, \quad x^* = Q^{\frac{1}{2}} w^* = \frac{Qg}{\sqrt{g^\top Q g}}$$
$$g^\top x^* = \frac{g^\top Q g}{\sqrt{g^\top Q g}} = \sqrt{g^\top Q g}$$

Support function



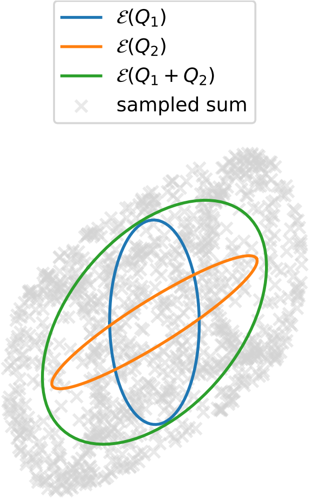
- Any non-empty compact convex set $\mathcal{S} \subset \mathbb{R}^n$ can be defined via its support function:

$$V(g) = \max_{x \in \mathbb{R}^n} \; g^\top x \quad \text{s.t.} \quad x \in \mathcal{S}$$

- Important tool for analysis of convex sets
- For ellipsoid:

$$\begin{aligned} V(g) &= \max_{x \in \mathbb{R}^n} \; g^\top x \quad \text{s.t.} \quad x \in \mathcal{E}(Q) \\ &= \sqrt{g^\top Q g} \end{aligned}$$

Sum of two ellipsoids



- Minkowski sum

$$\begin{aligned} &\mathcal{E}(Q_1, c_1) + \mathcal{E}(Q_2, c_2) \\ := &\{x_1 + x_2 \mid x_1 \in \mathcal{E}(Q_1, c_1),\ x_2 \in \mathcal{E}(Q_2, c_2)\} \end{aligned}$$

- not an ellipsoid (in general)

Overapproximating sum of ellipsoids by ellipsoid

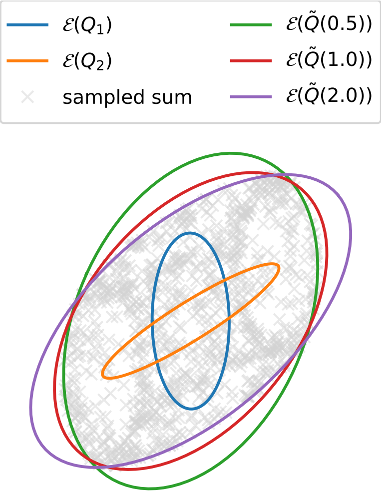
- Aim: find Q such that $\mathcal{E}(Q) \supseteq \mathcal{E}(Q_1) + \mathcal{E}(Q_2)$
- More general: Find Q such that $\mathcal{E}(Q) \supseteq \sum_{k=1}^{K} \mathcal{E}(Q_k)$
- Construct family of outer approximations parametrized by $\alpha \in \mathbb{R}_{++}^K$

$$Q(\alpha) = \left(\sum_{k=1}^{K} \alpha_k \right) \sum_{k=1}^{K} \frac{1}{\alpha_k} Q_k \qquad \Rightarrow \qquad \mathcal{E}(Q(\alpha)) \supseteq \sum_{k=1}^{K} \mathcal{E}(Q_k) \quad \forall \alpha \in \mathbb{R}_{++}^K$$

- Possible to assume / require $\sum_{k=1}^{K} \alpha_k = 1$ w.l.o.g.
- Parametrized outer approximation is tight (but not complete)

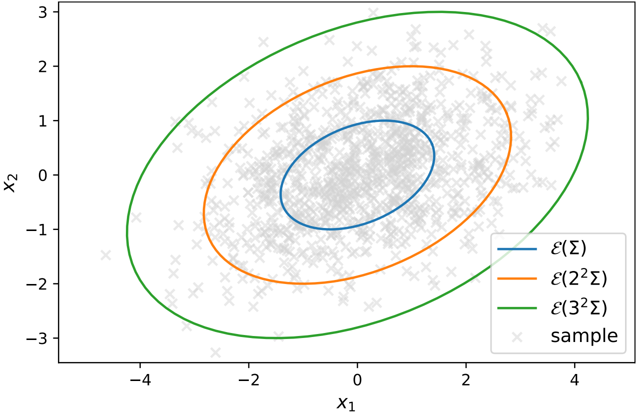
$$\bigcap_{\alpha \in \mathbb{R}_{++}^K} \mathcal{E}(Q(\alpha)) = \sum_{k=1}^{K} \mathcal{E}(Q_k)$$

Overapproximating sum of ellipsoids by ellipsoid (cont.)



- Consider $K = 2$
  - $Q(\alpha) = \frac{1}{\alpha_1} Q_1 + \frac{1}{\alpha_2} Q_2$ with $\alpha_1 + \alpha_2 = 1$
  - Reparametrize: $\alpha_2 = 1 - \alpha_1$, $\beta = \frac{1}{1-\alpha_1} > 0$
  - $\tilde{Q}(\beta) = (1 + \frac{1}{\beta}) Q_1 + (1 + \beta) Q_2$
- In general: Choose α according to some criterion
  - e.g., such that $\mathcal{E}(Q(\alpha))$ has minimal size
  - or such that approximation is tight in a given direction $p \in \mathbb{R}^n$ (approximation touches true sum)
  - $\rightarrow$ convex optimization problem

Ellipsoids in multivariate normal distribution

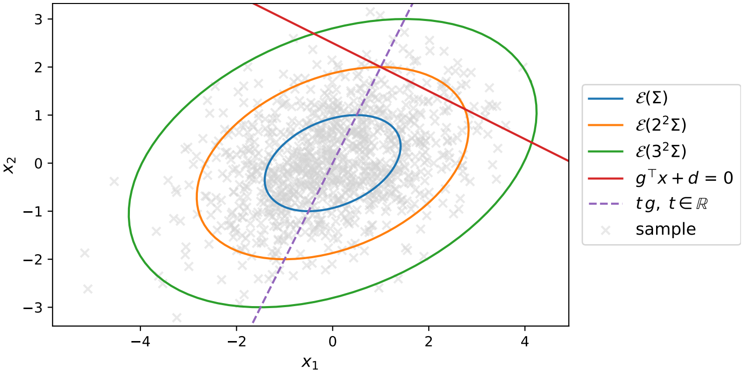


- Consider n-dimensional normal distribution $\mathcal{N}(0, \Sigma)$ with variance $\Sigma \in \mathbb{S}^n_+$
- Probability density function (pdf) has ellipsoidal level lines
- Ellipsoidal $s\sigma$ confidence regions

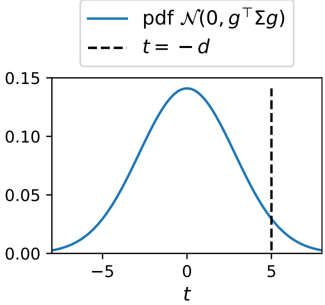
$$P_{x\sim\mathcal{N}(0,\Sigma)}(x \in \mathcal{E}(s^2\Sigma)) = f(s^2, n)$$

($f(s^2, n)$ is cumulative density of χ^2_n dist.)

$f(s^2, n)$	$n = 1$	$n = 2$	$n = 3$	$n = 5$
$s = 1$	0.683	0.393	0.199	0.037
$s = 2$	0.954	0.865	0.739	0.451
$s = 3$	0.997	0.989	0.971	0.891



$$
\begin{aligned}
&P_{x\sim\mathcal{N}(0,\Sigma)}(g^\top x + d \leq 0) = ?\\
&= P_{t\sim\mathcal{N}(0,g^\top\Sigma g)}(t + d \leq 0)\\
&= P_{t\sim\mathcal{N}(0,g^\top\Sigma g)}(t \leq -d)
\end{aligned}
$$



Approximate robust optimization

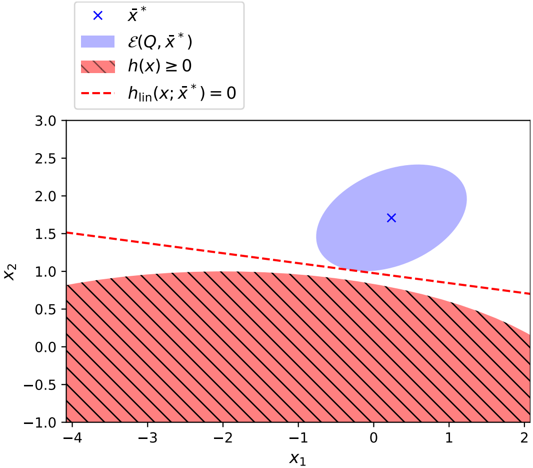
$$
\begin{aligned}
& \min_{\bar{x} \in \mathbb{R}^n} f(\bar{x}) \quad \text{s.t.} \quad 0 \geq h(x) \quad \forall x \in \mathcal{E}(Q, \bar{x}) \qquad (h : \mathbb{R}^n \to \mathbb{R}) \\
= & \min_{\bar{x} \in \mathbb{R}^n} f(\bar{x}) \quad \text{s.t.} \quad 0 \geq \max_{x \in \mathcal{E}(Q, \bar{x})} h(x) \\
\approx & \min_{\bar{x} \in \mathbb{R}^n} f(\bar{x}) \quad \text{s.t.} \quad 0 \geq \max_{x \in \mathcal{E}(Q, \bar{x})} h(\bar{x}) + \nabla h(\bar{x})^\top (x - \bar{x}) \\
= & \min_{\bar{x} \in \mathbb{R}^n} f(\bar{x}) \quad \text{s.t.} \quad 0 \geq h(\bar{x}) + \max_{\Delta x \in \mathcal{E}(Q)} \nabla h(\bar{x})^\top \Delta x \\
= & \min_{\bar{x} \in \mathbb{R}^n} f(\bar{x}) \quad \text{s.t.} \quad 0 \geq h(\bar{x}) + \sqrt{\nabla h(\bar{x})^\top Q \nabla h(\bar{x})}
\end{aligned}
$$

Approximate robust optimization – example

$$\min_{\bar{x} \in \mathbb{R}^2} \quad \bar{x}^\top \bar{x}$$
$$\text{s.t.} \quad x \notin \mathcal{E}(V, c) \qquad \forall x \in \mathcal{E}(Q, \bar{x})$$

$$\min_{\bar{x} \in \mathbb{R}^2} \quad \bar{x}^\top \bar{x}$$
$$\text{s.t.} \quad (x - c)^\top V^{-1}(x - c) \geq 1 \quad \forall x \in \mathcal{E}(Q, \bar{x})$$

Obtain approximate solution $\bar{x}^*$ by robustifying against linearized constraint



Approximate robust optimization – chance constraint variation

$$
\begin{aligned}
& \min_{\bar{x} \in \mathbb{R}^n} \; f(\bar{x}) \quad \text{s.t.} \quad P_{x \sim \mathcal{N}(\bar{x},Q)}\{h(x) \leq 0\} \geq \bar{p} \\
\approx\; & \min_{\bar{x} \in \mathbb{R}^n} \; f(\bar{x}) \quad \text{s.t.} \quad P_{x \sim \mathcal{N}(\bar{x},Q)}\{h(\bar{x}) + \nabla h(\bar{x})^\top (x - \bar{x}) \leq 0\} \geq \bar{p} \\
=\; & \min_{\bar{x} \in \mathbb{R}^n} \; f(\bar{x}) \quad \text{s.t.} \quad P_{\Delta x \sim \mathcal{N}(0,Q)}\{h(\bar{x}) + \nabla h(\bar{x})^\top \Delta x \leq 0\} \geq \bar{p} \\
=\; & \min_{\bar{x} \in \mathbb{R}^n} \; f(\bar{x}) \quad \text{s.t.} \quad h(\bar{x}) + \gamma(\bar{p}) \sqrt{\nabla h(\bar{x})^\top Q \nabla h(\bar{x})} \leq 0
\end{aligned}
$$

$$
\begin{gathered}
\gamma(\bar{p}) = 1 \Leftrightarrow \bar{p} = 0.841 \\
\gamma(\bar{p}) = 2 \Leftrightarrow \bar{p} = 0.977 \\
\gamma(\bar{p}) = 3 \Leftrightarrow \bar{p} = 0.999
\end{gathered}
$$

(cumulative density of normal dist. up to the $\gamma(\bar{p})\sigma$ level)

Uncertain linear dynamical systems

$$x_{k+1} = Ax_k + Bu_k + \Gamma w_k$$

- Stochastic setting

$$x_k \sim \mathcal{N}(\bar{x}_k, P_k),\ w_k \sim \mathcal{N}(0, \Sigma)$$
$$\Rightarrow x_{k+1} \sim \mathcal{N}(\underbrace{A\bar{x}_k + Bu_k}_{\bar{x}_{k+1}}, \underbrace{AP_kA^\top + \Gamma\Sigma\Gamma^\top}_{P_{k+1}})$$

  - Uncertainty set still ellipsoidal :)

- Robust setting

$$x_k \in \mathcal{E}(P_k, \bar{x}_k),\ w_k \in \mathcal{E}(\Sigma)$$
$$\Rightarrow x_{k+1} \in \mathcal{E}(AP_kA^\top, \underbrace{A\bar{x}_k + Bu_k}_{\bar{x}_{k+1}}) + \mathcal{E}(\Gamma\Sigma\Gamma^\top)$$

  - Uncertainty set not ellipsoidal :(

Uncertain linear dynamical systems – robust case

$$x_{k+1} = Ax_k + Bu_k + \Gamma w_k$$

$$x_k \in \mathcal{E}(P_k, \bar{x}_k),\ w_k \in \mathcal{E}(\Sigma) \Rightarrow x_{k+1} \in \mathcal{X}_{k+1}, \quad \text{not ellipsoidal}$$

- Option 1: Ellipsoidal overapproximation $P_{k+1}(\alpha) \supseteq \mathcal{X}_{k+1}$
  - choose arbitrary $\alpha \Rightarrow$ probably bad quality approximation
  - choose best α according to some relevant criterion $\Rightarrow$ optimization problem
- Option 2: Modify assumption on initial uncertainty

$$\begin{bmatrix} x_k \\ w_k \end{bmatrix} \in \mathcal{E}\left(\begin{bmatrix} P_k & 0 \\ 0 & \Sigma \end{bmatrix}, \begin{bmatrix} \bar{x}_k \\ 0 \end{bmatrix}\right) \quad \Rightarrow \quad x_{k+1} \in \mathcal{E}(\underbrace{AP_kA_k^\top + \Gamma\Sigma\Gamma^\top}_{P_{k+1}}, \underbrace{A\bar{x}_k + Bu_k}_{\bar{x}_{k+1}})$$

  - For N-step prediction assume $(x_0, w_0, \ldots, w_{N-1}) \in \mathcal{E}(\operatorname{diag}(P_0, \Sigma, \ldots, \Sigma), (\bar{x}_0, 0, \ldots, 0))$
  - Justification of assumption?

Nonlinear dynamical systems

- Consider a stochastic nonlinear dynamical system

$$x_0 = \bar{\bar{x}}_0, \qquad x_{k+1} = f_k(x_k, u_k, w_k), \qquad k = 0, \dots, N-1. \tag{1}$$

- Noise $w = (w_0, \dots, w_{N-1})$ drawn from ellipsoid $w \in \mathcal{E}(\sigma^2 I)$.
  - uncertainty scaling parameter $\sigma \geq 0$
  - $w \in \mathcal{E}(\sigma^2 I)$ instead of $w \in \mathcal{E}(\sigma^2 W)$ w.l.o.g.
- We are interested in robust constraint satisfaction for all possible trajectories

$$h(x_k, u_k) \leq 0, \quad \forall x_k \in \mathcal{X}_k(u), \quad k = 0, \dots, N, \tag{2}$$

$$h(x_N) \leq 0, \quad \forall x_N \in \mathcal{X}_N(u), \tag{3}$$

where $\mathcal{X}_k(u)$, $k = 0, \dots, N$, is the set of all reachable states at k given controls u.

Approximations

- Model uncertainty tube by ellipsoids around nominal trajectory $\bar{x}$, $\bar{u}$

$$\bar{x}_0 = \bar{\bar{x}}_0, \qquad \bar{x}_{k+1} = f_k(\bar{x}_k, \bar{u}_k, 0), \qquad k = 0, \dots, N-1, \tag{4}$$

$$x_k \in \mathcal{E}(P_k, \bar{x}_k), \qquad k = 0, \dots, N. \tag{5}$$

- Propagate ellipsoids according to dynamics linearized at $\bar{x}$, $\bar{u}$

$$P_0 = 0, \qquad P_{k+1} = A_k P_k A_k^\top + \sigma^2 \Gamma_k \Gamma_k^\top, \qquad k = 0, \dots, N-1. \tag{6}$$

where $A_k = \frac{\partial f_k}{\partial x_k}(\bar{x}_k, \bar{u}_k, 0), \qquad \Gamma_k = \frac{\partial f_k}{\partial w_k}(\bar{x}_k, \bar{u}_k, 0)$.

Simplifications (cont.)

- Approximate robust constraint satisfaction through linearization (componentwise)

$$
\begin{aligned}
0 \geq\ & h_k^i(x_k, \bar{u}_k) && \forall x_k \in \mathcal{E}(P_k, \bar{x}_k) \\
\geq\ & \max_{x_k \in \mathcal{E}(P_k, \bar{x}_k)} h_k^i(x_k, \bar{u}_k) \\
\approx\ & h_k^i(\bar{x}_k, \bar{u}_k) + \sqrt{\nabla_x h_k^i(\bar{x}_k, \bar{u}_k)^\top P_k \nabla_x h_k^i(\bar{x}_k, \bar{u}_k)}, && i = 1, \ldots, n_{h_k}, \\
& && k = 0, \ldots, N-1, \\
0 \geq\ & h_N^i(\bar{x}_N) + \sqrt{\nabla_x h_N^i(\bar{x}_N)^\top P_N \nabla_x h_N^i(\bar{x}_N)}, && i = 1, \ldots, n_{h_N}.
\end{aligned}
$$

- Variation: Single-chance constraints
  - Interpret P_k as variance of normal distribution
  - Multiply back-off by $\gamma(\bar{p})$ to ensure satisfaction of **this specific** constraint with at least probability $\bar{p}$ (approximately)
  - Probability that **no** constraint is violated is lower

Open-loop Robustified NMPC problem

$$\min_{\bar{x}, \bar{u}, P} \quad \sum_{k=0}^{N-1} l(\bar{x}_k, \bar{u}_k) + E(\bar{x}_N) \tag{8a}$$

$$\text{s.t.} \quad \bar{x}_0 = \bar{\bar{x}}_0, \tag{8b}$$

$$\bar{x}_{k+1} = f_k(\bar{x}_k, \bar{u}_k, 0), \quad k = 0, \dots, N-1, \tag{8c}$$

$$P_0 = 0, \tag{8d}$$

$$P_{k+1} = A_k P_k A_k^\top + \sigma^2 \Gamma_k \Gamma_k^\top, \quad k = 0, \dots, N-1, \tag{8e}$$

$$0 \geq h_k^i(\bar{x}_k, \bar{u}_k) + \sqrt{\nabla_x h_k^i(\bar{x}_k, \bar{u}_k)^\top P_k \nabla_x h_k^i(\bar{x}_k, \bar{u}_k)} \quad \begin{array}{l} i = 1, \dots, n_h, \\ k = 0, \dots, N-1, \end{array}, \tag{8f}$$

$$0 \geq h_N^i(\bar{x}_N) + \sqrt{\nabla_x h_N^i(\bar{x}_N)^\top P_N \nabla_x h_N^i(\bar{x}_N)} \quad i = 1, \dots, n_{h_N}, \tag{8g}$$

where $P = (P_0, \dots, P_N)$

Some references

- Material for this talk
  - B. Houska. *Robust Optimization of Dynamic Systems* , PhD thesis, KU Leuven, 2011
  - J. Gillis. *Practical methods for approximate robust periodic optimal control of nonlinear mechanical systems*, PhD thesis, KU Leuven, 2015
  - ... and a lot of wikipedia / stackoverflow / random slides
- To at least mention our most recent results and closely related ones
  - L. Hewing, J. Kabzan, and M. N. Zeilinger. *Cautious model predicitive control using gaussian process regression*. IEEE Transaction on Control Systems Technology, 28(6):2736–2743, 2020
  - X. Feng, S. Di Cairano, R. Quirynen. *Inexact Adjoint-based SQP Algorithm for Real-Time Stochastic Nonlinear MPC*, Proceedings of the IFAC World Congress, 2020
  - A. Zanelli, J. Frey, **F. Messerer**, M. Diehl. *Zero-Order Robust Nonlinear Model Predictive Control with Ellipsoidal Uncertainty Sets*, IFAC-NMPC 2021
  - **F. Messerer**, M. Diehl. *An Efficient Algorithm for Tube-based Robust Nonlinear Optimal Control with Optimal Linear Feedback*, 2021