Documentation:
Policy synthesis via formal abstraction

December 14, 2017

Contents

Chapter 1

Do abstraction of LTI system

1.1 Computation of simulation relation

Define LTI system as

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k, \qquad w_k \sim \mathcal{N}(0, \Sigma) \\ y_k &= Cx_k \end{aligned} \tag{1.1}$$

with

- x state of size n
- u input of size m
- A matrix of size $n \times n$
- B matrix of size $n \times m$
- y the output (used to compare accuracy)
- C output matrix of size $q \times n$
- Σ diagonal matrix

These stochastic transitions (1.33) can be abstracted to a finite state model with states $s \in S = 1, 2, ...,$. Each state s is associated to a representative point $x_s \in \mathbb{R}^n$ and associated to a cell $\Delta_s = \{x_s\} \oplus \prod_i^n [-d_i, d_i]$. Further it has transitions

$$t_{grid}(s'|s, u) = \hat{t}\left(\Delta_{s'} \mid x_s, u\right) \tag{1.2}$$

where $\hat{t}$ is the stochastic transition kernel associated with (1.33).

As written in the paper, the difference between the concrete and abstract system evolves over time as follows

$$x_+ - \tilde{x}_+ = (A + BK)(x - \tilde{x}) + \mathbf{r} \tag{1.3}$$

with

- $\mathbf{r}$ in a polytope, i.e., $\mathbf{r} \in \mathcal{V}(r_i)$, the polytope generated from vertices r_i.

Consider a set defined as

$$\mathcal{R} := \{(\tilde{x}, x) \mid (x - \tilde{x})^T M (x - \tilde{x}) \leq \epsilon^2\} \tag{1.4}$$

1.1.1 Optimize $\mathcal{R}$ for given grid $d_1, d_2, \ldots, d_3$

Objective: Design M, K and ϵ such that if $(\tilde{x}, x) \in \mathcal{R}$ then also

$$\{(x_+ - \tilde{x}_+) | \text{ s.t. (1.3) } \forall \mathbf{r} \in \mathcal{V}(r_i)\} \subseteq \mathcal{R}.$$

More over for all $(\tilde{x}, x) \in \mathcal{R}$ it should hold that $d(\tilde{y}, y) \leq \epsilon$. The latter can be expressed as $C^T C \preceq M$. The former can be written with matrix inequalities as

$$\begin{aligned} (x_+ - \tilde{x}_+)^T M (x_+ - \tilde{x}_+) &\leq \epsilon^2 \\ ((A + BK)(x - \tilde{x}) + \mathbf{r})^T M ((A + BK)(x - \tilde{x}) + \mathbf{r}) &\leq \epsilon^2 \end{aligned}$$

Hence we get something of this form

$$(x - \tilde{x})^T M (x - \tilde{x}) \leq \epsilon^2 \implies ((A + BK)(x - \tilde{x}) + \mathbf{r})^T M ((A + BK)(x - \tilde{x}) + \mathbf{r}) \leq \epsilon^2$$

S-procedure[a]
The implications

$$x^T F_1 x + 2g_1^T x + h_1 \leq 0 \implies x^T F_2 x + 2g_2^T x + h_2 \leq 0 \tag{1.5}$$

holds if and only if there exists $\lambda \geq 0$ such that

$$\lambda \begin{bmatrix} F_1 & g_1 \\ g_1^T & h_1 \end{bmatrix} - \begin{bmatrix} F_2 & g_2 \\ g_2^T & h_2 \end{bmatrix} \succeq 0 \tag{1.6}$$

[a] https://en.wikipedia.org/wiki/S-procedure

Using the S-procedure we get

$$(x-\tilde{x})^T(A+BK)^T M(A+BK)(x-\tilde{x}) + 2\mathbf{r}^T M(A+BK)(x-\tilde{x}) + \mathbf{r}^T M\mathbf{r} \le \epsilon^2 \tag{1.7}$$

$$\lambda \begin{bmatrix} M & 0 \\ 0 & -\epsilon^2 \end{bmatrix} - \begin{bmatrix} (A+BK)^T M(A+BK) & (A+BK)^T M\mathbf{r} \\ \mathbf{r}^T M(A+BK) & \mathbf{r}^T M\mathbf{r} - \epsilon^2 \end{bmatrix} \succeq 0 \tag{1.8}$$

$$\begin{bmatrix} \lambda M - ((A+BK)^T M(A+BK)) & -(A+BK)^T M\mathbf{r} \\ -\mathbf{r}^T M(A+BK) & (1-\lambda)\epsilon^2 - \mathbf{r}^T M\mathbf{r} \end{bmatrix} \succeq 0 \tag{1.9}$$

$$\begin{bmatrix} \lambda M & 0 \\ 0 & (1-\lambda)\epsilon^2 \end{bmatrix} - \begin{bmatrix} ((A+BK)^T M(A+BK)) & (A+BK)^T M\mathbf{r} \\ \mathbf{r}^T M(A+BK) & \mathbf{r}^T M\mathbf{r} \end{bmatrix} \succeq 0 \tag{1.10}$$

$$\begin{bmatrix} \lambda M & 0 \\ 0 & (1-\lambda)\epsilon^2 \end{bmatrix} - \begin{bmatrix} (A+BK)^T M \\ \mathbf{r}^T M \end{bmatrix} M^{-1} \begin{bmatrix} (A+BK)^T M \\ \mathbf{r}^T M \end{bmatrix}^T \succeq 0 \tag{1.11}$$

$$\begin{bmatrix} \lambda M & 0 & (A+BK)^T M \\ 0 & (1-\lambda)\epsilon^2 & \mathbf{r}^T M \\ M(A+BK) & M\mathbf{r} & M \end{bmatrix} \succeq 0 \tag{1.12}$$

$$\begin{bmatrix} \lambda M^{-1} & 0 & M^{-1}(A+BK)^T \\ 0 & (1-\lambda)\epsilon^2 & \mathbf{r}^T \\ (A+BK)M^{-1} & \mathbf{r} & M^{-1} \end{bmatrix} \succeq 0 \tag{1.13}$$

$$\begin{bmatrix} \lambda M^{-1} & 0 & M^{-1}(A+BK)^T \\ 0 & (1-\lambda)\epsilon^2 & r_i^T \\ (A+BK)M^{-1} & r_i & M^{-1} \end{bmatrix} \succeq 0, \ \forall r_i \tag{1.14}$$

$$\tag{1.15}$$

Remark that this implies that $1-\lambda \ge 0$ hence $1 \ge \lambda \ge 0$. And remark that

The objective to find a minimal ϵ can be expressed as follows

$$\mathbf{Objective}: \min_{M_{inv},L} \epsilon^2 \tag{1.16}$$

$$\begin{bmatrix} \lambda M_{inv} & 0 & M_{inv}A^T + L^T B^T \\ 0 & (1-\lambda)\epsilon^2 & r_i^T \\ AM_{inv} + BL & r_i & M_{inv} \end{bmatrix} \succeq 0 \tag{1.17}$$

$$\begin{bmatrix} M_{inv} & M_{inv}C^T \\ CM_{inv} & I \end{bmatrix} \succeq 0 \tag{1.18}$$

with $LM = K$ and $M^{-1} = M_{inv}$. This has been implemented as function $eps\_err()$ in python.

Verify that Polytope $\mathcal{V}(r_i)$ is in relation.

$\mathcal{V}(r_i)$ is in relation $\mathcal{R} := \{(\tilde{x}, x) \mid (x-\tilde{x})^T M(x-\tilde{x}) \le \epsilon^2\}$ if for all r_i it holds that

$$r_i^T M r_i \le \epsilon^2$$

.

Plot simulation relation

Input

$$\mathcal{R} := \{x \mid x^T M_\epsilon x \leq 1\}$$

1. Compute $M_\epsilon^{1/2} = U\Sigma^{1/2}$ with singular value decomposition $M_\epsilon = U\Sigma V^T$
2. Switch variable
$$\mathcal{R} := \{(\tilde{x}, x) \mid z^T z \leq 1 \text{ with } z = M_\epsilon^{1/2} x\}$$
3. compute outline given angle α
$$z(\alpha) = \begin{bmatrix} \cos(\alpha) \\ \sin(\alpha) \end{bmatrix}$$
remark $z^T z = 1$. then $x(\alpha) = \Sigma^{-1/2} U^T z(\alpha)$.

1.1.2 Optimise gridding for 2d models

For 2 D models a routine *tune_dratio* finds the optimal gridding ratio.

1.2 LTI with different noise sources

Define the **concrete** LTI system as

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k \qquad w_k \sim \mathcal{N}(0, \Sigma) \\ y_k &= Cx_k \end{aligned} \tag{1.19}$$

with 1. x state of size n 2. u input of size m 3. A matrix of size $n \times n$ 4. B matrix of size $n \times m$ 5. y the output (used to compare accuracy) 6. C output matrix of size $q \times n$ Suppose that an **abstract** LTI system has been given as

$$\begin{aligned} \tilde{x}_{k+1} &= A\tilde{x}_k + B\tilde{u}_k + \tilde{w}_k \qquad \tilde{w}_k \sim \mathcal{N}(0, \Sigma) \\ y_k &= Cx_k \end{aligned} \tag{1.20}$$

To lift the two systems, we consider the existence of the following combined system

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + B_w s_k \\ \tilde{x}_{k+1} &= A\tilde{x}_k + B\tilde{u}_k + \tilde{B}_w s_k \qquad s_k \sim \mathcal{N}(0, \Sigma) \end{aligned} \tag{1.21}$$

Given $u_k = \tilde{u}_k + K(x_k - \tilde{x}_k)$, the state difference evolves as

$$\begin{aligned} x_{k+1} - \tilde{x}_{k+1} &= (A + BK)(x_k - \tilde{x}_k) + (B_w - \tilde{B}_w)s_k \\ s_k &\sim \mathcal{N}(0, \Sigma) \end{aligned} \tag{1.22}$$

Given that Σ is a diagonal matrix, we can easily compute the probability that s_k is in a hypercube, i.e., $s_k \in \prod_i [-a_i, a_i]$, with associated probability $1-\delta = \prod_i \mathcal{N}([-a_i, a_i]|0, \sigma_i)$. Hence to quantify a simulation relation we consider the $1 - \delta$ invariance of transitions

$$x_{k+1} - \tilde{x}_{k+1} = (A + BK)(x_k - \tilde{x}_k) + (B_w - \tilde{B}_w)\mathbf{s} \quad \forall \mathbf{s} \in \prod_i [-a_i, a_i] \tag{1.23}$$

In combination with the gridding we get that

$$\begin{aligned} x_{k+1} - \tilde{x}_{k+1} = (A + BK)(x_k - \tilde{x}_k) + (B_w - \tilde{B}_w)\mathbf{s} + \mathbf{r} \\ \forall \mathbf{s} \in \prod_i [-a_i, a_i], \text{ and } \forall \mathbf{r} \in \prod_i [-d_i, d_i], \end{aligned} \tag{1.24}$$

Not implemented

1.3 Kalman filtered innovation models

Consider a Gaussian LTI system:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_t + w_k, \\ z_k &= Cx_k + Du_k + v_k. \end{aligned} \tag{1.25}$$

with $w_k \sim \mathcal{N}(0, \mathcal{W})$ and $v_k \sim \mathcal{N}(0, \mathcal{V})$.

At $k = 0$, we know $x_0 \sim \rho$ with $\rho := \mathcal{N}(x_\rho, P_\rho)$. Thus, before receiving a measurement z_0, the distribution of the belief is defined as $\mathcal{N}(x_{0|-}, P_{0|-})$

$$\hat{x}_{0|-} := x_\rho \tag{1.26}$$

$$P_{0|-} := P_\rho \tag{1.27}$$

After receiving the measurement z_0, this is updated to $\mathcal{N}(\hat{x}_{0|0}, P_{0|0})$

$$\hat{x}_{0|0} := x_\rho + L_0(z_0 - Cx_\rho) \tag{1.28}$$

$$P_{0|0} := (I - L_0C)P_\rho(I - L_0C)^T + L_0\mathcal{V}L_0^T \tag{1.29}$$

$$\text{with } L_0 = P_\rho C^T \left(CP_\rho C^T + \mathcal{V}\right)^{-1} \tag{1.30}$$

We represent the belief state $\mathcal{N}(\hat{x}_{0|0}, P_{0|0})$ by $b_0 := (\hat{x}_{0|0}, P_{0|0}) \in \mathbb{R}^n \times \mathbb{S}^n$.

The dynamics of the Kalman filter are given as

$$
\begin{aligned}
\textbf{Predict} \qquad \hat{x}_{k|k-1} &= A\hat{x}_{k-1|k-1} + Bu_{k-1} \\
P_{k|k-1} &= AP_{k-1|k-1}A^T + \mathcal{W} \\
\textbf{Update} \qquad e_k &= z_k - C\hat{x}_{k|k-1} \\
S_k &= CP_{k|k-1}C^T + \mathcal{V} \\
L_k &= P_{k|k-1}C^T S_k^{-1} \\
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + L_k e_k \\
P_{k|k} &= (I - L_k C)P_{k|k-1}
\end{aligned}
$$

Joseph Formula

$$P_{k|k} = (I - L_k C)P_{k|k-1}(I - L_k C_k)^T + L_k \mathcal{V}_k L_k^T$$

Observability based

$$P_{k|k}^{-1} = P_{k|k-1}^{-1} + C_k^T \mathcal{V}_k^{-1} C_k$$

Though the covariance of the belief state is defined as

$$P_{k|k} = (I - L_k C)P_{k|k-1}(I - L_k C_k)^T + L_k \mathcal{V}_k L_k^T,$$

The update equations for $P_{k|k-1}$ are more well know:

$$P_{k+1|k} = (A - K_k C)P_{k|k-1}(A - K_k C_k)^T + K_k \mathcal{V}_k K_k^T + \mathcal{W}$$

with $K_k = AL_k$.

Hence, the belief state is updated as

$$\hat{x}_{k|k} = A\hat{x}_{k-1|k-1} + Bu_{k-1} + L_k e_k \tag{1.31}$$

$$P_{k|k} = f(P_{k-1|k-1}) \tag{1.32}$$

We now want to model the random variable $s_k = L_k e_k$. We know that s_k evolves as a zero mean Gaussian distributed stochastic process. Further

$$
\begin{aligned}
\mathbf{E}[s_k] &= 0 \\
\mathbf{E}[s_k s_k^T] &= L_k \mathbf{E}[e_k e_k^T] L_k^T, \text{ and } \mathbf{E}[e_k e_k^T] = S_k \\
e_k &= C\left(x_k - \hat{x}_{k|k-1}\right) + v_k \\
\mathbf{E}[e_k e_k^T] &= CP_{k|k-1}C^T + \mathcal{V} \\
\mathbf{E}[s_k s_k^T] &= L_k S_k L_k^T, \\
\mathbf{E}[s_k s_k^T] &= P_{k|k-1}C^T S_k^{-1} CP_{k|k-1}, \\
\mathbf{E}[s_k s_k^T] &= P_{k|k-1}C^T \left(CP_{k|k-1}C^T + \mathcal{V}\right)^{-1} CP_{k|k-1}, \\
\mathbf{E}[s_k s_k^T] &= P_{k|k-1} - P_{k|k}
\end{aligned}
$$

Consider a LTI system

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k \\ z_k &= Cx_k + Du_k + v_k \end{aligned} \tag{1.33}$$

with $x \in \mathbb{R}^n$ with stochastic disturbances $w_t \sim \mathcal{N}(0, \mathcal{W})$, and $v_t \sim \mathcal{N}(0, \mathcal{V})$. (1.33) defines a MDP with state space $\mathbb{X} = \mathbb{R}^n$, initial distribution $\rho := \mathcal{N}(x_\rho, P_\rho)$, control inputs $u_t \in \mathbb{R}^m$, and transition kernel t defined based on (1.33). This is a partially observable MDP that can only be observed via $z_t \in \mathbb{R}^q$.

Before receiving a measurement z_0, the initial state is distributed as $\mathcal{N}(x_{0|-}, P_{0|-})$, with $\hat{x}_{0|-} := x_\rho$ and $P_{0|-} := P_\rho$. After receiving the measurement z_0, this is updated to

$$\begin{aligned} \hat{x}_{0|0} &:= x_\rho + L_0(z_0 - Cx_\rho), \\ P_{0|0} &:= (I - L_0C)P_\rho(I - L_0C)^T + L_0\mathcal{V}L_0^T, \\ &\text{with } L_0 = P_\rho C^T \left(CP_\rho C^T + \mathcal{V}\right)^{-1}, \end{aligned}$$

with $\mathbb{P}(x_t \in \cdot \,|\, \rho, z_0) := \mathcal{N}(\hat{x}_{0|0}, P_{0|0})$. This probability distribution defines a belief state as $b_0 := (\hat{x}_{0|0}, P_{0|0}) \in \mathbb{R}^n \times \mathbb{S}^n$. The belief space $\mathbb{X}_b$ is a finite dimensional space and can be parameterized. For example, let $\mathcal{G}$ denote the Gaussian belief space of dimension n, i.e. the space of Gaussian probability measures over $\mathbb{R}^n$. For brevity, we identify the Gaussian measures with their finite parametrization, mean and covariance matrix. Thus, $\mathbb{X}_b = \mathbb{R}^n \times \mathbb{S}^n$.

The dynamics of $b_k := (\hat{x}_{k|k}, P_{k|k})$ are defined via the Kalman filter, that is

$$\begin{aligned} &\textit{predict:} && \hat{x}_{k|k-1} = A\hat{x}_{k-1|k-1} + Bu_{k-1} \\ & && P_{k|k-1} = AP_{k-1|k-1}A^T + \mathcal{W}, \\ &\textit{update:} && \hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k\left(z_k - C\hat{x}_{k|k-1}\right) \\ & && P_{k|k} = (I - L_kC)P_{k|k-1} \end{aligned}$$

with $L_k = P_{k|k-1}C^T \left(CP_{k|k-1}C^T + \mathcal{V}\right)^{-1}$.

This defines a belief MDP with stochastic transitions of the belief state given as

$$\hat{x}_{k|k} = A\hat{x}_{k-1|k-1} + Bu_{k-1} + P_{k|k-1}C^T s_k \tag{1.34}$$

$$P_{k|k} = f(P_{k-1|k-1}) \tag{1.35}$$

with $e_k \sim \mathcal{N}(0, S_k^{-1})$ and $S_k = \left(CP_{k|k-1}C^T + \mathcal{V}\right)$.

As a first simplification, we can replace the stochastic transitions in (1.34) by

$$\hat{x}_k = A\hat{x}_{k-1} + B\hat{u}_{k-1} + \bar{P}C^T\hat{s}_k, \tag{1.36}$$

with $\hat{s}_k \sim \mathcal{N}(0, \hat{S}_{inv})$ and $\hat{S}_{inv} \preceq S_k^{-1}$ for all k.

The computational implementation is as follows:

$$\textbf{objective:} \quad \min_{W \succeq 0, s_{inv}} \ \text{trace}(W) \tag{1.37}$$

$$\text{s.t.} \quad W \succeq S_k^{-1} - S_{inv} \succeq 0 \tag{1.38}$$

And (1.38) is equivalent to

$$W + S_{inv} - S_k^{-1} \succeq 0, \quad S_k^{-1} - S_{inv} \succeq 0 \tag{1.39}$$

$$\begin{bmatrix} W + S_{inv} & I \\ I & S_k \end{bmatrix} \succeq 0, \quad \begin{bmatrix} S_{inv}^{-1} & I \\ I & S_k^{-1} \end{bmatrix} \succeq 0 \tag{1.40}$$

with $S_k = \left(CP_{k|k-1}C^T + \mathcal{V}\right)$

$$\begin{bmatrix} W + S_{inv} & I \\ I & \left(CP_{k|k-1}C^T + \mathcal{V}\right) \end{bmatrix} \succeq 0, \quad S_{inv}^{-1} - \left(CP_{k|k-1}C^T + \mathcal{V}\right) \succeq 0 \tag{1.41}$$

Given $P^- \preceq P_{k|k-1} \preceq P^+$

$$\begin{bmatrix} W + S_{inv} & I \\ I & \left(CP^-C^T + \mathcal{V}\right) \end{bmatrix} \succeq 0, \quad \left(CP^+C^T + \mathcal{V}\right)^{-1} - S_{inv} \succeq 0 \tag{1.42}$$

Note $W \succeq \left(CP^-C^T + \mathcal{V}\right)^{-1} - S_{inv}$. Hence the final solution is

$$W = \left(CP^-C^T + \mathcal{V}\right)^{-1} - \left(CP^+C^T + \mathcal{V}\right)^{-1} \tag{1.43}$$

$$S_{inv} = \left(CP^+C^T + \mathcal{V}\right)^{-1} \tag{1.44}$$

These stochastic transitions (1.36) can then be further abstracted to a finite state model $\hat{\mathcal{B}}$ with states $s \in S = 1, 2, ...,$. Each state s is associated to a representative point $x_s \in \mathbb{X}_b$ and associated to a cell $\Delta_s = \{x_s\} \oplus \prod_n [-d, d]$. Further the absstract system has transitions

$$t_{grid}(s'|s, u) = \hat{t}\left(\Delta_{s'} \mid x_s, u\right) \tag{1.45}$$

where $\hat{t}$ is the stochastic transition kernel associated with (1.36).

Consider a simulation relation defined as

$$\begin{aligned} \mathcal{R} := \big\{ &(s, b_k) | (\hat{x}_{k|k} - x_s)^T M (\hat{x}_{k|k} - x_s) \leq \epsilon, \\ &P^- \preceq P_{k|k} \preceq P^+ \text{ with } b_k = (\hat{x}_{k|k}, P_{k|k}) \big\}, \end{aligned} \tag{1.46}$$

and an interface

$$\mathcal{U}_v(\hat{u}, \hat{x}, \hat{x}_{|}) := K(\hat{x}_{|} - \hat{x}) + \hat{u}$$

for some matrices M, K, P^+, P^-.

We can quantify the difference between $\mathcal{B}$ and $\hat{\mathcal{B}}$ via (1.46) by verifying that for all $(\hat{x}_k, \hat{x}_{k|k}) \in \mathcal{R}$ with probability at least $1 - \delta$ it holds that $(\hat{x}_{k+1}, \hat{x}_{k+1|k+1}) \in \mathcal{R}$. Consider a choice for the lifted stochastic transitions for (1.36) and (1.49), denoted $\mathbb{W}_x((\hat{x}_k, \hat{x}_{k|k}) \in$

$\cdot|\hat{u}_{k-1}, \hat{x}_{k-1}, \hat{x}_{k-1|k-1})$, based on the combined stochastic difference equation given as

$$\begin{aligned}\hat{x}_{k+1} &= A\hat{x}_k + B\hat{u}_k + \bar{P}C^T\hat{s}_{k+1},\\ \hat{x}_{k+1|k+1} &= A\hat{x}_{k|k} + Bu_k + \bar{P}C^T(\hat{s}_{k+1} + s^{\Delta}_{k+1})\\ &\quad + \Delta_{k+1}(\hat{s}_{k+1} + s^{\Delta}_{k+1})\end{aligned}$$

with $\Delta_k := (P_{k|k-1}C^T - \bar{P}C^T)$ and with $\hat{s}_k \sim \mathcal{N}(0, \hat{S}_{inv})$ and $s^{\Delta}_k \sim \mathcal{N}(0,\ S_k^{-1} - \hat{S}_{inv})$.

We can now choose the lifted stochastic transition kernel $\mathbb{W}_t$ for the concrete belief MDP $\mathcal{B}$ and the abstracted finite MDP $\hat{\mathcal{B}}$ as follows. Denote $b = (\hat{x}_|, P)$ and $b_+ = (\hat{x}_{+|+}, P_+)$, then $\mathbb{W}_t$ is computed as

$$\begin{aligned}&\mathbb{W}_t((s_+, b_+) \in \cdot\,|\hat{u}, s, b)\\ &:= \begin{cases} \mathbb{W}_x((\Delta_{s_+}, \hat{x}_{+|+}) \in \cdot\,|\hat{u}, x_s, \hat{x}_|) & \text{for } P_+ = f(P)\\ 0 & \text{else}\end{cases}\end{aligned}$$

For this choice of $\mathbb{W}_x$, the difference expression in (1.46) evolves as

$$\begin{aligned}\hat{x}_{k+1|k+1} - \hat{x}_{k+1} &= (A + BK)(\hat{x}_{k|k} - \hat{x}_{k-1})\\ &\quad + \bar{P}C^T s^{\Delta}_{k+1} + \Delta_{k+1}(\hat{s}_{k+1} + s^{\Delta}_{k+1})\end{aligned} \tag{1.47}$$

with $\Delta_{k+1} := (P_{k+1|k}C^T - \bar{P}C^T)$, and with $\hat{s}_{k+1} \sim \mathcal{N}(0, \hat{S}_{inv})$ and $s^{\Delta}_{k+1} \sim \mathcal{N}(0,\ S^{-1}_{k+1} - \hat{S}_{inv})$. For all $\hat{x}_{k+1}$, there exists $\mathbf{r} \in \prod_n[-d, d]$ such that $\hat{x}_{k+1} - \mathbf{r} \in \{x_s | s \in S\}$. Therefore we can write the update of the difference expression as

$$\begin{aligned}\hat{x}_{+|+} - \hat{x}_{s_+} &= (A + BK)(\hat{x}_| - \hat{x}_s) + \mathbf{r}\\ &\quad + \bar{P}C^T s^{\Delta}_{k+1} + \Delta_{k+1}(\hat{s}_{k+1} + s^{\Delta}_{k+1}).\end{aligned} \tag{1.48}$$

Given that $(\hat{x}_| - \hat{x}_s)$ and $\mathbf{r}$ belongs to a bounded set, we can bound the influence of the noise terms s^{Δ}_{k+1} and $\hat{s}_{k+1}$ with respect to a probability at least $1 - \delta$ for which the update is always in $\mathcal{R}$ cf. (1.46).

$$\hat{x}_{+|+} - \hat{x}_{s_+} = (A + BK)(\hat{x}_| - \hat{x}_s) + \mathbf{r}$$
$$+\bar{P}C^T s^{\Delta}_{k+1} + \Delta_{k+1} s^{\Delta}_{k+1} + \Delta_{k+1}\hat{s}_{k+1}. \tag{1.49}$$

We want to find an upper bound for the random variable $\Delta_{k+1}\hat{s}_{k+1}$. This random variable has Gaussian distribution with covariance $\Delta_{k+1} S_{inv} \Delta^T_{k+1}$. Hence we look for the minimal S_Δ (with respect to the trace (or determinant?))such that $S_\Delta \succeq \Delta_{k+1} S_{inv} \Delta^T_{k+1}$. This is equivalent to

$$S_\Delta - \Delta_{k+1} S_{inv} \Delta^T_{k+1} \succeq 0 \tag{1.50}$$
$$\begin{bmatrix} S_\Delta & (P_{k+1|k} - \bar{P})C^T \\ C(P_{k+1|k} - \bar{P}) & \left(CP^+C^T + \mathcal{V}\right) \end{bmatrix} \succeq 0 \tag{1.51}$$

Write $P_{k+1|k} - \bar{P}$ as $H^+ - H^- = P_{k+1|k} - \bar{P}$ with minimal matrices $H^+ \succeq 0$ and $H^- \succeq 0$ (if $xH^+x > 0$ then $xH^-x = 0$, and if $xH^-x > 0$ then $xH^+x = 0$,). Assume that $P^- \preceq \bar{P} \preceq P^+$, then based on $P^- - \bar{P} \preceq P_{k|k-1} - \bar{P} \preceq P^+ - \bar{P}$ it follows that $H^- \preceq \bar{P} - P^-$ and $H^+ \preceq P^+ - \bar{P}$.

$$\begin{bmatrix} S_\Delta & (H^+ - H^-)C^T \\ C(H^+ - H^-) & \left(CP^+C^T + \mathcal{V}\right) \end{bmatrix} \succeq 0 \tag{1.52}$$
$$\begin{bmatrix} S_\Delta - H^+ - H^- & 0 \\ 0 & \left(CP^+C^T + \mathcal{V}\right) - C(H^+ + H^-)C^T \end{bmatrix} \tag{1.53}$$
$$+ \begin{bmatrix} I \\ C \end{bmatrix} H^+ \begin{bmatrix} I \\ C \end{bmatrix}^T + \begin{bmatrix} I \\ -C \end{bmatrix} H^- \begin{bmatrix} I \\ -C \end{bmatrix}^T \succeq 0 \tag{1.54}$$

We can see that $\left(CP^+C^T + \mathcal{V}\right) - C(H^+ + H^-)C^T \succeq 0$ always holds. Therefore $S_\Delta - H^+ - H^- \succeq 0$ is a sufficient condition. Since $xH^+x > 0$ then $xH^-x = 0$, and if $xH^-x > 0$ then $xH^+x = 0$, we design S_Δ to be minimal and such that $S_\Delta \succeq \bar{P} - P^-$ and $S_\Delta \succeq P^+ - \bar{P}$.

Find an upper bound on the the random variable $P_{k+1|k}C^T s^{\Delta}_{k+1}$. This random variable has Gaussian distribution with covariance $P_{k+1|k}C^T(S_k^{-1} - \hat{S}_{inv})CP_{k+1|k}$. Hence we look for the minimal S_Δ (with respect to the trace (or determinant?))such that $W_\Delta \succeq P_{k+1|k}C^T W C P_{k+1|k}$.

$$\begin{bmatrix} W_\Delta & P_{k+1|k}C^T \\ CP_{k+1|k} & W^{-1} \end{bmatrix} \succeq 0 \tag{1.55}$$
$$\begin{bmatrix} W_\Delta & P_{k+1|k}C^T W \\ WCP_{k+1|k} & W \end{bmatrix} \succeq 0 \tag{1.56}$$
$$\begin{bmatrix} W_\Delta + P_{k+1|k} & 0 \\ 0 & W + WCP_{k+1|k}C^T W \end{bmatrix} - \begin{bmatrix} I \\ -WC \end{bmatrix} P_{k+1|k} \begin{bmatrix} I \\ -WC \end{bmatrix}^T \succeq 0 \tag{1.57}$$

Infeasible: A sufficient condition follows as

$$\begin{bmatrix} W_\Delta + P^- & 0 \\ 0 & W + WCP^-C^T W \end{bmatrix} - \begin{bmatrix} I \\ -WC \end{bmatrix} P^+ \begin{bmatrix} I \\ -WC \end{bmatrix}^T \succeq 0 \tag{1.58}$$
$$\begin{bmatrix} W_\Delta + P^- - P^+ & P^+C^T W \\ WCP^+ & W + WC(P^- - P^+)C^T W \end{bmatrix} \succeq 0 \tag{1.59}$$

As an alternative $W = \left(CP^-C^T + \mathcal{V}\right)^{-1} - \left(CP^+C^T + \mathcal{V}\right)^{-1}$ together with $W_\Delta \succeq P_{k+1|k}C^TWCP_{k+1|k}$ gives

$$W_\Delta - P_{k+1|k}C^T\left(\left(CP^-C^T + \mathcal{V}\right)^{-1} - \left(CP^+C^T + \mathcal{V}\right)^{-1}\right)CP_{k+1|k} \succeq 0 \tag{1.60}$$

$$W_\Delta + P_{k+1|k}C^T\left(CP^+C^T + \mathcal{V}\right)^{-1}CP_{k+1|k} - P_{k+1|k}C^T\left(CP^-C^T + \mathcal{V}\right)^{-1}CP_{k+1|k} \succeq 0 \tag{1.61}$$

$$\begin{bmatrix} W_\Delta + P_{k+1|k}C^T\left(CP^+C^T + \mathcal{V}\right)^{-1}CP_{k+1|k} & P_{k+1|k}C^T \\ CP_{k+1|k} & CP^-C^T + \mathcal{V} \end{bmatrix} \succeq 0 \tag{1.62}$$

$$\begin{bmatrix} W_\Delta + P_{k+1|k}C^T\left(CP^+C^T + \mathcal{V}\right)^{-1}CP_{k+1|k} - P_{k+1|k} & 0 \\ 0 & \mathcal{V} \end{bmatrix} + \begin{bmatrix} I \\ C \end{bmatrix} P_{k+1|k} \begin{bmatrix} I \\ C \end{bmatrix} \succeq 0 \tag{1.63}$$

$$\begin{bmatrix} W_\Delta + P_{k+1|k}C^T\left(CP^+C^T + \mathcal{V}\right)^{-1}CP_{k+1|k} - P_{k+1|k} & 0 \\ 0 & \mathcal{V} \end{bmatrix} + \begin{bmatrix} I \\ C \end{bmatrix} P^- \begin{bmatrix} I \\ C \end{bmatrix} \succeq 0 \tag{1.64}$$

Find $P_{k+1|k}C^T\left(CP^+C^T + \mathcal{V}\right)^{-1}CP_{k+1|k} \succeq M$ go over P_+ instead of over P_-

$$C^T\left(CP^+C^T + \mathcal{V}\right)^{-1}C - P_{k+1|k}^{-1}MP_{k+1|k}^{-1} \succeq 0 \tag{1.65}$$

$$\begin{bmatrix} C^T\left(CP^+C^T + \mathcal{V}\right)^{-1}C & P_{k+1|k}^{-1} \\ P_{k+1|k}^{-1} & M^{-1} \end{bmatrix} \succeq 0 \tag{1.66}$$

$$\begin{bmatrix} C^T\left(CP^+C^T + \mathcal{V}\right)^{-1}C + P_{k+1|k}^{-1} & 0 \\ 0 & M^{-1} + P_{k+1|k}^{-1} \end{bmatrix} - \begin{bmatrix} I \\ -I \end{bmatrix} P_{k+1|k}^{-1} \begin{bmatrix} I \\ -I \end{bmatrix}^T \succeq 0 \tag{1.67}$$

$$\begin{bmatrix} C^T\left(CP^+C^T + \mathcal{V}\right)^{-1}C + P_{k+1|k}^{-1} & 0 \\ 0 & M^{-1} + P_{k+1|k}^{-1} \end{bmatrix} - \begin{bmatrix} I \\ -I \end{bmatrix} P_-^{-1} \begin{bmatrix} I \\ -I \end{bmatrix}^T \succeq 0 \tag{1.68}$$

$$\begin{bmatrix} C^T\left(CP^+C^T + \mathcal{V}\right)^{-1}C + P_{k+1|k}^{-1} - P_-^{-1} & P_-^{-1} \\ P_-^{-1} & M^{-1} + P_{k+1|k}^{-1} - P_-^{-1} \end{bmatrix} \succeq 0 \tag{1.69}$$

$$\tag{1.70}$$

$$\begin{bmatrix} W_\Delta & (\bar{P} + H^+ - H^-)C^TW \\ WC(\bar{P} + H^+ - H^-) & W \end{bmatrix} \succeq 0 \tag{1.71}$$

$$\begin{bmatrix} W_\Delta + H^+ + H^- & \bar{P}C^TW \\ WC\bar{P} & W + WCH^+C^TW + WCH^-C^TW \end{bmatrix} \tag{1.72}$$

$$- \begin{bmatrix} I \\ -WC \end{bmatrix} H^+ \begin{bmatrix} I \\ -WC \end{bmatrix}^T - \begin{bmatrix} I \\ WC \end{bmatrix} H^- \begin{bmatrix} I \\ WC \end{bmatrix}^T \succeq 0 \tag{1.73}$$

1.4 Model-reduction + ...

Chapter 2

Non-Gaussian systems

Consider the gridding of a stochastic process. For non-gaussian systems it makes sense to use some of the well known measures between probability measures.

The **total variation distance** $\delta(P, Q)$ is the most promising,

$$\delta(P, Q) = \sup\{|P(A) - Q(A)| A \in \Sigma \text{ is a measurable event.}\} \tag{2.1}$$

Via Pinskers inequality we have that

$$\delta(P, Q) \leq \sqrt{1/2 D_{KL}(P||Q)}. \tag{2.2}$$

where the latter is the Kullback-Leibler divergence

Chapter 3

Shrinking and expanding polytopes

3.1 Based on ellipsoid sets

Consider polytope composed of the intersection of half spaces. We develop the reasoning first in 3D and then expend it to multidimensional vectors. Given a plane defined as

$$ax + by + cz = d,$$

this defines a half space as

$$\{(x, y, z) \mid ax + by + bz \leq d\}.$$

Consider the case where $a^2 + b^2 + c^2 = 1$, then the vector (a, b, c) defines the normal vector on the plane. The plane is shifted with vector $r_0 = (ad, bd, cd)$.

$$\{(x, y, z) \mid a(x - ad) + b(y - bd) + c(z - cd) \leq 0\}.$$

We can now shift it as follows r_0 as $(d + \epsilon)(a, b, c)$

In conclusion we can expand the half place with $\|(x, y, z)\| \leq \epsilon$

$$\{(x, y, z) \mid ax + by + bz \leq d + \epsilon\}.$$

Similarly, we can shrink it with $\|(x, y, z)\| \leq \epsilon$

$$\{(x, y, z) \mid ax + by + bz \leq d - \epsilon\}.$$

Remark that for this $\|(a, b, c)\| = 1$.

To expand with $r^T M r \leq \epsilon^2$, define $r = M^{-1/2}\tilde{r}$, then $\tilde{r} = M^{1/2}r$. By transforming the half space with $M^{1/2}$, and normalize it we can do the above trick again.