Model order reduction of stochastic linear systems

Tobias Damm

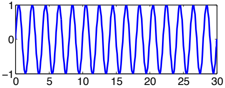
Department of Mathematics
TU Kaiserslautern

Durham, August 2017

1st Lecture: Basics

Stable linear system (deterministic)

$u \in L^2(\mathbb{R}_+ \to \mathbb{R}^m)$ → $y \in L^2(\mathbb{R}_+ \to \mathbb{R}^p)$

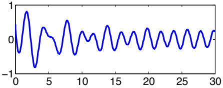


State space

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned}$$

Frequency domain

$$G(s) = C(sI - A)^{-1}B + D$$



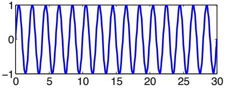
- $x \in \mathbb{R}^n$ is the state vector, $A \in \mathbb{R}^{n\times n}$ with $\sigma(A) \subset \mathbb{C}_-$
- $\mathbb{L} : u \mapsto y,\ L^2 \to L^2$ is the input-output operator.
- $G : \mathbb{C} \to \mathbb{C}^{p\times m}$ is the transfer function, here $n \gg p, m$

  Frequency response: $e^{i\omega t}u_0 \mapsto e^{i\omega t}y_0$ with $y_0 = G(i\omega)u_0$

  H^∞-norm: $\|\mathbb{L}\| = \sup_{\|u\|_{L^2}=1} \|y\|_{L^2} = \max_{\omega\in\mathbb{R}} \|G(i\omega)\|_2 = \|G\|_{H^\infty}$

Stable linear system (stochastic)

$u \in L^2_w(\mathbb{R}_+ \to \mathbb{R}^m)$



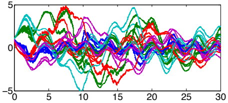
State space

$$dx = Ax\,dt + Nx\,dw + Bu\,dt$$

$$y = Cx + Du$$

Frequency domain ???

$y \in L^2_w(\mathbb{R}_+ \to \mathbb{R}^p)$



Stability: $\sigma(A \otimes I + I \otimes A + N \otimes N) \subset \mathbb{C}_-$

Input-output operator: $\mathbb{L} : u \mapsto y,\ L^2_w \to L^2_w.$

H^∞-type norm: $\|\mathbb{L}\| = \sup_{\|u\|_{L^2_w}=1} \|y\|_{L^2_w}$

Deterministic and stochastic linear systems

Deterministic

$$\dot{x} = Ax + Bu$$
$$y = Cx$$

Deterministic and stochastic linear systems

Deterministic with perturbed parameters

$$\dot{x} = (A + \mu(t)N)x + Bu$$
$$y = Cx$$

Deterministic and stochastic linear systems

Noisy parameters: $\mu(t) = \dot{w} =$ white noise

$$\dot{x} = (A + \dot{w}N)x + Bu$$
$$y = Cx$$

Deterministic and stochastic linear systems

Stochastic

$$
\begin{aligned}
dx &= (Ax + Bu)\,dt + Nx\,dw \\
y &= Cx
\end{aligned}
$$

Here

- $A, N \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{p\times n}$.
- w is a Wiener process
- L^2_w: space of L^2-functions x, nonanticipating w.r.t. w and

$$
\|x\|^2_{L^2_w} = \int_0^\infty E\|x(t)\|^2\,dt \; < \; \infty \,,
$$

where E denotes expectation.

Deterministic and stochastic linear systems

General stochastic

$$dx = (Ax + Bu)\,dt + \sum_{j=1}^{\nu} N_j x\,dw_j$$

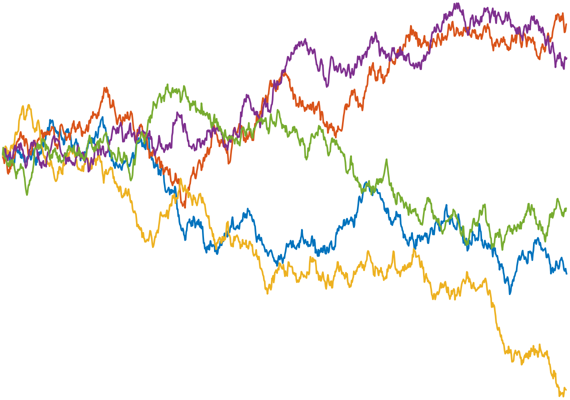
$$y = Cx$$

Here

- $A, N_j \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$.
- w_j independent Wiener processes
- L^2_w: space of L^2-functions x, nonanticipating w.r.t. $w_1, \ldots, w_\nu$ and

$$\|x\|^2_{L^2_w} = \int_0^\infty E\|x(t)\|^2\,dt \;<\; \infty\,.$$

Wiener process



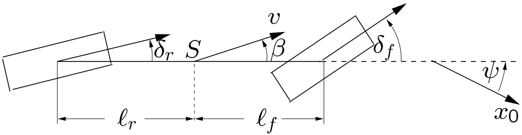
Independent normally distributed increments $w(t_2) - w(t_1)$,

$$E\Big((w(t_2) - w(t_1))^2\Big) = t_2 - t_1$$

Note: E.g. $w(2t)$ is anticipating w.r.t. $w(t)$.

Example: Car-steering (modified from Ackermann 1993)

Single-track model of a vehicle, by lumping front and rear wheels:



β sideslip angle

δ_f, δ_r steering angles

$r = \dot{\psi}$ yaw rate

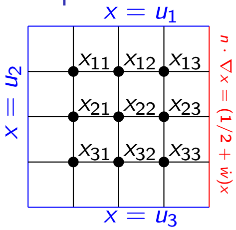
State-space model:

$$\begin{bmatrix} \dot{\beta} \\ \dot{r} \end{bmatrix} = \underbrace{\begin{bmatrix} -\mu \frac{c_r + c_f}{mv} & \mu \frac{c_r \ell_r - c_f \ell_f}{mv^2} - 1 \\ \mu \frac{c_r \ell_r + c_f \ell_f}{J} & -\mu \frac{c_r \ell_r^2 + c_f \ell_f^2}{Jv} \end{bmatrix}}_{=: A_\mu} \begin{bmatrix} \beta \\ r \end{bmatrix} + \mu \underbrace{\begin{bmatrix} \frac{c_f}{mv} & \frac{c_r}{mv} \\ \frac{c_f \ell_f}{J} & -\frac{c_r \ell_r}{J} \end{bmatrix}}_{=: B_\mu} \begin{bmatrix} \delta_r \\ \delta_f \end{bmatrix}$$

Adhesion coefficient $0.15 < \mu < 1$ (icy/wet/dry road)
Assumed to be stochastic around $\mu = 0.5$:

$$\begin{aligned} A_\mu &= A_{0.5} + \dot{w} N_A \\ B_\mu &= B_{0.5} + \dot{w} N_B \end{aligned}$$

Example: Stochastic heat equation



$$\dot{x} = \Delta x$$

$$n \cdot \nabla x = (1/2 + \dot{w})x \,, \qquad \text{right boundary}$$

Finite difference discretization

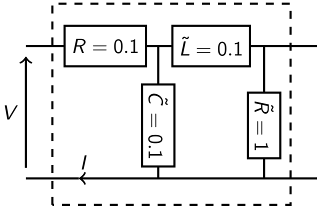
$$\Delta x_{ij} \approx -\frac{1}{h^2}\left(4x_{ij} - x_{i+1,j} - x_{i,j+1} - x_{i-1,j} - x_{i,j-1}\right)$$

$$x_{0,j} = u_1, \quad x_{i,0} = u_2, \quad x_{4,j} = u_3$$

$$x_{i4} \approx x_{i3} - hu(1/2 + \dot{w})x_{i3}$$

$$y = h^2 \sum_{i,j} x_{ij}$$

Example: An electrical ladder network



$n/2$ sections

Here $\tilde{L}^{-1} = L^{-1} + \dot{w}$

- [Gugercin/Antoulas 2004]
- [Ugrinovskii/Petersen 1999]

$$A = \begin{bmatrix} \frac{-1}{\tilde{C}R} & \frac{-1}{\tilde{C}} & 0 & 0 & 0 & 0 \\ \frac{1}{L} & \frac{-R\tilde{R}}{L(R+\tilde{R})} & \frac{-\tilde{R}}{L(R+\tilde{R})} & 0 & 0 & 0 \\ 0 & \frac{\tilde{R}}{\tilde{C}(R+\tilde{R})} & \frac{-1}{\tilde{C}(R+\tilde{R})} & \frac{-1}{\tilde{C}} & 0 & 0 \\ 0 & 0 & \frac{1}{L} & \frac{-R\tilde{R}}{L(R+\tilde{R})} & \frac{-\tilde{R}}{L(R+\tilde{R})} & 0 \\ 0 & 0 & 0 & \frac{\tilde{R}}{\tilde{C}(R+\tilde{R})} & \frac{-1}{\tilde{C}(R+\tilde{R})} & \frac{-1}{\tilde{C}} \\ 0 & 0 & 0 & 0 & \frac{1}{L} & \frac{-\tilde{R}}{L} \end{bmatrix}, N = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & \frac{-R\tilde{R}}{R+\tilde{R}} & \frac{-\tilde{R}}{R+\tilde{R}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \frac{-R\tilde{R}}{R+\tilde{R}} & \frac{-\tilde{R}}{R+\tilde{R}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -\tilde{R} \end{bmatrix}$$

$$B = \begin{bmatrix} \frac{1}{\tilde{C}R} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T, C = \begin{bmatrix} -\frac{1}{R} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Example: SPDEs

Frequently considered: Stochastic heat and wave equation

$$\dot{z}(t,x) = \Delta z(t,x) + N(z)\dot{w}(t)$$

$$\ddot{z}(t,x) = \Delta z(t,x) + N(z,\dot{z})\dot{w}(t)$$

Stochastic linear control systems, early references

- Wonham. Optimal stationary control of a linear system with state-dependent noise. SICON, 1967.
- Kleinman. On the stability of linear stochastic systems. IEEE TAC, 1969.
- Sagirow. Stochastic Methods in the Dynamics of Satellites. 1970.
- Haussmann. Optimal stationary control with state and control dependent noise. SICON 1971
- Willems & Willems. Feedback stabilizability for stochastic systems with state and control depending noise. Automatica, 1976.
- Bernstein & Hyland. Optimal projection equations for reduced-order modeling, estimation and control of linear systems with multiplicative noise. JOTA 1988.
- Hinrichsen & Pritchard. Stochastic H_∞. SICON, 1998.
- TD. Rational Matrix Equations in Stochastic Control, 2004

Some theory on linear SDEs

Fundamental solution

Homogenous equation:

$$dx = Ax\,dt + Nx\,dw\ , \quad x(\tau) = x_0$$

Fundamental solution: $\Phi(t,\tau)$ for $t \geq \tau$: $x(t) = \Phi(t,\tau)x_0$.

Inhomogenous equation:

$$dx = (Ax + Bu)\,dt + Nx\,dw\ , \quad x(\tau) = x_0$$

Variation of constants:

$$x(t) = \Phi(t,\tau)x_0 + \int_\tau^t \Phi(t,s)Bu(s)\,ds$$

Mean square stability

The system

$$dx = Ax\,dt + Nx\,dw\ , \quad x(0) = x_0$$

is called (asymp). mean square stable, if $\forall x_0 \in \mathbb{R}^n$:

$$E(\|x(t)\|^2) \stackrel{t\to\infty}{\to} 0.$$

Equivalently

$$E(x(t)x(t)^T) \stackrel{t\to\infty}{\to} 0\ .$$

Itô's rule and Lyapunov condition

Set $P(t) = E(x(t)x(t)^T)$. Then

$$\begin{aligned} dP &= E\Big((x+dx)(x+dx)^T - xx^T\Big) \\ &= E\Big(x \cdot dx^T + dx \cdot x^T + dx \cdot dx^T\Big) \\ &= E\Big(xx^T(A^T\,dt + N^T\,dw) + (A\,dt + N\,dw)xx^T\Big) \\ &\quad + E\Big((A\,dt + N\,dw)xx^T(A^T\,dt + N^T\,dw)\Big) \end{aligned}$$

Note: $E(dt) = dt$, $E(dw) = 0$, $E(dw \cdot dw) = dt$ [1]. Hence

$$\dot{P} = PA^T + AP + NPN^T$$

$$\rightsquigarrow \sigma(A \otimes I + I \otimes A + N \otimes N) \overset{!}{\subset} \mathbb{C}_-.$$

[1] because $E(w(t+\Delta t) - w(t))^2 = \Delta t$

Lyapunov equation and stability

$$X \mapsto \mathcal{L}_A(X) = A^T X + XA$$

Theorem: The following are equivalent

(a) The linear system $\dot{x} = Ax$ is asymptotically stable.

(b) $\sigma(A) \subset \mathbb{C}_-$

(c) $\exists Y > 0 : \exists X > 0 : A^T X + XA = -Y$

(d) $\forall Y > 0 : \exists X > 0 : A^T X + XA = -Y$

Here: $X > 0$ means that X is symmetric positive definite.

Lyapunov equation and stability

$$X \mapsto \mathcal{L}_A(X) = A^T X + XA$$

$$X \mapsto \Pi_N(X) = N^T X N$$

Theorem: The following are equivalent ([Khasminskii, 1980])

(a) $dx = Ax\,dt + Nx\,dw$ is asymptotically mean square stable.

(b) $\sigma(A \otimes I + I \otimes A + N \otimes N) \subset \mathbb{C}_-$

(b') $\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_-$

(b") $\sigma(\mathcal{L}_A) \subset \mathbb{C}_-$ and $\rho(\mathcal{L}_A^{-1}\Pi_N) < 1$

(c) $\exists Y > 0 : \exists X > 0 : A^T X + XA + N^T X N = -Y$

(d) $\forall Y > 0 : \exists X > 0 : A^T X + XA + N^T X N = -Y$

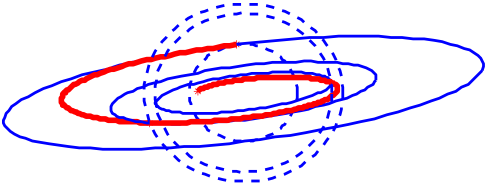
Here: $X > 0$ means that X is symmetric positive definite.

Interpretation of Lyapunov conditions

$$(\mathcal{L}_A + \Pi_N)(X) = -Y \ , \quad X > 0, Y < 0 \ .$$

(i) X defines norm $\|x\|_X = \sqrt{x^* X x}$

$$\frac{d}{dt} E\|x(t)\|_X^2 = -E(x(t)^* Y x(t)) < 0$$



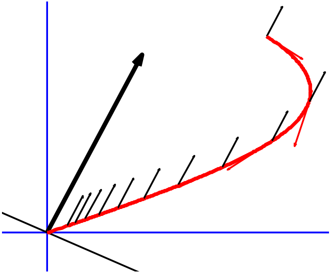
Interpretation of Lyapunov conditions

$$(\mathcal{L}_A + \Pi_N)(X) = -Y \ , \quad X > 0, Y < 0 \ .$$

(ii) $P(t) = E(x(t)x(t)^T) \geq 0$ satisfies $\dot{P} = (\mathcal{L}_{A^T} + \Pi_{N^T})(P)$.

$\mathcal{L}_{A^T} + \Pi_{N^T}$ generates stable positive (semi)group

$$\frac{d}{dt}\langle X, P\rangle = \langle X, (\mathcal{L}_A + \Pi_N)^*(P)\rangle = \langle -Y, P\rangle < 0 \ .$$



Gramians

Lyapunov equation and Gramians

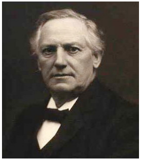
1857–1918

$$A^T Q + QA = -C^T C$$

(observability Gramian)

$$AP + PA^T = -BB^T$$

(controllability Gramian)



1850–1916

Recall energy functionals

$$E_c(x_0) = \min_{x(-\infty)=0, x(0)=x_0} \|u\|_{L^2}^2 = x_0^T P^{-1} x_0$$

$$E_o(x_0) = \|y(\cdot, x_0)\|_{L^2}^2 = x_0^T Q x_0$$

Lyapunov equation and Gramians

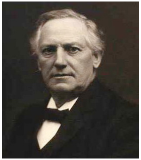
1857–1918

$$A^T Q + QA = -C^T C$$

(observability Gramian)

$$AP + PA^T = -BB^T$$

(controllability Gramian)



1850–1916

Recall energy functionals

$$\begin{aligned} E_c(x_0) &= \min_{x(-\infty)=0, x(0)=x_0} \|u\|_{L^2}^2 &= x_0^T P^{-1} x_0 \\ E_o(x_0) &= \|y(\cdot, x_0)\|_{L^2}^2 &= x_0^T Q x_0 \end{aligned}$$

Note:

$$AP + PA^T = -BB^T \quad \Longleftrightarrow \quad A^T P^{-1} + P^{-1} A = -P^{-1} BB^T P^{-1}$$

Gramians for stochastic system

Observability:

$$\begin{aligned} A^T Q + QA &= -C^T C \\ \rightsquigarrow \quad A^T Q + QA + N^T QN &= -C^T C \end{aligned}$$

Controllability:

- Type I

$$\begin{aligned} AP + PA^T &= -BB^T \\ \rightsquigarrow \quad AP + PA^T + NPN^T &= -BB^T \end{aligned}$$

- Type II

$$\begin{aligned} A^T P^{-1} + P^{-1} A &= -P^{-1} BB^T P^{-1} \\ \rightsquigarrow \quad A^T P^{-1} + P^{-1} A + N^T P^{-1} N &= -P^{-1} BB^T P^{-1} \end{aligned}$$

LMI formulation

Type I: $$\begin{bmatrix} PA^T + AP + BB^T & NP \\ PN^T & -P \end{bmatrix} \leq 0$$

Type II: $$\begin{bmatrix} PA^T + AP + BB^T & PN^T \\ NP & -P \end{bmatrix} \leq 0$$

Balancing and truncation

Let Gramians Q and P (type I or II) be given.

Factorize $P = LL^T$ (e.g. Cholesky) and
$L^T Q L = U\Sigma^2 U^T$ (spectral decomposition)

and consider a state-space transformation with $T = LU\Sigma^{-1/2}$.

Equivalent system

$$\tilde{A} = T^{-1}AT = \begin{bmatrix} A_{11} & \star \\ \star & \star \end{bmatrix}, \ \tilde{N} = T^{-1}NT = \begin{bmatrix} N_{11} & \star \\ \star & \star \end{bmatrix},$$

$$\tilde{B} = T^{-1}B = \begin{bmatrix} B_1 \\ \star \end{bmatrix}, \ \tilde{C} = CT = \begin{bmatrix} C_1 & \star \end{bmatrix}.$$

with Gramians $\tilde{P} = \tilde{Q} = \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} = \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix}$

Balanced Truncation: Deterministic and Stochastic

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned} \quad \rightsquigarrow \quad \begin{aligned} dx &= Ax\,dt + Nx\,dw + Bu\,dt \\ y &= Cx \end{aligned}$$

$$\begin{aligned} \mathcal{L}_A Q &= -C^T C \\ \mathcal{L}_{A^T}(P) &= -BB^T \end{aligned} \quad \rightsquigarrow \quad \begin{aligned} (\mathcal{L}_A + \Pi_N)(Q) &= -C^T C \\ (\mathcal{L}_{A^T} + \Pi_{N^T})(P) &= -BB^T \end{aligned}$$

or

$$(\mathcal{L}_A + \Pi_N)(P^{-1}) = -P^{-1}BB^T P^{-1}$$

Balancing: $P = Q = \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix}$, $A = \begin{bmatrix} A_{11} & \star \\ \star & \star \end{bmatrix}$ etc.

$$\begin{aligned} dx_r &= A_{11}x_r\,dt + N_{11}x_r\,dw + B_1 u\,dt \\ y_r &= C_1 x_r \end{aligned} \qquad \mathbb{L}_r : u \mapsto y_r$$

Motivation for Type I Gramian

Type I Gramian: Bilinear systems

Consider bilinear system

$$\dot{x} = Ax + Nxu + Bu$$
$$y = Cx$$

Then

$$A^T Q + QA + N^T QN = -C^T C$$
$$AP + PA^T + NQN^T = -BB^T$$

define (bilinear) reachability and controllability Gramian.
[Ruberti/Isidori/d'Alessandro, 1972], [AlBayat/Bettayeb 1993]

In particular: Ker Q is unobservable, Ker P is unreachable.
But: Energy interpretation not so obvious. [Benner, D. 2011]

Type I Gramian: Integral expressions

Consider fundamental solution $\Phi(t,0)$ of

$$dx = Ax\,dt + Nx\,dw\ .$$

If the system is mean-square stable, then

$$Q = E\int_0^\infty \Phi(t,0)^T C^T C \Phi(t,0)\,dt$$
$$P = E\int_0^\infty \Phi(t,0) B B^T \Phi(t,0)^T\,dt$$

solve

$$A^T Q + QA + N^T QN = -C^T C\ ,\quad AP + PA^T + NPN^T = -BB^T$$

Type I Gramian: Snapshots and empirical Gramian

Write $\Phi(t,0) = \Phi_\omega(t,0) = \Phi_\omega(t)$, for $\omega \in \Omega$ probability space. Then

$$P = \int_\Omega \int_0^\infty \Phi_\omega(t) B B^T \Phi_\omega(t)^T \, dt \, d\omega$$

Let $u(t) = \delta(t) I$ (Dirac) such that $x_\omega(t) = \Phi_\omega(t) B$.
Consider snapshot matrix

$$S = [\Phi_{\omega_1}(t_1)B, \ldots, \Phi_{\omega_1}(t_k)B, \ldots, \Phi_{\omega_\ell}(t_1)B, \ldots, \Phi_{\omega_\ell}(t_k)B]$$

Then $P \approx cSS^T$, $c \in \mathbb{R}$ (empirical Gramian).

Type I Gramian: Energy functionals

Consider

$$dx = Ax\,dt + Nx\,dw + Bu\,dt\ , \quad y = Cx$$

with the Gramians

$$A^T Q + QA + N^T QN = -C^T C\ , \quad AP + PA^T + NPN^T = -BB^T$$

Energy functionals

$$E_o(x_0) \;:=\; \|y(\cdot, x_0)\|^2_{L^2_w} \;=\; x_0^T Q x_0$$

$$E_c(x_0) \;:=\; \min_{\substack{u \in L^2_W(-\infty,0) \\ x(-\infty,u)=0,\, E(x(0,u))=x_0}} \|u\|^2_{L^2_w} \;\geq\; x_0^T P^{-1} x_0$$

Type I Gramian: More on the control energy

Consider the input-to-state mapping $u \mapsto x$ with

$$x(t) = \int_0^t \Phi(t,s)Bu(s)\,ds$$

Now allow for $u \in L^2$, possibly anticipating to Wiener process. Let $P > 0$ of type I. Then

$$E_c(x_0) \geq \tilde{E}_c(x_0) := \min_{\substack{u \in L^2(-\infty,0) \\ x(-\infty,u)=0,\, E(x(0,u))=x_0}} E\|u\|^2_{L^2} = x_0^T P^{-1} x_0$$

Minimizing control is anticipating:

$$u(t) = B^T \Phi(0,t) P^{-1} x_0, \;\; t < 0$$

Type I Gramian: A Hankel-operator interpretation

Given this x_0 and $u_- \in L^2(]-\infty,0])$ define $y_+ \in L^2([0,\infty[)$ via

$$y_+(t) = \Phi(t,0)x_0 = \int_{-\infty}^{0} \Phi(t,s)Bu_-(s)\,ds\ , \quad t \geq 0\ .$$

Hankel operator

$$\mathcal{H} : L^2(]-\infty,0]) \to L^2([0,\infty[)\ , \quad u_- \mapsto y_+\ .$$

Assume $Q = P = \mathrm{diag}(\sigma_1,\ldots,\sigma_n)$.
For $j = 1,\ldots,n$ define orthonormal L^2-functions via

$$u_j(s) = \sigma_j^{1/2} B^T \Phi(0,s)^T \Sigma^{-1} e_j\ , \quad s < 0$$
$$y_j(s) = \sigma_j^{-1/2} C\Phi(s,0)e_j\ , \quad s > 0\ .$$

Then we have the SVD: $\boxed{\mathcal{H} : u_- \mapsto \sum_{j=1}^{n} \sigma_j y_j \langle u_j, u_- \rangle\ .}$ f

Why type II ???

$\rightsquigarrow$ We want an H^∞-error bound...

$\rightsquigarrow$ See also Martin Redmann's talk!

2nd Lecture: Analysis

Recap: Gramians for stochastic system

$$dx = (Ax + Bu)\,dt + Nx\,dw\ , \quad y = Cx\ .$$

Observability: $A^TQ + QA + N^TQN = -C^TC$

Controllability

- Type I: $AP + PA^T + NPN^T = -BB^T$
- Type II: $A^TP^{-1} + P^{-1}A + N^TP^{-1}N = -P^{-1}BB^TP^{-1}$

Our original naming

- Type I: 'energy Gramian'
- Type II: 'H^∞-Gramian'

But energy interpretation not entirely consistent.

Recap: Gramians for stochastic system

$$dx = (Ax + Bu)\,dt + Nx\,dw\ , \qquad y = Cx\ .$$

Observability: $A^TQ + QA + N^TQN = -C^TC$

Controllability

- Type I: $AP + PA^T + NPN^T \leq -\,BB^T$
- Type II: $A^TP^{-1} + P^{-1}A + N^TP^{-1}N \leq -\,P^{-1}BB^TP^{-1}$

Our original naming

- Type I: 'energy Gramian'
- Type II: 'H^∞-Gramian'

But energy interpretation not entirely consistent.

Recap: LMI formulation

Type I: $$\begin{bmatrix} PA^T + AP + BB^T & NP \\ PN^T & -P \end{bmatrix} \leq 0$$

Type II: $$\begin{bmatrix} PA^T + AP + BB^T & PN^T \\ NP & -P \end{bmatrix} \leq 0$$

Deterministic case: Properties of balanced truncation

- The reduced system is asymptotically stable.
- H^∞-Norm $\sigma_{r+1} \leq \|G - G_r\|_{H^\infty} \leq 2(\sigma_{r+1} + \ldots + \sigma_n)$

Stochastic case: Questions

Recall

$$\mathcal{L}_A : \quad X \mapsto A^T X + XA \qquad \Pi_N : \quad X \mapsto N^T X N \qquad \begin{cases} \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n} \\ \text{symmetric} \mapsto \text{symmetric} \end{cases}$$

Stability

$$\boxed{\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_- \overset{?}{\Rightarrow} \sigma(\mathcal{L}_{A_{11}} + \Pi_{N_{11}}) \subset \mathbb{C}_-}$$

Error bound

$$\boxed{\|\mathbb{L} - \mathbb{L}_r\| \overset{?}{\leq} 2(\sigma_{r+1} + \ldots + \sigma_n)}$$

Stochastic H^{∞}

Stochastic bounded real lemma

Theorem: Consider the mean square stable system

$$\begin{aligned} dx &= Ax\,dt + Nx\,dw + Bu\,dt \\ y &= Cx \end{aligned}$$

Then for all $u \in L^2_w$, also $y =: \mathbb{L}u \in L^2_w$.

The induced norm $\|\mathbb{L}\|$ is the infimum over all γ so that $\exists X < 0$:

$$\mathcal{R}_\gamma(X) := A^T X + XA + N^T XN - C^T C - \frac{1}{\gamma^2} XBB^T X > 0$$

[Hinrichsen/Pritchard, 1998]

Example (for H^∞-norm)

$$(A, N, B, C) = \left(\begin{bmatrix} -1 & 0 \\ 0 & -a^2 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \end{bmatrix} \right), \quad a > 1$$

BRL: For given $\gamma > 0$ find $X = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} < 0$ such that

$$\begin{aligned} 0 &< A^T X + XA + N^T XN - C^T C - \frac{1}{\gamma^2} XBB^T X \\ &= \begin{bmatrix} -2x_1 + x_3 - \gamma^{-2} x_1^2 & -(a^2+1)x_2 - \gamma^{-2} x_1 x_2 \\ -(a^2+1)x_2 - \gamma^{-2} x_1 x_2 & -2a^2 x_3 - \gamma^{-2} x_2^2 - 1 \end{bmatrix} \end{aligned}$$

$\rightsquigarrow$ w.l.o.g. $x_2 = 0$

Example (for H^∞-norm)

$$(A, N, B, C) = \left(\begin{bmatrix} -1 & 0 \\ 0 & -a^2 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \end{bmatrix} \right), \ a > 1$$

BRL: For given $\gamma > 0$ find $X = \begin{bmatrix} x_1 & 0 \\ 0 & x_3 \end{bmatrix} < 0$ such that

$$\begin{aligned} 0 &< A^T X + XA + N^T XN - C^T C - \frac{1}{\gamma^2} XBB^T X \\ &= \begin{bmatrix} -2x_1 + x_3 - \gamma^{-2} x_1^2 & 0 \\ 0 & -2a^2 x_3 - 1 \end{bmatrix} \end{aligned}$$

Example (for H^∞-norm)

$$(A, N, B, C) = \left(\begin{bmatrix} -1 & 0 \\ 0 & -a^2 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \end{bmatrix} \right), \ a > 1$$

BRL: For given $\gamma > 0$ find $X = \begin{bmatrix} x_1 & 0 \\ 0 & x_3 \end{bmatrix} < 0$ such that

$$0 < A^T X + XA + N^T XN - C^T C - \frac{1}{\gamma^2} XBB^T X$$
$$= \begin{bmatrix} -2x_1 + x_3 - \gamma^{-2} x_1^2 & 0 \\ 0 & -2a^2 x_3 - 1 \end{bmatrix}$$

$x_3 < -\frac{1}{2a^2}$ and

$$0 > x_1^2 + 2\gamma^2 x_1 - \gamma^2 x_3 = (x_1 + \gamma^2)^2 - \gamma^2(\gamma^2 + x_3)$$
$$> (x_1 + \gamma^2)^2 - \gamma^2(\gamma^2 - \tfrac{1}{2a^2}) \, .$$

$$\Rightarrow \quad \|\mathbb{L}\| = \inf\{\gamma > 0 \mid \gamma^2 > \frac{1}{2a^2}\} = \frac{1}{\sqrt{2}a}$$

A counter-example (error bound)

$$(A, N, B, C) = \left(\begin{bmatrix} -1 & 0 \\ 0 & -a^2 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \end{bmatrix} \right), \ a > 1$$

Balancing transformation: $x = S\xi$ with $S = \begin{bmatrix} 2a^2 & 0 \\ 0 & 1/2 \end{bmatrix}^{1/4}$

Then $P = Q = \text{diag}(\sigma_1, \sigma_2)$ with $\sigma_1 = \frac{1}{\sqrt{8}a}$, $\sigma_2 = \frac{1}{\sqrt{8}a^2}$

Reduction to order 1 yields $A_{11} = -1$ (asymp. stable) but $C_1 = 0$. Thus

$$\|\mathbb{L} - \mathbb{L}_1\| = \|\mathbb{L}\| = 2a\sigma_2 > 2\sigma_2$$

Even worse: $\frac{\|\mathbb{L} - \mathbb{L}_1\|}{\sigma_2}$ can be arbitrarily large!

Stochastic case: Questions

$$\text{Set} \begin{cases} \mathcal{L}_A : & X \mapsto A^T X + XA \\ \Pi_N : & X \mapsto N^T X N \end{cases} , \quad \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$$

Stability

$$\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_- \overset{?}{\Rightarrow} \sigma(\mathcal{L}_{A_{11}} + \Pi_{N_{11}}) \subset \mathbb{C}_-$$

No H^∞-error bound

$$\|\mathbb{L} - \mathbb{L}_r\| \overset{\text{in general}}{\not\leq} c(\sigma_{r+1} + \ldots + \sigma_n) \text{ for any } c \in \mathbb{R}$$

Preservation of stability

Preservation of stability

Theorem 1: Let $A, N \in \mathbb{R}^{n\times n}$ with $\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_-$,

$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} > 0$ with $\sigma(\Sigma_1) \cap \sigma(\Sigma_2) = \emptyset$, so that

$$A\Sigma + \Sigma A^T + N\Sigma N^T \leq 0$$
$$A^T\Sigma + \Sigma A + N^T\Sigma N \leq 0$$

Partition $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, $N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$.

Then

$$\sigma(\mathcal{L}_{A_{11}} + \Pi_{N_{11}}) \subset \mathbb{C}_-$$

Sketch of proof later. [Benner/D/Redmann/Rodriguez-Cruz (2014)]

Theory of generalized Lyapunov operators

Matrix spaces

Let

- $\mathcal{S}^n = \{X \in \mathbb{R}^{n\times n} \mid X = X^T\}$ symmetric matrices
- $\mathcal{S}^n_+ = \{X \in \mathcal{S}^n \mid X > 0\}$ PSD-cone

Then $\mathcal{S}^n$ is an ordered real vector space with cone $\mathcal{S}^n_+$

Scalar products:

$\mathbb{R}^{n\times m}$: $\langle X, Y\rangle = \operatorname{trace}(X^T Y) = \operatorname{trace}(Y^T X)$.

$\mathcal{S}^n$: in particular $\langle X, Y\rangle = \operatorname{trace}(XY) = \operatorname{trace}(YX)$

Adjoint operators $\mathcal{L}_A^* = \mathcal{L}_{A^T}$, $\Pi_N^* = \Pi_{N^T}$

- If $X, Y \in \mathcal{S}^n_+$, then $\langle X, Y\rangle \geq 0$ and $\langle X, Y\rangle = 0 \iff XY = 0$.
- $\langle X, (\mathcal{L}_A + \Pi_N)(X)\rangle = (\operatorname{vec} X)^T \Big(A \otimes I + I \otimes A + N \otimes N\Big) \operatorname{vec} X$

Resolvent positive operators on $\mathcal{S}^n$

Definition: $T : \mathcal{S}^n \to \mathcal{S}^n$, linear, is called

- *positive* $(T \geq 0)$ $\iff$ $T(\mathcal{S}^n_+) \subset \mathcal{S}^n_+$
- *resolvent positive* $\iff$ $\forall \alpha \gg 0$: $(\alpha I - T)^{-1} \geq 0$

$\iff$ $\forall t \geq 0$: $e^{Tt} \geq 0$

Resolvent positive operators on $\mathcal{S}^n$

Definition: $T : \mathcal{S}^n \to \mathcal{S}^n$, linear, is called

- *positive* $(T \geq 0) \iff T(\mathcal{S}^n_+) \subset \mathcal{S}^n_+$
- *resolvent positive* $\iff \forall \alpha \gg 0\colon (\alpha I - T)^{-1} \geq 0$
  $\iff \forall t \geq 0\colon e^{Tt} \geq 0$

Facts:

(i) If Π is positive, then $\mathcal{L}_A + \Pi$ is resolvent positive.

(ii) T and $-T$ are resolvent positive $\iff \exists A : T = \mathcal{L}_A$

Remark: $\mathcal{L}_A$ generates positive group, i.e. reversible behaviour
$\mathcal{L}_A + \Pi_N$ models irreversible behaviour

Aside: Exotic resolvent positive operators

The mapping $T : \mathcal{S}^3 \to \mathcal{S}^3$ given by

$$T(X) = \begin{bmatrix} x_{22} - x_{11} & -\frac{5}{2}x_{12} & -\frac{5}{2}x_{13} \\ -\frac{5}{2}x_{21} & 4(x_{33} - x_{22}) & -10x_{23} \\ -\frac{5}{2}x_{31} & -10x_{32} & 4(x_{11} - x_{33}) \end{bmatrix},$$

but cannot be written in the form $\mathcal{L}_A + \Pi$
[Kuzma, Omladič, Šivic, Teichmann, 2015]

Other names

For linear $T : \mathcal{S}^n \to \mathcal{S}^n$ the following are equivalent:

(i) T is resolvent positive.

(ii) $\exp(tT) \geq 0$ for all $t \geq 0$ exponentially positive

(iii) $\forall X \in \partial\mathcal{S}^n_+ : \exists V \in \partial\mathcal{S}^n_+ : \langle X, V \rangle = 0 \ \& \ \langle T(X), V \rangle \geq 0$ quasimonotonic

(iv) $X \in \partial\mathcal{S}^n_+ : V \in \partial\mathcal{S}^n_+ :, \langle X, V \rangle = 0 \Rightarrow \langle T(X), V \rangle \geq 0$ cross positive

(v) $T \in \mathrm{cl}\{T_0 - \alpha I \mid T_0 \geq 0, \alpha \in \mathbb{R}\}$ essentially positive

[Schneider/Vidyasagar 1970, Elsner 1974 Arendt 1987, Berman/Neumann/Stern 1989]

Spectral properties

Theorem [Perron-Frobenius / Krein-Rutman]

Let $T : \mathcal{S}^n \to \mathcal{S}^n$, $\rho(T) = \max |\sigma(T)|$, $\alpha(T) = \max \operatorname{Re} \sigma(T)$

- T positive $\Rightarrow \exists V \in \mathcal{S}^n_+$: $T(V) = \rho(T)V$
- T res. pos. $\Rightarrow \exists V \in \mathcal{S}^n_+$: $T(V) = \alpha(T)V$

Spectral properties

Theorem [Perron-Frobenius / Krein-Rutman]

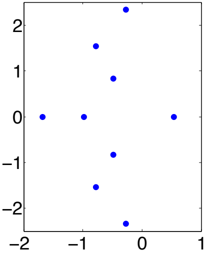
Let $T : \mathcal{S}^n \to \mathcal{S}^n$, $\rho(T) = \max|\sigma(T)|$, $\alpha(T) = \max \operatorname{Re} \sigma(T)$

- T positive $\Rightarrow \exists V \in \mathcal{S}^n_+ \colon\ T(V) = \rho(T)V$
- T res. pos. $\Rightarrow \exists V \in \mathcal{S}^n_+ \colon\ T(V) = \alpha(T)V$

Examples:

- $Av = (\alpha + \beta i)v$, $V = vv^* + \bar{v}\bar{v}^* \in \mathcal{S}^n_+$
  $\Rightarrow \mathcal{L}_A(V) = 2\alpha V$
- `A,N=randn(3)` $\rightsquigarrow$ see figure

Generically, rk $V = n$, but all ranks possible.



A General Lyapunov Theorem

Theorem [Hans Schneider, 1965]
If T is resolvent positive, the following are equivalent:

- $\exists X > 0 : T(X) < 0$
- $\sigma(T) \subset \mathbb{C}_-$
- $-T^{-1} > 0$
- If $T = \mathcal{L}_A + \Pi$ with $\Pi \geq 0$
  then $\sigma(\mathcal{L}_A) \subset \mathbb{C}_-$ and $\rho(\mathcal{L}_A^{-1}\Pi) < 1$.

Proof: Preservation of stability

Preservation of stability

Theorem 1: Let $A, N \in \mathbb{R}^{n\times n}$ with $\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_-$,

$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} > 0$ with $\sigma(\Sigma_1) \cap \sigma(\Sigma_2) = \emptyset$, so that

$$A\Sigma + \Sigma A^T + N\Sigma N^T \leq 0$$
$$A^T\Sigma + \Sigma A + N^T\Sigma N \leq 0$$

Partition $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, $N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$.

Then

$$\sigma(\mathcal{L}_{A_{11}} + \Pi_{N_{11}}) \subset \mathbb{C}_-$$

Proof: Stability of the reduced system

Let $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, $N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$, $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} > 0$,

$\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_-$ and $\sigma(\Sigma_1) \cap \sigma(\Sigma_2) = \emptyset$

$$A\Sigma + \Sigma A^T + N\Sigma N^T = -BB^T \,, \qquad A^T\Sigma + \Sigma A + N^T\Sigma N = -C^T C$$

Proof: Stability of the reduced system

Let $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, $N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$, $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} > 0$,

$\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_-$ and $\sigma(\Sigma_1) \cap \sigma(\Sigma_2) = \emptyset$

$A\Sigma + \Sigma A^T + N\Sigma N^T = -BB^T$, $\qquad A^T\Sigma + \Sigma A + N^T\Sigma N = -C^TC$

In particular:

$$A_{11}\Sigma_1 + \Sigma_1 A_{11}^T + N_{11}\Sigma_1 N_{11}^T = -B_1 B_1^T - N_{12}\Sigma_2 N_{12}^T \qquad (\star)$$

Proof: Stability of the reduced system

Let $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, $N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$, $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} > 0$,

$\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_-$ and $\sigma(\Sigma_1) \cap \sigma(\Sigma_2) = \emptyset$

$A\Sigma + \Sigma A^T + N\Sigma N^T = -BB^T$, $\quad A^T\Sigma + \Sigma A + N^T\Sigma N = -C^TC$

In particular:

$$A_{11}\Sigma_1 + \Sigma_1 A_{11}^T + N_{11}\Sigma_1 N_{11}^T = -B_1 B_1^T - N_{12}\Sigma_2 N_{12}^T \qquad (\star)$$

Assume: $\sigma(\mathcal{L}_{A_{11}} + \Pi_{N_{11}}) \not\subset \mathbb{C}_-$

Then $\exists \alpha_1 \geq 0$, $V_1 \geq 0$: $A_{11}^T V_1 + V_1 A_{11} + N_{11}^T V_1 N_{11} = \alpha_1 V_1$.

Proof: Stability of the reduced system

Let $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, $N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$, $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} > 0$,

$\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_-$ and $\sigma(\Sigma_1) \cap \sigma(\Sigma_2) = \emptyset$

$A\Sigma + \Sigma A^T + N\Sigma N^T = -BB^T$, $A^T\Sigma + \Sigma A + N^T\Sigma N = -C^T C$

In particular:

$$A_{11}\Sigma_1 + \Sigma_1 A_{11}^T + N_{11}\Sigma_1 N_{11}^T = -B_1 B_1^T - N_{12}\Sigma_2 N_{12}^T \qquad (\star)$$

Assume: $\sigma(\mathcal{L}_{A_{11}} + \Pi_{N_{11}}) \not\subset \mathbb{C}_-$

Then $\exists \alpha_1 \geq 0,\ V_1 \geq 0$: $A_{11}^T V_1 + V_1 A_{11} + N_{11}^T V_1 N_{11} = \alpha_1 V_1$.

$$0 \geq \langle (\star), V_1 \rangle = \alpha_1 \underbrace{\langle \Sigma_1, V_1 \rangle}_{>0} \geq 0 \Rightarrow \alpha_1 = 0,\ B_1^T V_1 = 0,\ N_{12}^T V_1 = 0$$

Proof: Stability of the reduced system

Let $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, $N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$, $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} > 0$,

$\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_-$ and $\sigma(\Sigma_1) \cap \sigma(\Sigma_2) = \emptyset$

$A\Sigma + \Sigma A^T + N\Sigma N^T = -BB^T$, $\qquad A^T\Sigma + \Sigma A + N^T\Sigma N = -C^TC$

In particular:

$$A_{11}\Sigma_1 + \Sigma_1 A_{11}^T + N_{11}\Sigma_1 N_{11}^T = -B_1 B_1^T - N_{12}\Sigma_2 N_{12}^T \qquad (\star)$$

Assume: $\sigma(\mathcal{L}_{A_{11}} + \Pi_{N_{11}}) \not\subset \mathbb{C}_-$

Then $\exists \alpha_1 \geq 0, V_1 \geq 0$: $A_{11}^T V_1 + V_1 A_{11} + N_{11}^T V_1 N_{11} = \alpha_1 V_1$.

$$0 \geq \langle (\star), V_1 \rangle = \alpha_1 \underbrace{\langle \Sigma_1, V_1 \rangle}_{>0} \geq 0 \Rightarrow \alpha_1 = 0,\ B_1^T V_1 = 0,\ N_{12}^T V_1 = 0$$

First observations:

- Reduced system *almost stable*.
- Further proof would be easier, if $V_1 > 0$

Orthogonal Transformation

$A_{11}^T V_1 + V_1 A_{11} + N_{11}^T V_1 N_{11} = 0$ with $V_1 \geq 0$ of max. rank

Let $V_1 = [V_{11}, V_{12}] \, \mathrm{diag}(D_1, 0)[V_{11}, V_{12}]^T$, $[V_{11}, V_{12}]$ orthogonal

Easy observation: $\mathrm{Im}\, V_{11} = \mathrm{Im}\, V_1$ is invariant under A_{11}^T and N_{11}^T.

Orthogonal Transformation

$A_{11}^T V_1 + V_1 A_{11} + N_{11}^T V_1 N_{11} = 0$ with $V_1 \geq 0$ of max. rank

Let $V_1 = [V_{11}, V_{12}]\,\mathrm{diag}(D_1, 0)[V_{11}, V_{12}]^T$, $[V_{11}, V_{12}]$ orthogonal

Easy observation: $\mathrm{Im}\, V_{11} = \mathrm{Im}\, V_1$ is invariant under A_{11}^T and N_{11}^T.

Proof: If $V_1 z = 0$, then

$$z^* N_{11}^T V_1 N_{11} z = z^T (A_{11}^T V_1 + V_1 A_{11} + N_{11}^T V_1 N_{11}) z = 0$$

$\Rightarrow V_1 N_{11} z = 0 \ldots \Rightarrow V_1 A_{11} z = 0$
$\Rightarrow$ Ker V_1 is A_{11}- and N_{11}-invariant
$\Rightarrow$ Im V_1 is A_{11}^T- and N_{11}^T-invariant

Orthogonal Transformation

$A_{11}^T V_1 + V_1 A_{11} + N_{11}^T V_1 N_{11} = 0$ with $V_1 \geq 0$ of max. rank

Let $V_1 = [V_{11}, V_{12}] \,\mathrm{diag}(D_1, 0) [V_{11}, V_{12}]^T$, $[V_{11}, V_{12}]$ orthogonal

Easy observation: $\mathrm{Im}\ V_{11} = \mathrm{Im}\ V_1$ is invariant under A_{11}^T and N_{11}^T.

Transformation with $\left[\begin{array}{c|c} [V_{11}, V_{12}] & 0 \\ \hline 0 & I \end{array}\right]$:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \mapsto \begin{bmatrix} \tilde{A}_{11} & 0 & \tilde{A}_{13} \\ \tilde{A}_{21} & \tilde{A}_{22} & \tilde{A}_{23} \\ \tilde{A}_{31} & \tilde{A}_{32} & \tilde{A}_{33} \end{bmatrix}, \quad \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ \tilde{B}_2 \\ \tilde{B}_3 \end{bmatrix}$$

$$\begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \mapsto \begin{bmatrix} \tilde{N}_{11} & 0 & 0 \\ \tilde{N}_{21} & \tilde{N}_{22} & \tilde{N}_{23} \\ \tilde{N}_{31} & \tilde{N}_{32} & \tilde{N}_{33} \end{bmatrix}, \quad \begin{bmatrix} C_1{}^T \\ C_2^T \end{bmatrix} \mapsto \begin{bmatrix} \tilde{C}_1{}^T \\ \tilde{C}_2^T \\ \tilde{C}_3^T \end{bmatrix}$$

$$\begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \mapsto \begin{bmatrix} \tilde{\Sigma}_{11} & \tilde{\Sigma}_{12} & 0 \\ \tilde{\Sigma}_{21} & \tilde{\Sigma}_{22} & 0 \\ 0 & 0 & \tilde{\Sigma}_{33} \end{bmatrix}$$

Let us omit the tilde

A technical computation

$$\Sigma A^T + A\Sigma + N\Sigma N^T = -BB^T \quad (1)$$
$$A^T\Sigma + \Sigma A + N^T\Sigma N = -C^T C \quad (2)$$

Selected block components:

$$0 = \Sigma_{11}A_{11}^T + A_{11}\Sigma_{11} + N_{11}\Sigma_{11}N_{11}^T$$
$$0 = \Sigma_{21}A_{11}^T + N_{21}\Sigma_{11}N_{11}^T + A_{21}\Sigma_{11} + N_{22}\Sigma_{21}N_{11}^T + A_{22}\Sigma_{21}$$

$$\begin{aligned} -C_1^T C_1 &= A_{11}^T\Sigma_{11} + \Sigma_{11}A_{11} + N_{11}^T\Sigma_{11}N_{11} + \Sigma_{12}A_{21} + A_{21}^T\Sigma_{21} \\ &\quad + N_{21}^T\Sigma_{21}N_{11} + N_{11}^T\Sigma_{12}N_{21} + N_{21}^T\Sigma_{22}N_{21} + N_{31}^T\Sigma_{33}N_{31} \end{aligned}$$

A technical computation

$$\Sigma A^T + A\Sigma + N\Sigma N^T = -BB^T \quad (1)$$
$$A^T\Sigma + \Sigma A + N^T\Sigma N = -C^T C \quad (2)$$

Selected block components:

$$0 = \Sigma_{11}A_{11}^T + A_{11}\Sigma_{11} + N_{11}\Sigma_{11}N_{11}^T$$
$$0 = \Sigma_{21}A_{11}^T + N_{21}\Sigma_{11}N_{11}^T + A_{21}\Sigma_{11} + N_{22}\Sigma_{21}N_{11}^T + A_{22}\Sigma_{21}$$

$$\begin{aligned}-C_1^T C_1 &= A_{11}^T\Sigma_{11} + \Sigma_{11}A_{11} + N_{11}^T\Sigma_{11}N_{11} + \Sigma_{12}A_{21} + A_{21}^T\Sigma_{21} \\ &+ N_{21}^T\Sigma_{21}N_{11} + N_{11}^T\Sigma_{12}N_{21} + N_{21}^T\Sigma_{22}N_{21} + N_{31}^T\Sigma_{33}N_{31}\end{aligned}$$

$$\langle C_1^T C_1 + N_{21}^T\Sigma_{22}N_{21} + N_{31}^T\Sigma_{33}N_{31}, \Sigma_{11}\rangle = 2\langle \Sigma_{21}, -A_{21}\Sigma_{11} - N_{21}\Sigma_{11}N_{11}^T\rangle$$

A technical computation

$$\Sigma A^T + A\Sigma + N\Sigma N^T = -BB^T \qquad (1)$$

$$A^T\Sigma + \Sigma A + N^T\Sigma N = -C^T C \qquad (2)$$

Selected block components:

$$0 = \Sigma_{11} A_{11}^T + A_{11}\Sigma_{11} + N_{11}\Sigma_{11} N_{11}^T$$

$$0 = \Sigma_{21} A_{11}^T + N_{21}\Sigma_{11} N_{11}^T + A_{21}\Sigma_{11} + N_{22}\Sigma_{21} N_{11}^T + A_{22}\Sigma_{21}$$

$$-C_1^T C_1 = A_{11}^T\Sigma_{11} + \Sigma_{11} A_{11} + N_{11}^T\Sigma_{11} N_{11} + \Sigma_{12} A_{21} + A_{21}^T\Sigma_{21} + N_{21}^T\Sigma_{21} N_{11} + N_{11}^T\Sigma_{12} N_{21} + N_{21}^T\Sigma_{22} N_{21} + N_{31}^T\Sigma_{33} N_{31}$$

$$\langle C_1^T C_1 + N_{21}^T\Sigma_{22} N_{21} + N_{31}^T\Sigma_{33} N_{31}, \Sigma_{11}\rangle = 2\langle \Sigma_{21}, -A_{21}\Sigma_{11} - N_{21}\Sigma_{11} N_{11}^T\rangle$$

A technical computation

$$\Sigma A^T + A\Sigma + N\Sigma N^T = -BB^T \quad (1)$$

$$A^T\Sigma + \Sigma A + N^T\Sigma N = -C^T C \quad (2)$$

Selected block components:

$$0 = \Sigma_{11}A_{11}^T + A_{11}\Sigma_{11} + N_{11}\Sigma_{11}N_{11}^T$$

$$0 = \Sigma_{21}A_{11}^T + N_{21}\Sigma_{11}N_{11}^T + A_{21}\Sigma_{11} + N_{22}\Sigma_{21}N_{11}^T + A_{22}\Sigma_{21}$$

$$-C_1^T C_1 = A_{11}^T\Sigma_{11} + \Sigma_{11}A_{11} + N_{11}^T\Sigma_{11}N_{11} + \Sigma_{12}A_{21} + A_{21}^T\Sigma_{21} + N_{21}^T\Sigma_{21}N_{11} + N_{11}^T\Sigma_{12}N_{21} + N_{21}^T\Sigma_{22}N_{21} + N_{31}^T\Sigma_{33}N_{31}$$

$$\langle C_1^T C_1 + N_{21}^T\Sigma_{22}N_{21} + N_{31}^T\Sigma_{33}N_{31}, \Sigma_{11}\rangle = 2\langle \Sigma_{21}, -A_{21}\Sigma_{11} - N_{21}\Sigma_{11}N_{11}^T\rangle$$

$$\boxed{0 \le \langle \Sigma_{21}, \Sigma_{21}A_{11}^T + N_{22}\Sigma_{21}N_{11}^T + A_{22}\Sigma_{21}\rangle =: \langle \Sigma_{21}, T_{21}(\Sigma_{21})\rangle}$$

Field of values

Set $A_1 = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}$, $N_1 = \begin{bmatrix} N_{11} & 0 \\ N_{21} & N_{22} \end{bmatrix}$, $\Sigma_1 = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$

By construction $\Big(\underbrace{(\mathcal{L}_{A_1} + \Pi_{N_1}) + (\mathcal{L}_{A_1} + \Pi_{N_1})^*}_{=:T_1}\Big)(\Sigma_1) \leq 0$

whence $\sigma(T_1) \subset]-\infty, 0]$

$\Rightarrow$ field of values of T_1 in $]-\infty, 0]$

Field of values

Set $A_1 = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}$, $N_1 = \begin{bmatrix} N_{11} & 0 \\ N_{21} & N_{22} \end{bmatrix}$, $\Sigma_1 = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$

By construction $\Big(\underbrace{(\mathcal{L}_{A_1} + \Pi_{N_1}) + (\mathcal{L}_{A_1} + \Pi_{N_1})^*}_{=:T_1}\Big)(\Sigma_1) \leq 0$

whence $\sigma(T_1) \subset]-\infty, 0]$
$\Rightarrow$ field of values of T_1 in $]-\infty, 0]$

Can show

$$0 \leq \langle \Sigma_{21}, T_{21}(\Sigma_{21}) \rangle = \langle T_1(\left[\begin{smallmatrix} 0 & 0 \\ \Sigma_{21} & 0 \end{smallmatrix}\right]), \left[\begin{smallmatrix} 0 & 0 \\ \Sigma_{21} & 0 \end{smallmatrix}\right] \rangle \leq 0$$

Field of values

Set $A_1 = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}$, $N_1 = \begin{bmatrix} N_{11} & 0 \\ N_{21} & N_{22} \end{bmatrix}$, $\Sigma_1 = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$

By construction $\Big(\underbrace{(\mathcal{L}_{A_1} + \Pi_{N_1}) + (\mathcal{L}_{A_1} + \Pi_{N_1})^*}_{=:T_1}\Big)(\Sigma_1) \leq 0$

whence $\sigma(T_1) \subset]-\infty, 0]$
$\Rightarrow$ field of values of T_1 in $]-\infty, 0]$

Can show

$$0 \leq \langle \Sigma_{21}, T_{21}(\Sigma_{21}) \rangle = \langle T_1(\left[\begin{smallmatrix} 0 & 0 \\ \Sigma_{21} & 0 \end{smallmatrix}\right]), \left[\begin{smallmatrix} 0 & 0 \\ \Sigma_{21} & 0 \end{smallmatrix}\right] \rangle \leq 0$$

$\Rightarrow\ T_{21}(\Sigma_{21}) = 0,\ N_{21} = 0,\ N_{31} = 0,\ C_1 = 0$

Field of values

Set $A_1 = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}$, $N_1 = \begin{bmatrix} N_{11} & 0 \\ N_{21} & N_{22} \end{bmatrix}$, $\Sigma_1 = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$

By construction $\Big(\underbrace{(\mathcal{L}_{A_1} + \Pi_{N_1}) + (\mathcal{L}_{A_1} + \Pi_{N_1})^*}_{=:T_1}\Big)(\Sigma_1) \leq 0$

whence $\sigma(T_1) \subset]-\infty, 0]$
$\Rightarrow$ field of values of T_1 in $]-\infty, 0]$

Can show

$$0 \leq \langle \Sigma_{21}, T_{21}(\Sigma_{21}) \rangle = \langle T_1(\left[\begin{smallmatrix} 0 & 0 \\ \Sigma_{21} & 0 \end{smallmatrix}\right]), \left[\begin{smallmatrix} 0 & 0 \\ \Sigma_{21} & 0 \end{smallmatrix}\right] \rangle \leq 0$$

$\Rightarrow \; T_{21}(\Sigma_{21}) = 0,\ N_{21} = 0,\ N_{31} = 0,\ C_1 = 0$

Moreover $A_{21}\Sigma_{11} + \Sigma_{22}A_{21} = 0 \Rightarrow A_{21} = 0$

Complete decoupling

Now we know: $A_1 = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$, $N_1 = \begin{bmatrix} N_{11} & 0 \\ 0 & N_{22} \end{bmatrix}$

Necessarily

$$\sigma(A_{11} \otimes I + I \otimes A_{11} + N_{11} \otimes N_{11}) \subset \overline{\mathbb{C}_-}$$
$$\sigma(A_{22} \otimes I + I \otimes A_{22} + N_{22} \otimes N_{22}) \subset \mathbb{C}_-$$

(Otherwise the eigenvector V_1 could have been chosen of larger rank)

Can show: $T_{21} : X \mapsto A_{22}X + XA_{11}^T + N_{22}XN_{11}^T$ is nonsingular.

$$T(\Sigma_{21}) = 0 \quad \Rightarrow \quad \Sigma_1 = \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{bmatrix}$$

Some more zeros

$$A = \begin{bmatrix} A_{11} & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \quad N = \begin{bmatrix} N_{11} & 0 & 0 \\ 0 & N_{22} & N_{23} \\ 0 & N_{32} & N_{33} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & 0 & 0 \\ 0 & \Sigma_{22} & 0 \\ 0 & 0 & \Sigma_{33} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ B_2 \\ B_3 \end{bmatrix}, \quad C^T = \begin{bmatrix} 0 \\ C_2^T \\ C_3^T \end{bmatrix}$$

Some more zeros

$$A = \begin{bmatrix} A_{11} & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \quad N = \begin{bmatrix} N_{11} & 0 & 0 \\ 0 & N_{22} & N_{23} \\ 0 & N_{32} & N_{33} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & 0 & 0 \\ 0 & \Sigma_{22} & 0 \\ 0 & 0 & \Sigma_{33} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ B_2 \\ B_3 \end{bmatrix}, \quad C^T = \begin{bmatrix} 0 \\ C_2^T \\ C_3^T \end{bmatrix}$$

The south-west blocks of

$$A\Sigma + \Sigma A^T + N\Sigma N^T = -BB^T$$
$$A^T\Sigma + \Sigma A + N\Sigma N^T = -C^T C$$

simplify to $$0 = \Sigma_{33} A_{13}^T + A_{31}\Sigma_{11}\,, \quad 0 = A_{13}^T \Sigma_{11} + \Sigma_{33} A_{31}$$

Some more zeros

$$A = \begin{bmatrix} A_{11} & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \quad N = \begin{bmatrix} N_{11} & 0 & 0 \\ 0 & N_{22} & N_{23} \\ 0 & N_{32} & N_{33} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & 0 & 0 \\ 0 & \Sigma_{22} & 0 \\ 0 & 0 & \Sigma_{33} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ B_2 \\ B_3 \end{bmatrix}, \quad C^T = \begin{bmatrix} 0 \\ C_2^T \\ C_3^T \end{bmatrix}$$

The south-west blocks of

$$A\Sigma + \Sigma A^T + N\Sigma N^T = -BB^T$$
$$A^T\Sigma + \Sigma A + N\Sigma N^T = -C^T C$$

simplify to $\quad 0 = \Sigma_{33}A_{13}^T + A_{31}\Sigma_{11}, \quad 0 = A_{13}^T\Sigma_{11} + \Sigma_{33}A_{31}$

$$\Rightarrow \begin{cases} -\Sigma_{33}A_{31}\Sigma_{11} & = \Sigma_{33}^2 A_{13}^T = A_{13}^T \Sigma_{11}^2 \\ -\Sigma_{33}A_{13}^T\Sigma_{11} & = A_{31}\Sigma_{11}^2 = \Sigma_{33}^2 A_{31} \end{cases} .$$

Some more zeros

$$A = \begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{bmatrix}, \quad N = \begin{bmatrix} N_{11} & 0 & 0 \\ 0 & N_{22} & N_{23} \\ 0 & N_{32} & N_{33} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & 0 & 0 \\ 0 & \Sigma_{22} & 0 \\ 0 & 0 & \Sigma_{33} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ B_2 \\ B_3 \end{bmatrix}, \quad C^T = \begin{bmatrix} 0 \\ C_2^T \\ C_3^T \end{bmatrix}$$

The south-west blocks of

$$A\Sigma + \Sigma A^T + N\Sigma N^T = -BB^T$$
$$A^T\Sigma + \Sigma A + N\Sigma N^T = -C^T C$$

simplify to $\quad 0 = \Sigma_{33} A_{13}^T + A_{31}\Sigma_{11}, \quad 0 = A_{13}^T \Sigma_{11} + \Sigma_{33} A_{31}$

$$\Rightarrow \begin{cases} -\Sigma_{33} A_{31} \Sigma_{11} & = \Sigma_{33}^2 A_{13}^T = A_{13}^T \Sigma_{11}^2 \\ -\Sigma_{33} A_{13}^T \Sigma_{11} & = A_{31} \Sigma_{11}^2 = \Sigma_{33}^2 A_{31} \end{cases} \quad \Rightarrow \quad \begin{cases} A_{13} = 0 \\ A_{31} = 0 \end{cases}.$$

Final step

Consider

$$A = \begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{bmatrix}, \quad N = \begin{bmatrix} N_{11} & 0 & 0 \\ 0 & N_{22} & N_{23} \\ 0 & N_{32} & N_{33} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ B_2 \\ B_3 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & 0 & 0 \\ 0 & \Sigma_{22} & 0 \\ 0 & 0 & \Sigma_{33} \end{bmatrix}, \quad \Sigma_0 := \begin{bmatrix} \Sigma_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

From

$$A\Sigma + \Sigma A^T + N\Sigma N^T = -BB^T$$

we get

$$A\Sigma_0 + \Sigma_0 A^T + N\Sigma_0 N^T = 0$$

contradicting

$$\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_-$$

A Cauchy-Schwarz-type inequality (needed in the proof)

$\rho(T) = \max_{\lambda\in\sigma(T)} |\lambda|$, $\alpha(T) = \max_{\lambda\in\sigma(T)} \operatorname{Re}\lambda$

Proposition 1 Let

$$\Pi_L(X) = \sum_{j=1}^{\nu} L_j X L_j^T \ , \quad \Pi_M(Y) = \sum_{j=1}^{\nu} M_j Y M_j^T \ , \quad \Pi_{LM}(Z) = \sum_{j=1}^{\nu} L_j Z M_j^T$$

Then $\rho(\Pi_{LM})^2 \le \rho(\Pi_L)\rho(\Pi_M)$.

Proposition 2 Let

$$T_1(X) = K_1X + XK_1^T + \Pi_L(X) \ , \quad T_2(Y) = K_2Y + YK_2^T + \Pi_M(Y) \ ,$$
$$T_{12}(Z) = K_1Z + ZK_2^T + \Pi_{LM}(Z) \ .$$

Then $\alpha(T_{12}) \le \frac{1}{2}(\alpha(T_1) + \alpha(T_2))$.

Message so far

Stochastic BT based on

$$\begin{aligned} AP + PA^T + NPN^T &= -BB^T \\ A^T Q + QA + N^T QN &= -C^T C \end{aligned}$$

- ☺ Preservation of asymptotic stability
- ☹ H^∞-error bound

Message so far

Stochastic BT based on

$$\boxed{\begin{array}{rcl} AP + PA^T + NPN^T & = & -BB^T \\ A^TQ + QA + N^TQN & = & -C^TC \end{array}}$$

- ☺ Preservation of asymptotic stability
- ☹ H^∞-error bound

Other observations:

- ☺ H^2-error bound (Benner/Redmann)
- ☺ Good numerical results

Type II Gramian

Stability and H^∞-error bound: Type II-Gramian

Theorem 2: Keep Q but replace P by solution of

$$P^{-1}A + A^T P^{-1} + N^T P^{-1} N \leq -P^{-1}BB^T P^{-1} \,. \tag{3}$$

Balance and truncate as before s.t. $Q = P = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$. Then reduced system is asymptotically mean-square stable and

$$\|\mathbb{L} - \mathbb{L}_r\| \leq 2(\sigma_{r+1} + \ldots + \sigma_n)$$

Remark:

- ☹ In general no solution for (3) with equality and $P > 0$.
- ☺ Can replace (3) by LMI $\begin{bmatrix} PA^T + AP + BB^T & PN^T \\ NP & -P \end{bmatrix} \leq 0$, which is feasible under given stability assumption.
- ☺ Gives correct error bound in our example.
- ☹ Need to solve LMI.

Continuation of Example

$$(A, N, B, C) = \left(\begin{bmatrix} -1 & 0 \\ 0 & -a^2 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \end{bmatrix} \right)$$

- $P = \begin{bmatrix} 1 + \sqrt{1-p} & 0 \\ 0 & p \end{bmatrix}^{-1} > 0$ satisfies type II LMI if $0 < p \leq 1$
- $Q = \begin{bmatrix} \frac{1}{4a^2} & 0 \\ 0 & \frac{1}{2a^2} \end{bmatrix}$ and $\mathbb{L}_1 = 0 \Rightarrow \|\mathbb{L} - \mathbb{L}_1\| = \frac{1}{\sqrt{2}a}$ as before
- $\sigma_2^2 = \min \sigma(PQ) = \frac{1}{4a^2(1+\sqrt{1-p})} > \frac{1}{8a^2} = \frac{1}{4}\|\mathbb{L} - \mathbb{L}_1\|^2$
- $\|\mathbb{L} - \mathbb{L}_1\| = 2\sigma_2$ for $p \to 0$

Existence of Type II Gramian

Lemma: Assume that $dx = Ax\,dt + Nx\,dw$ is asymptotically mean-square-stable. Then inequality

$$P^{-1}A + A^T P^{-1} + N^T P^{-1} N \leq -P^{-1} B B^T P^{-1} \ .$$

is solvable with $P > 0$.
Proof: By generalized Lyapunov Theorem:

$$\forall Y < 0 : \exists \tilde{P} > 0 : A^T \tilde{P}^{-1} + \tilde{P}^{-1} A + N^T \tilde{P}^{-1} N = Y \ .$$

Then $P = \varepsilon^{-1} \tilde{P}$, for sufficiently small $\varepsilon > 0$, satisfies

$$A^T P^{-1} + P^{-1} A + N^T P^{-1} N = \varepsilon Y < -\varepsilon^2 \tilde{P}^{-1} B B^T \tilde{P}^{-1} = -P^{-1} B B^T P^{-1} \ .$$

Sketch of the proof for the error bound

Consider

$$dx = Ax\,dt + Nx\,dw + Bu\,dt\ .$$

In partitioned form we have

$$\begin{aligned} dx_1 &= (A_{11}x_1 + A_{12}x_2)\,dt + (N_{11}x_1 + N_{12}x_2)\,dw + B_1 u\,dt \\ dx_2 &= (A_{21}x_1 + A_{22}x_2)\,dt + (N_{21}x_1 + N_{22}x_2)\,dw + B_2 u\,dt \\ y &= C_1x_1 + C_2x_2 \end{aligned}$$

The reduced system obtained by truncation is

$$\begin{aligned} dx_r &= A_{11}x_r + N_{11}x_r\,dw + B_1 u\,dt \\ y_r &= C_1x_r \end{aligned}$$

Sketch of the proof for the error bound

Assuming $x(0) = 0$ and $x_r(0) = 0$, show for all $T \geq 0$

$$\begin{aligned}\int_0^T \mathcal{E}\left(\|y - y_r\|^2\right) dt \;&\leq\; 2\int_0^T \mathcal{E}\left(x_2^T \Sigma_2 (A_{21}x_r + B_2 u)\right) dt \\ &+ 2\int_0^T \mathcal{E}\left((N_{21}x_r)^T \Sigma_2 (2N_{21}x_1 + 2N_{22}x_2)\right) dt\end{aligned}$$

and

$$\begin{aligned}4\int_0^T \mathcal{E}\left(\|u\|^2\right) dt \;&\geq\; 2\int_0^T \mathcal{E}\left(x_2^T \Sigma_2^{-1} (A_{21}x_r + B_2 u)\right) dt \\ &+ 2\int_0^T \mathcal{E}\left((N_{21}x_r)^T \Sigma_2^{-1} (2N_{21}x_1 + 2N_{22}x_2)\right) dt\,.\end{aligned}$$

If $\Sigma_2 = \sigma_{\min} I$, then multiplication of second equation with $\sigma_{\min}^2$ gives the estimate. Thus proceed step by step.

Sketch of the proof for the error bound

$$\begin{aligned} 4\int_0^T \mathcal{E}\left(\|u\|^2\right)\,dt \;&\geq\; 2\int_0^T \mathcal{E}\left(x_2^T \Sigma_2^{-1}(A_{21}x_r + B_2 u)\right)\,dt \\ &+ 2\int_0^T \mathcal{E}\left((N_{21}x_r)^T \Sigma_2^{-1}(2N_{21}x_1 + 2N_{22}x_2)\right)\,dt\,. \end{aligned} \qquad (\star)$$

The inequality

$$A^T\Sigma^{-1} + \Sigma^{-1}A + N\Sigma^{-1}N^T \;\leq\; -\Sigma^{-1}BB^T\Sigma^{-1}\,.$$

implies

$$\begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^T \begin{bmatrix} 0 & \Sigma^{-1} \\ \Sigma^{-1} & 0 \end{bmatrix} \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} + \begin{bmatrix} N\Sigma^{-1}N & 0 \\ 0 & 0 \end{bmatrix}^T \leq \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}^T$$

Multiply from left and right by $\begin{bmatrix} x_1+x_r \\ x_2 \\ 2u \end{bmatrix}$ to get $(\star)$.

Alternative

Stochastic Bounded Real Lemma

$\|\mathbb{L}\|$ is the infimum over all γ so that $\exists X < 0$:

$$\mathcal{R}_\gamma(X) := A^T X + XA + N^T XN - C^T C - \frac{1}{\gamma^2} XBB^T X > 0$$

Apply to
$$dx_e = A_e x_e\, dt + N_e x_e\, dw + B_e u\, dt \ ,$$
$$y_e = C_e x_e = y - y_r \ ,$$

where

$$x_e = \begin{bmatrix} x_1 \\ x_2 \\ x_r \end{bmatrix}, \quad A_e = \begin{bmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22} & 0 \\ 0 & 0 & A_{11} \end{bmatrix},$$
$$N_e = \begin{bmatrix} N_{11} & N_{12} & 0 \\ N_{21} & N_{22} & 0 \\ 0 & 0 & N_{11} \end{bmatrix}, \quad B_e = \begin{bmatrix} B_1 \\ B_2 \\ B_1 \end{bmatrix},$$
$$C_e = \begin{bmatrix} C_1 & C_2 & -C_1 \end{bmatrix} \ .$$

$$\rightsquigarrow X \sim \operatorname{diag}(\Sigma_1, 2\Sigma_2, \sigma_\nu^2 \Sigma_1^{-1}) > 0$$

3rd Lecture: Computations

Summary: Two types of controllability Gramian

Type I: $$PA^T + AP + NPN^T = -BB^T$$

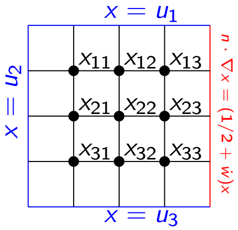
$$i.e. \begin{bmatrix} PA^T + AP + BB^T & NP \\ PN^T & -P \end{bmatrix} \le 0$$

Type II: $$\begin{bmatrix} PA^T + AP + BB^T & PN^T \\ NP & -P \end{bmatrix} \le 0$$

Type	I	II
Def. of P	Matrix equation	LMI
Stability?	Yes, Thm. 1	Yes, Thm. 2
H^2-bound?	Yes, [Redmann& Benner 2014]	Yes, [Redmann 2015]
H^∞-bound?	No, counter-example	Yes, Thm. 2
comput. cost	medium	high (via LMI)

Numerical examples

Numerical example: Stochastic heat equation



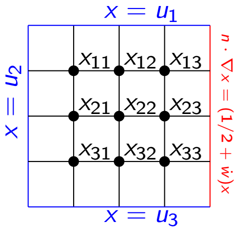
$$n \cdot \nabla x = (1/2 + \dot{w})x\,, \quad \dot{x} = \Delta x$$

$$\Delta x_{ij} \approx -\frac{1}{h^2}\left(4x_{ij} - x_{i+1,j} - x_{i,j+1} - x_{i-1,j} - x_{i,j-1}\right)$$

$x_{14} \approx x_{13} - hu(1/2 + \dot{w})x_{13}, \ldots$

10×10 grid, $C = \frac{1}{100}[1, \ldots, 1]$, $u \equiv \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

Numerical example: Stochastic heat equation



$$n \cdot \nabla x = (1/2 + \dot{w})x \,, \quad \dot{x} = \Delta x$$

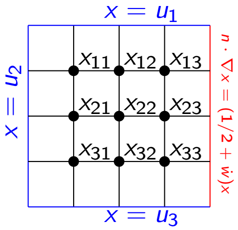
$$\Delta x_{ij} \approx -\frac{1}{h^2}\left(4x_{ij} - x_{i+1,j} - x_{i,j+1} - x_{i-1,j} - x_{i,j-1}\right)$$

$$x_{14} \approx x_{13} - hu(1/2 + \dot{w})x_{13}, \ \ldots$$

10×10 grid, $C = \frac{1}{100}[1, \ldots, 1]$, $u \equiv \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

Type II Gramian P via semidefinite programming/LMI-solver

Numerical example: Stochastic heat equation



$$n \cdot \nabla x = (1/2 + \dot{w})x\ , \quad \dot{x} = \Delta x$$

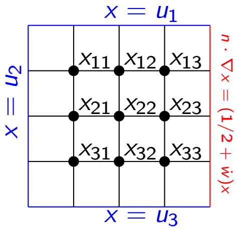
$$\Delta x_{ij} \approx -\frac{1}{h^2}\left(4x_{ij} - x_{i+1,j} - x_{i,j+1} - x_{i-1,j} - x_{i,j-1}\right)$$

$$x_{14} \approx x_{13} - hu(1/2 + \dot{w})x_{13}, \ \ldots$$

10×10 grid, $C = \frac{1}{100}[1, \ldots, 1]$, $u \equiv \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

	$2\sum_{j=11}^{100} \sigma_j$	$\lVert \mathbb{L} - \mathbb{L}_{10} \rVert$	$2\sum_{j=21}^{100} \sigma_j$	$\lVert \mathbb{L} - \mathbb{L}_{20} \rVert$
I	$4.66e-06$	$9.30e-06$	$2.00e-09$	$9.65e-09$
II	$1.75e-05$	$4.83e-06$	$1.72e-08$	$9.70e-09$

Numerical example: Stochastic heat equation



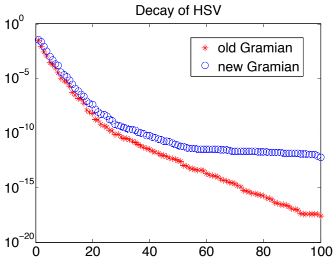
$$n \cdot \nabla x = (1/2 + \dot{w})x \,, \quad \dot{x} = \Delta x$$

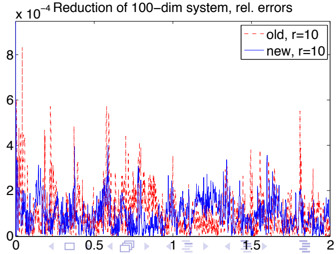
$$\Delta x_{ij} \approx -\frac{1}{h^2}\left(4x_{ij} - x_{i+1,j} - x_{i,j+1} - x_{i-1,j} - x_{i,j-1}\right)$$

$$x_{14} \approx x_{13} - hu(1/2 + \dot{w})x_{13}, \ldots$$

10×10 grid, $C = \frac{1}{100}[1, \ldots, 1]$, $u \equiv \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

	$2\sum_{j=11}^{100} \sigma_j$	$\|\mathbb{L} - \mathbb{L}_{10}\|$	$2\sum_{j=21}^{100} \sigma_j$	$\|\mathbb{L} - \mathbb{L}_{20}\|$
I	$4.66e-06$	$9.30e-06$	$2.00e-09$	$9.65e-09$
II	$1.75e-05$	$4.83e-06$	$1.72e-08$	$9.70e-09$





Another heat equation by courtesy of Martin Redmann

$$\frac{\partial X(t,\zeta)}{\partial t} = \Delta X(t,\zeta) + 1_{[\frac{\pi}{4},\frac{3\pi}{4}]^2}(\zeta)u(t) + \mathrm{e}^{-|\zeta_1-\frac{\pi}{2}|-\zeta_2} X(t-,\zeta)\frac{\partial M(t)}{\partial t},$$

$$\frac{\partial X(t,\zeta)}{\partial \mathbf{n}} = 0, \quad t \geq 0,\ \zeta \in \partial[0,\pi]^2,$$

$$X(0,\zeta) \equiv 0, \qquad M(t) = w(t) - (\Pi(t) - t).$$

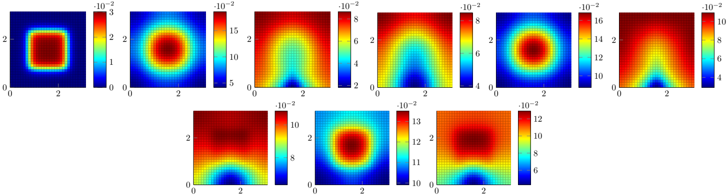


Figure : $u(t) = 1 \cdot \mathrm{e}^{w(t)} +$ "stabilising part".

$$Y(t) = \frac{4}{3\pi^2} \int_{[0,\pi]^2 \setminus [\frac{\pi}{4},\frac{3\pi}{4}]^2} X(t,\zeta) d\zeta.$$

Another heat equation by courtesy of Martin Redmann

A semi-discretised and stabilised version of the heat equation:

$$
\begin{aligned}
dx(t) &= [A_S x(t) + B\tilde{u}(t)]dt + Nx(t-)dM(t), \\
y(t) &= Cx(t).
\end{aligned}
$$

The error between y and the output of the ROM $\tilde{y}_{BT}$:

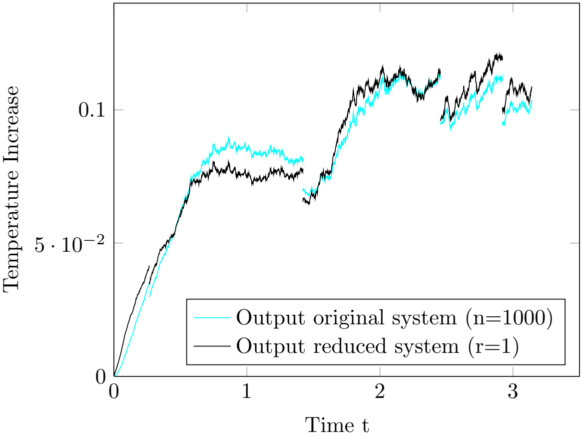
$$
\sup_{t\in[0,\pi]} \mathbb{E} \left\| y(t) - \tilde{y}_{BT}(t) \right\|_{\mathbb{R}^p} \leq \mathcal{E} \left\| \tilde{u} \right\|_{L^2_T} .
$$

For $u_1(t) = \frac{\sqrt{2}}{\pi} w(t)$ and $u_2(t) = \sqrt{\frac{2}{1-e^{-2\pi}}}\, e^{-t}$, $t \in [0, \pi]$, we obtain

Dim. ROM	Exact Error ($\tilde{u} = u_1$)	Exact Error ($\tilde{u} = u_2$)	Bound $\mathcal{E}$
8	$3.4778 \cdot 10^{-6}$	$2.3816 \cdot 10^{-6}$	$3.8971 \cdot 10^{-5}$
4	$1.0199 \cdot 10^{-4}$	$3.7347 \cdot 10^{-4}$	$7.2362 \cdot 10^{-4}$
2	$1.3258 \cdot 10^{-3}$	$1.4083 \cdot 10^{-3}$	$3.8652 \cdot 10^{-3}$
1	$3.9404 \cdot 10^{-3}$	0.0103	0.0335

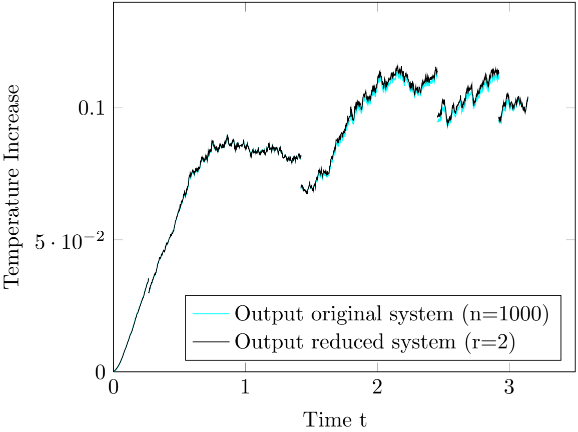
Another heat equation by courtesy of Martin Redmann

Choosing $r = 1, 2, 3$ and the control $\tilde{u}(t) = 1 \cdot e^{w(t)}$ yields



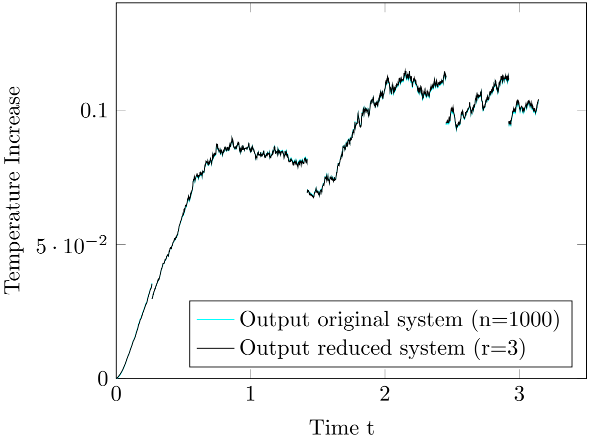
Another heat equation by courtesy of Martin Redmann

Choosing $r = 1, 2, 3$ and the control $\tilde{u}(t) = 1 \cdot \mathrm{e}^{w(t)}$ yields

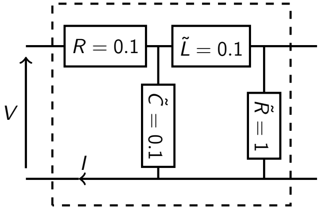


Another heat equation by courtesy of Martin Redmann

Choosing $r = 1, 2, 3$ and the control $\tilde{u}(t) = 1 \cdot \mathrm{e}^{w(t)}$ yields



Numerical example: Electrical ladder network



$n/2$ sections
Here $\tilde{L}^{-1} = L^{-1} + \dot{w}$

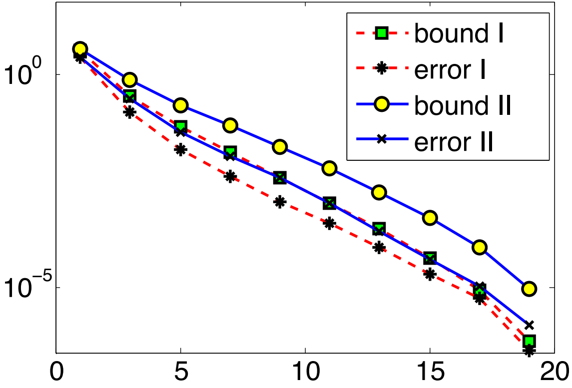
- [Gugercin/Antoulas 2004]
- [Ugrinovskii/Petersen 1999]

$$A = \begin{bmatrix} \frac{-1}{\tilde{C}R} & \frac{-1}{\tilde{C}} & 0 & 0 & 0 & 0 \\ \frac{1}{L} & \frac{-R\tilde{R}}{L(R+\tilde{R})} & \frac{-\tilde{R}}{L(R+\tilde{R})} & 0 & 0 & 0 \\ 0 & \frac{\tilde{R}}{\tilde{C}(R+\tilde{R})} & \frac{-1}{\tilde{C}(R+\tilde{R})} & \frac{-1}{\tilde{C}} & 0 & 0 \\ 0 & 0 & \frac{1}{L} & \frac{-R\tilde{R}}{L(R+\tilde{R})} & \frac{-\tilde{R}}{L(R+\tilde{R})} & 0 \\ 0 & 0 & 0 & \frac{\tilde{R}}{\tilde{C}(R+\tilde{R})} & \frac{-1}{\tilde{C}(R+\tilde{R})} & \frac{-1}{\tilde{C}} \\ 0 & 0 & 0 & 0 & \frac{1}{L} & \frac{-\tilde{R}}{L} \end{bmatrix}, N = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & \frac{-R\tilde{R}}{R+\tilde{R}} & \frac{-\tilde{R}}{R+\tilde{R}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \frac{-R\tilde{R}}{R+\tilde{R}} & \frac{-\tilde{R}}{R+\tilde{R}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -\tilde{R} \end{bmatrix}$$

$$B = \left[\, \tfrac{1}{\tilde{C}R} \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \,\right]^T, C = \left[\, -\tfrac{1}{R} \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \,\right].$$

Numerical example: Electrical ladder network

- $n = 20,\ r = 2, 4, \ldots, 18.$
- For both types of Gramians we compare the error $\|\mathbb{L}_r - \mathbb{L}\|$ and the alleged bound $2 \sum_{j=r+1}^{n} \sigma_j$



Computational issues

Numerical solution of generalized Lyapunov equations

$$\mathcal{L}_A(X) + \Pi(X) := AX + XA^T + \sum N_j X N_j^T = -Y$$

Lyapunov operator: $\mathcal{L}_A(X) = AX + XA^T$

Positive operator: $\Pi(X) = \sum N_j X N_j^T$

Convergent iteration: $X_{k+1} = -\mathcal{L}_A^{-1}\Pi(X_k) - \mathcal{L}_A^{-1}(Y)$

$\Rightarrow \mathcal{L}_A^{-1}$ can be used as preconditioner.

Krylov subspace methods

Why Krylov subspace methods?

$$X_{k+1} = -\mathcal{L}_A^{-1}\Pi(X_k) - \mathcal{L}_A^{-1}(Y)$$

Instead of the iterates $X_0, X_1, \ldots, X_k$ use an „optimal“ linear combination $\tilde{X}_k = \sum_{j=0}^{k} a_j X_j$.

- If X_k converges, then so does $\tilde{X}_k$.
- A good preconditioner is essential.
- Inversion of $\mathcal{L}_A$ can be too expensive.
- In the following: Approximation of $\mathcal{L}_A^{-1}$ via ADI.
- ADI: alternate direction implicit

Idea of ADI-iteration

e.g. Wachspress, Smith, Penzl, Li:

$$
\begin{aligned}
XA^T + AX &= -Y \\
\Longleftrightarrow \quad (A - pI)X(A - pI)^T &= (A + pI)X(A + pI)^T + 2pY
\end{aligned}
$$

yields the fixed point equation

$$
X = \frac{A + pI}{A - pI} X \frac{A^T + pI}{A^T - pI} + 2p \frac{1}{(A - pI)} Y \frac{1}{(A - pI)^T}
$$

Idea of ADI-iteration

e.g. Wachspress, Smith, Penzl, Li:

$$\begin{aligned} XA^T + AX &= -Y \\ \iff \quad (A - pI)X(A - pI)^T &= (A + pI)X(A + pI)^T + 2pY \end{aligned}$$

yields the fixed point equation

$$X = \prod_j \frac{A + p_j I}{A - p_j I} \, X \, \prod_j \frac{A^T + p_j I}{A^T - p_j I} + \prod_j \frac{\sqrt{2p_j}}{A - p_j I} \, Y \, \prod_j \frac{\sqrt{2p_j}}{A^T - p_j I}$$

Idea of ADI-iteration

e.g. Wachspress, Smith, Penzl, Li:

$$\begin{aligned} XA^T + AX &= -Y \\ \iff \quad (A - pI)X(A - pI)^T &= (A + pI)X(A + pI)^T + 2pY \end{aligned}$$

yields the fixed point equation

$$X = \prod_j \frac{A + p_j I}{A - p_j I} X \prod_j \frac{A^T + p_j I}{A^T - p_j I} + \prod_j \frac{\sqrt{2p_j}}{A - p_j I} Y \prod_j \frac{\sqrt{2p_j}}{A^T - p_j I}$$

Choice of parameters: For $\varepsilon < 1$, find $p_1, \ldots, p_\ell$:

$$\rho\Big(\prod_j \frac{A + p_j I}{A - p_j I}\Big) \leq \max_{\lambda \in D} \prod_j \left| \frac{\lambda - p_j}{\lambda + p_j} \right| \overset{!}{\leq} \varepsilon, \qquad \sigma(A) \subset D$$

Classical problem, solution known e.g. for real spectra.

ADI-preconditioner for generalized Lyapunov

Also: $AX + XA^T + \Pi(X) = Y \iff$

$$X = (A - pI)^{-1}\Big((A + pI)X(A + pI)^T + 2p(\Pi(X) - Y)\Big)(A - pI)^{-T}$$

Preconditioned fixed point equation

Iteration: Choose $p_1, \ldots, p_\ell$ (e.g. according to Wachspress).

$$\begin{aligned}
X_1 &= (A - p_1 I) \setminus [(A + p_1 I)X_0(\ldots)^T + 2p_1(\Pi(X_0) - Y)]/(A - p_1 I)^T \\
X_2 &= (A - p_2 I) \setminus [(A + p_1 I)X_1(\ldots)^T + 2p_2(\Pi(X_0) - Y)]/(A - p_2 I)^T \\
&\vdots \\
X_\ell &= (A - p_\ell I) \setminus [(A + p_\ell I)X_{\ell-1}(\ldots)^T + 2p_\ell(\Pi(X_0) - Y)]/(A - p_\ell I)^T
\end{aligned}$$

$X_0 \mapsto X_\ell$: Another preconditioner.
Cheaper than $\mathcal{L}_A^{-1}$ e.g. if A sparse. Usually $\ell \approx 4$ suffices.

ADI-preconditioner for generalized Lyapunov

Also: $AX + XA^T + \Pi(X) = Y \iff$

$$X = (A - pI)^{-1}\Big((A + pI)X(A + pI)^T + 2p(\Pi(X) - Y)\Big)(A - pI)^{-T}$$

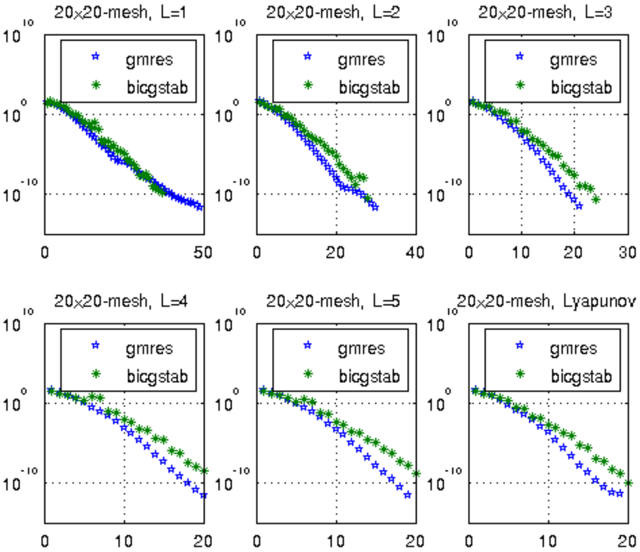
Preconditioned fixed point equation

Iteration: Choose $p_1, \ldots, p_\ell$ (e.g. according to Wachspress).

$$\begin{aligned}
X_1 &= (A - p_1 I) \setminus [(A + p_1 I)X_0(\ldots)^T + 2p_1(\Pi(X_0) - Y)]/(A - p_1 I)^T \\
X_2 &= (A - p_2 I) \setminus [(A + p_1 I)X_1(\ldots)^T + 2p_2(\Pi(X_1) - Y)]/(A - p_2 I)^T \\
&\vdots \\
X_\ell &= (A - p_\ell I) \setminus [(A + p_\ell I)X_{\ell-1}(\ldots)^T + 2p_\ell(\Pi(X_{\ell-1}) - Y)]/(A - p_\ell I)^T
\end{aligned}$$

$X_0 \mapsto X_\ell$: Another preconditioner.
Cheaper than $\mathcal{L}_A^{-1}$ e.g. if A sparse. Usually $\ell \approx 4$ suffices.

Convergence histories for Poisson example



Low rank methods

Recently: Low rank methods for generalized Lyapunov/Sylvester

- Benner/Breiten, *Low rank methods for a class of generalized Lyapunov equations and related issues*, 2013
- Shank/Simoncini/Szyld, *Efficient low-rank solutions of generalized Lyapunov equations*, 2013
- Kressner/Sirković, *Greedy low-rank methods for solving general linear matrix equations*, 2014.
- Jarlebring/Mele/Palitta/Ringh: *Krylov methods for low-rank commuting generalized Sylvester equations*, 2017

Computation of Type II Gramian

Solve $\begin{bmatrix} PA^T + AP + BB^T & PN^T \\ NP & -P \end{bmatrix} \leq 0$

So far only via LMI-solver... ☹

Additional optimization criterion:

$$\text{minimize} \quad \text{trace}\, P$$

Computation of Type II Gramian: `mincx`

$$\text{Solve} \quad \begin{bmatrix} PA^T + AP + BB^T & PN^T \\ NP & -P \end{bmatrix} \leq 0$$

So far only via LMI-solver... ☹

e.g. from MATLAB Robust Control Toolbox

```
setlmis([])
X=lmivar(1,[n,1])
lmiterm([1 1 1 X],A,1,'s')
lmiterm([1 1 1 0],B*B')
lmiterm([1 1 2 X],1,N','s')
lmiterm([1 2 2 X],-1,1)
LMISYS = getlmis;
q = mat2dec(LMISYS,Q);
[~,xopt] = mincx(LMISYS,q,[1e-13,0,0,0,0]);
P = dec2mat(LMISYS,xopt,X);
```

Computation of Type II Gramian: Sedumi

Solve $$\begin{bmatrix} PA^T + AP + BB^T & PN^T \\ NP & -P \end{bmatrix} \leq 0$$

So far only via LMI-solver... ☹

e.g. using YALMIP (and e.g. Sedumi)

```
X=sdpvar(n,n);
F=[X>=0,[A*X+X*A'+B*B',X*N';N*X -X]<=0];
ops = sdpsettings('solver','sedumi',...
            'sedumi.eps',tol,'verbose',0);
optimize(F,trace(X),ops);
P=value(X);
```

Empirical complexity $\approx O(n^6)$.

Some references on LMI-methods

- Boyd/El Ghaoui/Feron/Balakrishnan. *Linear Matrix Inequalities in Systems and Control Theory*, SIAM, (1994).
- Nesterov/Nemirovski, *Interior Point Polynomial Methods in Convex Programming: Theory and Applications*, SIAM, (1994).
- Nemirovski/Gahinet, *The Projective Method for Solving Linear Matrix Inequalities*, Proc. Amer. Contr. Conf., (1994).
- Vandenberghe/Boyd, *A primal-dual potential reduction method for problems involving matrix inequalities*, Math.Programming, (1995).
- Sturm, Using SeDuMi 1.02, *A MATLAB toolbox for optimization over symmetric cones*. Optim.Methods Softw. (1999).
- Löfberg, *YALMIP: A toolbox for modeling and optimization in MATLAB*, Proceedings of the CACSD, (2004).

Computation of stochastic H^{∞}-norm

Stochastic H^∞-norm: Riccati equation and LMI

Problem: Compute stochastic H^∞-norm only using Riccati.

$$\mathcal{R}_\gamma(X) = A^T X + XA + N^T XN - C^T C - \gamma^{-2} XBB^T X$$

By Bounded Real Lemma:

$$\boxed{\|L\| = \inf\{\gamma > 0 \mid \exists \text{ stabilizing } X < 0 : \mathcal{R}_\gamma(X) = 0\}}$$

Stochastic H^∞-norm: Riccati equation and LMI

Problem: Compute stochastic H^∞-norm only using Riccati.

$$\mathcal{R}_\gamma(X) = A^T X + XA + N^T XN - C^T C - \gamma^{-2} XBB^T X$$

By Bounded Real Lemma:

$$\boxed{\|L\| = \inf\{\gamma > 0 \mid \exists \text{ stabilizing } X < 0 : \mathcal{R}_\gamma(X) = 0\}}$$

The case $\gamma = \|L\|$

Assume that (A, N) is mean-square stable, (A, B) controllable

The following are equivalent

(H^∞) $\|\mathbb{L}\| \leq \gamma$

(ARE) $\exists$ (largest) $X_+ \leq 0$ with $\mathcal{R}_\gamma(X_+) = 0$

(LMI) $\exists X \leq 0$: $\begin{bmatrix} A^T X + XA + N^T XN - C^T C & XB \\ B^T X & \gamma^2 I \end{bmatrix} \geq 0$

Computation of stochastic H^∞-norm: mincx

Solve $\quad \exists X \leq 0 : \begin{bmatrix} A^T X + XA + N^T XN - C^T C & XB \\ B^T X & \gamma^2 I \end{bmatrix} \geq 0$

e.g. using MATLAB Robust Control Toolbox

```
  setlmis([])
X = lmivar(1,[n,1]);g = lmivar(1,[1,1]);
lmiterm([1 1 1 X],N',N);
lmiterm([1 1 1 X],A',1,'s');
lmiterm([1 1 1 0],C'*C);
lmiterm([1 1 2 X],1,B,'s');
lmiterm([1 2 2 g],-1,1);
lmisys = getlmis;
c = mat2dec(lmisys,zeros(n),1);
options = [tol,0,0,0,1];
copt = mincx(lmisys,c,options);
gamma = sqrt(copt)
```

Computation of stochastic H^∞-norm: Sedumi

Solve $\quad \exists X \leq 0 : \begin{bmatrix} A^T X + XA + N^T XN - C^T C & XB \\ B^T X & \gamma^2 I \end{bmatrix} \geq 0$

e.g. using YALMIP (and e.g. Sedumi)

```
  P=sdpvar(n,n);gamma=sdpvar(1);
F=[P<=0,[A'*P+P*A+N'*P*N-C'*C,P*B-C'*D;
      B'*P-D'*C gamma*eye(size(B,2))-D'*D]>=0];
ops = sdpsettings('solver','sedumi',
 'sedumi.eps',tol,'verbose',0);
optimize(F,gamma,ops);
gamma=sqrt(value(gamma))
```

Solution of stochastic Riccati equation

$$\mathcal{R}_\gamma(X) = A^T X + XA + N^T XN - C^T C - \gamma^{-2} XBB^T X$$

Derivative of $\mathcal{R}_\gamma$ at X: $(\mathcal{R}_\gamma)'_X = \mathcal{L}_{A-\gamma^{-2}BB^T X} + \Pi_N =: \mathcal{L}_{A_X} + \Pi_N$

Theorem: Let $\gamma > \|\mathbb{L}\|$ and consider the Newton iteration

$$X_{k+1} = X_k - (\mathcal{R}_\gamma)'_{X_k}{}^{-1}(\mathcal{R}(X_k))$$

- If $\sigma((\mathcal{R}_\gamma)'_{X_0}) \subset \mathbb{C}_-$, then X_k converges monotonically to X_+.
- $\forall k \geq 1 : \sigma((\mathcal{R}_\gamma)'_{X_k}) \subset \mathbb{C}_-$, $\mathcal{R}_\gamma(X_k) \leq 0$, $X_k \geq X_{k+1}$

- In particular $X_0 = 0$ is suitable
- Idea: Bisection approach.

Basic algorithm

Algorithm 1 *Computation of stochastic H^∞-norm*

Choose $\gamma_0 < \|\mathbb{L}\| < \gamma_1$, $k_{\max}$, tol
repeat
  Set $\gamma = \frac{\gamma_0+\gamma_1}{2}$, $X_0 = 0$
  repeat
    $\alpha = \max \operatorname{Re} \sigma((\mathcal{R}_\gamma)'_{X_k})$
    if $\alpha < 0$ **then**
      $X_{k+1} = X_k - (\mathcal{R}_\gamma)'_{X_k}{}^{-1}(\mathcal{R}(X_k))$
    end if
  until convergence **or** $k = k_{\max}$ **or** $\alpha \geq 0$
  if convergence **then**
    $\gamma_1 = \gamma$,
  else
    $\gamma_0 = \gamma$
  end if
until $\gamma_1 - \gamma_0 < \text{tol}$

Some implementation issues

- Newton step
- Stability test
- Choosing γ_0 and γ_1
- Can we do better than bisection?

Some implementation issues

- Newton step $X_{k+1} = X_k + H$
  Solve generalized Lyapunov equation

$$A_{X_k}^T H + HA_{X_k} + N^T HN = -\mathcal{R}_\gamma(X_k)$$
$$\iff H = -\mathcal{L}_{A_{X_k}}^{-1}(\Pi_N(H) + \mathcal{R}_\gamma(X_k))$$

  $\leadsto$ use `bicgstab`.
  Recently e.g. [Benner & Breiten 13], [Shank, Simoncini, Szyld 14]
- Stability test
- Choosing γ_0 and γ_1
- Can we do better than bisection?

Some implementation issues

- Newton step
- Stability test Check if $\alpha = \max \operatorname{Re} \sigma((\mathcal{R}_\gamma)'_{X_k}) < 0$
  - First check if $\sigma(A_{X_k}) \subset \mathbb{C}_-$
  - Then check if $\rho(\mathcal{L}_{A_{X_k}}^{-1} \Pi_N) < 1$ by power method:

$$P_0 = I, \quad P_{k+1} = -\mathcal{L}_{A_X}^{-1} \Pi_N(P_k), \quad \rho_k = \frac{\operatorname{trace}(P_k P_{k+1})}{\operatorname{trace}(P_k P_k)}$$

- Choosing γ_0 and γ_1
- Can we do better than bisection?

Some implementation issues

- Newton step
- Stability test
- Choosing γ_0 and γ_1

  Fact: $\|\mathbb{L}_{\text{det}}\| \leq \|\mathbb{L}_{\text{stoch}}\|$

$$\rightsquigarrow \quad \gamma_0 = \max_{s \in i\mathbb{R}} \|C(sI - A)^{-1}B\|_2$$

$$\text{try} \quad \gamma_1 = 2^k \gamma_0 \stackrel{!}{>} \|\mathbb{L}\|, k = 1, 2, \ldots$$

- Can we do better than bisection?

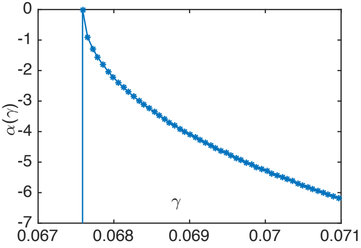
Some implementation issues

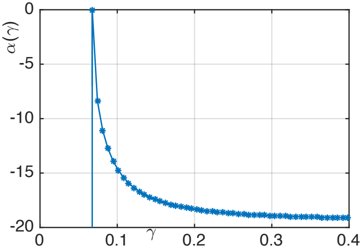
- Newton step
- Stability test
- Choosing γ_0 and γ_1
- Can we do better than bisection?

  Idea: Let $X_+(\gamma)$ = Riccati solution for $\gamma > \|\mathbb{L}\|$. Then

$$\alpha(\gamma) := \max \operatorname{Re} \sigma((\mathcal{R}_\gamma)'_{X_+(\gamma)}) < 0 \quad \text{and} \quad \alpha(\gamma) \overset{\gamma \to \|\mathbb{L}\|}{\longrightarrow} 0$$

But zero-crossing hard to predict:

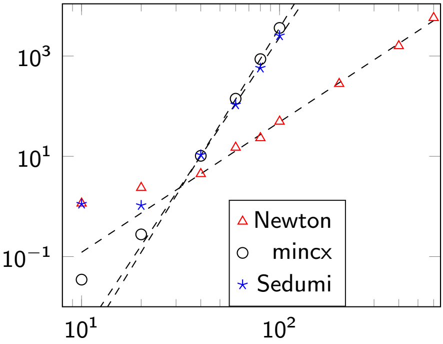




LMI vs. Newton: Random data

Table: Averaged computing times (in sec) for random systems.

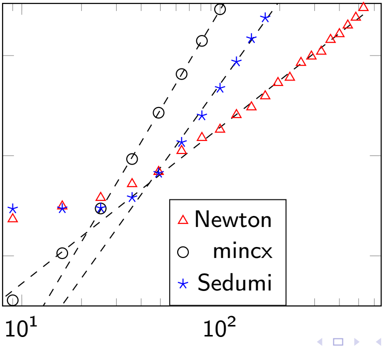
n	10	20	40	60	80	100	200	400	$O(n^s)$
`mincx`	0.03	0.3	10	141	867	3635	-	-	$s \approx 6.4$
Sedumi	1.1	1.1	11	107	577	2550	-	-	$s \approx 5.9$
Newton	1.2	2.4	5	15	23	50	280	1600	$s \approx 2.7$



LMI vs. Newton: Heat equation

Table: Computing times (in sec) for discretized heat equation.

n	16	49	100	144	196	256	400	$O(n^s)$
`mincx`	0.1	72	8418	-	-	-	-	$s \approx 6.6$
Sedumi	0.9	4.5	223	2191	13219	-	-	$s \approx 5.5$
Newton	1	4.8	33	93	285	720	2699	$s \approx 3.1$



Final words

- Model order reduction for stochastic systems widely open topic
- Balanced truncation with different types of Gramians
- Central results can be extended but technically involved
- Need to study generalized Lyapunov operators
- Computations should avoid general purpose LMI-solvers

Some references

- Benner/D., *Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems*, SICON, (2011)
- Benner/Redmann, *Model reduction for stochastic systems.* Stoch. Partial Differ. Equ. Anal. Comput., (2015)
- Benner/Redmann/Rodriguez-Cruz/D., *Positive operators and stable truncation*, LAA, (2016)
- Benner/Redmann/Rodriguez-Cruz/D., *Dual pairs of generalized Lyapunov inequalities and balanced truncation of stochastic linear systems*, IEEE TAC, (2017)
- Redmann/Benner, Peter, *An $\mathcal{H}_2$-type error bound for balancing-related model order reduction of linear systems with Lévy noise*, SCL, (2017)