5. Distributed Control, Quadratic Invariance and Sparsity Invariance

Yang Zheng

Postdoc, Harvard University

May 7, 2020

Outline

1. Recap: optimal control and its convex formulations
2. Quadratic Invariance (QI), and its special cases
3. Beyond QI: Sparsity Invariance
4. QI in finite horizon and Gradient Dominance
5. Summary and future work

Outline

1. Recap: optimal control and its convex formulations
2. Quadratic Invariance (QI), and its special cases
3. Beyond QI: Sparsity Invariance
4. QI in finite horizon and Gradient Dominance
5. Summary and future work

Optimal control

- General optimal control formulation

$$\min_{\mathbf{K}} \quad f(\mathbf{P}, \mathbf{K})$$
$$\text{subject to} \quad \mathbf{K} \text{ internally stabilizes } \mathbf{P}.$$

  where $f(\mathbf{P}, \mathbf{K})$ defines a certain performance index.

- Specifically

Frequency-domain formulation

$$\min_{\mathbf{K}} \quad \|\mathbf{T}_{zw}\|$$
$$\text{subject to} \quad \mathbf{K} \in \mathcal{C}_{\text{stab}},$$

where

$$\mathbf{T}_{zw} = \mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^{-1}\mathbf{P}_{21}.$$

State-space formulation

$$\min \left\| \left[\begin{array}{cc|c} A + B_2 D_k C_2 & B_2 C_k & B_1 + B_2 D_k D_{21} \\ B_k C_2 & A_k & B_k D_{21} \\ \hline C_1 + D_{12} D_k C_2 & D_{12} C_k & D_{11} + D_{12} D_k D_{21} \end{array}\right] \right\|$$

$$\text{s.t.} \quad \begin{bmatrix} A + B_2 D_k C_2 & B_2 C_k \\ B_k C_2 & A_k \end{bmatrix} \text{ is stable.}$$

Input-output parameterization

Consider the closed-loop responses from (δ_y, δ_u) to (y, u):

$$\begin{bmatrix} I & -\mathbf{P}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{Y} & \mathbf{W} \\ \mathbf{U} & \mathbf{Z} \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix}, \tag{1a}$$

$$\begin{bmatrix} \mathbf{Y} & \mathbf{W} \\ \mathbf{U} & \mathbf{Z} \end{bmatrix} \begin{bmatrix} -\mathbf{P}_{22} \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix}, \tag{1b}$$

$$\mathbf{Y}, \mathbf{U}, \mathbf{W}, \mathbf{Z} \in \mathcal{RH}_\infty. \tag{1c}$$

Theorem (Input-output parameterization)

The set of all internally stabilizing controllers can be represented as

$$\mathcal{C}_{stab} = \{\mathbf{K} = \mathbf{U}\mathbf{Y}^{-1} \mid \mathbf{Y}, \mathbf{U}, \mathbf{W}, \mathbf{Z} \textit{ are in the affine subspace } (1a)\text{-}(1c)\}.$$

$$\begin{aligned} \min_{\mathbf{K}} \quad & \|\mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^{-1}\mathbf{P}_{21}\| \\ \text{subject to} \quad & \mathbf{K} \in \mathcal{C}_{\text{stab}}, \end{aligned}$$

$$\begin{aligned} \min_{\mathbf{Y},\mathbf{U},\mathbf{W},\mathbf{Z}} \quad & \|\mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{U}\mathbf{P}_{21}\| \\ \text{subject to} \quad & (1a) - (1c). \end{aligned}$$

System-level parameterization

Consider the closed-loop responses from $(\boldsymbol{\delta}_x, \boldsymbol{\delta}_y)$ to $(\boldsymbol{x}, \boldsymbol{u})$:

$$\begin{bmatrix} sI - A & -B_2 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{N} \\ \mathbf{M} & \mathbf{L} \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix}, \tag{2a}$$

$$\begin{bmatrix} \mathbf{R} & \mathbf{N} \\ \mathbf{M} & \mathbf{L} \end{bmatrix} \begin{bmatrix} sI - A \\ -C_2 \end{bmatrix} = \begin{bmatrix} I \\ 0 \end{bmatrix}, \tag{2b}$$

$$\mathbf{R}, \mathbf{M}, \mathbf{N} \in \mathcal{RH}_\infty, \quad \mathbf{L} \in \mathcal{RH}_\infty. \tag{2c}$$

Theorem (System-level parameterization)

For strictly proper plants, the set of all internally stabilizing controllers can be represented as

$\mathcal{C}_{stab} = \{\mathbf{K} = \mathbf{L} - \mathbf{M}\mathbf{R}^{-1}\mathbf{N} \mid \mathbf{R},\ \mathbf{M},\ \mathbf{N},\ \mathbf{L}$ *are in the affine subspace* (2a)-(2c)$\}$.

System-level synthesis

$$\min_{\mathbf{R},\mathbf{M},\mathbf{N},\mathbf{L}} \quad \left\| \begin{bmatrix} C_1 & D_{12} \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{N} \\ \mathbf{M} & \mathbf{L} \end{bmatrix} \begin{bmatrix} B_1 \\ D_{21} \end{bmatrix} + D_{11} \right\|$$

subject to (2a) − (2c).

Youla parameterization

- **Classical Optimal control**

$$\begin{aligned} \min_{\mathbf{K}} \quad & \|\mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{K}(I - \mathbf{G}\mathbf{K})^{-1}\mathbf{P}_{21}\| \\ \text{subject to} \quad & \mathbf{K} \text{ internally stabilizes } \mathbf{G}. \end{aligned}$$

- We have the following equivalence

$$\mathcal{C}_{\mathsf{stab}} = \{\mathbf{K} = (\mathbf{V}_r - \mathbf{M}_r\mathbf{Q})(\mathbf{U}_r - \mathbf{N}_r\mathbf{Q})^{-1} \mid \mathbf{Q} \in \mathcal{RH}_\infty\},$$

  where $\mathbf{Q}$ is denoted as the Youla parameter.

- **Convex reformulation in Youla**

$$\begin{aligned} \min_{\mathbf{Q}} \quad & \|\mathbf{T}_{11} + \mathbf{T}_{12}\mathbf{Q}\mathbf{T}_{21}\| \\ \text{subject to} \quad & \mathbf{Q} \in \mathcal{RH}_\infty, \end{aligned}$$

  where $\mathbf{T}_{11} = \mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{V}_r\mathbf{M}_l\mathbf{P}_{21}, \mathbf{T}_{12} = -\mathbf{P}_{12}\mathbf{M}_r$, and $\mathbf{T}_{21} = \mathbf{M}_l\mathbf{P}_{21}$.

Optimal state-feedback control

Consider the state-space formulation:

$$\min \left\| \left[\begin{array}{cc|c} A + B_2 D_k C_2 & B_2 C_k & B_1 + B_2 D_k D_{21} \\ B_k C_2 & A_k & B_k D_{21} \\ \hline C_1 + D_{12} D_k C_2 & D_{12} C_k & D_{11} + D_{12} D_k D_{21} \end{array} \right] \right\|$$

$$\text{s.t.} \begin{bmatrix} A + B_2 D_k C_2 & B_2 C_k \\ B_k C_2 & A_k \end{bmatrix} \text{ is stable.}$$

$$\min_{D_k} \left\| \left[\begin{array}{c|c} A + B_2 D_k & B_1 \\ \hline C_1 + D_{12} D_k & D_{11} \end{array} \right] \right\|$$

$$\text{s.t.} \quad A + B_2 D_k \text{ is stable.}$$

- $\mathcal{H}_2$ cost function

$$\begin{aligned} \min_{P,X,Z} \quad & \text{trace}(Z) \\ \text{s.t.} \quad & (AP + B_2 X) + (AP + B_2 X)^\mathsf{T} + B_1 B_1^\mathsf{T} \prec 0, \\ & \begin{bmatrix} Z & C_1 P + D_{12} X \\ (C_1 P + D_{12} X)^\mathsf{T} & P \end{bmatrix} \succ 0, \end{aligned}$$

  where the optimal $\mathcal{H}_2$ optimal state feedback gain is recovered by $D_k = XP^{-1}$.

- LMI formulations are available for the output feedback case.

Classical distributed control

- Canonical distributed control formulation

$$\begin{aligned} \min_{\mathbf{K}} \quad & f(\mathbf{P}, \mathbf{K}) \\ \text{subject to} \quad & \mathbf{K} \text{ internally stabilizes } \mathbf{P} \\ & \mathbf{K} \in \mathcal{S}. \end{aligned}$$

  where $f(\mathbf{P}, \mathbf{K})$ defines a certain performance index and $\mathcal{S}$ denotes a subspace constraint.

- Two typical formulations in the literature

$$\begin{aligned} \min_{\mathbf{K}} \quad & \|\mathbf{T}_{zw}\| \\ \text{subject to} \quad & \mathbf{K} \in \mathcal{C}_{\text{stab}} \cap \mathcal{S}, \end{aligned}$$

where

$$\mathbf{T}_{zw} = \mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^{-1}\mathbf{P}_{21}.$$

$$\begin{aligned} \min_{D_k} \quad & \left\| \left[\begin{array}{c|c} A + B_2 D_k & B_1 \\ \hline C_1 + D_{12} D_k & D_{11} \end{array} \right] \right\| \\ \text{s.t.} \quad & A + B_2 D_k \text{ is stable} \\ & D_k \in \mathcal{S}. \end{aligned}$$

Outline

1. Recap: optimal control and its convex formulations
2. Quadratic Invariance (QI), and its special cases
3. Beyond QI: Sparsity Invariance
4. QI in finite horizon and Gradient Dominance
5. Summary and future work

Classical distributed control formulation

- A canonical problem is to minimize a norm of the closed-loop map subject to a subspace constraint

$$\begin{aligned} \min_{\mathbf{K}} \quad & \|\mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^{-1}\mathbf{P}_{21}\| \\ \text{subject to} \quad & \mathbf{K} \in \mathcal{C}_{\text{stab}}, \\ & \mathbf{K} \in \mathcal{S}, \end{aligned} \tag{3}$$

  where $\mathcal{S}$ is a subspace demoting sparsity or delay constraints on the controller.

- After applying the change of variables in Youla, input-output, or system-level parameterizations, we need to introduce the following non-convex constraint on the decision variables

$$\begin{aligned} (\mathbf{V}_r - \mathbf{M}_r\mathbf{Q})(\mathbf{U}_r - \mathbf{N}_r\mathbf{Q})^{-1} &\in \mathcal{S}, \\ \mathbf{U}\mathbf{Y}^{-1} &\in \mathcal{S}, \\ \mathbf{L} - \mathbf{M}\mathbf{R}^{-1}\mathbf{N} &\in \mathcal{S}. \end{aligned}$$

Quadratic Invariance

Definition (Quadratic Invariance (QI))

Given a plant $\mathbf{P}_{22}$ and a subspace $\mathcal{S}$. The subspace $\mathcal{S}$ is called **quadratically invariant** under $\mathbf{P}_{22}$ if

$$\mathbf{K}\mathbf{P}_{22}\mathbf{K} \in \mathcal{S}, \qquad \forall \mathbf{K} \in \mathcal{S}.$$

Cayley–Hamilton theorem: for any $n \times n$ matrix A, we have the following identity:

$$p(A) = c_0 + c_1 A + \ldots + c_n A^n = 0.$$

Theorem

Define $\mathbf{U} = \mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^{-1}$. *If* $\mathcal{S}$ *is quadratically invariant under* $\mathbf{P}_{22}$*, then*

$$\mathbf{K} \in \mathcal{S} \Longleftrightarrow \mathbf{U} \in \mathcal{S}.$$

$\Rightarrow$ Observe that

$$(I - \mathbf{P}_{22}\mathbf{K})^{-1} = \alpha_0 + \alpha_1 (I - \mathbf{P}_{22}\mathbf{K}) + \ldots + \alpha_{p-1}(I - \mathbf{P}_{22}\mathbf{K})^{p-1}$$

$\Leftarrow$ Observe that $\mathbf{U} = \mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^{-1}$ leads to $\mathbf{K} = \mathbf{U}(I + \mathbf{P}_{22}\mathbf{U})^{-1}$.

QI with the IOP

Theorem (QI with the IOP)

If $\mathcal{S}$ is QI under $\mathbf{P}_{22}$, then

1. *We have*

$$\mathcal{C}_{stab} \cap \mathcal{S} = \{\mathbf{K} = \mathbf{U}\mathbf{Y}^{-1} \mid \mathbf{Y}, \mathbf{U}, \mathbf{W}, \mathbf{Z} \text{ are in } (1a)\text{-}(1c), \mathbf{U} \in \mathcal{S}\}.$$

2. *Problem* (3) *can be equivalently formulated as a convex problem*

$$\begin{aligned} \min_{\mathbf{Y},\mathbf{U},\mathbf{W},\mathbf{Z}} \quad & \|\mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{U}\mathbf{P}_{21}\| \\ \text{subject to} \quad & (1a) - (1c), \\ & \mathbf{U} \in \mathcal{S}. \end{aligned}$$

Proof: From the affine constraint, we have

$$\mathbf{Y} - \mathbf{P}_{22}\mathbf{U} = I \quad \Rightarrow \quad \mathbf{Y} = I + \mathbf{P}_{22}\mathbf{U}.$$

Then, we have

$$\mathbf{K} = \mathbf{U}(I + \mathbf{P}_{22}\mathbf{U})^{-1} \quad \Leftrightarrow \quad \mathbf{U} = \mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^{-1}.$$

QI with the SLP

Corollary (QI with the SLP)

If $\mathcal{S}$ is QI under $\mathbf{P}_{22}$, then

1. *We have*

$$\mathcal{C}_{stab} \cap \mathcal{S} = \{\mathbf{K} = \mathbf{L} - \mathbf{M}\mathbf{R}^{-1}\mathbf{N} \mid \mathbf{R},\ \mathbf{M},\ \mathbf{N},\ \mathbf{L} \text{ are in (2a)-(2c)}, \mathbf{L} \in \mathcal{S}\}.$$

2. *Problem* (3) *can be equivalently formulated as a convex problem*

$$\begin{aligned} \min_{\mathbf{R},\mathbf{M},\mathbf{N},\mathbf{L}} \quad & \left\| \begin{bmatrix} C_1 & D_{12} \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{N} \\ \mathbf{M} & \mathbf{L} \end{bmatrix} \begin{bmatrix} B_1 \\ D_{21} \end{bmatrix} + D_{11} \right\| \\ \textit{subject to} \quad & \text{(2a)} - \text{(2c)}, \\ & \mathbf{L} \in \mathcal{S}. \end{aligned}$$

QI with the Youla

Corollary (QI with Youla)

If $\mathcal{S}$ is QI under $\mathbf{P}_{22}$, then

1. *We have*

$$\mathcal{C}_{stab} \cap \mathcal{S} = \{\mathbf{K} = (\mathbf{V}_r - \mathbf{M}_r\mathbf{Q})(\mathbf{U}_r - \mathbf{N}_r\mathbf{Q})^{-1} \mid \mathbf{Q} \in \mathcal{RH}_\infty \quad (\mathbf{V}_r - \mathbf{M}_r\mathbf{Q})\mathbf{M}_l \in \mathcal{S}, \}.$$

2. *Problem* (3) *can be equivalently formulated as a convex problem*

$$\begin{aligned} \min_{\mathbf{Q}} \quad & \|\mathbf{T}_{11} + \mathbf{T}_{12}\mathbf{Q}\mathbf{T}_{21}\| \\ \textit{subject to} \quad & (\mathbf{V}_r - \mathbf{M}_r\mathbf{Q})\mathbf{M}_l \in \mathcal{S} \\ & \mathbf{Q} \in \mathcal{RH}_\infty. \end{aligned}$$

Summary: Quadratic invariance (QI)

Youla

$$\min_{\mathbf{Q}} \quad \|\mathbf{T}_{11} + \mathbf{T}_{12}\mathbf{Q}\mathbf{T}_{21}\|$$
$$\text{subject to} \quad \mathbf{Q} \in \mathcal{RH}_\infty,$$
$$(\mathbf{V}_r - \mathbf{M}_r\mathbf{Q})(\mathbf{U}_r - \mathbf{N}_r\mathbf{Q})^{-1} \in \mathcal{S}$$

$$(\mathbf{V}_r - \mathbf{M}_r\mathbf{Q})\mathbf{M}_l \in \mathcal{S}$$

IOP

$$\min_{\mathbf{Y},\mathbf{U},\mathbf{W},\mathbf{Z}} \quad \|\mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{U}\mathbf{P}_{21}\|$$
$$\text{subject to} \quad (1a) - (1c).$$
$$\mathbf{U}\mathbf{Y}^{-1} \in \mathcal{S}$$

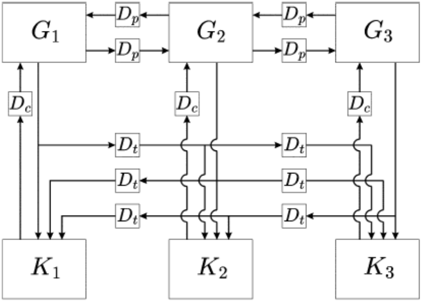
$$\mathbf{U} \in \mathcal{S}$$

SLS

$$\min_{\mathbf{R},\mathbf{M},\mathbf{N},\mathbf{L}} \quad \left\| \begin{bmatrix} C_1 & D_{12} \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{N} \\ \mathbf{M} & \mathbf{L} \end{bmatrix} \begin{bmatrix} B_1 \\ D_{21} \end{bmatrix} + D_{22} \right\|$$
$$\text{subject to} \quad (2a) - (2c)$$
$$\mathbf{L} - \mathbf{M}\mathbf{R}^{-1}\mathbf{N} \in \mathcal{S}.$$

$$\mathbf{L} \in \mathcal{S}$$

Typical QI cases: delay patterns



- Transmission delay $t \geq 0$, propagation delay $p \geq 0$, and computational delay $c \geq 0$. We define: $\mathbf{K} \in \mathcal{S}$ if and only if

$$\mathbf{K} = \begin{bmatrix} D_c H_{11} & D_{t+c} H_{12} & \dots & D_{(n-1)t+c} H_{1n} \\ D_{t+c} H_{21} & D_c H_{22} & \dots & D_{(n-2)t+c} H_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ D_{(n-1)t+c} H_{n1} & D_{(n-2)t+c} H_{n2} & \dots & D_c H_{nn} \end{bmatrix}$$

Typical QI cases: delay patterns

- The plant dynamics are

$$\mathbf{G} = \begin{bmatrix} A_{11} & D_p A_{12} & \dots & D_{(n-1)p} A_{1n} \\ D_p A_{21} & A_{22} & \dots & D_{(n-2)p} A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ D_{(n-1)p} A_{n1} & D_{(n-2)p} A_{n2} & \dots & A_{nn} \end{bmatrix}$$

Theorem

Suppose that $\mathbf{G}$ *and* $\mathcal{S}$ *are defined as above, then* $\mathcal{S}$ *is QI with respect to* $\mathbf{G}$ *if and only if*

$$t \leq p + \frac{c}{n-1}.$$

A nice physical interpretation: The constraint is QI *if the controllers can communicate faster than the dynamics propagate*, i.e., $t \leq p$.

Typical QI cases: symmetric constraints

- When the plant is symmetric, the constraint of symmetric controllers is naturally QI.
- In particular, we have the following result:

Theorem

Suppose $\mathbb{H}^n = \{A \in \mathbb{C}^{n\times n} \mid A = A^*\}$, *and*

$$\mathcal{S} = \{\mathbf{K} \in \mathcal{R}_p \mid K(j\omega) \in \mathbb{H}^n, \forall \omega \in \mathbb{R}\}.$$

If $\mathbf{G} \in \mathcal{R}_p$ *with* $\mathbf{G}(j\omega) \in \mathbb{H}^n$, *then* $\mathcal{S}$ *is QI with respect to* $\mathbf{G}$.

- Rotkowitz, Michael, and Sanjay Lall. "A characterization of convex problems in decentralized control." IEEE transactions on Automatic Control 50.12 (2005): 1984-1996.
- Lessard, Laurent, and Sanjay Lall. "Convexity of decentralized controller synthesis." IEEE Transactions on Automatic Control 61.10 (2015): 3122-3127.
- Lessard, Laurent, and Sanjay Lall. "Quadratic invariance is necessary and sufficient for convexity." Proceedings of the 2011 American Control Conference. IEEE, 2011.

Typical QI cases: sparsity constraints

- Suppose $A^{\text{bin}} \in \{0,1\}^{m \times n}$ is a binary matrix. We define the subspace

$$\text{Sparse}(A^{\text{bin}}) = \{\mathbf{B} \in \mathcal{R}_p \, | \mathbf{B}_{ij}(\omega) = 0, \text{ for all}, i, j, \text{such that } A^{\text{bin}}_{ij} = 0, \\ \text{for almost all } \omega \in \mathbb{R}\}.$$

- Also, if $\mathbf{B} \in \mathcal{R}_p$, let $A^{\text{bin}} = \text{Pattern}(\mathbf{B})$ be the binary matrix given by

$$A^{\text{bin}}_{ij} = \begin{cases} 0, & \text{if } \mathbf{B}_{ij}(j\omega) = 0 \text{ for almost all } \omega \in \mathbb{R}, \\ 1, & \text{otherwise.} \end{cases}$$

Theorem

Suppose $\mathcal{S} = \text{Sparse}(K^{bin})$, and let $G^{bin} = \text{Pattern}(\mathbf{G})$. Then, the following condition are equivalent:

1. *$\mathcal{S}$ is QI with respect to $\mathbf{G}$.*
2. *$\mathbf{KGJ} \in \mathcal{S}, \forall \mathbf{K}, \mathbf{J} \in \mathcal{S}$.*
3. *$K^{bin} G^{bin} K^{bin} \leq K^{bin}$.*

Typical QI cases: sparsity constraints

A negative result:

Perfectly decentralized control is never QI except for the trivial case where no subsystem affects any other.

Corollary

Suppose there exists i, j with $i \neq j$ such that $\mathbf{G}_{ij} \neq 0$. Suppose K^{bin} is diagonal and $\mathcal{S} = Sparse(K^{bin})$. Then, $\mathcal{S}$ is not QI under $\mathbf{G}$.

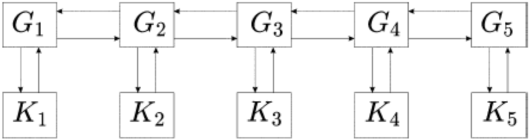


Figure: Fully decentralized control

Non-QI cases: perturb $\mathcal{S}$

- **Closest Superset:** Solve the following binary optimization problem

$$\begin{aligned} \min_{Z} \quad & \|Z\|_0 \\ \text{subject to} \quad & Z G^{\text{bin}} Z \leq Z \\ & K^{\text{bin}} \leq Z, \end{aligned}$$

  It is proved that this problem admits a polynomial time solution as follow

$$Z_0 = K^{\text{bin}}, \qquad Z_{m+1} = Z_m + Z_m G^{\text{bin}} Z_m, \quad m \geq 0$$

  which will converge within finite iterations.

- **Closest subset:** Solve the following binary optimization problem

$$\begin{aligned} \max_{Z} \quad & \|Z\|_0 \\ \text{subject to} \quad & Z G^{\text{bin}} Z \leq Z \\ & Z \leq K^{\text{bin}}. \end{aligned}$$

  There is no known efficient algorithms to solve the problem above.

- M. Rotkowitz and N. Martins. On the nearest quadratically invariant information constraint.IEEE Transactions on Automatic Control, 57(5):1314–1319, 2011.

Non-QI cases: perturb P_{22}

A dual approach

- Approximate the plant dynamics $\mathbf{G}$, which can be combined with robust control to provide **a suboptimality guarantee**.

$$\begin{aligned} \min_{\mathbf{G}_0} \quad & \|\mathbf{G}_0 - \mathbf{G}\|_\infty \\ \text{subject to} \quad & K^{\text{bin}} \cdot \text{Pattern}(\mathbf{G}_0) \cdot K^{\text{bin}} \leq K^{\text{bin}}. \end{aligned}$$

- This problem is equivalent to

$$\begin{aligned} \max_{G_0} \quad & \|G_0\|_0 \\ \text{subject to} \quad & G_0 \leq \text{Pattern}(\mathbf{G}) \\ & K^{\text{bin}} G_0 K^{\text{bin}} \leq K^{\text{bin}}. \end{aligned}$$

- Unlike the nearest QI subset approach, the constraint above is linear in the decision version G_0. The Problem above admits a globally optimal solution.

Outline

Sparsity invariance

Given a binary matrix S, the pair of binary matrices T, R satisfies a property of sparsity invariance (SI) with respect to S if

$$\begin{gathered}\mathbf{Y} \in \text{Sparse}(T) \text{ and } \mathbf{X} \in \text{Sparse}(R) \\ \Downarrow \\ \mathbf{Y}\mathbf{X}^{-1} \in \text{Sparse}(S).\end{gathered} \tag{4}$$

Applications in distributed control

$$\begin{aligned}\min_{\mathbf{Y},\mathbf{U},\mathbf{W},\mathbf{Z}} \quad & \|\mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{U}\mathbf{P}_{21}\| \\ \text{subject to} \quad & (1a) - (1c). \\ & \mathbf{U} \in \text{Sparse}(T), \\ & \mathbf{Y} \in \text{Sparse}(R)\end{aligned}$$

$$\begin{aligned}\min_{P,X,Z} \quad & \text{trace}(Z) \\ \text{s.t.} \quad & (AP + B_2X) + (AP + B_2X)^{\mathsf{T}} + B_1B_1^{\mathsf{T}} \prec 0, \\ & \begin{bmatrix} Z & C_1P + D_{12}X \\ (C_1P + D_{12}X)^{\mathsf{T}} & P \end{bmatrix} \succ 0, \\ & X \in \text{Sparse}(T), P \in \text{Sparse}(R).\end{aligned}$$

The notion of QI is irrelevant of static controller design!

Sparsity invariance

Main result: all characterizations of sparsity invariance

Given a binary matrix $S \subseteq \{0,1\}^{m\times n}$, the following are equivalent

- S is SI under (T, R),
- $T \leq S$, and $TR^{n-1} \leq S$,

where we use a binary matrix to represent a sparse space, *i.e.*, $T \in \{0,1\}^{m\times n}$.

Proof sketch: Cayley-Hamilton theorem

$$X^{-1} = \lambda_0 I + \lambda_1 X + \lambda_2 X^2 + \ldots + \lambda_{n-1} X^{n-1},$$

Two facts

- $\forall X \in \text{Sparse}(R)$ and any integer $r \in \mathbb{R}$, we have $X^r \in \text{Sparse}(R^r)$,
- $\forall X \in \text{Sparse}(R)$ and $Y \in \text{Sparse}(T)$, we have $YX \in \text{Sparse}(TR)$,

Combining the facts and the Cayley Hamilton theorem:
$\forall Y \in \text{Sparse}(T), X \in \text{Sparse}(R)$, we have

$$\begin{aligned} YX^{-1} &= \lambda_0 Y + \lambda_1 YX + \lambda_2 YX^2 + \ldots + \lambda_{n-1} YX^{n-1} \\ &\in \text{Sparse}(T) + \text{Sparse}(TR) \ldots + \text{Sparse}(TR^{n-1}). \end{aligned}$$

Sparsity invariance

Computation: maximizing the number of nonzeros in R

- Given a binary matrix $T \subseteq S \in \{0,1\}^{m \times n}$, consider

$$\begin{aligned} \max_{R \in \{0,1\}^{n \times n}} \quad & \|R\|_0 \\ \text{subject to} \quad & TR^{n-1} \leq S \end{aligned}$$

- Consider another binary optimization problem

$$\begin{aligned} \max_{R \in \{0,1\}^{n \times n}} \quad & \|R\|_0 \\ \text{subject to} \quad & TR^{n-1} \leq T \end{aligned}$$

- The problem above admits an analytical solution R^* (Algorithm 1)

$$R^*_{jk} = \begin{cases} 0 & \text{if } \exists i = 1, \ldots, m, \text{such that } T_{ik} = 0, T_{ij} = 1 \\ 1 & \text{otherwise} \end{cases}$$

Proof: we have $TR^{n-1} \leq T \Leftrightarrow TR \leq T$.

Connection with QI

Theorem

Let $\Delta = \text{Pattern}(\mathbf{G})$ and let $R_S^\star$ be the binary matrix generated by Our algorithm with $T = S$. The following statements are equivalent.

i) *$\text{Sparse}(S)$ is QI with respect to $\mathbf{G}$.*

ii) *$R^\star \geq I_p + \Delta S$, where $R^\star$ is generated by our algorithm with $T = S$.*

Beyond QI for sparsity constraints

1. SI is guaranteed to return a globally optimal solution when QI holds,
2. SI may still recover globally optimal solutions when QI does *not* hold,
3. SI can be guaranteed to return a solution at least as good as the closest QI subset approach.

Outline

Distributed control in finite horizon

- We consider time-varying linear systems in discrete-time

$$x_{t+1} = A_t x_t + B_t u_t + w_t \,,$$
$$y_t = C_t x_t + v_t \,,$$

- Consider a planning problem for the next N steps:

$$\mathbf{A} = \text{blkdg}(A_0, \ldots, A_N),\ \mathbf{B} = \begin{bmatrix} \text{blkdg}(B_0, \ldots, B_{N-1}) \\ 0_{n \times mN} \end{bmatrix},\ \mathbf{C} = \text{blkdg}(C_0, \ldots, C_N)\,,$$

$$\mathbf{x} = \begin{bmatrix} x_0^\mathsf{T} & \ldots & x_N^\mathsf{T} \end{bmatrix}^\mathsf{T}, \qquad \mathbf{y} = \begin{bmatrix} y_0^\mathsf{T} & \ldots & y_N^\mathsf{T} \end{bmatrix}^\mathsf{T}, \quad \mathbf{u} = \begin{bmatrix} u_0^\mathsf{T} & \ldots & u_{N-1}^\mathsf{T} \end{bmatrix}^\mathsf{T},$$

$$\mathbf{w} = \begin{bmatrix} x_0^\mathsf{T} & w_0^\mathsf{T} & \ldots & w_{N-1}^\mathsf{T} \end{bmatrix}^\mathsf{T}, \quad \mathbf{v} = \begin{bmatrix} v_0^\mathsf{T} & \ldots & v_N^\mathsf{T} \end{bmatrix}^\mathsf{T},$$

and the block-down shift matrix $\mathbf{Z} = \begin{bmatrix} 0_{1 \times N} & 0 \\ I_N & 0_{N \times 1} \end{bmatrix} \otimes I_n$.

- we can write the system dynamics compactly as

$$\mathbf{x} = \mathbf{ZAx} + \mathbf{ZBu} + \mathbf{w}, \qquad \mathbf{y} = \mathbf{Cx} + \mathbf{v},$$

leading to

$$\mathbf{x} = \mathbf{P}_{11}\mathbf{w} + \mathbf{P}_{12}\mathbf{u}\,, \qquad \mathbf{y} = \mathbf{Cx} + \mathbf{v}$$

where $\mathbf{P}_{11} = (I - \mathbf{ZA})^{-1}$ and $\mathbf{P}_{12} = (I - \mathbf{ZA})^{-1}\mathbf{ZB}$.

Distributed control in finite horizon

- We consider linear output-feedback policies

$$u_t = K_{t,0}y_0 + K_{t,1}y_1, + \ldots, K_{t,t}y_t, \qquad t = 0, 1, \ldots, N-1.$$

- Compactly, we write

$$\mathbf{u} = \mathbf{K}\mathbf{y}, \quad \mathbf{K} \in \mathcal{K}\,,$$

where $\mathcal{K}$ is a subspace in $\mathbb{R}^{mN \times p(N+1)}$ encoding a certain sparsity for distributed control.

Distributed control in finite horizon

$$\begin{aligned} \min \quad & J(\mathbf{K}) := \mathbb{E}_{\mathbf{w},\mathbf{v}}\left[\sum_{t=0}^{N-1}\left(y_t^\mathsf{T} M_t y_t + u_t^\mathsf{T} R_t u_t\right) + y_N^\mathsf{T} M_N y_N\right], \\ \text{subject to} \quad & \mathbf{K} \in \mathcal{K} \end{aligned} \tag{5}$$

QI and gradient dominance

Let $d \in \mathbb{N}$ be the dimension of $\mathcal{K}$, and the columns of $P \in \mathbb{R}^{mpN(N+1)\times d}$ be a basis of the subspace $\{\text{vec}(\mathbf{K})|\ \forall \mathbf{K} \in \mathcal{K}\}$. Define the function $f : \mathbb{R}^d \to \mathbb{R}$ as $f(z) := J(\text{vec}^{-1}(Pz))$. Then, (5) is equivalent to the *unconstrained* problem

$$\min_{z \in \mathbb{R}^d} f(z)\,.$$

Quadratic Invariance

$$\mathbf{KCP}_{12}\mathbf{K} \in \mathcal{K}, \quad \forall \mathbf{K} \in \mathcal{K}\,.$$

Theorem (Gradient Dominance)

Let $\mathcal{K}$ be QI with respect to $\mathbf{CP}_{12}$. For any $\delta > 0$ define the sublevel set $\mathcal{G}_\delta = \{z \in \mathbb{R}^d \mid\ f(z) - J^\star \leq \delta\}$. Then, the following statements hold.

1. *$\mathcal{G}_\delta$ is compact, and $f(z)$ has a unique stationary point.*
2. *$f(z)$ admits a local gradient dominance constant $\mu_\delta > 0$ over $\mathcal{G}_\delta$*

$$\mu_\delta(f(z) - J^\star) \leq \|\nabla f(z)\|_2^2\,,\ \forall z \in \mathcal{G}_\delta\,.$$

QI and gradient dominance: proof

- Closed-loop systems:

$$\begin{bmatrix}\mathbf{y}\\ \mathbf{u}\end{bmatrix} = \begin{bmatrix}\mathbf{C}(I-\mathbf{P}_{12}\mathbf{K}\mathbf{C})^{-1}\mathbf{P}_{11} & (I-\mathbf{C}\mathbf{P}_{12}\mathbf{K})^{-1}\\ \mathbf{K}\mathbf{C}(I-\mathbf{P}_{12}\mathbf{K}\mathbf{C})^{-1}\mathbf{P}_{11} & \mathbf{K}(I-\mathbf{C}\mathbf{P}_{12}\mathbf{K})^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{w}\\ \mathbf{v}\end{bmatrix}.$$

- Youla parameterization in finite horizon

$$\mathbf{Q} = \mathbf{K}(I-\mathbf{C}\mathbf{P}_{12}\mathbf{K})^{-1}, \qquad \mathbf{Q}\in\mathcal{S} \Leftrightarrow \mathbf{K}\in\mathcal{S} \qquad \text{QI}$$

- A change of variables leads to a strongly convex problem in $\mathbf{Q}$.

$$f(z) \quad \overset{q=h(z)}{\Longleftrightarrow} \quad g(q)$$

- Apply the chain rule of differentiation $\nabla f(z) = \mathbf{J}_h \cdot \nabla g(h(z))$

$$\begin{aligned} f(z) - f^* = g(h(z)) - g^* &\leq \mu_g \|\nabla g\,(h(z))\,\|_2^2\,,\\ &= \mu_g \|\mathbf{J}_h^{-1}\cdot\nabla f\,(z)\|_2^2\\ &\leq \mu_g \|\mathbf{J}_h^{-1}\|_F^2 \|\nabla f(z)\|_2^2\,, \quad \forall z\in\mathbb{R}^d\,. \end{aligned}$$

Outline

1. Recap: optimal control and its convex formulations
2. Quadratic Invariance (QI), and its special cases
3. Beyond QI: Sparsity Invariance
4. QI in finite horizon and Gradient Dominance
5. Summary and future work

Summary

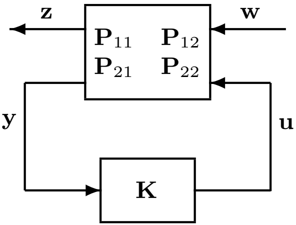


Figure: Interconnection of the plant $\mathbf{P}$ and controller $\mathbf{K}$

- **Lecture 1**: Problem formulation in both frequency domain and time-domain, well-posedness, internal stability

Frequency-domain formulation

$$\min_{\mathbf{K}} \quad \|\mathbf{T}_{zw}\|$$

subject to $\quad \mathbf{K} \in \mathcal{C}_{\text{stab}},$

where

$$\mathbf{T}_{zw} = \mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^{-1}\mathbf{P}_{21}.$$

State-space formulation

$$\min \left\| \left[\begin{array}{cc|c} A + B_2 D_k C_2 & B_2 C_k & B_1 + B_2 D_k D_{21} \\ B_k C_2 & A_k & B_k D_{21} \\ \hline C_1 + D_{12} D_k C_2 & D_{12} C_k & D_{11} + D_{12} D_k D_{21} \end{array}\right] \right\|$$

s.t. $\begin{bmatrix} A + B_2 D_k C_2 & B_2 C_k \\ B_k C_2 & A_k \end{bmatrix}$ is stable.

Summary

Lecture 2: Convex reformulation in frequency domain

- External transfer matrix characterization of internal stability

Two useful facts:

- The set of stable matrices $\{A \in \mathbb{R}^{n\times n} \mid A \text{ is stable}\}$ is non-convex, but finite-dimensional;
- The set of stable transfer matrices $\{\mathbf{T}(s) \mid \mathbf{T}(s) \in \mathcal{RH}_\infty\}$ is convex, but infinite-dimensional;

- System-level parameterization and input-output parameterization
- Some applications in learning-based control: robust stability

$$(A, B) \to (\hat{A}, \hat{B}), \qquad \mathbf{G} \to \hat{\mathbf{G}}$$

Summary

Lecture 3: Youla parameterization and disturbance-based implementation

- Youla for open-loop stable plants, and internal model principle: The controller $\mathbf{K} = \mathbf{Q}(I + \mathbf{G}\mathbf{Q})^{-1}$ can be implemented in a disturbance-based form:

$$\beta = \mathbf{y} - \mathbf{G}\mathbf{u},$$
$$\mathbf{u} = \mathbf{Q}\beta.$$

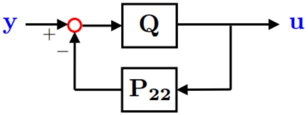


Figure: Internal model principle, where $\mathbf{P}_{22} := \mathbf{G}$.

- Disturbance-based feedback (in both infinite-horizon and finite horizon) $\rightarrow$ convexity
- Doubly co-prime factorization, its computation, and its feedback interpretation

Summary

Lecture 4: Convex formulation in state-space domain

- Hardy space $\mathcal{H}_2$ and $\mathcal{H}_\infty$, and their LMI characterizations
- Change of variables $K = YX^{-1}$ and optimal control in state feedback.
- We consider static state feedback $u = D_k x$, and the controller synthesis problem becomes

$$\begin{aligned} \min_{D_k} \quad & \left\| \left[\begin{array}{c|c} A + B_2 D_k & B_1 \\ \hline C_1 + D_{12} D_k & D_{11} \end{array} \right] \right\| \\ \text{subject to} \quad & A + B_2 D_k \text{ is stable.} \end{aligned}$$

- LMI for the $\mathcal{H}_2$ case

$$\begin{aligned} \min_{P,X,Z} \quad & \operatorname{trace}(Z) \\ \text{subject to} \quad & (AP + B_2 X) + (AP + B_2 X)^\mathsf{T} + B_1 B_1^\mathsf{T} \prec 0, \\ & \begin{bmatrix} Z & C_1 P + D_{12} X \\ (C_1 P + D_{12} X)^\mathsf{T} & P \end{bmatrix} \succ 0, \end{aligned}$$

Summary

Lecture 5: Distributed control, Quadratic Invariance, and Sparsity Invariance

- Classical formulation

$$\begin{aligned} \min_{\mathbf{K}} \quad & \|\mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^{-1}\mathbf{P}_{21}\| \\ \text{subject to} \quad & \mathbf{K} \in \mathcal{C}_{\text{stab}}, \\ & \mathbf{K} \in \mathcal{S}, \end{aligned}$$

- Quadratic Invariance allows for equivalent convex reformulation in frequency domain

$$\mathbf{K}\mathbf{P}_{22}\mathbf{K} \in \mathcal{S}, \forall \mathbf{K} \in \mathcal{S}$$

- QI is independent of controller parameterizations, and works for Youla, SLP, and IOP.
- Non-QI cases: 1) perturb the constraints (QI subset or QI superset); 2) perturb the dynamics (Robust control)
- Beyond QI for sparsity constraints: Sparsity Invariance

$$\mathbf{Y} \in \text{Sparse}(T) \text{ and } \mathbf{X} \in \text{Sparse}(R)$$
$$\Downarrow$$
$$\mathbf{Y}\mathbf{X}^{-1} \in \text{Sparse}(S).$$

Possible research topics

- Other closed-loop parameterization, and mismatches in equality constraints.

$$(zI - A_2)\mathbf{R} - B_2\mathbf{M} = I + \mathbf{\Delta}_1, \qquad \mathbf{Y} - \mathbf{G}\mathbf{U} = I + \mathbf{\Delta}_2.$$

- Robust control, sub-optimality, sample complexity analysis for learning LQG controller (an output feedback version of the Corse-ID procedure);

$$\mathbf{G} = \mathbf{G}_0 + \mathbf{\Delta}, \|\mathbf{\Delta}\|_\infty < \gamma$$

- Online learning of distributed controller with/without QI constraints: gradient descent over the structured Youla parameter?
- Safe learning and robust control with stability guarantees