Designing controllers for unknown systems using data - lecture 2

Claudio De Persis

Institute of Engineering and Technology
J.C. Willems Center for Systems and Control

university of groningen

PhD summer school – Complex networks: dynamics and control
Mathematics Institute, University of Warwick, July 3 - 5 2019

Joint work with Pietro Tesi (Università di Firenze)

Outline

Lecture 1

- Linear control
- Data-based system representation

Lecture 2

- **Direct data-based control design**
- **Extensions**

Direct data-based control design

Closed-loop data-based representation

The closed-loop data-based representation enables the design of controllers without the intermediate step of estimating the model.

Why should we care about this different solution if we already know how to solve the problem?

Because

- We want a **systematic** method that **might** be extendible to those systems (nonlinear, hybrid) for which SI is hard
- It is intellectually stimulating
- We skip one step (the SI one)

Closed-loop data-based representation

Arrange the closed loop system as

$$A + BK = \begin{bmatrix} B & A \end{bmatrix} \begin{bmatrix} K \\ I \end{bmatrix}$$

By the Fundamental Lemma and Rouché-Capelli Theorem

$$\begin{bmatrix} K \\ I \end{bmatrix} = \begin{bmatrix} U_{0,1,T} \\ X_{0,T} \end{bmatrix} G_K, \quad \text{for some } G_K$$

Hence

$$A + BK = \begin{bmatrix} B & A \end{bmatrix} \begin{bmatrix} U_{0,1,T} \\ X_{0,T} \end{bmatrix} G_K = X_{1,T} G_K$$

Data-based parametrization of the closed-loop

Theorem System

$$x(k+1) = Ax(k) + Bu(k)$$

in closed-loop with a state feedback $u = Kx$ has the following equivalent representation

$$x(k+1) = X_{1,T} G_K x(k)$$

where G_K is a $T \times n$ matrix satisfying

$$\begin{bmatrix} K \\ I_n \end{bmatrix} = \begin{bmatrix} U_{0,1,T} \\ X_{0,T} \end{bmatrix} G_K$$

One can then **look for** G_K such that the **closed-loop matrix** $X_{1,T} G_K$ has desired properties

Detour: LMIs

A **linear matrix inequality** (LMI) is an inequality

$$F(y) := F_0 + F_1 y_1 + \ldots + F_N y_N \prec 0$$

where

- $y_1, \ldots, y_N$ are the decision variables in $\mathbb{R}$
- $F_0, F_1, \ldots, F_N$ are **symmetric** matrices
- $F(y) \prec 0$ means that $F(y)$ is **negative definite**

Checking an LMI or optimizing a convex function over a constraint defined by an LMI is a **convex optimization problem**

C. Scherer and S. Weiland, "Linear matrix inequalities in control," notes for a course of the Dutch Institute of Systems and Control, 2004.

Detour: LMIs

LMIs often appear as **functions of matrix variables**. Consider

$$\begin{bmatrix} X_0 Y & X_1 Y \\ Y^\top X_1^\top & X_0 Y \end{bmatrix} \prec 0 \qquad \begin{array}{l} Y \in \mathbb{R}^{m_1 \times m_2} \text{ decision variable} \\ X_0 Y \text{ symmetric} \end{array}$$

Since inequality is homogeneous in Y there is no loss of generality to consider the **nonstrict** inequality

$$\begin{bmatrix} X_0 Y & X_1 Y \\ Y^\top X_1^\top & X_0 Y \end{bmatrix} \preceq -I$$

Detour: LMIs

LMIs often appear as **functions of matrix variables**. Consider

$$\begin{bmatrix} X_0 Y & X_1 Y \\ Y^\top X_1^\top & X_0 Y \end{bmatrix} \preceq -I \qquad \begin{matrix} Y \in \mathbb{R}^{m_1 \times m_2} \text{ decision variable} \\ X_0 Y \text{ symmetric} \end{matrix}$$

Let $Y_1, \ldots, Y_N$ be a basis of the vector space $\mathbb{R}^{m_1 \times m_2}$ and let

$$Y = \sum_j y_j Y_j, \quad y_j \in \mathbb{R}$$

Then

$$-I \succeq \begin{bmatrix} X_0 Y & X_1 Y \\ Y^\top X_1^\top & X_0 Y \end{bmatrix} = \sum_j y_j \begin{bmatrix} X_0 Y_j & X_1 Y_j \\ Y_j^\top X_1^\top & X_0 Y_j \end{bmatrix}$$

is an LMI

Direct data-based stabilization

Problem Find G_K such that the closed-loop system

$$x(k+1) = X_{1,T} G_K x(k)$$

is asymptotically stable

A necessary and sufficient condition is the existence of matrices P and G_K such that

$$P \succ 0, \quad X_{1,T} G_K P G_K^\top X_{1,T}^\top - P \prec 0, \quad I = X_{0,T} G_K$$

If they exist, then the control gain is given by

$$K = U_{0,1,T} G_K$$

These are not LMIs because of the product $G_K P G_K^\top$.

Direct data-based stabilization

To overcome this drawback, we consider the change of variables

$$Y = G_K P$$

to obtain

$$P \succ 0, \quad X_{1,T} Y P^{-1} Y^\top X_{1,T}^\top - P \prec 0, \quad P = X_{0,T} Y$$

which, by Schur's complement, is equivalent to

$$\begin{bmatrix} X_{0,T} Y & Y^\top X_{1,T}^\top \\ X_{1,T} Y & X_{0,T} Y \end{bmatrix} \succ 0$$

The solution to the LMI returns Y. The control gain is obtained via

$$\begin{array}{rl} K = & U_{0,1,T} G_K \\ Y = & G_K P \\ P = & X_{0,T} Y \end{array} \quad \Rightarrow \quad K = U_{0,1,T} Y (X_{0,T} Y)^{-1}$$

Direct data-based stabilization

Theorem Any matrix Y satisfying

$$\begin{bmatrix} X_{0,T}\, Y & X_{1,T}\, Y \\ Y^\top X_{1,T}^\top & X_{0,T}\, Y \end{bmatrix} \succ 0$$

is such that

$$K = U_{0,1,T}\, Y (X_{0,T}\, Y)^{-1}$$

is a stabilizing state-feedback gain for system

$$x(k+1) = Ax(k) + Bu(k)$$

Converse result if K is a stabilizing state-feedback gain for the system, then it can be written as $K = U_{0,1,T}\, Y (X_{0,T}\, Y)^{-1}$

Example

Data-based stabilization of the linearized predator-prey model

State response to PE input from experiment

$$X_{0,5} = \begin{bmatrix} 0.4027 & 0.3478 & 0.3571 & 0.3216 & 0.2362 \\ 0.4448 & 1.1451 & 1.7499 & 2.3708 & 2.9301 \end{bmatrix}$$

$$X_{1,5} = \begin{bmatrix} 0.3478 & 0.3571 & 0.3216 & 0.2362 & 0.1541 \\ 1.1451 & 1.7499 & 2.3708 & 2.9301 & 3.3409 \end{bmatrix}$$

Solve for Y

```
cvx_begin sdp
variable Y(T,n)
[X0T*Y X1T*Y; Y'*X1T' X0T*Y]>=eye(2*n);
cvx_end
```

which returns

$$Y = \begin{bmatrix} 27.4724 & -20.8515 \\ -25.5235 & -8.8555 \\ -1.6399 & -2.0356 \\ 5.3938 & 3.6399 \\ 0.1696 & 18.8019 \end{bmatrix}$$

S. Boyd. "Solving semidefinite programs using cvx," http://stanford.edu/class/ee363/notes/lmi-cvx.pdf

Example

Data-based stabilization of the linearized predator-prey model

Feedback gain

$$K = U_{0,1,5} Y (X_{0,5} Y)^{-1} = \begin{bmatrix} -8.2995 & -1.2512 \end{bmatrix}$$

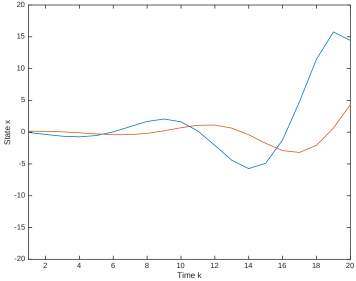

Figure: Unforced solution $u(k) = 0$

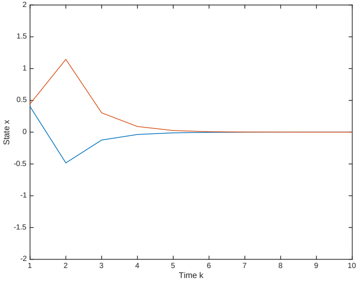

Figure: Solution under data-based feedback $u(k) = Kx(k)$

Spectral radius data-based controlled system $\rho(A + BK) = 0.5666$

Extensions

Optimality - LQR

LQR – Design $u(0), u(1), u(2), \ldots$ that **minimizes**

$$J_\infty(x_0, u) := \sum_{k=0}^{\infty} (x(k) Q x(k) + u(k)^\top R u(k)), \quad Q \succeq 0, R \succ 0$$

for the system $x(k+1) = Ax(k) + Bu(k)$, $x(0) = x_0$

There exists a unique solution given by

$$u_\star := K_\star x, \quad K_\star := -(R + B^\top P B)^{-1} B^\top P A$$

with P the **stabilizing** solution of the DARE

$$A^\top P A - P - A^\top P B (R + B^\top P B)^{-1} B^\top P A + Q = 0,$$

`[-Kstar,P,e] = dlqr(A,B,Q,R) returns the solution Kstar,P to the LQR problem`

Data-based solution to LQR

A reformulation of LQR as an optimization problem

$$\min_{K,P,X}\ \mathrm{trace}\,(QP) + \mathrm{trace}\,(X)$$

$$\text{subject to}$$

$$\begin{cases} (A+BK)P(A+BK)^\top - P + I_n \preceq 0 \\ P \succeq I_n \\ X - R^{1/2}KPK^\top R^{1/2} \succeq 0 \end{cases}$$

E. Feron, V. Balakrishnan, S. Boyd, L. El Ghaoui, "Numerical methods for H_2 related problems," in 1992 American Control Conference, pp. 2921–2922.

Data-based solution to LQR

A similar change of coordinates as before leads to the semidefinite program

$$\min_{Y,X} \ \mathrm{trace}\left(QX_{0,T}Y\right) + \mathrm{trace}\left(X\right)$$

$$\text{subject to}$$

$$\begin{cases} \begin{bmatrix} X & R^{1/2}U_{0,1,T}Y \\ Y^\top U_{0,1,T}^\top R^{1/2} & X_0 Y \end{bmatrix} \succeq 0 \\ \begin{bmatrix} X_{0,T}Y - I_n & X_{1,T}Y \\ Y^\top X_{1,T}^\top & X_{0,T}Y \end{bmatrix} \succeq 0 \end{cases}$$

solvable via efficient numerical algorithms (cvx)

The resulting optimal gain matrix is given by

$$K_\star = U_{0,1,T}Y(X_{0,T}Y)^{-1}$$

which coincides with the DARE-based solution

$$K_\star = -(R + B^\top PB)^{-1}B^\top PA$$

Data-based solution to LQR

```
cvx_begin sdp
    variable Q(T,n)
    variable X(m,m) symmetric
    minimize ( trace(Q*X0*Y) +trace(X) )
    [X, sqrtm(R)*U0*Y; Y'*U0'*sqrtm(R)', X0*Y] >= 0
    [X0*Y-eye(n), X1*Y; Y'*X1', X0*Y] >= 0
cvx_end

K = U0*Y*(inv(X0*Y));
```

Q-learning and LQR

Algorithm 1 The Q-learning algorithm applied to the LQR problem

Guess initial stabilizing gain K_0
Set initial time $k = 0$
for $i = 0$ to ∞ **do**
  for $j = 1$ to N **do**
    Apply $u(k) = K_i x(k) + e(k)$, $e(k)$ PE "exploration signal"
    Estimate $K_i(j)$ using RLS and I/O measurements
    $k = k + 1$
  end for
  Set $K_{i+1} = K_i(N)$
end for

There exists an estimation interval N such that the algorithm generates a sequence $\{K_i \,:\, i = 0, 1, 2, \ldots\}$ such that $\lim_{i\to\infty} \|K_i - K_\star\| = 0$

S.J. Bradtke, B.E. Ydstie and A.G. Barto. Adaptive linear quadratic control using policy iteration. Proceedings of the 1994 American Control Conference, 3475–3479, 1994.
J.C.H. Watkins and P. Dayan. Q-learning. Machine learning, 8(3-4):279–292, 1992.

Noisy measurements

$$
\begin{aligned}
x(k+1) &= \quad Ax(k) + Bu(k) \\
\zeta(k) &= \quad x(k) + w(k) \qquad k = 0, 1, 2, \ldots
\end{aligned}
$$

where w is an unknown measurement noise

Experiment

- Consider a PE input $u_{[0,T-1]}$ of order $n+t$ with $t=1$
- Apply it to the system and collect the **measured** (hence, **noisy**) state response in the $n \times T$ matrix

$$
Z_{0,T} = X_{0,T} + W_{0,T}
$$

where

$$
\begin{aligned}
X_{0,T} &= \quad \begin{bmatrix} x(0) & x(1) & \ldots & x(T-1) \end{bmatrix} \\
W_{0,T} &= \quad \begin{bmatrix} w(0) & w(1) & \ldots & w(T-1) \end{bmatrix}
\end{aligned}
$$

Noisy measurements

Theorem Let

$$R_{0,T}R_{0,T}^\top \preceq \gamma Z_{1,T}Z_{1,T}^\top \qquad \text{"signal-to-noise ratio"}$$

for some $\gamma \in (0,1)$, where $R_{0,T} := AW_{0,T} - W_{1,T}$.

Any matrix Y and scalar $\alpha > 0$ satisfying $\gamma < \alpha^2/(4+2\alpha)$ and

$$\begin{bmatrix} Z_{0,T}\,Y - \alpha Z_{1,T}Z_{1,T}^\top & Z_{1,T}\,Y \\ Y^\top Z_{1,T}^\top & Z_{0,T}\,Y \end{bmatrix} \succeq 0 \quad \begin{bmatrix} I_T & Y \\ Y^\top & Z_{0,T}\,Y \end{bmatrix} \succeq 0$$

is such that

$$K = U_{0,1,T}\,Y(X_{0,T}\,Y)^{-1}$$

is a stabilizing state-feedback gain for system

$$x(k+1) = Ax(k) + Bu(k)$$

In practice, search for the feasible solution maximizing α

Output feedback

Minimal SISO space representation with output measurements

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) & x(k) \in \mathbb{R}^n,\ u(k) \in \mathbb{R} \\ y(k) &= Cx(k) & y(k) \in \mathbb{R},\ k = 0, 1, 2, \dots \end{aligned}$$

Experiment

- Consider an input $u_{[-n,T-1]}$, $T \geq 2n+1$, with $u_{[0,T-1]}$ PE of order $2n+1$
- At time $k = -n$ from the initial condition $x(-n)$, apply $u_{[-n,T-1]}$ to the system and collect the **measured output** response in the $(2n+1) \times T$ matrix

$$\left[\begin{array}{c} U_{0,1,T} \\ \hline \hat{X}_{0,T} \end{array}\right] := \left[\begin{array}{c} U_{0,1,T} \\ \hline Y_{-n,n,T} \\ U_{-n,n,T} \end{array}\right]$$

- By the Fundamental Lemma and Key Reachability Lemma

$$\operatorname{rank} \begin{bmatrix} U_{0,1,T} \\ \hat{X}_{0,T} \end{bmatrix} = 2n+1$$

Output feedback

Theorem Any matrix $\mathcal{Y}$ satisfying

$$\begin{bmatrix} \hat{X}_{0,T}\,\mathcal{Y} & \hat{X}_{1,T}\mathcal{Y} \\ \mathcal{Y}^\top \hat{X}_{1,T}^\top & \hat{X}_{0,T}\,\mathcal{Y} \end{bmatrix} \succ 0,$$

is such that the **dynamic** controller

$$\begin{aligned} \xi(k+1) &= \begin{bmatrix} -c_n & 1 & 0 & \dots & 0 \\ -c_{n-1} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -c_1 & 0 & 0 & \dots & 0 \end{bmatrix} \xi(k) + \begin{bmatrix} d_n \\ d_{n-1} \\ \vdots \\ d_2 \\ d_1 \end{bmatrix} y(k) \\ u(k) &= \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \end{bmatrix} \xi(k) \end{aligned}$$

with coefficients given by

$$\mathcal{K} := \begin{bmatrix} d_1 & \dots & d_n - c_1 & \dots & -c_n \end{bmatrix} = U_{0,1,T}\mathcal{Y}(\hat{X}_{0,T}\mathcal{Y})^{-1}$$

stabilizes the system.

Converse result any stabilizing output feedback controller can be given the form above with coefficients $\mathcal{K} = U_{0,1,T}\mathcal{Y}(\hat{X}_{0,T}\mathcal{Y})^{-1}$

Conclusions

A systematic method for the direct design of control policies of linear systems under persistence of excitation

- Stabilization, LQR, output feedback, MIMO systems
- One-shot experiment of duration $n+1$ (or $2n+1$ – output feedback)
- Robustness to noise (stabilization in first approximation of nonlinear systems possible!)
- LMIs and SDPs are ubiquitous in control and these results can be extended in several ways

Future work

- Complex dynamics nonlinear and hybrid
- Large-scale systems experiments are performed in open-loop, problems with large-scale unstable systems (large-scale systems require large T)
- Connections with machine learning

Reference

De Persis, Tesi (2019). On persistency of excitation and formulas for data-driven control. arXiv:1802.08457

Thank you!