BALANCING USING POLYNOMIAL ALGEBRA

Jan C. Willems

ESAT-SCD (SISTA), University of Leuven, Belgium

To Anders Lindquist

on the occasion of his 60-th birthday.

THEME

!! Given a representation of a dynamical system,

find a representation of a reduced model !!

THEME

!! Given a representation of a dynamical system,

find a representation of a reduced model !!

Algorithm: **parameters of model $\mapsto$ parameters of reduced model**

THEME

!! Given a representation of a dynamical system,

find a representation of a reduced model !!

Algorithm: **parameters of model $\mapsto$ parameters of reduced model**

For example, **model: discrete-time impulse response**

reduced model: balanced reduced model

Algorithm: **SVD of Hankel matrix.**

THEME

!! Given a representation of a dynamical system,

find a representation of a reduced model !!

Algorithm: parameters of model $\mapsto$ parameters of reduced model

For example, model: discrete-time impulse response

reduced model: balanced reduced model

Algorithm: SVD of Hankel matrix.

For example, model: transfer function

reduced model: balanced reduced model

Algorithm: ???

THE SYSTEM

For simplicity, (today) only: **SISO systems & classical I/O balancing**

THE SYSTEM

For simplicity, (today) only: **SISO systems & classical I/O balancing**

System $\cong$ $p, q \in \mathbb{R}[\xi]$, degree$(q) \leq$ degree$(p) =: \mathtt{n} \rightsquigarrow$

$$\boxed{p(\tfrac{d}{dt})y = q(\tfrac{d}{dt})u,}$$

relating the input $u : \mathbb{R} \to \mathbb{R}$ **to the output** $y : \mathbb{R} \to \mathbb{R}$.

Behavior:

$$\mathfrak{B}_{(p,q)} := \{(u, y) \in \mathcal{L}_2^{\text{loc}}(\mathbb{R}, \mathbb{R}^2) \times \mathfrak{L}_2^{\text{loc}}(\mathbb{R}, \mathbb{R}) \mid \text{diff. eq'n holds}\}.$$

CONTROLLABILITY & OBSERVABILITY

Well-known: $\mathfrak{B}_{(p,q)}$ is controllable if and only if p and q are co-prime.

CONTROLLABILITY & OBSERVABILITY

Well-known: $\mathfrak{B}_{(p,q)}$ is controllable if and only if p and q are co-prime.

Controllability $\Leftrightarrow$ $\exists$ image representation for $\mathfrak{B}_{(p,q)}$:

$$u = p(\frac{d}{dt})\ell, \quad y = q(\frac{d}{dt})\ell,$$

$$\mathfrak{Im}_{(p,q)} := \{(u,y) \in \mathcal{L}_2^{\mathrm{loc}}(\mathbb{R},\mathbb{R}^2) \mid \exists \ell : \mathbb{R} \to \mathbb{R} : \text{ diff. eq'n holds}\}$$

is *exactly* equal to $\mathfrak{B}_{(p,q)}$.

CONTROLLABILITY & OBSERVABILITY

Well-known: $\mathfrak{B}_{(p,q)}$ is controllable if and only if p and q are co-prime.

Controllability $\Leftrightarrow$ $\exists$ image representation for $\mathfrak{B}_{(p,q)}$:

$$u = p(\frac{d}{dt})\ell, \quad y = q(\frac{d}{dt})\ell,$$

$$\mathfrak{Im}_{(p,q)} := \{(u,y) \in \mathcal{L}_2^{\text{loc}}(\mathbb{R},\mathbb{R}^2) \mid \exists \ell : \mathbb{R} \to \mathbb{R} : \text{ diff. eq'n holds}\}$$

is *exactly* equal to $\mathfrak{B}_{(p,q)}$.

Co-primeness of p and q $\Rightarrow$

controllability of $\mathfrak{B}_{(p,q)}$ & observability of $\mathfrak{Im}_{(p,q)}$

observability means: for every $(u,y) \in \mathfrak{Im}_{(p,q)} = \mathfrak{B}_{(p,q)}, \exists(!) \ \ell$.

STATE

Any set of polynomials $\{x_1, x_2, \ldots, x_{\mathrm{n}}\}$ that form a basis for $\mathbb{R}_{\mathrm{n}-1}[\xi]$

$\Rightarrow$ a **minimal state representation** of $\mathfrak{B}_{(p,q)}$ with state

$$x = (x_1(\frac{d}{dt})\ell, x_2(\frac{d}{dt})\ell, \ldots, x_{\mathrm{n}-1}(\frac{d}{dt})\ell).$$

The associated system matrices are the (unique) solution matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ of the following system of linear equations in $\mathbb{R}^{\mathrm{n}}[\xi]$:

$$\begin{bmatrix} \xi x_1(\xi) \\ \xi x_2(\xi) \\ \vdots \\ \xi x_{\mathrm{n}}(\xi) \\ q(\xi) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_1(\xi) \\ x_2(\xi) \\ \vdots \\ x_{\mathrm{n}}(\xi) \\ p(\xi) \end{bmatrix}.$$

BALANCING

In the context of the state construction through an image representation, being balanced becomes a property of the polynomials $x_1, x_2, \ldots, x_n$.

The central problem of this paper is:

Choose the polynomials $x_1, x_2, \ldots, x_n$ so that this

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

is balanced.

QDF's

The real two-variable polynomial

$$\Phi(\zeta,\eta) = \Sigma_{k,k'} \Phi_{k,k'} \zeta^{k} \eta^{k'}$$

induces the map

$$w \in \mathfrak{C}^{\infty}(\mathbb{R},\mathbb{R}) \;\mapsto\; \Sigma_{k,k'} \left(\frac{d^{k}}{dt^{k}} w\right) \Phi_{k,k'} \left(\frac{d^{k'}}{dt^{k'}} w\right) \quad \in \mathfrak{C}^{\infty}(\mathbb{R},\mathbb{R}),$$

called a ***a quadratic differential form*** **(QDF), denoted as** Q_{Φ}.

THE CONTROLLABILITY GRAMIAN

We will consider the controllability and observability gramians as QDF's acting on the latent variable ℓ of the image representation.

THE CONTROLLABILITY GRAMIAN

We will consider the controllability and observability gramians as QDF's acting on the latent variable ℓ of the image representation.

The *controllability gramian* Q_K is defined as:

THE CONTROLLABILITY GRAMIAN

We will consider the controllability and observability gramians as QDF's acting on the latent variable ℓ of the image representation.

The *controllability gramian Q_K* is defined as:

Let $\ell \in \mathfrak{C}^\infty(\mathbb{R}, \mathbb{R})$ and define $Q_K(\ell)$ by

$$Q_K(\ell)(0) := \text{infimum} \int_{-\infty}^{0} |p(\frac{d}{dt})\ell'(t)|^2 \, dt,$$

infimum over all $\ell' \in \mathfrak{C}^+(\mathbb{R}, \mathbb{R})$ that join the 'fixed' future ℓ at $t = 0$, i.e., such that $\ell(t) = \ell'(t)$ for $t \geq 0$.

THE OBSERVABILITY GRAMIAN

The ***observability gramian*** Q_W **is defined as:**

THE OBSERVABILITY GRAMIAN

The *observability gramian* Q_W is defined as:

Let $\ell \in \mathfrak{C}^\infty(\mathbb{R}, \mathbb{R})$ and define $Q_W(\ell)$ by

$$Q_W(\ell)(0) := \int_0^\infty |q(\frac{d}{dt})\ell'(t)|^2 \, dt,$$

where $\ell' \in \mathcal{D}(\mathbb{R}, \mathbb{R})$ is such that

(i) $\ell|_{(-\infty,0)} = \ell'|_{-\infty,0)}$,

(ii) $(p(\frac{d}{dt})\ell', q(\frac{d}{dt})\ell') \in \mathfrak{B}_{(p,q)}$,

(iii) $p(\frac{d}{dt})\ell'(t)|_{(0,\infty)} = 0.$

THE OBSERVABILITY GRAMIAN

The *observability gramian* Q_W is defined as:

Let $\ell \in \mathfrak{C}^\infty(\mathbb{R}, \mathbb{R})$ and define $Q_W(\ell)$ by

$$Q_W(\ell)(0) := \int_0^\infty |q(\frac{d}{dt})\ell'(t)|^2 \, dt,$$

where $\ell' \in \mathcal{D}(\mathbb{R}, \mathbb{R})$ is such that

$$\text{(i)} \qquad \ell|_{(-\infty,0)} = \ell'|_{-\infty,0)},$$

$$\text{(ii)} \qquad (p(\frac{d}{dt})\ell', q(\frac{d}{dt})\ell') \in \mathfrak{B}_{(p,q)},$$

$$\text{(iii)} \qquad p(\frac{d}{dt})\ell'(t)|_{(0,\infty)} = 0.$$

ℓ' smoothly continues ℓ at $t = 0$ such that $u|_{(0,\infty)} = p(\frac{d}{dt})\ell'|_{(0,\infty)} = 0$.

COMPUTATION of K and W

Given $\mathfrak{B}_{(p,q)}$, p, q **co-prime,** $\mathrm{degree}(q) \leq \mathrm{degree}(p) =: \mathtt{n}$, p **Hurwitz.**

COMPUTATION of K and W

Given $\mathfrak{B}_{(p,q)}$, p, q **co-prime,** $\text{degree}(q) \leq \text{degree}(p) =: \mathtt{n}$, p **Hurwitz.**

The controllability gramian and the observability gramian are QDF's,
Q_K **and** Q_W, **with** $K, W \in \mathbb{R}[\zeta, \eta]$. **They can be computed as follows**

COMPUTATION of K and W

Given $\mathfrak{B}_{(p,q)}$, p, q **co-prime,** $\mathrm{degree}(q) \leq \mathrm{degree}(p) =: \mathtt{n}$, p **Hurwitz.**

The controllability gramian and the observability gramian are QDF's,

Q_K **and** Q_W**, with** $K, W \in \mathbb{R}[\zeta, \eta]$. **They can be computed as follows**

$$K(\zeta,\eta) = \frac{p(\zeta)p(\eta) - p(-\zeta)p(-\eta)}{\zeta + \eta}$$

COMPUTATION of K and W

Given $\mathfrak{B}_{(p,q)}$, p, q co-prime, $\text{degree}(q) \leq \text{degree}(p) =: \mathtt{n}$, p Hurwitz.

The controllability gramian and the observability gramian are QDF's, Q_K and Q_W, with $K, W \in \mathbb{R}[\zeta, \eta]$. They can be computed as follows

$$K(\zeta,\eta) = \frac{p(\zeta)p(\eta) - p(-\zeta)p(-\eta)}{\zeta + \eta}$$

$$W(\zeta,\eta) = \frac{p(\zeta)f(\eta) + f(\zeta)p(\eta) - q(\zeta)q(\eta)}{\zeta + \eta}$$

with $f \in \mathbb{R}_{\mathtt{n}-1}[\xi]$ the (unique) solution of the Bezout-type equation

$$p(\xi)f(-\xi) + f(\xi)p(-\xi) - q(\xi)q(-\xi) = 0.$$

BALANCED STATE REPRESENTATION

The minimal state repr. with polynomials $(x_1, x_2, \ldots, x_n)$ is *balanced* if

BALANCED STATE REPRESENTATION

The minimal state repr. with polynomials $(x_1, x_2, \ldots, x_n)$ is *balanced* if

(i) for ℓ_k such that $x_{k'}(\frac{d}{dt})\ell_k(0) = \delta_{kk'}$ ($\delta_{kk'}$: Kronecker delta):

$$Q_K(\ell_k)(0) = \frac{1}{Q_W(\ell_k)(0)}$$

state components that are difficult to reach are also difficult to observe.

BALANCED STATE REPRESENTATION

The minimal state repr. with polynomials $(x_1, x_2, \ldots, x_n)$ is *balanced* if

(i) for ℓ_k such that $x_{k'}(\frac{d}{dt})\ell_k(0) = \delta_{kk'}$ ($\delta_{kk'}$: Kronecker delta):

$$Q_K(\ell_k)(0) = \frac{1}{Q_W(\ell_k)(0)}$$

state components that are difficult to reach are also difficult to observe.

(ii) The state components are ordered so that 'easiest to reach first':

$$0 < Q_K(\ell_1)(0) \leq Q_K(\ell_2)(0) \leq \cdots \leq Q_K(\ell_n)(0),$$

and hence 'easiest to observe' first:

$$Q_W(\ell_1)(0) \geq Q_W(\ell_2)(0) \geq \cdots \geq Q_W(\ell_n)(0) > 0.$$

It is a standard result from linear algebra (see Gantmacher, chapter 9) that there exist polynomials

$$(x_1^{\mathrm{bal}}, x_2^{\mathrm{bal}}, \ldots, x_{\mathrm{n}}^{\mathrm{bal}})$$

that form a basis for $\mathbb{R}_{\mathrm{n}-1}[\xi]$, and real numbers

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\mathrm{n}} > 0$$

such that K and W are factored as

$$\boxed{K(\zeta, \eta) = \Sigma_{k=1}^{\mathrm{n}} \sigma_{\mathrm{k}}^{-1} x_{\mathrm{k}}^{\mathrm{bal}}(\zeta) x_{\mathrm{k}}^{\mathrm{bal}}(\eta)}$$

$$\boxed{W(\zeta, \eta) = \Sigma_{\mathrm{k}=1}^{\mathrm{n}} \sigma_{\mathrm{k}} \; x_{\mathrm{k}}^{\mathrm{bal}}(\zeta) x_{\mathrm{k}}^{\mathrm{bal}}(\eta)}$$

The σ_{k}'s are uniquely defined by K and W, the $x_{\mathrm{k}}^{\mathrm{bal}}$ 'almost'.

THEOREM: These σ_k's are the Hankel singular values of $\mathfrak{B}_{(p,q)}$ and

$$u = p(\frac{d}{dt})\ell, y = q(\frac{d}{dt})\ell,$$

$$x^{\mathrm{bal}} = (x_1^{\mathrm{bal}}(\frac{d}{dt})\ell, x_2^{\mathrm{bal}}(\frac{d}{dt})\ell, \ldots, x_n^{\mathrm{bal}}(\frac{d}{dt})\ell)$$

is a balanced state space representation of $\mathfrak{B}_{(p,q)}$.

THEOREM: These σ_k's are the Hankel singular values of $\mathfrak{B}_{(p,q)}$ and

$$u = p(\frac{d}{dt})\ell, y = q(\frac{d}{dt})\ell,$$

$$x^{\mathrm{bal}} = (x_1^{\mathrm{bal}}(\frac{d}{dt})\ell, x_2^{\mathrm{bal}}(\frac{d}{dt})\ell, \ldots, x_n^{\mathrm{bal}}(\frac{d}{dt})\ell)$$

is a balanced state space representation of $\mathfrak{B}_{(p,q)}$.

The balanced system matrices: sol'n of the following linear equations in $\mathbb{R}^n[\xi]$:

$$\begin{bmatrix} \xi x_1^{\mathrm{bal}}(\xi) \\ \xi x_2^{\mathrm{bal}}(\xi) \\ \vdots \\ \xi x_n^{\mathrm{bal}}(\xi) \\ q(\xi) \end{bmatrix} = \begin{bmatrix} A^{\mathrm{bal}} & B^{\mathrm{bal}} \\ C^{\mathrm{bal}} & D^{\mathrm{bal}} \end{bmatrix} \begin{bmatrix} x_1^{\mathrm{bal}}(\xi) \\ x_2^{\mathrm{bal}}(\xi) \\ \vdots \\ x_n^{\mathrm{bal}}(\xi) \\ p(\xi) \end{bmatrix}.$$

ALGORITHM

DATA: $p, q \in \mathbb{R}[\xi]$, co-prime, degree$(q) \leq$ degree$(p) := \mathtt{n}$, p Hurwitz.

COMPUTE:

1. $K \in \mathbb{R}[\zeta, \eta]$,
2. $f \in \mathbb{R}_{\mathtt{n}-1}[\xi]$ and $W \in \mathbb{R}[\zeta, \eta]$,
3. $(x_1^{\mathrm{bal}}, x_2^{\mathrm{bal}}, \ldots, x_{\mathtt{n}}^{\mathrm{bal}})$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\mathtt{n}} > 0$ by the expansions: $K(\zeta, \eta) = \Sigma_{k=1}^{\mathtt{n}} \sigma_{\mathrm{k}}^{-1} x_{\mathrm{k}}^{\mathrm{bal}}(\zeta) x_{\mathrm{k}}^{\mathrm{bal}}(\eta)$, $W(\zeta, \eta) = \Sigma_{\mathrm{k}=1}^{\mathtt{n}} \sigma_{\mathrm{k}} \; x_{\mathrm{k}}^{\mathrm{bal}}(\zeta) x_{\mathrm{k}}^{\mathrm{bal}}(\eta)$,
4. the balanced system matrices $\begin{bmatrix} A^{\mathrm{bal}} & B^{\mathrm{bal}} \\ C^{\mathrm{bal}} & D^{\mathrm{bal}} \end{bmatrix}$ by solving the linear eq'ns

OUTPUT: a balanced state representation of $\mathfrak{B}_{(p,q)}$.

REMARKS

1. Model reduction by balanced truncation follows.

REMARKS

2. These algorithms open up the possibility to involve ‘fast’ polynomial computations in order to obtain a balanced representation.

REMARKS

3. The reduction algorithms solve linear equations in $\mathbb{R}_{n-1}[\xi]$ 'approximately'.

Suggests other (say, least squares) methods than simple truncation.

REMARKS

4. Instead of computing the σ_{k}'s and the $x_{\mathrm{k}}^{\mathrm{bal}}$'s by the factorization of K, W, we can also proceed by evaluating K and W at $\mathtt{n}$ distinct points $\lambda_1, \lambda_2, \cdots, \lambda_{\mathrm{n}} \in \mathbb{C}$.

Define

$$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

$$\begin{aligned} K_\Lambda &= \left[K(\lambda_k^*, \lambda_{k'})\right]_{k=1,\ldots,n}^{k'=1,\ldots,n} \\ W_\Lambda &= \left[W(\lambda_k^*, \lambda_{k'})\right]_{k=1,\ldots,n}^{k'=1,\ldots,n} \end{aligned}$$

$$\begin{aligned} X_\Lambda &= \left[x_k^{\mathrm{bal}}(\lambda_{k'})\right]_{k=1,\ldots,n}^{k'=1,\ldots,n} \\ \Sigma &= \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n) \end{aligned}$$

Define

$$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

$$\begin{aligned} K_\Lambda &= \left[K(\lambda_k^*, \lambda_{k'})\right]_{k=1,\ldots,n}^{k'=1,\ldots,n} \\ W_\Lambda &= \left[W(\lambda_k^*, \lambda_{k'})\right]_{k=1,\ldots,n}^{k'=1,\ldots,n} \end{aligned}$$

$$\begin{aligned} X_\Lambda &= \left[x_k^{\mathrm{bal}}(\lambda_{k'})\right]_{k=1,\ldots,n}^{k'=1,\ldots,n} \\ \Sigma &= \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n) \end{aligned}$$

There holds

$$K_\Lambda = X_\Lambda^* \Sigma^{-1} X_\Lambda, W_\Lambda = X_\Lambda^* \Sigma X_\Lambda.$$

This implies that X_Λ and Σ can be computed directly from K_Λ, W_Λ.

This implies that X_Λ and Σ can be computed directly from K_Λ, W_Λ.

Once X_Λ is known, the matrices of the balanced state representation

$$\begin{bmatrix} A^{\mathrm{bal}} & B^{\mathrm{bal}} \\ C^{\mathrm{bal}} & D^{\mathrm{bal}} \end{bmatrix}$$

is readily computed.

K_Λ follows immediately from evaluation of p at the λ_k's.

K_{Λ} follows immediately from evaluation of p at the λ_k's.

Unfortunately, in order to compute W_{Λ} we have to solve for f.

K_Λ follows immediately from evaluation of p at the λ_k's.

Unfortunately, in order to compute W_Λ we have to solve for f.

However, if we take for the λ_k's the roots of p, assumed distinct,

then f is not needed,

and a very explicit expression for both K and W is obtained.

K_Λ follows immediately from evaluation of p at the λ_k's.

Unfortunately, in order to compute W_Λ we have to solve for f.

However, if we take for the λ_k's the roots of p, assumed distinct,

then f is not needed,

and a very explicit expression for both K and W is obtained.

In this case

$$K_\Lambda = -\left[\frac{p(-\lambda_k^*)p(-\lambda_{k'})}{\lambda_k^* + \lambda_{k'}}\right]_{k=1,\ldots,n}^{k'=1,\ldots,n}$$

$$W_\Lambda = -\left[\frac{q(\lambda_k^*)q(\lambda_{k'})}{\lambda_k^* + \lambda_{k'}}\right]_{k=1,\ldots,n}^{k'=1,\ldots,n}$$

Balancing and model reduction: $\rightsquigarrow$ the pencil

$$\left[\frac{p(-\lambda_{k}^{*})p(-\lambda_{k'})}{\lambda_{k}^{*}+\lambda_{k'}}\right]_{k=1,\ldots,n}^{k'=1,\ldots,n} \quad ; \quad \left[\frac{q(\lambda_{k}^{*})q(\lambda_{k'})}{\lambda_{k}^{*}+\lambda_{k'}}\right]_{k=1,\ldots,n}^{k'=1,\ldots,n}$$

Balancing and model reduction: $\rightsquigarrow$ **the pencil**

$$\left[\frac{p(-\lambda_{k}^{*})p(-\lambda_{k'})}{\lambda_{k}^{*}+\lambda_{k'}}\right]_{k=1,\ldots,n}^{k'=1,\ldots,n} \quad ; \quad \left[\frac{q(\lambda_{k}^{*})q(\lambda_{k'})}{\lambda_{k}^{*}+\lambda_{k'}}\right]_{k=1,\ldots,n}^{k'=1,\ldots,n}$$

5. Heuristic: evaluate K, W at less than $\mathtt{n}$ points, obtain reduced model.

Balancing and model reduction: $\rightsquigarrow$ **the pencil**

$$\left[\frac{p(-\lambda_{k}^{*})p(-\lambda_{k'})}{\lambda_{k}^{*}+\lambda_{k'}}\right]_{k=1,\ldots,n}^{k'=1,\ldots,n} \quad ; \quad \left[\frac{q(\lambda_{k}^{*})q(\lambda_{k'})}{\lambda_{k}^{*}+\lambda_{k'}}\right]_{k=1,\ldots,n}^{k'=1,\ldots,n}$$

5. Heuristic: evaluate K, W at less than $\mathtt{n}$ points, obtain reduced model.

6. Suggests algorithms to fit the reduced order transfer function with the original transfer function at privileged points of the complex plane.

$$(p, q) \longmapsto (p_{\mathbf{reduced}}, q_{\mathbf{reduced}})??$$

More info, copy sheets? **Surf to**

`http://www.esat.kuleuven.ac.be/~jwillems`

or write to me at `Jan.Willems@esat.kuleuven.ac.be`

Thank you!