Finding the Characteristics: Radial Basis Function Interpolation for Parametric Model Order Reduction

Nils Hornung (Fraunhofer SCAI), Sara Grundel (MPI Magdeburg)

2nd Workshop on Model Reduction of Complex Dynamical Systems
December 12, 2013
Definition and stability

Let

\[ A \in \mathbb{R}^{d \times d}, \quad B \in \mathbb{R}^d, \quad C \in \mathbb{R}^{1 \times d} \]

A linear time-invariant (LTI) system

\[ \Sigma : \begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} \] (1)

is called stable if A has eigenvalues only in the left half plane.
Model order reduction methods try to find a reduced LTI system

\[ \hat{\Sigma} : \begin{cases} \dot{x}(t) = \hat{A}x(t) + \hat{B}u(t) \\ \hat{y}(t) = \hat{C}x(t) \end{cases} \tag{2} \]

where

- \( r \ll d \)
- \( \hat{A} \in \mathbb{R}^{r \times r}, \hat{B} \in \mathbb{R}^r, \hat{C} \in \mathbb{C}^{1 \times r} \)

and \( \hat{A} \) has eigenvalues only in the left half plane.
The input-output map $y(u)$ of (1) is characterized by the transfer function

$$H : \mathbb{C} \rightarrow \mathbb{C}, \quad H(\omega) = C(\omega I - A)^{-1}B$$

in frequency space. $\hat{H}$ is defined accordingly for (2).
Error estimate

Let \( y(t) \) and \( \hat{y}(t) \) be the output of (1) and (2). Then the error of \( y(t) \) is bounded by

\[
\max_{t>0} |y(t) - \hat{y}(t)| \leq \|H - \hat{H}\|_{\mathcal{H}_2}\|u\|_{\mathcal{L}_2},
\]

where the \( \mathcal{H}_2 \)-norm is defined as

\[
\|H - \hat{H}\|_{\mathcal{H}_2}^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(\omega) - \hat{H}(\omega)|^2 d\omega
\]

\( \mathcal{H}_2 \) optimality

Given

- a stable dynamical system (1),
- a reduced order system (2).

If the reduced system (2) minimizes \( \|H - \hat{H}\|_{\mathcal{H}_2} \), it Hermite interpolates (1) at its mirror poles \( \sigma_1, \ldots, \sigma_r \).
Petrov-Galerkin projection

Let

- $r$ fixed, $\sigma_1, \ldots, \sigma_r$ given
- $V, W$ such that

$$
(\sigma_i I - A)^{-1} B \in \text{span}(V)
$$
$$
(\sigma_i I - A)^{-T} C^T \in \text{span}(W)
$$
$$
V^T W = I
$$

Then the reduced order model by Petrov-Galerkin projection

$$
\hat{A} = V^T AW, \quad \hat{B} = V^T B, \quad \hat{C} = CW
$$

Hermite interpolates (1) at $\sigma_1, \ldots, \sigma_r$.

Remark

- $\hat{A}$ is unique up to matrix similarity
Iterative Rational Krylov Algorithm (IRKA)

- **Problem:** Find $\sigma_1, \ldots, \sigma_r$ for (1)
- **Solution by IRKA:** Local optimum
  - Initial $\sigma_1, \ldots, \sigma_r$ given
  - Fixed-point iteration
  - Locally convergent if local optimum is attractive (e.g. for state-space-symmetric systems)
Given a compact domain $\Omega \subset \mathbb{R}^n$. Let

- $A$, $B$ and $C$ as in (1)
- $A$, $B$ and $C$ depend (smoothly) on some $p \in \Omega$

Then

- $A(p)$, $B(p)$ and $C(p)$ define a parametric LTI system

$$\Sigma : \begin{cases} \dot{x}(t) = A(p) x(t) + B(p) u(t), \\ y(t) = C(p) x(t). \end{cases}$$

- Each value of $p$ defines an LTI system, which can be reduced as before
Parametric systems

Parametric LTI system

Transfer function of a parametrized LTI system for different choices of $p$ (elastic beam):

\[ 20 \log_{10}(\|H_p(s)\|) \]

- $p=1$
- $p=0.83$
- $p=1.14$

Nils Hornung (Fraunhofer SCAI), Sara Grundel (MPI Magdeburg)
Parametric systems

Parametric model order reduction

- **Goal**: Fast computation of $\hat{A}(p), \hat{B}(p), \hat{C}(p) \forall p$
- **General ideas**:
  - Relax $\mathcal{H}_2$-optimality slightly
  - Apply well-established approximation methods
    ...such as radial basis function interpolation
- to be effective, *smoothness* is absolutely essential!
Approximation of parametric dependency

Candidates for approximation are

- $\hat{A}(p)$, $\hat{B}(p)$, $\hat{C}(p)$
Approximation of parametric dependency

Candidates for approximation are

- $\hat{A}(p), \hat{B}(p), \hat{C}(p) \Rightarrow$ non-unique, matrix similarity!
Approximation of parametric dependency

Candidates for approximation are

- $\hat{A}(p), \hat{B}(p), \hat{C}(p)$ → non-unique, matrix similarity!
- $\sigma_1(p), \ldots, \sigma_r(p)$
Approximation of parametric dependency

Candidates for approximation are

- \( \hat{A}(p), \hat{B}(p), \hat{C}(p) \) \( \sim \) non-unique, matrix similarity!
- \( \sigma_1(p), \ldots, \sigma_r(p) \) \( \sim \) eigenvalue crossings and splittings, non-smooth!

Imaginary parts of two eigenvalues of a matrix depending on two parameters:
Approximation of parametric dependency

Candidates for approximation are

\( \hat{A}(p), \hat{B}(p), \hat{C}(p) \sim \) non-unique, matrix similarity!

\( \sigma_1(p), \ldots, \sigma_r(p) \sim \) eigenvalue crossings and splittings, non-smooth!

Imaginary parts of two eigenvalues of a matrix depending on two parameters:

\[ \prod_{i=1}^{r} (s - \sigma_i(p)) \sim \text{smooth enough?} \]
Parametric systems

Smoothness of the characteristic polynomial

Let

- \( \pi \) map a matrix to its characteristic polynomial
- \( Q \) map a polynomial to its coefficients
- \( \lambda \) map a matrix to its eigenvalues

Then

\[
\hat{A}(\cdot) \in C^\infty(\mathbb{R}^n; \mathbb{R}^{r \times r}) \Rightarrow Q(\pi(\hat{A}(\cdot))) \in C^\infty(\mathbb{R}^n; \mathbb{R}^{r+1})
\]

\( \hat{A} \) stable \( \Rightarrow \)

\[
\begin{cases}
Q(\pi(\hat{A}(p))) \geq 0 \\
\Re \lambda(\hat{A}(p)) \leq 0
\end{cases}
\]

\( \forall P \in \mathbb{R}^{r \times r}, \det P \neq 0 \) : \[
\begin{cases}
\pi(\hat{A}(p)) = \pi(P \hat{A}(p) P^{-1}) \\
\lambda(\hat{A}(p)) = \lambda(P \hat{A}(p) P^{-1})
\end{cases}
\]
Smoothness of the characteristic polynomial

Let

- \( \rho \) map a set of \( r \) roots to their polynomial
- \( Q \) map the resulting polynomial to its coefficients

Then

- \( Q \) is linear, hence \( Q^{-1} \), too
- \( \rho^{-1} \) maps a polynomial to its roots
  - *closed form representations* for \( r \leq 5 \)
  - computation unstable for \( r > 5 \)
Smoothness of the characteristic polynomial

Let $r \leq 5$. Assume IRKA converges

- locally
- to a local optimum
- returns $\Sigma(p) = (\sigma_1(p), ..., \sigma_r(p))$

Moreover, assume that a perturbation of $p$ is small enough to not leave the region of

- convergence
- attraction to the local minimum

Then

- $f = Q \circ \rho \circ \Sigma(\cdot)$ is smooth
- standard $\textit{RBF interpolation}$ is applicable
Smoothness of the characteristic polynomial

Assume IRKA converges as before, \( r \leq 5 \). Moreover, assume again that a perturbation of \( p \) is small enough to not leave the region of convergence.

- convergence
- attraction to the local minimum

Let

\[
\tilde{f} \approx f = Q \circ \rho \circ \Sigma(\cdot)
\]

\[
\tilde{\Sigma} = \rho^{-1} \circ Q^{-1} \circ \tilde{f}
\]

Then \( \tilde{\Sigma} \)

- approximates the results of IRKA
- can be computed stably

\(~\) find those smooth regions!
Let $f = Q \circ \rho \circ \Sigma(\cdot)$.

- We are looking for *discontinuities* of $f(p)$
- Simple *criterion* for $k$-means or spectral clustering (Ng et al.): tuple $(p, f(p))$

~ How to determine the *number of clusters*?
Reproducing Kernels

Definition

Let

- $\Omega \subset \mathbb{R}^n$ a domain
- $F$ a class of functions $f : \Omega \to \mathbb{C}$ that form a Hilbert space $\mathcal{H}$ with inner product $(\cdot, \cdot)$

The function $\kappa : \Omega \times \Omega \to \mathbb{C}$ is called **reproducing kernel** if

$$\forall y \in \Omega : \kappa(\cdot, y) \in F,$$

$$\forall f \in F, y \in \Omega : f(y) = (f(\cdot), \kappa(\cdot, y)) \quad \text{(reproducing property)}.$$
Reproducing Kernels

Properties

Let $\xi_i \in \mathbb{C}$, $x_i, x, y, z \in \Omega$, $i, j = 1, \ldots, N, N \in \mathbb{N}$ arbitrary

- **Positive definiteness**

  $$\sum_{i,j} \xi_i \xi_j \kappa(x_j, x_i) \geq 0$$

- $\kappa(y, z) = (\kappa(x, z), \kappa(x, y))$, $\kappa(x, y) = \overline{\kappa(y, x)}$, $\kappa(x, x) \geq 0$, ...
Reproducing Kernels

Given: $\mathcal{H}(\Omega)$ with inner product $(\cdot, \cdot)$

**Existence**

Necessary and sufficient condition: A *continuous evaluation functional*

$$\delta_x : \mathcal{H} \rightarrow \mathbb{C}, \ f \rightarrow f(x)$$

exists on $\mathcal{H}$

**Uniqueness**

- Assumption: A reproducing kernel $\kappa$ exists for $\mathcal{H}$

Then the reproducing kernel $\kappa$ of $\mathcal{H}$ is *unique* and, therefore, characterizes $\mathcal{H}$. 
Given

- $\kappa : \Omega \times \Omega \to \mathbb{C}$, positive definite
- $F = \text{span}\{\kappa(\cdot, x) : x \in \Omega\}$

Moreover, define

$$(f, g)_\kappa \equiv \sum_{i,j} \alpha_i \beta_j \kappa(x_j, x_i)$$

for arbitrary $f, g \in F$ with

- $f = \sum_i \alpha_i \kappa(\cdot, x_i)$
- $g = \sum_j \beta_j \kappa(\cdot, x_j)$

Then

- $\mathcal{H} = \text{cl}F$ with respect to $\|f\|^2_\kappa \equiv (f, f)_\kappa$ has reproducing kernel $\kappa$
- $\mathcal{H}$ is called the \textit{native space} of $\kappa$
Reproducing Kernels

Examples

Let \( x, y \in \Omega = \mathbb{R}^n \).

- **Positive definite functions**
  \[
  \kappa(x, y) = \phi(x - y), \quad \text{invariant to } T(n)
  \]

- **Radial basic functions (RBF)**
  \[
  \kappa(x, y) = \phi(\|x - y\|_2), \quad \text{invariant to } SE(n)
  \]
Reproducing Kernels

RBF examples

Let $\epsilon > 0$, $\tau > n/2$. Denote by

- $K_{\nu}$ the modified Bessel function of 2nd kind,
- $\mathcal{F}f$ the Fourier transform of $f$.

Popular RBF choices are

**Sobolev splines**

$$
\phi(x) = \frac{K_{\tau-n/2}(\|x\|_2)\|x\|_2^{\tau-n/2}}{2^{\tau-1}\Gamma(\tau)}, \quad \mathcal{H} = W_2^{\tau}(\mathbb{R}^n)
$$

**Gaussians**

$$
\phi(x) = e^{-\epsilon^2\|x\|_2^2}, \quad \mathcal{H} = \left\{ f \in L^2(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n) : e^{\frac{\|\cdot\|_2^2}{8\epsilon^2}} \mathcal{F} f \in L^2(\mathbb{R}^n) \right\}
$$
Reconstruction by symmetric interpolation

RBF interpolation

Given a function \( f \in \mathcal{H} \), select

- sampling \( X = \{x_1, \cdots, x_N\} \subset \Omega, N = |X| < \infty \)
- ansatz

\[
\tilde{f}(x) = \sum_{i=1}^{N} \xi_i \kappa(x, x_i).
\]

Then \( \tilde{f} \) is an interpolant to \( f \) on \( X \) if \((\xi_1, \cdots, \xi_N)\) is a solution of

\[
\forall j = 1 \ldots N : \tilde{f}(x_j) = f(x_j).
\] (3)

\( \sim \) offline phase (sampling, IRKA) \( \leftrightarrow \) online phase (metamodel, reduced model)
Reconstruction by symmetric interpolation

Given $f, \tilde{f}, X$ as before.

**Optimality of RBF interpolation**

- $\forall \tilde{s} \in \{s \in \mathcal{H} : (3)\} : \|\tilde{f}\|_k \leq \|\tilde{s}\|_k$
- $\forall \tilde{s} \in \{\sum_i \xi_i \mathcal{K}(\cdot, x_i) : \xi_i \in \mathbb{C}\} : \|f - s\|_k \leq \|f - \tilde{s}\|_k$

Define the *fill-distance* of $X$ as

$$h \equiv \sup_{y \in \Omega} \max_{x \in X} \|x - y\|_2$$

**Sampling inequalities**

Let

- $\alpha$ a multi-index
- $\sigma$ the sampling order

Then $\exists C_1 > 0 : \|D^\alpha f\|_{L^q(\Omega)} \leq C_1 \left( h^\sigma \|f\|_k + h^{-|\alpha|} \|f(X)\|_{\ell^{\infty}(\mathbb{R}^{|X|})} \right)$

**Error estimates**

Assume a continuous embedding of $\mathcal{H}$ into $W^p_2, 0 < p < \infty$.

Then $\exists C > 0 : \|f - \tilde{f}\|_{L^q(\Omega)} \leq Ch^{p-n\max(0, \frac{1}{2} - \frac{1}{q})}\|f\|_k$
Remarks

- Gaussians, multi-quadrics: *spectral* approximation orders
- *Sobolev* functions $\leftrightarrow$ ansatz with Gaussians: polynomial approximation orders
- *Conditionally* positive functions: polynomial detrending
- Native space *norm*: indicator for problems (e.g. discontinuities)

$\leadsto$ employ *Gaussians* (or multiquadrics)
$\leadsto$ use low-order polynomial *detrending*
$\leadsto$ determine *number of clusters* by norm of the native space
Reconstruction by symmetric interpolation

Medium size model

- **Reuse results** from offline phase
- Galerkin projection for system matrices in *affine form* (medium size)
- Project *medium-size model* to $\tilde{\Sigma}$ in online phase

For details, see Sara Grundel’s talks at MoRePas II, Nonlinear MOR Workshop, and Overton’s “60th birthday” Workshop.

$\sim$ *speed-up* without additional cost
Numerical results

Examples

▶ Parametric beam model \((d = 240)\)
▶ Anemometer \((d = 29,008, n = 1 \text{ and } n = 3)\)
▶ Synthetic model (to exhibit more challenging problems)
Numerical results

Timoshenko beam

Transfer function of a parametrized LTI system for different choices of $p$:
Numerical results

Anemometer (1D)

Transfer function of a parametrized LTI system for different choices of $p$:

![Graph showing the transfer function for different values of $p$.]
Synthetic example

Transfer function of a parametrized LTI system for different choices of $p$:

\[
\frac{10^{-2}}{10^{-1}} \frac{10^0}{10^1} \frac{10^2}{10^3}
\]

\[\begin{array}{c}
\text{frequency} \\
20 \log_{10} (\|H_p(s)\|)
\end{array}\]

- Blue: 0.1629
- Green: 0.3812
- Red: 0.4254
- Cyan: 0.7827
- Purple: 0.9265
Numerical results

Error evaluation, Timoshenko beam

$H_2$ error of the reduced parametrized system using IRKA (no interpolation), IRKA with RBF (interpolation), IRKA and medium-size model with RBF – three vs. five interpolation points:
$\mathcal{H}_2$ error of the reduced parametrized system of size 4 using IRKA (no interpolation), IRKA and medium-size model with RBF – five interpolation points:
Numerical results

Error evaluation, Anemometer (3D)

$\mathcal{H}_2$ error of the reduced parametrized system using IRKA (no interpolation), IRKA and medium-size model with RBF – different reduced sized ($r$) and number of interpolation points ($N$):

<table>
<thead>
<tr>
<th></th>
<th>$r = 4, N = 5$</th>
<th>$r = 8, N = 5$</th>
<th>$r = 8, N = 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBF-IRKA</td>
<td>$3.21 \times 10^{-5}$</td>
<td>$1 \times 10^{-6}$</td>
<td>$1 \times 10^{-8}$</td>
</tr>
<tr>
<td>IRKA</td>
<td>$3.19 \times 10^{-5}$</td>
<td>$3 \times 10^{-8}$</td>
<td>$2 \times 10^{-8}$</td>
</tr>
</tbody>
</table>
Numerical results

Error evaluation, synthetic example

$H_2$ error of the reduced parametrized system using IRKA (no interpolation), IRKA with RBF – several $p$:

\[ H_2 \text{ error} \]

![Graph showing $H_2$ error comparison between RBF-IRKA and IRKA for varying parameter $p$.]
Numerical results

Clustering, synthetic example

Eigenvalues $\Sigma(p)$ of the reduced system matrix, for $r = 4$ and several $p$ (dots):
Numerical results

Clustering, synthetic example

Coefficients $f(p)$ of the corresponding characteristic polynomial, for $r = 4$ and several $p$ (colored dots), and approximation $\tilde{f}(p)$ (black line):
Summary

Parametric model order reduction

- Parametric linear time-invariant systems
- $H_2$ optimal model order reduction (IRKA)
- *RBF interpolation* of $\Sigma(p)$ using coefficients of the characteristic polynomial
- *Clustering* guided by the norm of the reproducing kernel Hilbert space innate to a radial basis
- *Medium-size model* and projection to interpolated $\Sigma(p)$
- Numerical results (synthetic as well as simple practical test problems)
Future work

Open problems

- Stable root finding (minimum polynomial?)
- Nonlinear systems (bilinear systems)
- Transfer RBF error bounds to reduced model

Thank you for your attention!