Application of low-rank approximation for nonlinear system identification

Ivan Markovsky

VRIJE UNIVERSITEIT BRUSSEL

Outline

Polynomially time-invariant model class

System identification via low-rank approximation

Solution method based on bias correction

Outline

Polynomially time-invariant model class

System identification via low-rank approximation

Solution method based on bias correction

General setup: linearly parameterized discrete-time nonlinear systems

kernel: $R\Big(\underbrace{w(t), w(t-1), \ldots, w(t-\ell)}_{x(t)}\Big) = 0$

special case: input/output NARX system

$$\mathscr{B} = \Big\{ w = \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \mid y(t) = f\big(u(t), w(t-1), \ldots, w(t-\ell)\big) \Big\}$$

linearly parameterized model $\mathscr{B}_\theta$

$$R(x) = \sum \theta_i \phi_i(x) = \theta \phi(x),$$

ϕ — model structure

θ — parameter vector

Example: single-input single-output polynomially time-invariant model

ϕ is a vector of monomials $\phi_i := x_1^{\mathrm{n}_{i1}} \cdots x_{\mathrm{n}_x}^{\mathrm{n}_{ix}}$

the structure ϕ is defined by the degrees matrix

$$\phi \;\leftrightarrow\; \left[\mathrm{n}_{ij}\right] \in \mathbb{N}^{\mathrm{n}_\phi \times \mathrm{n}_x}$$

polynomially time-invariant (PTI) model class

$$\mathscr{P}_{\ell,\mathrm{n}} := \{\, \mathscr{B}_{\boldsymbol{\theta}} \mid \boldsymbol{\theta} \in \mathbb{R}^{\mathrm{n}_\theta} \,\}, \qquad \begin{array}{lll} \ell & — & \text{lag} \\ \mathrm{n} & := & \max_{i,j}\ \mathrm{n}_{ij} \end{array}$$

Our goal is to find PTI model from data:

$$\big(w(1),\dots,w(T)\big) \mapsto \mathscr{B} \in \mathscr{P}_{\ell,\mathrm{n}}$$

1. structure selection: find ϕ

2. parameter estimation: find θ

$$\begin{array}{lll} \text{minimize} & \text{over } \theta \text{ and } \widehat{w} & \|w - \widehat{w}\| \\ \text{subject to} & \widehat{w} \in \mathscr{B}_\theta & \end{array} \qquad \text{(NL SYSID)}$$

Outline

Polynomially time-invariant model class

System identification via low-rank approximation

Solution method based on bias correction

Link to low-rank approximation

$$w \in \mathscr{B}_{\boldsymbol{\theta}}$$

$$\Updownarrow$$

$$R\big(x(t)\big) = \theta^\top \phi\big(x(t)\big) = 0, \quad \text{for } t = 1, \ldots, T - \ell$$

$$\Updownarrow$$

$$\theta^\top \begin{bmatrix} \phi\big(x(1)\big) & \cdots & \phi\big(x(T-\ell)\big) \end{bmatrix} = 0$$

$$\Updownarrow$$

$$\operatorname{rank}\Big(\Phi(w)\Big) \leq n_\phi - 1$$

(NL SYSID) $\iff$ low-rank approximation

$$\begin{array}{lll} \text{minimize} & \text{over } \widehat{w} & \|w - \widehat{w}\|_2 \\ \text{subject to} & \text{rank}\big(\Phi(\widehat{w})\big) \leq n_\phi - 1 & \end{array} \qquad \text{(SLRA)}$$

non-convex optimization problem

there are no efficient solution methods

heuristic method: ignore the structure of $\Phi(\widehat{w})$

$$\text{minimize} \quad \text{over } \theta \neq 0 \quad \|\theta^\top \Phi(w)\|_2 \qquad \text{(LRA)}$$

Structure selection via sparsity regularization

select "large" model class $\mathscr{P}_{\ell,\mathrm{n}}$ and impose sparsity on θ

$$\text{minimize} \quad \text{over } \theta \neq 0 \quad \|\theta^\top \Phi(w)\|_2 + \gamma \|\theta\|_1$$

γ controls the sparsity level

- $\gamma = 0 \quad \rightsquigarrow \quad$ (LRA) $\quad \rightsquigarrow \quad$ full θ
- $\gamma \to \infty \quad \rightsquigarrow \quad \theta \to 0$

selected, so that # nonzero elements $=$ given number

Outline

Polynomially time-invariant model class

System identification via low-rank approximation

Solution method based on bias correction

Unstructured LRA is biased

the heuristic method ignoring the structure of $\Phi(\widehat{w})$

$$\text{minimize} \quad \text{over } \theta \neq 0 \quad \|\theta^\top \Phi(w)\|_2$$

is easy to compute, but biased ($\mathbf{E}(\theta) \neq \bar{\theta}$) in the EIV setup

$$w = \bar{w} + \widetilde{w}, \quad \text{where} \quad \bar{w} \in \bar{\mathscr{B}} \quad \text{and} \quad \widetilde{w} \sim \mathrm{N}(0, \sigma^2 I)$$

The bias can be corrected for

define $\quad \Psi := \Phi(w)\Phi^\top(w) \quad$ and $\quad \bar{\Psi} := \Phi(\bar{w})\Phi^\top(\bar{w})$

goal: construct "corrected" matrix Ψ_c, such that

$$\mathbf{E}(\Psi_c) = \bar{\Psi}$$

then solve

$$\text{minimize} \quad \text{over } \theta \neq 0 \quad \|\theta^\top \Psi_c(w)\|_2$$

Derivation of the correction

Hermite polynomials $h_k(x)$ have the property

$$\mathbf{E}\left(h_k(\bar{x}+\widetilde{x})\right) = \bar{x}^k, \quad \text{where} \quad \widetilde{x} \sim \mathrm{N}(0,\sigma^2) \qquad (*)$$

with $w = (u,y)$, the (i,j)th element of $\Psi = \Phi\Phi^\top$ is

$$\sum (\bar{u}+\widetilde{u})^{n_{u,i}+n_{u,j}} (\bar{y}+\widetilde{y})^{n_{y,i}+n_{y,j}}$$

then, by $(*)$

$$\phi_{\mathrm{c},ij} := \sum h_{n_{u,i}+n_{u,j}}(u) h_{n_{y,i}+n_{y,j}}(y)$$

has the desired property

$$\mathbf{E}(\psi_{\mathrm{c},ij}) = \sum \bar{u}^{n_{u,i}+n_{u,j}} \bar{y}^{n_{y,i}+n_{y,j}} =: \bar{\psi}_{ij}$$

Unbiased estimator

the corrected Ψ_c is an even polynomial in σ

$$\Psi_c(\sigma^2) = \Psi_{c,0} + \sigma^2 \Psi_{c,1} + \cdots + \sigma^{2n_\psi} \Psi_{c,n_\psi}$$

estimate: $\Psi_c(\sigma^2)\theta = 0$

computing simultaneously σ and θ is polynomial EVP

Summary

Polynomially time-invariant model class

System identification via low-rank approximation

Solution method based on bias correction