Structure Detection of Nonlinear Aeroelastic Systems with Application to Aeroelastic Flight Test Data: Part II

Sunil L. Kukreja and Martin J. Brenner

NASA Dryden Flight Research Center, Structural Dynamics Group, Edwards, California, USA
Outline

- Motivation
- Nonlinear Model Form
- Structure Detection
- Least Absolute Shrinkage and Selection Operator (LASSO)
- Objectives
- Results
  - Assess LASSO as a Structure Detection Tool: Simulated Nonlinear Models
  - Applicability to Complex Systems: F/A-18 Active Aeroelastic Wing Flight Test Data
- Conclusions
Motivation

- Parsimonious system description
- Black-box model
- Efficient control strategies
- Insight into functionality of system
Nonlinear Model Form

- Linear statistical model
  \[ z(n) = \sum_{j=1}^{p} \theta_j f(\varphi_j(n)) + e(n) \]
  - \( z \): observed system output
  - \( \theta_j \): unknown system parameter
  - \( \varphi_j \): regressor
  - \( e \): independent Gaussian variable, zero-mean, constant variance \( \sigma^2 \)
  - \( f \): nonlinear mapping

- Let \( \varphi \) be described as:
  \[ \varphi(n) = [1, z(n-1), \ldots, z(n-n_z), u(n), \ldots, u(n-n_u), e(n-1), \ldots, e(n-n_e)]^T \]
  - Special case \( f \) polynomial: \( u^2(n-3), u(n)u(n-1), z(n-1)z(n-2), u^2(n-1)z(n-2) \)
  - General case \( f \): wide variety of nonlinear functions such as a sigmoid
  - NARMAX
Structure Detection

- NARMAX models described by few terms

- Maximum number of candidate terms:

\[
P = \sum_{k=1}^{l} p_k + 1
\]

\[
p_k = \frac{p_{k-1}(n_z + n_u + n_e + k)}{k}, \quad p_0 = 1
\]

- Example: model of order: \( O = [4 \ 4 \ 4 \ 2] \Rightarrow p = 105 \) candidate terms
- The curse of dimensionality!

- Often leads to computationally intractable combinatorial optimisation problems
Several Fundamental Approaches to Structure Detection

- Exhaustive search
  - Every possible subset of the full model is considered
  - Requires large number of computations

- Covariance matrix, $P_\theta$
  - Based on input-output data and estimated residuals to assess parameter relevance
  - Parameter variance estimates often inaccurate when the number of candidate terms large

- Bootstrap method
  - Numerical procedure for estimating parameter statistics
  - For convergence: number of data points needed \textit{at least} 10 times square of initial number of candidate terms
• Least absolute shrinkage and selection operator (LASSO)

\[
\min_\theta \frac{1}{2} \|(Z - \Phi \theta)\|_2^2 + \lambda \|\theta\|_1
\]

- Least-squares like problem: addition of \(\ell_1\) penalty on parameter vector

• LASSO shrinks least-squares estimator towards 0, potentially sets \(\theta_j = 0\) for some \(j\)

• Regularisation parameter \(\mathbb{R} \ni \lambda = [\lambda_{min}, \ldots, \lambda_{max}]\) controls the trade-off between approximation error and sparseness

• LASSO behaves as a structure selection instrument
**Assumption 1.** Input signal is persistently exciting.

**Theorem 1.** If the excitation signal is persistently exciting, LASSO will have a unique optimum.

**Proof.**

(i) Since the excitation signal is persistently exciting implies $\Phi^T\Phi$ is positive definite

(ii) As a result the first term of

$$\min_{\theta} \frac{1}{2} \|(Z - \Phi\theta)\|^2_2 + \lambda \|\theta\|_1$$

is a strictly convex function.

(iii) Since the second term is convex, it follows that the sum is strictly convex and a unique optimiser is guaranteed.

$\square$
**Assumption 2.** Optimal regularisation parameter, $\lambda^*$, is known.

**Theorem 2.** If the excitation signal is persistently exciting and has a unique optimum, LASSO will converge to a unique global minimum.

*Proof.* Since

$$\min_{\theta} \frac{1}{2} \| (Z - \Phi \theta) \|_2^2 + \lambda \| \theta \|_1$$

is strictly a convex optimisation problem the solution will converge to a unique global minimum. \(\Box\)
Solution of LASSO

- Quadratic programming framework with slack variables

$$\min_x \frac{1}{2} x^T M x + c^T x \quad \text{such that } x_k \geq 0,$$

and where,

$$M = \begin{bmatrix} \Phi^T \Phi & -\Phi^T \Phi \\ -\Phi^T \Phi & \Phi^T \Phi \end{bmatrix}, \ c = \lambda 1 - \begin{bmatrix} \Phi^T Z \\ -\Phi^T Z \end{bmatrix}, \ x = \begin{bmatrix} \theta^+ \\ \theta^- \end{bmatrix}$$

- Model parameters: $$\theta = \theta^+ - \theta^-$$

- QP problem readily solved using standard optimisers

- Given suitable $$\lambda$$ general structure computation problem can be solved
Selection of Regularisation Parameter: $\lambda$

- Method of cross-validation to estimate prediction error

$$PE(\lambda) = E \left[ Z - \Phi \theta \right]^2$$

- Determined by numerically minimising the cross-validation error across a discrete set of logarithmically spaced $\lambda$ values

$$10^{\lambda_{min}} \leq \lambda \leq 10^{\lambda_{max}}$$

- Regularisation parameter, $\lambda$, is chosen to minimise this estimate
Ideal Selection of $\lambda$
Local Minima

Practical $\lambda$ Selection Problem

$10^{-4}$ $10^{0}$

$10^{0.155}$ $10^{0.154}$ $10^{0.153}$ $10^{0.152}$ $10^{0.151}$

$\lambda$

ICAHA 2006 – Beijing, China

Dryden Flight Research Center
Objectives

- Investigate LASSO as a structure detection tool
- Hypothesise useful for structure detection
- Performance evaluation
Simulated System

\[ z(n) = 0.4[u(n - 1) + u(n - 1)^2 + u(n - 1)^3] + 0.8z(n - 1) - 0.8e(n - 1) + e(n) \]

- Model order known: \( O = [1 1 1 3] \)
  - 35 candidate terms
  
  - True system has only 5 true terms
Simulations

• One thousand Monte-Carlo simulations
  – Input white, uniform distribution
  – Each output realisation had unique Gaussian distributed, white, zero-mean, noise sequence added

• Noise amplitude increased 5 dB increments, from 20 to 0 dB SNR

• $N_e = 667$ points for estimation and $N_v = 333$ for validation

• Regularisation parameter 1,000 logarithmically spaced $\lambda$’s: $10^{-10} \leq \lambda \leq 10^{1.5}$

• Compare LASSO with covariance matrix, $P_\theta$ approach
  – Parameters tested for significance at 95% confidence-level
Results Classified into Three Categories

1. Exact Model: A model which contains only true system terms.

2. Over-modelled: A model with all its true system terms plus spurious parameters and

3. Under-modelled: A model without all its true system terms. An under-modelled model may contain spurious terms as well.
Results: Selection Rate

LASSO Selection Rate

- Exact
- Over
- Under

SNR

P_θ Selection Rate

- Exact
- Over
- Under

SNR

ICAHA 2006 – Beijing, China

Dryden Flight Research Center
Results: Spurious Term Selection Rate

LASSO Spurious Terms: Over-Modelled Model

Mean vs. SNR graph showing the spurious term selection rate for the LASSO method in an over-modelled model.
• Assumed model order: \( O = [4 \ 4 \ 4 \ 3] \):
  
  – Fourth-order dynamics selected because many aeroelastic structures are well defined by a fourth-order LTI system
  – Third-order nonlinearity selected because models of higher nonlinear order can often be decomposed to second or third-order
  – Full model description 560 candidate terms

• 1,000 logarithmically spaced \( \lambda \)'s: \( \lambda_{\text{min}} = -10 \) and \( \lambda_{\text{max}} = 1.0 \)

• Estimation \( N_e = 5,200 \): right wing & cross-validation \( N_v = 5,200 \): left wing
Identification Data

Collective Aileron Position

Structural Accelerometer Response (Right Wing)
Computed Structure

- Contains 25 terms

\[ z(n) = \hat{\theta}_0 + \hat{\theta}_1 u(n - 1) + \hat{\theta}_2 u(n - 2) + \hat{\theta}_3 u(n - 4) \]
\[ + \hat{\theta}_4 u^2(n - 1) + \hat{\theta}_5 u^2(n - 2) + \hat{\theta}_6 u^2(n - 4) \]
\[ + \hat{\theta}_7 z(n - 1) + \hat{\theta}_8 z(n - 4) + \hat{\theta}_9 u^2(n - 1) z(n - 4) \]
\[ + \hat{\theta}_{10} u^2(n - 2) z(n - 1) + \hat{\theta}_{11} u^2(n - 4) z(n - 4) \]
\[ + \hat{\theta}_{12} z^3(n - 1) + \hat{\theta}_{13} z^3(n - 4) + \hat{\theta}_{14} \hat{\epsilon}(n - 1) \]
\[ + \hat{\theta}_{15} \hat{\epsilon}(n - 4) + \hat{\theta}_{16} u^2(n - 1) \hat{\epsilon}(n - 4) \]
\[ + \hat{\theta}_{17} u^2(n - 2) \hat{\epsilon}(n - 1) + \hat{\theta}_{18} u^2(n - 4) \hat{\epsilon}(n - 4) \]
\[ + \hat{\theta}_{19} z^2(n - 1) \hat{\epsilon}(n - 1) + \hat{\theta}_{20} z(n - 1) \hat{\epsilon}^2(n - 1) \]
\[ + \hat{\theta}_{21} \hat{\epsilon}^3(n - 1) + \hat{\theta}_{22} z^2(n - 4) \hat{\epsilon}(n - 4) \]
\[ + \hat{\theta}_{23} z(n - 4) \hat{\epsilon}^2(n - 4) + \hat{\theta}_{24} \hat{\epsilon}^3(n - 4) \].
Conclusions

- Novel approach for detecting the structure of highly over-parameterised nonlinear models in situations where other methods may be inadequate.

- Practical significance in the analysis of aircraft dynamics during envelope expansion and could lead to more efficient control strategies.

- Could allow greater insight into the functionality of various systems dynamics, by providing a quantitative model which is easily interpretable.
Acknowledgments

Supported by the National Academy of Sciences and the National Aeronautics and Space Administration (NASA), Dryden Flight Research Center, Aerostructures Branch (Grant Number: NASA-NASW-99027).