Project Background

- There is some error in Revolving Needle Driver (RND) robot
- Actual kinematic description does not match idealized version
- Reality may be complicated
- Must determine a more accurate model based on behavior
- Must do Systems Identification
  - White-box
  - Grey-box
  - Black-box
Paper selection and why?

• *Nonlinear black-box modeling in system identification - a unified overview*
• Jonas Sjöberg et al.
• Great starting point for those trying to identify parameters in unknown systems
• Broadly applicable
• Written for the user
Necessary background

- Systems ID can't be 'formalized, automated'
- Methods can be novel
Background continued

- Systems ID finds a relationship between input and output
- Must choose basis functions and parameters
What the authors did

- Summarize the field of systems ID with a common approach
- Explain major choices systems identifiers will face
- Provide examples which help answer structural issues (of function choice)
General Approach

Inputs and Outputs

\[ u' = [u(1) \ u(2) \ \ldots \ u(t)], \]
\[ y' = [y(1) \ y(2) \ \ldots \ y(t)]. \]

Looking for relationship

\[ y(t) = g(u^{t-1}, y^{t-1}) + v(t). \]
Approach continued

Relationship can be decomposed into parameters and function of past I/O

\[ g(u(t-1), y(t-1), \theta) = g(\varphi(t), \theta), \]

where

\[ \varphi(t) = \varphi(u(t-1), y(t-1)). \]

Function of past I/O pairs is regression vector
Approach continued

The mapping

\[ g(\varphi, \theta), \]

which for any given \( \theta \) goes from \( \mathbb{R}^d \) to \( \mathbb{R}^p \).

Can be written as a sum (family) of basis functions

\[ g(\varphi, \theta) = \sum \alpha_k g_k(\varphi). \]
Approach continued

Basis functions can be constructed from single 'mother basis function'

\[ g_k(\varphi) = \kappa(\varphi, \beta_k, \gamma_k) \quad \kappa(\beta_k(\varphi - \gamma_k)) \cdot \]

With scaling (directional) and offset terms determining region of support

Ex: \( k(x) = \cos(x) \) : Basis functions = Fourier
Basis function construction

Radial Construction

\[ g_k(\varphi) = g_k(\varphi, \beta_k, \gamma_k) = \kappa(\|\varphi - \gamma_k\|_{\beta_k}), \]

Support diminishes by scale factor with distance from offset

The most homogeneous choice
Basis function construction (cntd)

Ridge Construction

\[ g_k(\varphi) = g_k(\varphi, \beta_k, \gamma_k) = \kappa(\beta_k^T \varphi + \gamma_k), \quad \varphi \in \mathbb{R}^d. \]

Unbounded support in subspace along 'ridge' of scaling (direction) function

- These basis functions can yield recognizable structures
Basis function construction (cntd)

Tensor Product of $d$ (dimension) basis functions

\[ g_1(\varphi_1) \cdots g_d(\varphi_d). \]

Can behave very differently in arbitrary directions

Computationally expensive for high-dimension case
Model Quality

True model

\[ y(t) = g_0(\varphi(t)) + e(t) \]

Quality

\[ \bar{V}(\theta) = E \| y(t) - g(\varphi(t), \theta) \|^2 \]
\[ = \lambda + E \| g_0(\varphi(t)) - g(\varphi(t), \theta) \|^2 \]
Model Quality and Variance

Best fit (#parameters=m) minimizes variance

\[ \theta_*(m) = \arg\min_{\theta} \bar{V}(\theta) \]

Quality of a specific model with N I/O pairs

\[ E\bar{V}(\hat{\theta}_N) = V_*(m). \]
Bias and Variance

Decompose deviation from true model

\[ V_*(m) = EV(\hat{\theta}_N) \]
\[ = \lambda + E \| g_0(\varphi(t)) - g(\varphi(t), \hat{\theta}_N) \|^2 \]
\[ \approx \lambda + E \| g_0(\varphi) - g(\varphi, \theta_*(m)) \|^2 \]
\[ + E \| g(\varphi, \theta_*(m)) - g(\varphi, \hat{\theta}_N) \|^2. \]
What will help with bias?

Within a known model family:
- Increasing number of parameters will give better overall model quality
- Increasing number of basis functions will reduce variance

\[ \hat{\theta}_N \rightarrow \theta_*(m) \]

How will this affect variance in unknown model families with novel mother basis functions?
What will help with variance?

Variance is a function of # of parameters and # of regressor-output pairs (w/ error variance)

\[ E \| g(\varphi(t), \hat{\theta}_N) - g(\varphi(t), \theta_*(m)) \|^2 \approx \lambda \frac{m}{N} \]

Giving model quality succinctly

\[ V_*(m) = E \bar{V}(\hat{\theta}_N) = \lambda + \lambda \frac{m}{N} \]

\[ + E \| g_0(\varphi) - g(\varphi, \theta_*(m)) \|^2 \]

\[ = \bar{V}(\theta_*(m)) + \lambda \frac{m}{N} . \]
Advice

- Look at data
- Try and find physically intuitive explanation
- Pick efficient basis functions
- Do not add parameters 'spuriously'
- Do not add basis functions unless you are confident about the mother basis function
- Remember tradeoff between bias and variance
Significance of key result

- No new 'results'
- Significant due to comprehensive nature of paper - all described from a common framework
- Has great benefit for researchers in an array of fields - user focused
My assessment

- Important as a blueprint for Systems ID
- Only most general abstractions reviewed are applicable to RCM link parameter estimation

- Good
  - Broad
  - Written from users perspective

- Bad
  - Not all applicable
  - Desire to be broad leaves out many specific mathematical processes
Applicability

Suppose needle tip position in base coordinates is a function of 3 parameters: $\Theta_1$, $\Theta_3+\gamma$ (offset), $L$

- Rotation about Z
  \[ R_{b1} = \begin{pmatrix} \cos[\Theta_1] & -\sin[\Theta_1] & 0 \\ \sin[\Theta_1] & \cos[\Theta_1] & 0 \\ 0 & 0 & 1 \end{pmatrix} \]

- Rotation about X
  \[ R_{2t} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos[\Theta_3 + \gamma] & -\sin[\Theta_3 + \gamma] \\ 0 & \sin[\Theta_3 + \gamma] & \cos[\Theta_3 + \gamma] \end{pmatrix} \]

- Idealized model: $[xyz] = R_{b1} * R_{2t} *[0;-L;0]$
Accounting for additional parameter

- Rotation about Y
  
  \[
  R_{12} = \begin{pmatrix}
  \cos(\beta) & 0 & \sin(\beta) \\
  0 & 1 & 0 \\
  -\sin(\beta) & 0 & \cos(\beta)
  \end{pmatrix}
  \]

- Resulting in a new model: \([xyz]\) = \(R_{b1} \times R_{12} \times R_{2t} \times [0; -L; 0]\)
Next steps

- Future research in Systems ID should focus on expanding catalogue of basis functions
- This will move more problems from black-box to grey-box, white-box
- Future reading for our project should focus on grey-box systems identification
Conclusions

- Great paper!
- Very useful to many people
- Needs some interpretation to apply
- Desire to be broad excludes specific path
  - Kind of like variance vs bias

so,

- Questions?