

#### Integration of Physical Knowledge in Empirical Models – A New Approach to Regression Analysis

**G. Kronberger,** C. Haider, M. Kommenda, (University of Applied Sciences Upper Austria) S. Scheidel (AVL List GmbH), M. Kordon (AVL Commercial Driveline & Tractor Engineering GmbH)

12. Nov. 2019

# Summary

We include prior knowledge into empirical models by enforcing monotonicity over selected inputs

Method: Shape-constrained polynomial regression Application: modelling diesel engine emission Results:

- Model conforms to prior knowledge
- Extrapolation is plausible
- Slightly larger fitting errors





# Background

Empirical models and model-based optimization are widely known and applied in engineering.

Facts:

- No analytical model required
- Relatively low computational effort (compared to CFD, FEM)
- Large set of methods from statistical learning
- Broadly applicable
- Data availability and quality is essential
- Extrapolation is (almost) impossible or risky



### Example: friction systems





# A simple example

Empirical model for the maximum temperature of a disc brake





# Extrapolating...

The prediction does not conform to physical intuition.





#### How can we enforce monotonicity?





### Another example

Empirical model for diesel engine emissions





### A potential cause for the problem





### Further investigation





# Objectives

- Integrate vague knowledge into empirical models.
- The additional effort should be minimal.
- Better or equal model accuracy
- Low computational overhead
- Applicable to many methods for empirical modelling:
  - Polynomial regression, artificial neural networks, kernel regression, symbolic regression, ...



#### The concept of shape-constrained regression

 $f^* = \operatorname*{argmin}_{f \in \mathcal{F}} \mathcal{L}(f, X, \mathbf{y})$ 

 $\mathcal{L}(f, X, y)$  is the loss function (e.g. sum of squared errors)

 $\mathcal{F}$  is a model class e.g.:

- polynomials of given degree
- neural network architecture



#### The concept of shape-constrained regression

 $f^* = \operatorname*{argmin}_{f \in \mathcal{F}} \mathcal{L}(f, X, y)$ 

s.t.:  

$$l_f \qquad f(x_f) \qquad u_f$$
  
 $l_{Jac} \leq \nabla f(x_{Jac}) \leq u_{Jac}$   
 $l_{Hess} \qquad \nabla^2 f(x_{Hess}) \qquad u_{Hess}$ 

 $\mathcal{L}(f, X, y)$  is the loss function (e.g. sum of squared errors)

 $\mathcal{F}$  is a model class e.g.:

- polynomials of given degree
- neural network architecture

 $\nabla f$  is the vector of partial derivatives of f over all inputs

*f* must be differentiable



#### The concept of shape-constrained regression

 $f^* = \operatorname*{argmin}_{f \in \mathcal{F}} \mathcal{L}(f, X, y)$ 

s.t.:  

$$l_{f} \qquad f(x_{f}) \qquad u_{f}$$

$$l_{Jac} \leq \nabla f(x_{Jac}) \leq u_{Jac}$$

$$l_{Hess} \qquad \nabla^{2} f(x_{Hess}) \qquad u_{Hess}$$

$$\forall x_{f}, x_{Jac}, x_{Hess} \in \mathbb{R}^{d},$$

$$l_{x_{f}} \qquad x_{f} \qquad u_{x_{f}}$$

$$l_{x_{Jac}} \leq x_{Jac} \leq u_{x_{Jac}}$$

$$l_{x_{Hess}} \qquad u_{x_{Hess}} \qquad u_{x_{Hess}}$$

 $\mathcal{L}(f, X, y)$  is the loss function (e.g. sum of squared errors)

 $\mathcal{F}$  is a model class e.g.:

- polynomials of given degree
- neural network architecture

 $\nabla f$  is the vector of partial derivatives of f over all inputs

*f* must be differentiable



# The optimistic and pessimistic approach

#### Optimistic

- Evaluate constraint on finite sample from input space
- Easy to implement
- Suffers from curse of dimensionality
- Result can violate the constraints



#### Pessimistic

- Approximate constraint for the infinite set
- Feasible for certain model classes
- Result conforms to constraints
- Potentially rejects optimal solution





A polynomial is positive if it can be expressed as a sum of squares.

 $P(\boldsymbol{x},\boldsymbol{\theta}) = T_1(\boldsymbol{x},\boldsymbol{\theta}_1)^2 + \dots + T_n(\boldsymbol{x},\boldsymbol{\theta}_n)^2 \Rightarrow P(\boldsymbol{x}) \ge 0$ 

Derivatives of polynomials are polynomials.

 $\rightarrow$  Shape-constrained polynomial regression is a quadratic optimization problem with sum-of-squares constraints.



Toy-Problem:





Fitting a polynomial with degree 4:





Fitting a polynomial with degree 9 and constraints:





# Application to Emission Models



# Methodology

- 1. Gather data from test bench using DOE
  - 11 parameters
  - 3 targets
  - 1500 observations
- 2. Find and fit polynomial model
  - 3rd degree
  - two-way interactions
  - approx. 50 terms
- **3**. Add monotonicity constraint  $\frac{\partial}{\partial X_{11}} P(x) \ge 0$
- 4. Fit parameters with constraints

	30	X2	X0	ж	XS	35	30	X	X0	XD	3011	495,71	apt_12	Cf_tga
×		0.14 0.60 0.30 0.30 0.30 0.30 0.30 0.30 0.30		0.66 0.62 0.22 0.02 0.02 0.02 0.4 16 0.5 1		0.84 0.66 0.22 0.36 0.2 0.2 84 85 0.5 1	0.52 0.52 0.52 0.52 0.52 0.52 0.52 0.52	0.52 0.54 0.55 0.56 0.55 0.56 0.56 0.56 0.56 0.56	0.34 0.35 0.35 0.35 0.35 0.35 0.35 0.35 0.35	0.84 0.55 0.55 0.56 0.55 0.44 0.44 0.44 0.72 7		1 0.84 0.82 0.25 0.26 0.26 0.26 0.26 0.26 0.26 0.26 0.26	0.54 0.52 0.36 0.2 0.2 0.2 0.2 0.2 0.2 0.2 04 08 08 7	0.54 0.55 0.55 0.55 0.25 0.02 0.4 0.5 0.5
X	0,74 0,48 0,22 -0,04 -0,3 -0,05 0,22 0,52 0,52 0,55 0,55 0,55 0,55		0.74 0.44 0.22 0.94 0.3 0.22 0.52 0.52 0.54 0.44 0.3 0.55 0.56 0.58 1	0.74 0.48 0.22 0.04 0.3 0 0.2 84 08 03 1	0,74 0,40 0,22 0,84 0,30 0 0,2 84 88 0,3 1	0.74 0.46 0.25 0.05 0.02 0.02 0.02 0.01 0.03 0.03 0.03 0.03 0.03 0.03 0.03	0.34 0.22 0.04 0.22 0.04 0.34 0.34 0.34 0.34 0.34 0.34 0.34		0.22 0.22 -0.24 -0.34 -0.46 -0.46 -0.40 -0.22 -0.64 -0.40 -0	0.14 0.22 0.22 0.24 0.23 0.44 0.72 1	0.74 0.40 0.22 0.04 0.25 0.5 05 07 03 09 1			1 0,74 0,44 0,22 0,02 0,02 0,02 0,02 0,04 0,05 0,05
R	0.84 0.65 0.36 0.2 0.2 0.2 0.2 0.2 0.2 0.52 0.53 0.54 0.55 0.55 0.55 0.55 0.55 0.55 0.55	0.64 0.52 0.52 0.24 0.52 0.24 0.52 0.24 0.52 0.25 0.25 0.25 0.25 0.25 0.25 0.25	100 100 200 200 200 200 200 200 200 200	0.84 0.68 0.36 0.36 0.36 0.36 0.36 0.36 0.36 0.36	0.84 0.68 0.52 0.36 0.2 0.02 04 08 03 1	1 0.84 0.66 0.35 0.2 0 0.2 0.4 0.5 0.8 1	1 0.84 0.52 0.32 0.22 0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2 0	0.54 0.55 0.55 0.55 0.55 0.55 0.55 0.55	1 0.84 0.82 0.32 0.32 0.32 0.32 0.32 0.32 0.32 0.3	0.54 0.58 0.58 0.58 0.58 0.58 0.58 0.58 0.58	1 0.54 0.52 0.52 0.52 0.52 0.52 0.52 0.52 0.52		1 0.84 0.83 0.83 0.93 0.93 0.93 0.93 0.93 0.93 0.93 0.9	0.34 0.52 0.36 0.2 0 02 04 03 03
ж		0.8 0.6 0.2 0.3 0.3 0.2 0.3 0.22 0.3 0.22 0.3 0.22 0.3 0.22 0.3 0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4					0.8 0.4 0.2 0.2 0.2 0.2 0.2 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5		0.8 0.4 0.2 0.1 0.1 0.2 0.1 0.20 0.1 0.20 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.	0.8 0.4 0.2 0.3 0.50 0.06 0.4 0.72 1	0.8 0.4 0.2 0.5 04 07 08 05 1		0.8 0.6 0.4 0.2 0 0.2 54 06 58 1	
2	1 0.8 0.4 0.2 0.2 0.2 0.2 0.35 0.84 0.85 0.84 0.35 0.85 7										0.8 0.4 0.2 0.5 05 67 08 09 1	0.8 0.4 0.2 0 02 04 05 08 1		
86														
R	0.54 0.52 0.2 02 02 02 052 054 030 052 054 030	1 0.84 0.85 0.2 0.3 0.2 0.3 0.2 0.3 0.2 0.3 0.2 0.3 0.2 0.3 0.2 0.3 0.2 0.3 0.2 0.3 0.3 0.3 0.3 0.3 0.3 0.3 0.3	0.84 0.84 0.35 0.2 0.2 0.2 0.2 0.2 0.55 0.55 0.54 0.36 0.55 0.54 0.55 0.54 0.55 0.55 0.55 0.54 0.55 0.55	0.54 0.68 0.52 0.52 0.2 0.02 0.4 60 0.4	0,84 0,85 0,12 0,25 0,2 0,2 0,2 8,4 8,6 0,8 1				0.84 0.85 0.85 0.80 0.80 0.80 0.80 0.80 0.80	0.84 0.52 0.36 0.32 0.36 0.32 0.36 0.44 0.72 1	0.88 0.88 0.32 0.35 0.35 0.35 0.35 0.35 0.35 0.35 0.35	0.34 0.68 0.35 0.35 0.35 0.35 0.35 0.02 0.4 0.4 0.8 1	0.54 0.52 0.35 0.25 0.25 0.25 0.22 0.4 0.6 0.8 1	0,34 0,683 0,522 0,32 0,32 0,32 0,4 0,6 0,8
RX.	0.82 0.46 0.46 0.28 0.1 0.2 0.25 0.25 0.28 0.28 0.27 0.27 0.27 0.82 0.84 0.28 0.28 0.28 0.28 0.28 0.28 0.48 0.48 0.48 0.48 0.48 0.48 0.48 0.4	1 0.42 0.44 0.45 0.1 0.3 0.1 0.3 0.22 0.74 0.46 1 0.45 1 0.42 0.74 0.45 0.1 0.45 0.1 0.45 0.1 0.45 0.1 0.45 0.1 0.45 0.1 0.45 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1	1 0.42 0.44 0.45 0.33 0.1 0.2 0.52 0.54 0.52 0.54 1	0.82 0.46 0.46 0.1 0.02 04 85 03 1	0.52 0.46 0.28 0.7 0.02 0.4 05 0.5 1					0.80 0.80 0.86 0.86 0.38 0.3 0.3 0.3 0.3 0.3 0.52 0.85 0.44 0.72 1	0.82 0.54 0.28 0.28 0.28 0.28 0.28 0.28 0.28 0.28		0.82 0.54 0.48 0.28 0.28 0.28 0.28 0.28 0.28 0.20 0.4 0.6 0.8 1	0.82 0.46 0.23 0.1 0 0.2 0.4 0.6 0.8
6X	0.82 0.46 0.28 0.20 0.2 0.25 0.55 0.80 0.80 0.80 0.80 0.80 0.80	1 0.82 0.44 0.28 0.1 0.3 0.22 0.74 0.34 0.22 0.74 0.46 1	1 6.82 6.44 6.44 6.44 6.44 6.23 0.1 0.2 0.52 0.54 0.36 0.52 0.34 1	0.12 0.64 0.45 0.1 0.1 0.0 0.2 0.4 0.9 0.4 1	1 0.12 0.44 0.45 0.1 0 0.2 0.4 0.9 0.8 7		0.82 0.84 0.44 0.44 0.25 0.25 0.25 0.25 0.35 0.35 0.35 0.35 0.35 0.35 0.35 0.3				0.82 0.64 0.85 0.20 0.20 0.20 0.20 0.5 0.8 0.7 0.8 0.9 1	0.82 0.64 0.83 0.20 0.20 0.20 0.20 0.22 0.4 0.6 0.8 1	0.22 0.64 0.28 0.10 0.20 0.10 0.22 0.4 0.6 0.8 1	0.82 0.46 0.45 0.1 0.1 0.02 0.4 0.6 0.8
800	0.50 0.20 0.20 0.20 0.20 0.20 0.20 0.20	1 0.86 0.72 0.44 0.3 0.3 0.3 0.3 0.3 0.32 0.74 0.34 0.34 0.34 0.44 0.3 0.32 0.74 0.45 0.34 0.3	0.85 0.27 0.44 0.3 0.2 0.52 0.34 0.35 0.65 1		0.56 0.77 0.58 0.44 0.3 0 0.2 0.4 0.5 0.8 1	0.86 0.72 0.64 0.44 0.3 0 82 84 85 0.5 1	0.86 0.72 0.84 0.44 0.25 0.24 0.25 0.25 0.25 0.25 0.25 0.25 0.25 0.25	0.086 0.775 0.44 0.3 0.1 6.44 0.02 0.28 0.64 1		000 000 000 000 000 000 000 000 000 00	0.86 0.72 0.46 0.46 0.46 0.46 0.46 0.46 0.46 0.46		0.36 0.72 0.84 0.44 0.20 0.44 0.05 0.05 1	
XII	0.3 0.7 0.5 0.2 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5		0.9 0.9 0.7 0.8 0.5 0.2 0.8 0.5 0.2 0.5 0.2 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5	0.5 0.2 0.5 0.2 0.4 0.5 0.5 0.4 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5		1 0.9 0.0 0.7 0.5 0.5 0.5 0.2 0.4 0.6 0.3 1	0.9 0.7 0.5 0.5 0.2 0.2 0.2 0.2 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5	0.9 0.9 0.9 0.9 0.9 0.9 0.9 0.9 0.9 0.9	0.9 0.7 0.8 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5	0.9 0.9 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5			1 0.9 0.8 0.7 0.6 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5	03 03 07 05 0 82 04 66 03
April 71	0.8 0.4 0.2 0.2 0.20 0.55 0.04 0.000 0.000 1	0.8 0.4 0.2 0.4 0.3 0.5 0.4 0.5 0.4 0.5 0.4 0.5 0.4 0.5 0.4 0.5 0.4 0.5 0.4 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5	0.8 0.4 0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2				0.8 0.4 0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2		0.3 0.4 0.2 0.1 0.3 0.2 0.2 0.3 0.3 0.4 0.4 0.2 0.3 0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4	0.3 0.4 0.2 0.3 0.5 0.2 0.3 0.55 0.44 0.72 1				
44C/2		1 0.5 0.2 0.2 0.5 0.2 0.5 0.2 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5					1 0.8 0.4 0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2	0,0 0,0 0,0 0,0 0,0 0,0 0,0 0,0 0,0 0,0	0.8 0.4 0.2 0.1 0.2 0.2 0.2 0.2 0.0 0.4 0.4 0.4 0.4 0.2 0.2 0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4	0.4 0.4 0.2 0.3 0.4 0.4 0.2 0.3 0.58 0.48 0.48 0.48 0.48 0.48 0.48 0.48 0.4				1 03 04 04 0 0 02 04 05 03
April 73														400 200 0



### Model dependence plots for target Y1





### Detail of X1 and X11 interaction for Y1





### Model dependence plots for target Y2





#### Detail of X1 and X11 interaction for Y2





### Model dependence plots for target Y3





# Model accuracy and runtime

Variable	RMSE (no constraints)	RMSE (with constraints)	Run time (with constraints)
Y1	0.0228	0.0405	117 s
Y2	0.0444	0.0475	44.5 s
Y3	0.0509	0.0521	108 s

- Higher training error for shape-constrained regression
- Runtime: 1 2 minutes
  - Office PC
  - a single monotonicity constraint
  - ~ 1500 data points



# Conclusions

- Shape constraints can be used to improve polynomial regression models.
- Successful for monotonicity constraints.
- Concavity / convexity can be enforced similarly.
- Runtime depends mainly on the number of constraints.

Limitations:

- Works only for polynomial models.
- Improved extrapolation has not been quantified





https://symreg.at



#### Integration of Physical Knowledge in Empirical Models – A New Approach to Regression Analysis

**G. Kronberger,** C. Haider, M. Kommenda, (University of Applied Sciences Upper Austria) S. Scheidel (AVL List GmbH), M. Kordon (AVL Commercial Driveline & Tractor Engineering GmbH)