Gaussian Process Semi-Parametric Regression w/ Misspecified Models

Kyla Jones

kjones29@nd.edu

Nonlinear & Stochastic Optimization: Final Presentation

May 8, 2023

Motivating Example – The Simple Machine[1]

Ground Truth Model

$$\zeta(x) = \frac{\theta x}{1 + x/a}$$

Data

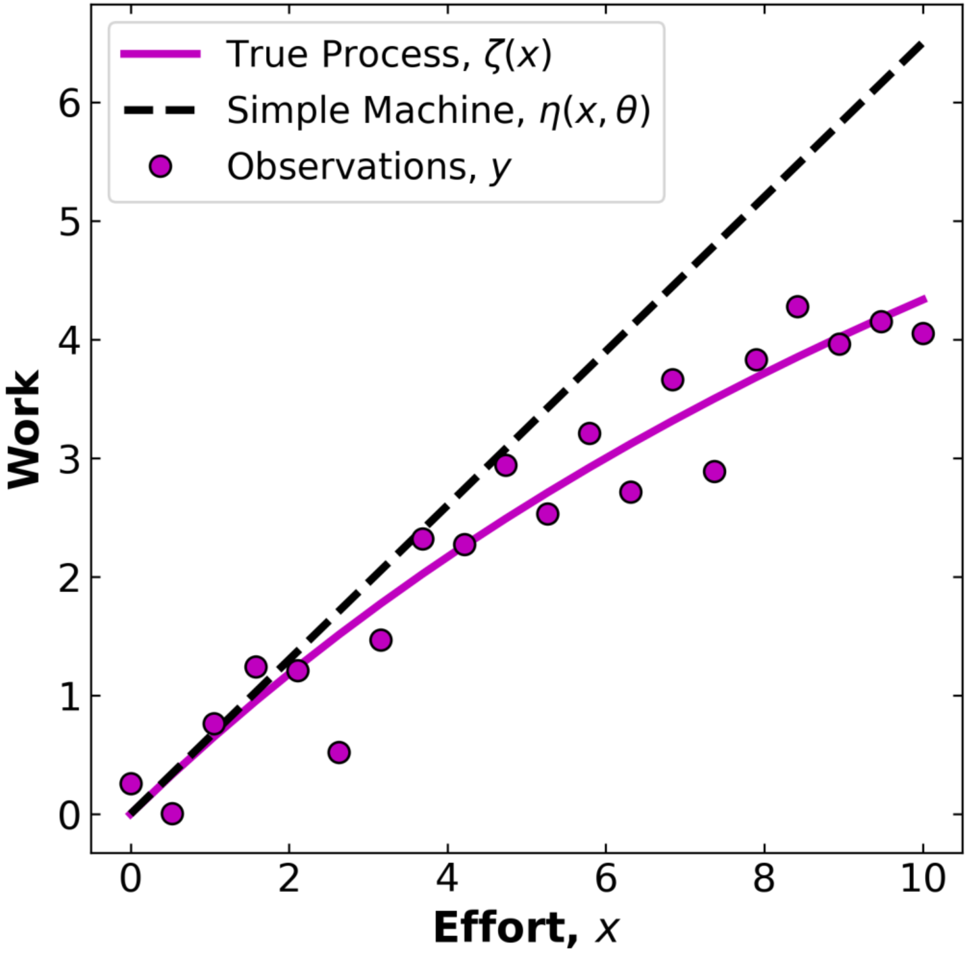
$$y = \zeta(x) + \epsilon, \qquad \epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$$

Misspecified Model

$$\eta(x, \theta) = \theta x$$

ζ: work
x: effort
θ: efficiency
a: friction factor
y: observations
ϵ: measurement error
σ_ϵ^2 : variance of error

How to learn model-form uncertainty ?



Kennedy & O'Hagan (KOH) model[2]:

$$y = \zeta(x) + \epsilon = \eta(x, \theta) + \delta(x) + \epsilon$$

$\delta(x) \sim \mathcal{GP}(0, \sigma_\delta^2 K(x, x'; \psi))$: model discrepancy
σ_δ^2: process variance
$K(\cdot,\cdot)$: correlation function
ψ: length scale parameters

Objective: Find $(\theta, \sigma_\epsilon^2, \sigma_\delta^2, \psi)$ that best reproduce y.

[1] Kennedy, M.C. & O'Hagan, A. (2001). *JRSSB* 63(3):425–464.
[2] Brynjarsdóttir, J. & O'Hagan, A. (2014). *Inverse Probl.* 30(11): 114007.

Gaussian Process Semi-Parametric Regression w/ Correlated Errors [3]

Parameter estimation as nested optimization:

$$\arg\min_{\boldsymbol{\omega}\in\Omega} \frac{1}{2}\log|\mathbb{V}(\boldsymbol{\omega})| + \frac{1}{2}\left(\mathbf{y} - \mathbf{x}\hat{\theta}(\boldsymbol{\omega})\right)^{\intercal}\mathbb{V}(\boldsymbol{\omega})^{-1}\left(\mathbf{y} - \mathbf{x}\hat{\theta}(\boldsymbol{\omega})\right) \longleftarrow \text{negative log-likelihood}$$

$$\text{s.t.}\quad \hat{\theta}(\boldsymbol{\omega}) = (\mathbf{x}^{\intercal}\mathbb{V}(\boldsymbol{\omega})^{-1}\mathbf{x})^{-1}\mathbf{x}^{\intercal}\mathbb{V}(\boldsymbol{\omega})\mathbf{y} \longleftarrow \text{generalized least squares estimator}$$

$$\mathbb{V}(\boldsymbol{\omega}) = \sigma_{\delta}^{2}\left[K\left(x_i, x_j, \psi\right)\right]_{i,j=1}^{n} + \sigma_{\epsilon}^{2}\mathbf{I} \longleftarrow \text{plug-in variance}$$

$$K\left(x_i, x_j; \psi\right) = \exp\left(-\left(\frac{x_i - x_j}{\psi}\right)^2\right) \longleftarrow \text{correlation function}$$

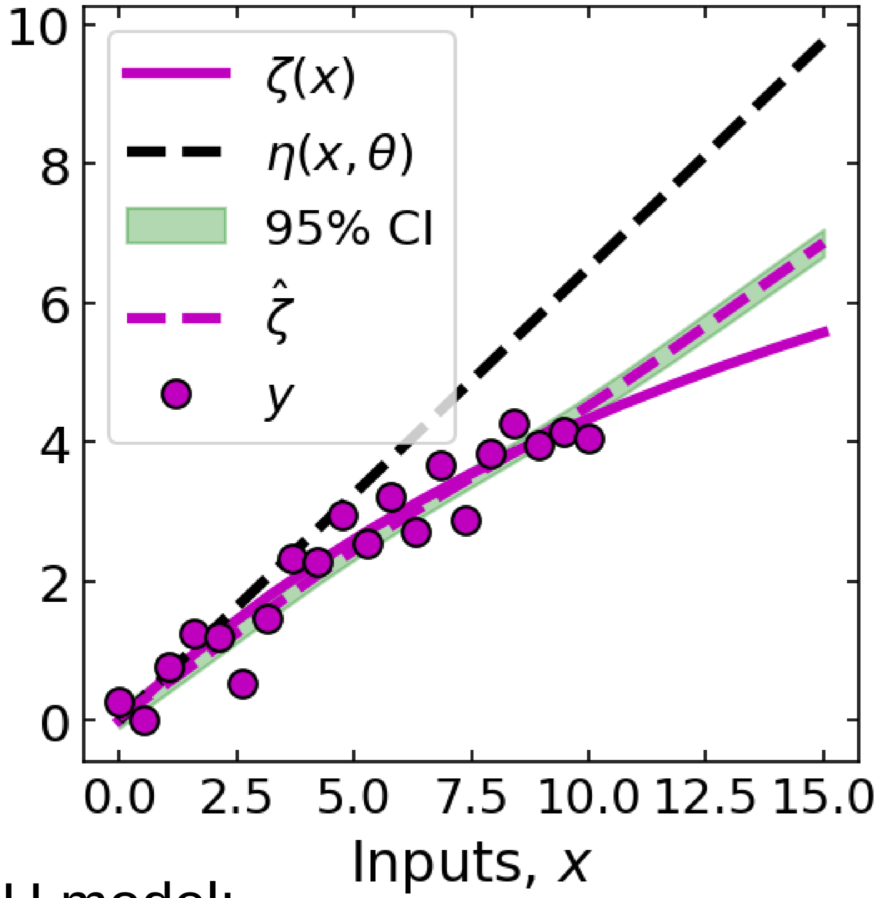
$$\mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^n \longleftarrow \text{data (given)}$$

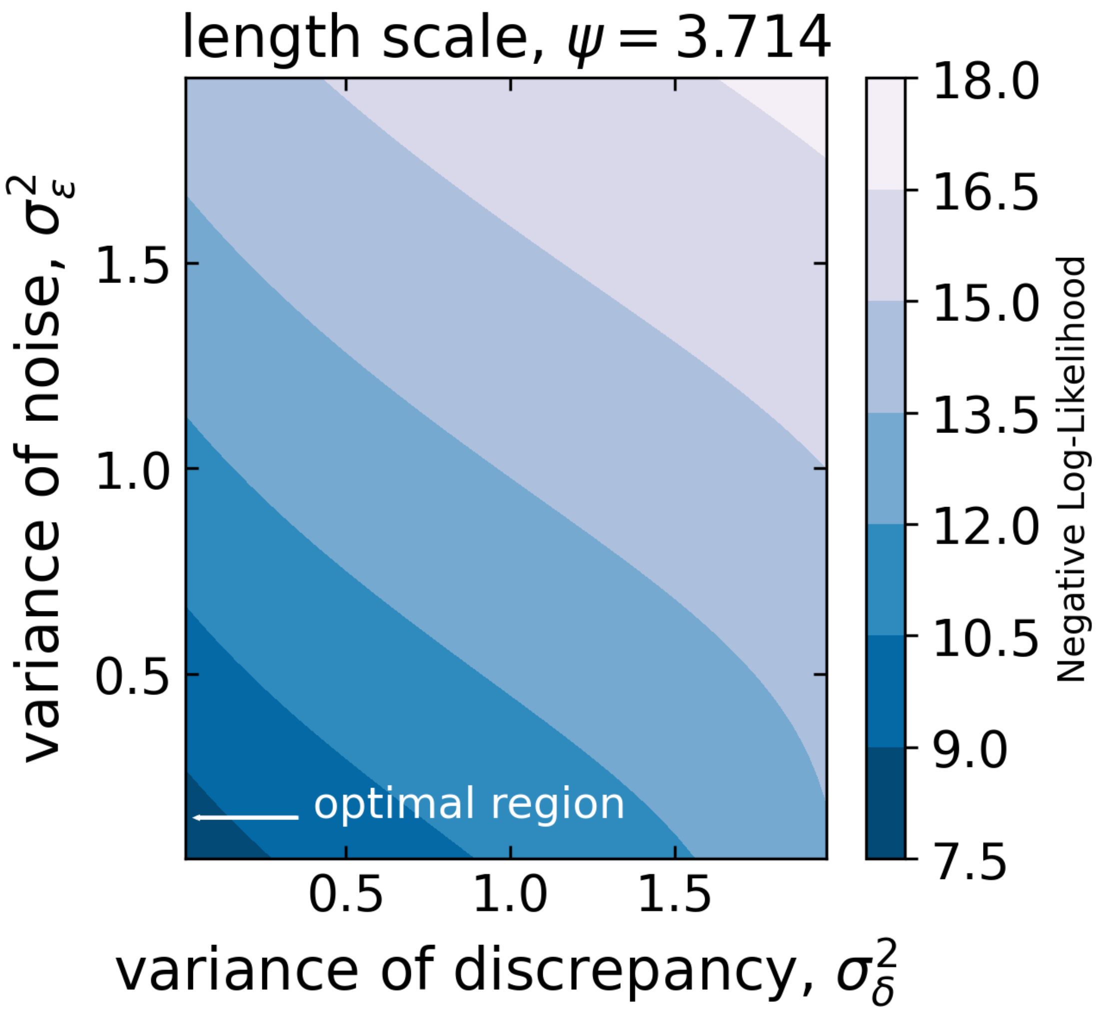
$$\boldsymbol{\omega} = \left[\sigma_{\epsilon}^{2}, \sigma_{\delta}^{2}, \psi\right] \in \Omega \subseteq \mathbb{R}_{+}^{3} \longleftarrow \text{nuisance parameters (decision variables)}$$

Problem size: 3 decision variables, 3 supporting equations, 44 data variables, 0 constraints

[3] He, H. & Severini, T.A. (2016). *CSDA* 103(3):316–329.

Results





KOH model:
(+) outperforms misspecified model
(-) is model better at interpolation than prediction

(-) nuisance parameters not identifiable ⇒ θ not identifiable
(-) different approach needed for parameter estimation

Thank you for your attention! References:

[1] Kennedy, Marc C. & O’Hagan, Anthony. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 63(3):425–464.

[2] Brynjarsdóttir, Jenný & O’Hagan, Anthony. (2014). Learning about model parameters: the importance of model discrepancy. *Inverse Problems* 30(11): 114007.

[3] He, Heping & Severini, Thomas A. (2016). A flexible approach to inference in semi-parametric regression models with correlated errors using Gaussian processes. *Computational Statistics & Data Analysis* 103(3):316–329.