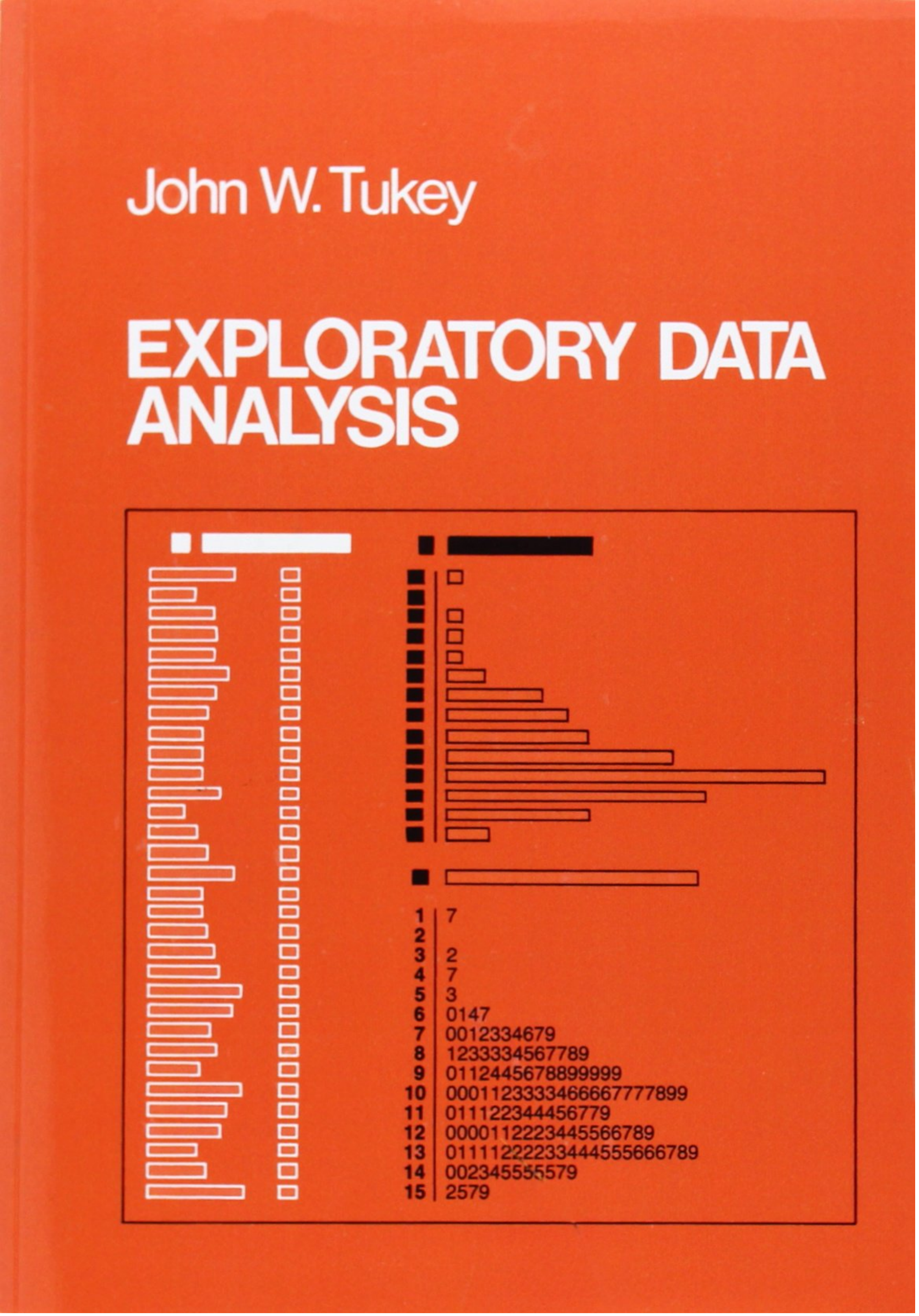
Approximation theoretic advice for supervised learning

PAUL CONSTANTINE
Assistant Professor
Department of Computer Science
University of Colorado, Boulder

`activesubspaces.org`
`@DrPaulynomial`

SLIDES AVAILABLE UPON REQUEST

DISCLAIMER: These slides are meant to complement the oral presentation. Use out of context at your own risk.



Even more understanding is *lost* if we consider each thing we can do to data *only* in terms of some set of very restrictive assumptions under which that thing is best possible—assumptions we *know we CANNOT check in practice.*

TRIGGER WARNING!

data

model

noise

parameter

error

uncertainty

overfit

Selected regression / approximation / UQ literature

My personal bibliography

Ghanem and Spanos, *Stochastic Finite Elements* (Springer, 1991)

Xiu and Karniadakis, *The Wiener-Askey polynomial chaos* (SISC, 2002)

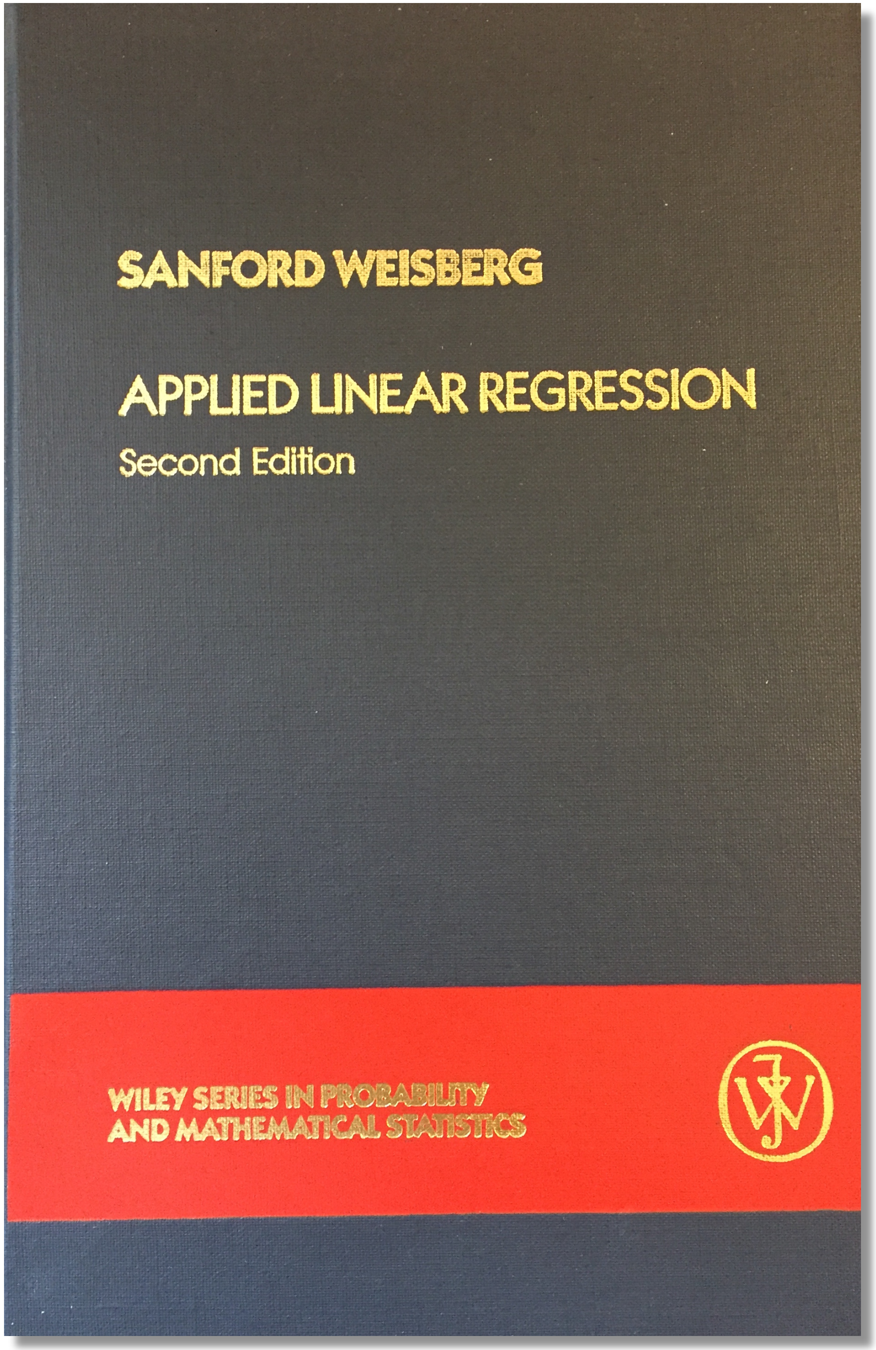
Nobile, Tempone, and Webster, *A sparse grid stochastic collocation method* (SINUM, 2008)

Gautschi, *Orthogonal Polynomials* (Oxford UP, 2004)

Koehler and Owen, *Computer experiments* (Handbook of Statistics, 1996)

Jones, *A taxonomy of global optimization methods based on response surfaces* (JGO, 2001)

Cook, *Regression Graphics* (Wiley, 1998)

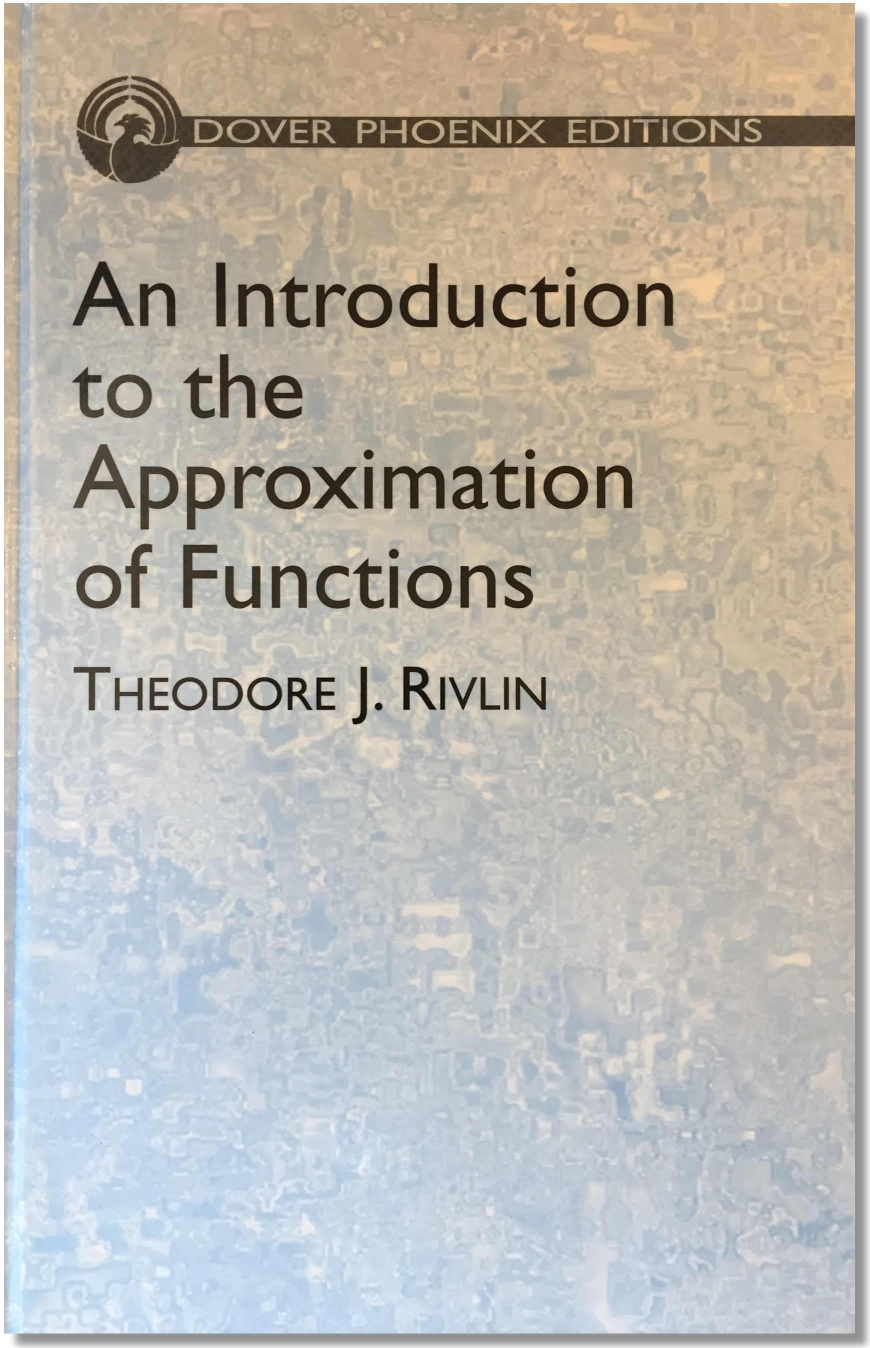


SANFORD WEISBERG

APPLIED LINEAR REGRESSION

Second Edition

WILEY SERIES IN PROBABILITY
AND MATHEMATICAL STATISTICS

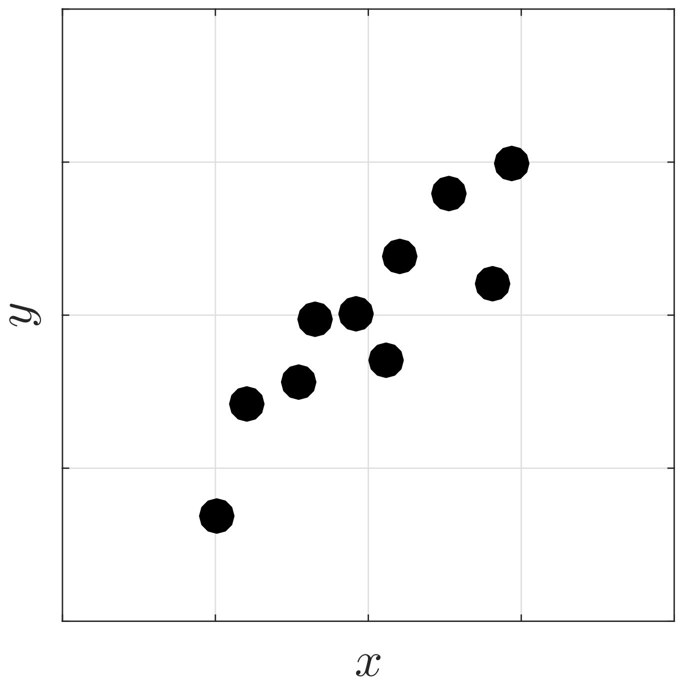


DOVER PHOENIX EDITIONS

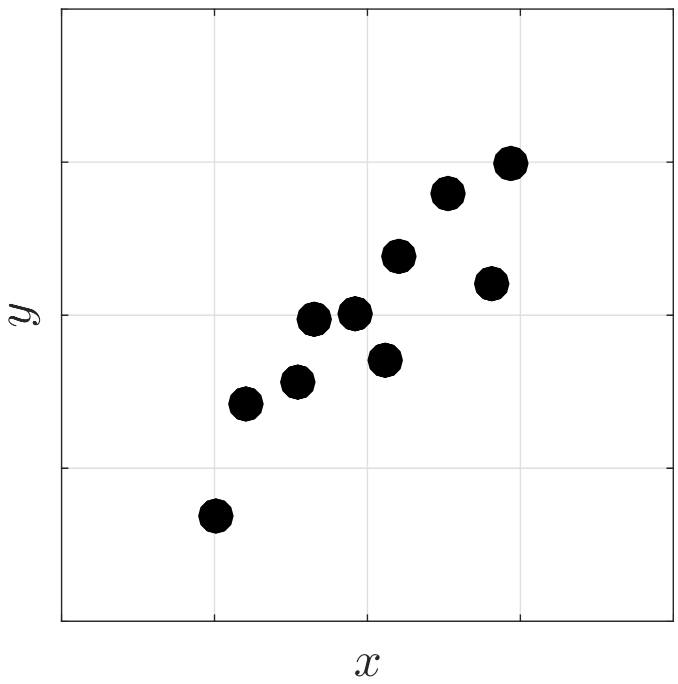
An Introduction
to the
Approximation
of Functions

THEODORE J. RIVLIN

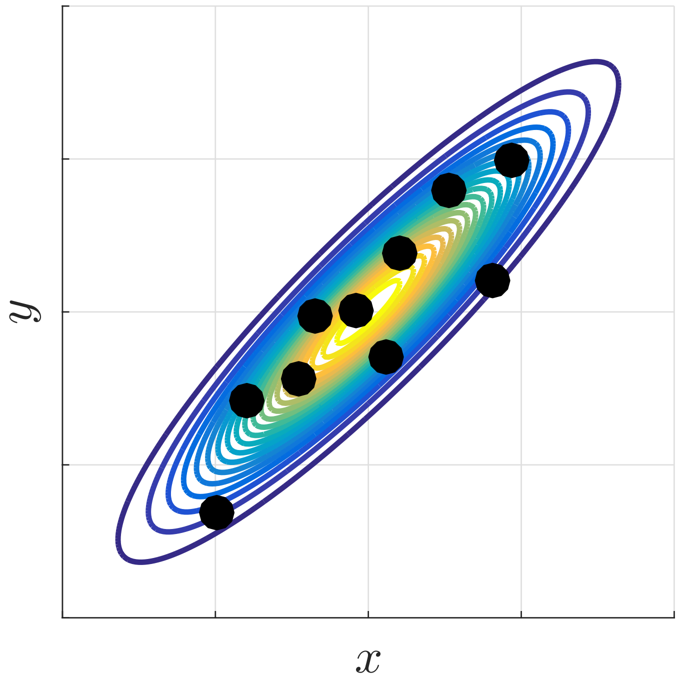
Regression



Approximation



Regression



GIVEN

i.i.d. samples $\{x_i, y_i\}$

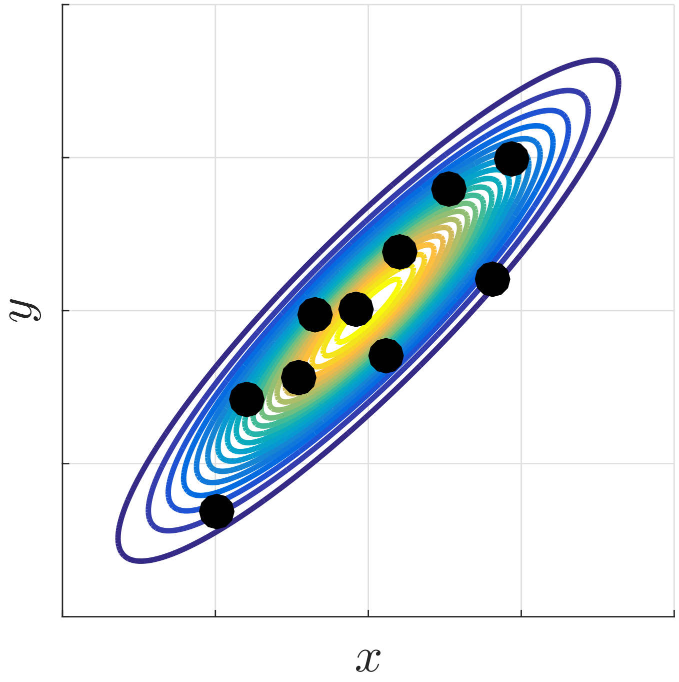
from **unknown** $\pi(x, y)$

GOAL

statistically characterize $y \,|\, x$

e.g., $\mathbb{E}[\, y \,|\, x \,],\ \mathrm{Var}[\, y \,|\, x \,]$

Regression



MODEL (e.g., polynomials)

$$y = \underbrace{p(x, \theta)}_{\mathbb{E}[y|x]} + \varepsilon \longleftarrow$$

modeled r.v.,
zero-mean,
independent of x

GIVEN

i.i.d. samples $\{x_i, y_i\}$

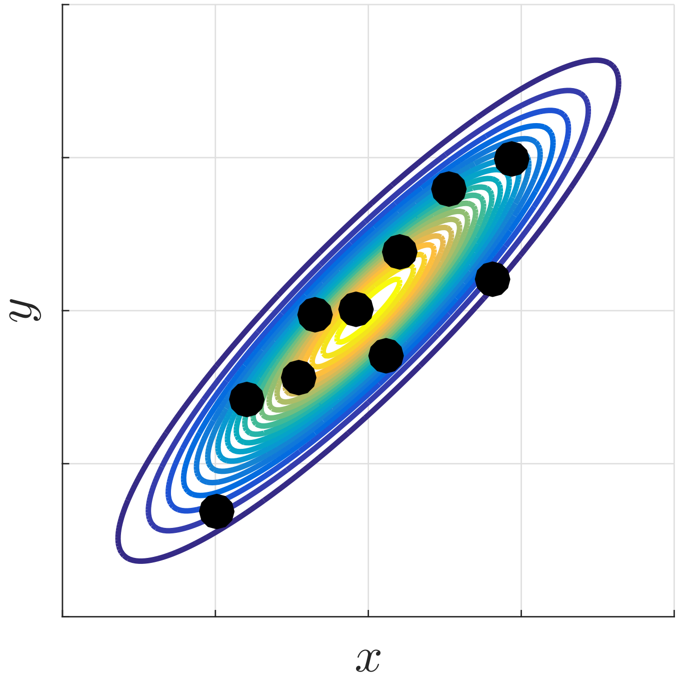
from **unknown** $\pi(x, y)$

GOAL

statistically characterize $y \,|\, x$

e.g., $\mathbb{E}[\, y \,|\, x \,]$, $\mathrm{Var}[\, y \,|\, x \,]$

Regression



GIVEN

i.i.d. samples $\{x_i, y_i\}$

from **unknown** $\pi(x, y)$

GOAL

statistically characterize $y \,|\, x$

e.g., $\mathbb{E}[\,y \,|\, x\,],\ \mathrm{Var}[\,y \,|\, x\,]$

MODEL (e.g., polynomials)

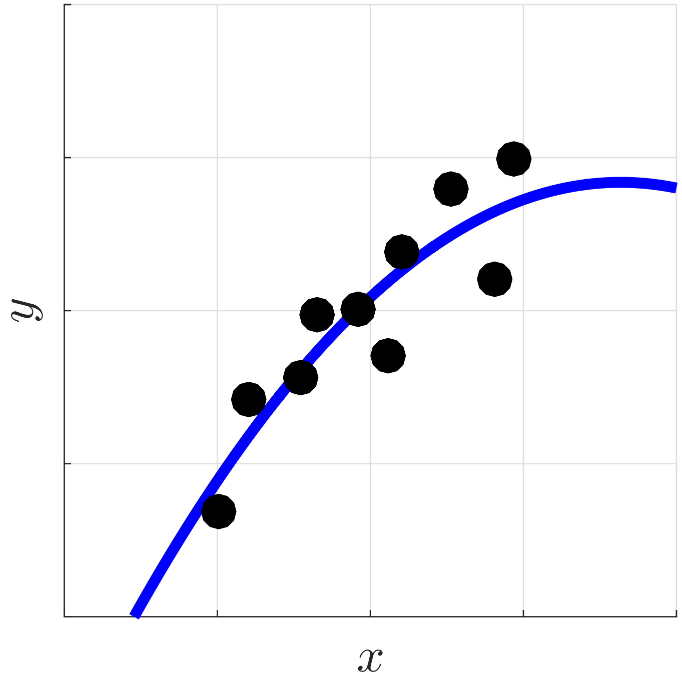
$$y \;=\; \underbrace{p(x,\theta)}_{\mathbb{E}[y|x]} + \varepsilon \;\longleftarrow$$

modeled r.v.,
zero-mean,
independent of x

FIT (e.g., max likelihood)

$$\hat{\theta} \;=\; \underset{\theta}{\mathrm{argmin}} \sum_i \big(y_i - p(x_i,\theta)\big)^2$$

Regression



GIVEN

i.i.d. samples $\{x_i, y_i\}$

from **unknown** $\pi(x, y)$

GOAL

statistically characterize $y \,|\, x$

e.g., $\mathbb{E}[\,y\,|\,x\,]$, $\mathrm{Var}[\,y\,|\,x\,]$

MODEL (e.g., polynomials)

$$y \;=\; \underbrace{p(x, \theta)}_{\mathbb{E}[y|x]} \;+\; \varepsilon \;\longleftarrow$$

modeled r.v.,
zero-mean,
independent of x

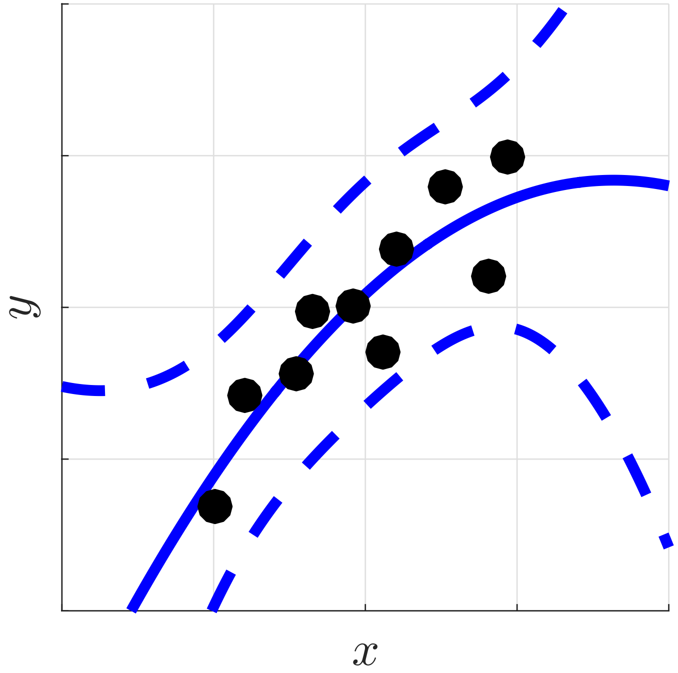
FIT (e.g., max likelihood)

$$\hat{\theta} \;=\; \underset{\theta}{\operatorname{argmin}} \sum_i \big(y_i - p(x_i, \theta)\big)^2$$

PREDICT

$$\mathbb{E}[\,y\,|\,x^*\,] \;\approx\; p(x^*, \hat{\theta}) \;=\; \hat{p}(x^*)$$

Regression



GIVEN

i.i.d. samples $\{x_i, y_i\}$

from **unknown** $\pi(x, y)$

GOAL

statistically characterize $y \,|\, x$

e.g., $\mathbb{E}[\,y\,|\,x\,]$, $\mathrm{Var}[\,y\,|\,x\,]$

MODEL (e.g., polynomials)

$$y \;=\; \underbrace{p(x, \theta)}_{\mathbb{E}[y|x]} + \varepsilon \longleftarrow$$

modeled r.v.,
zero-mean,
independent of x

FIT (e.g., max likelihood)

$$\hat{\theta} \;=\; \underset{\theta}{\operatorname{argmin}} \sum_i \big(y_i - p(x_i, \theta)\big)^2$$

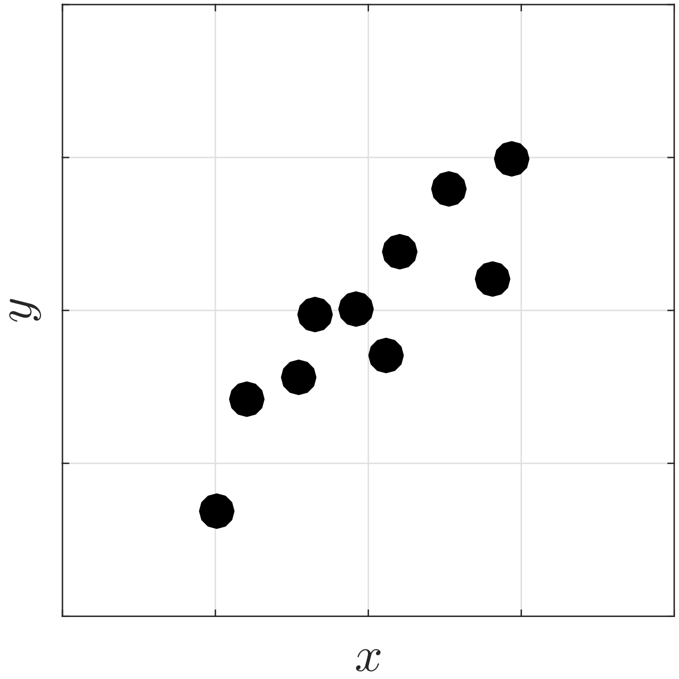
PREDICT

$$\mathbb{E}[\,y\,|\,x^*\,] \;\approx\; p(x^*, \hat{\theta}) \;=\; \hat{p}(x^*)$$

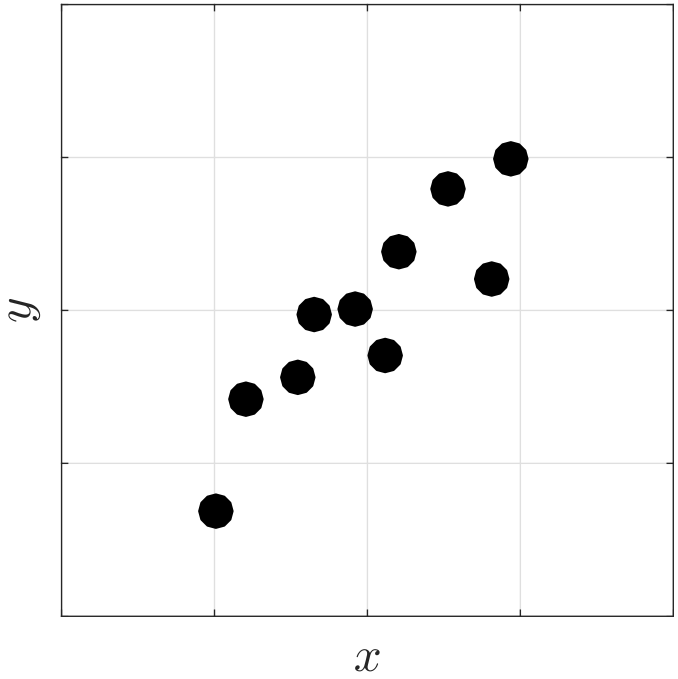
QUANTIFY UNCERTAINTY

$$\mathrm{Var}[\,y\,|\,x^*\,] \;\approx\; \text{“formula”}$$

Regression



Approximation



Approximation

Does a **unique**, **best** approximation exist?

$$p^* \;=\; \underset{p \in \mathcal{P}_n}{\operatorname{argmin}} \,\|\, p - f \,\|$$

polynomials

continuous function

How does the best error behave?

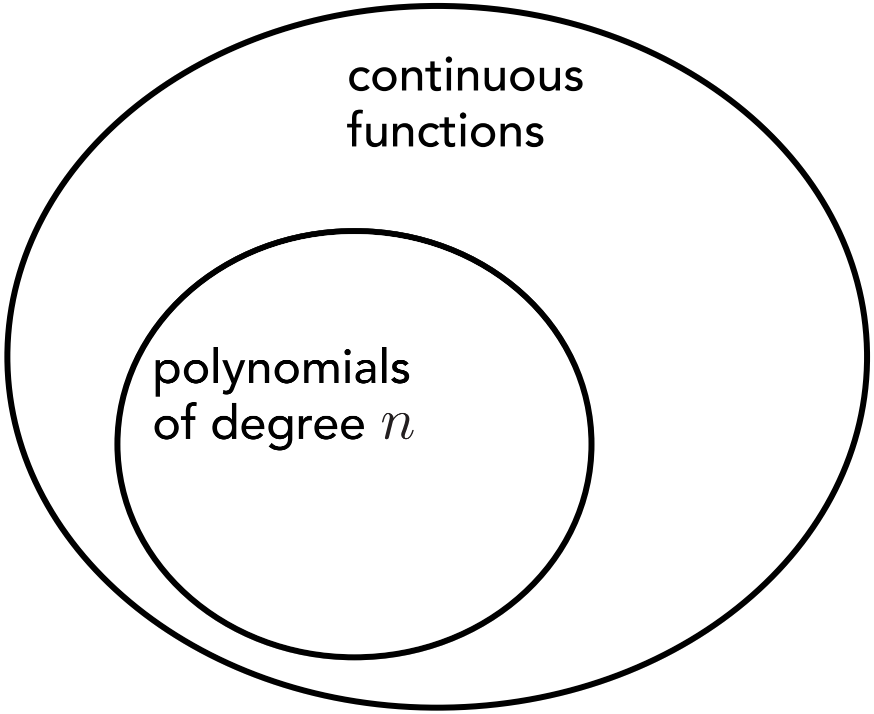
$$\|\, p^* - f \,\| \;=\; e^*(n)$$

Can we construct an approximation?

$$\boxed{\textbf{Algorithm:}\ \text{Given } f, \text{ compute } \hat{p}}$$

And analyze its error?

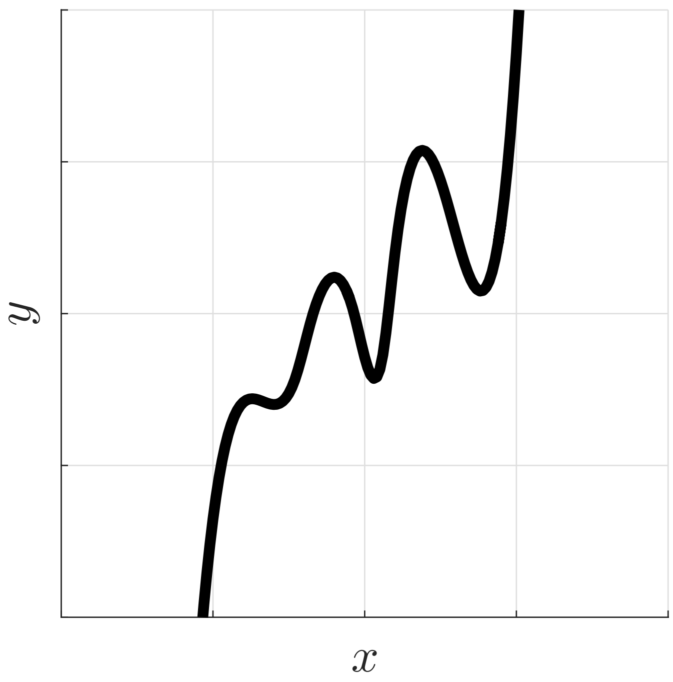
$$\|\, \hat{p} - f \,\| \;\le\; C\, e^*(n)$$



continuous functions

polynomials of degree n

Approximation



GIVEN

a function $f(x)$

a **known** density $\pi(x)$

GOAL

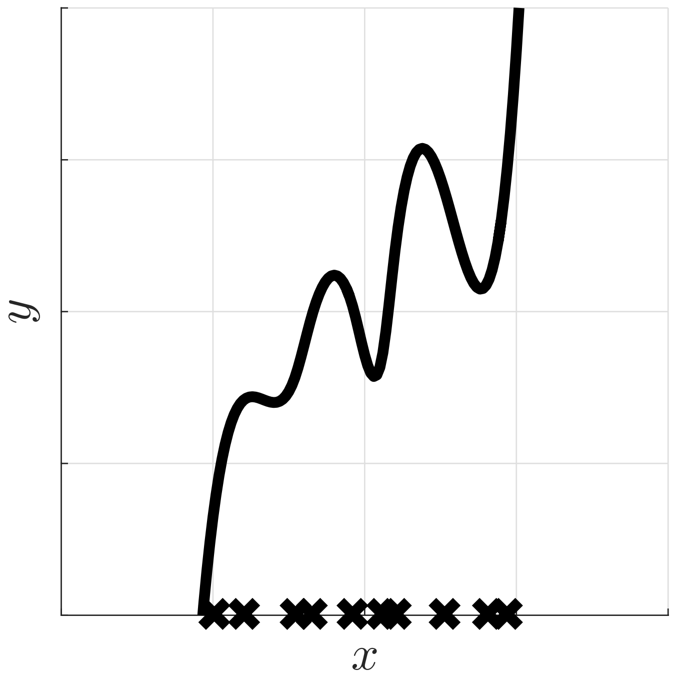
find $\hat{p}(x)$ such that

the error $||\hat{p} - f||$ is small

CONSTRUCTION

choose x_i

Approximation



GIVEN

a function $f(x)$

a **known** density $\pi(x)$

GOAL

find $\hat{p}(x)$ such that

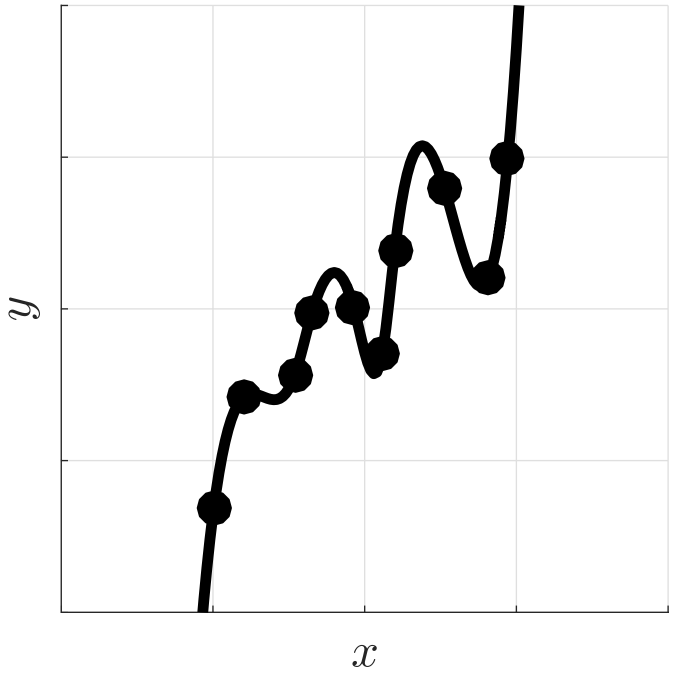
the error $||\hat{p} - f||$ is small

Approximation

CONSTRUCTION

choose x_i

compute $y_i = f(x_i)$



GIVEN

a function $f(x)$

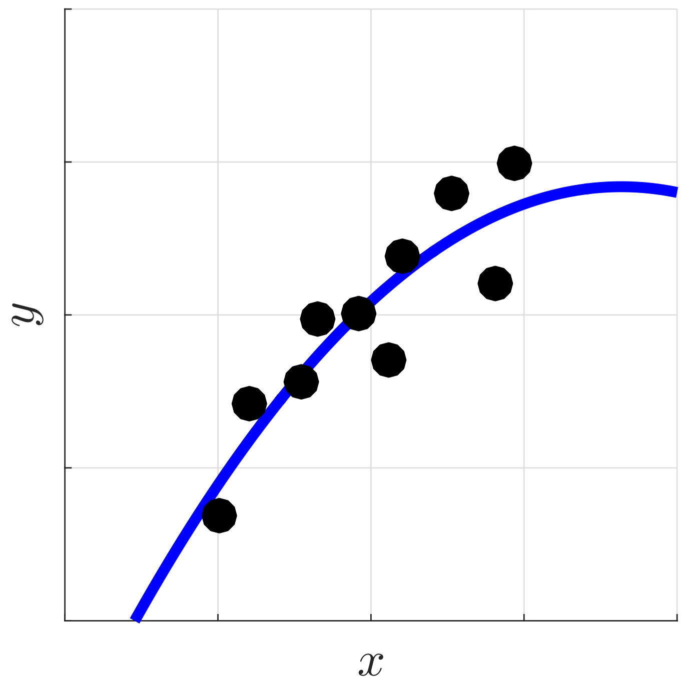
a **known** density $\pi(x)$

GOAL

find $\hat{p}(x)$ such that

the error $||\hat{p} - f||$ is small

Approximation



CONSTRUCTION

choose x_i

compute $y_i = f(x_i)$

fit $\hat{p} = \underset{p \in \mathcal{P}_n}{\operatorname{argmin}} \sum_i \left(y_i - p(x_i)\right)^2$

GIVEN

a function $f(x)$

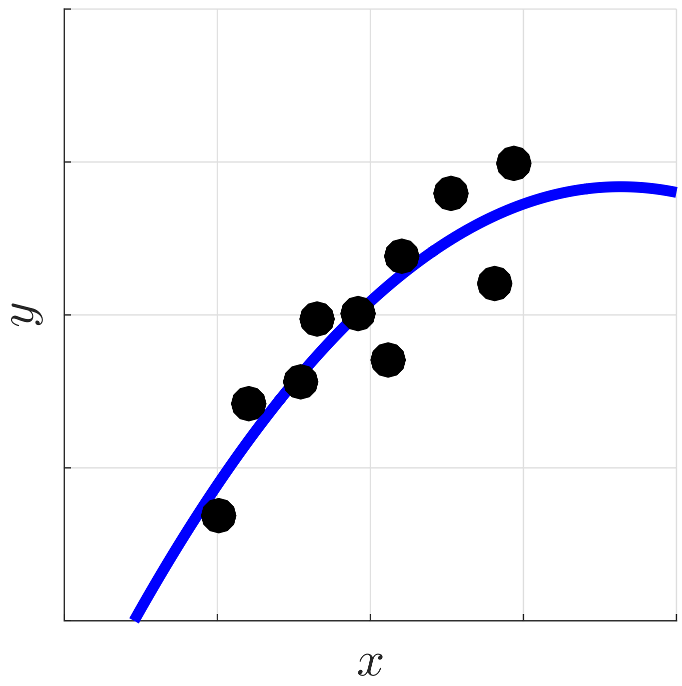
a **known** density $\pi(x)$

GOAL

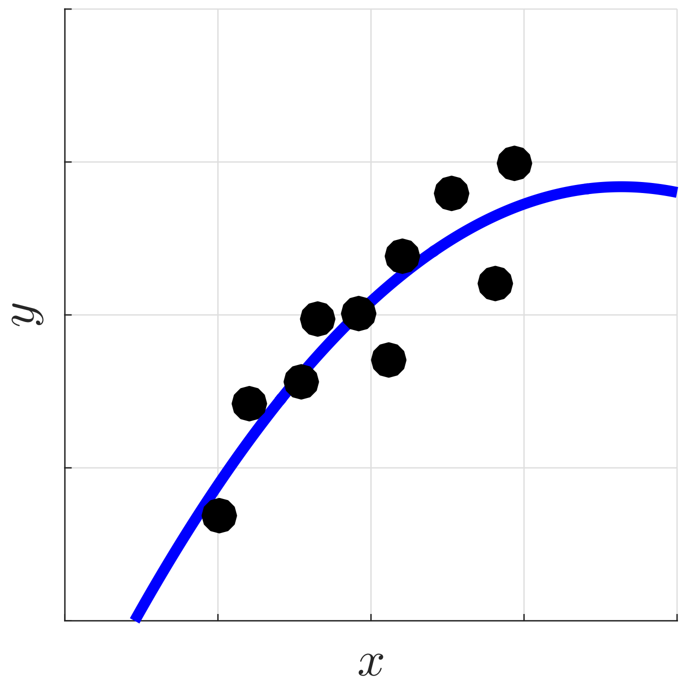
find $\hat{p}(x)$ such that

the error $\|\hat{p} - f\|$ is small

Regression



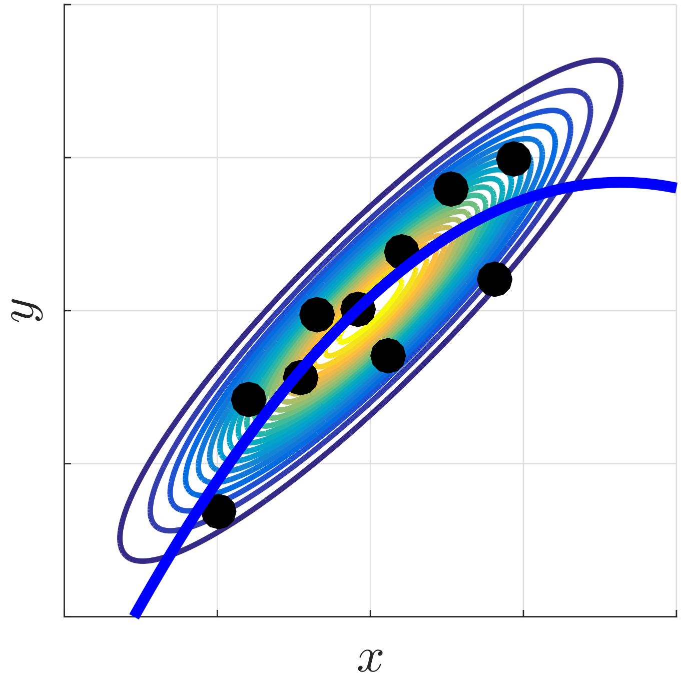
Approximation



$$\hat{\theta} \;=\; \underset{\theta}{\operatorname{argmin}} \sum_i \big(y_i - p(x_i, \theta)\big)^2$$

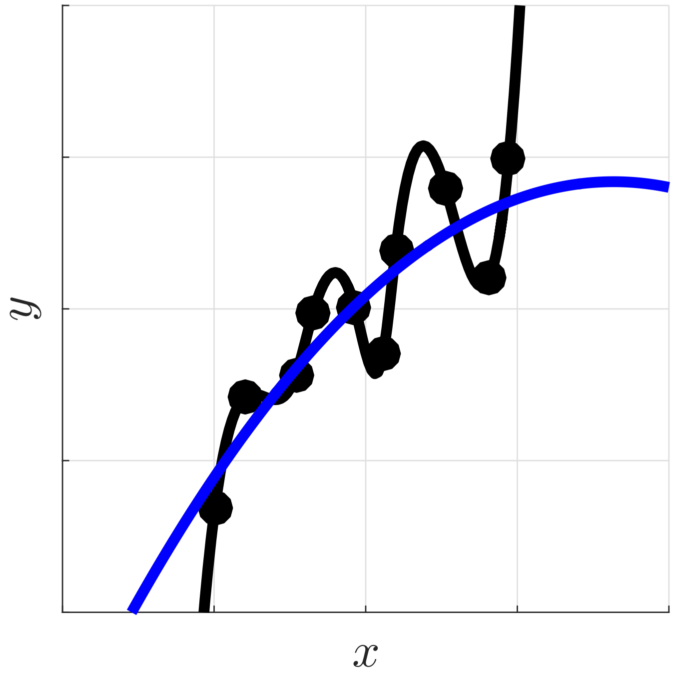
$$\hat{p} \;=\; \underset{p \,\in\, \mathcal{P}_n}{\operatorname{argmin}} \sum_i \big(y_i - p(x_i)\big)^2$$

Regression



$$\hat{\theta} = \operatorname*{argmin}_{\theta} \sum_{i} \big(y_i - p(x_i,\theta)\big)^2$$

Approximation



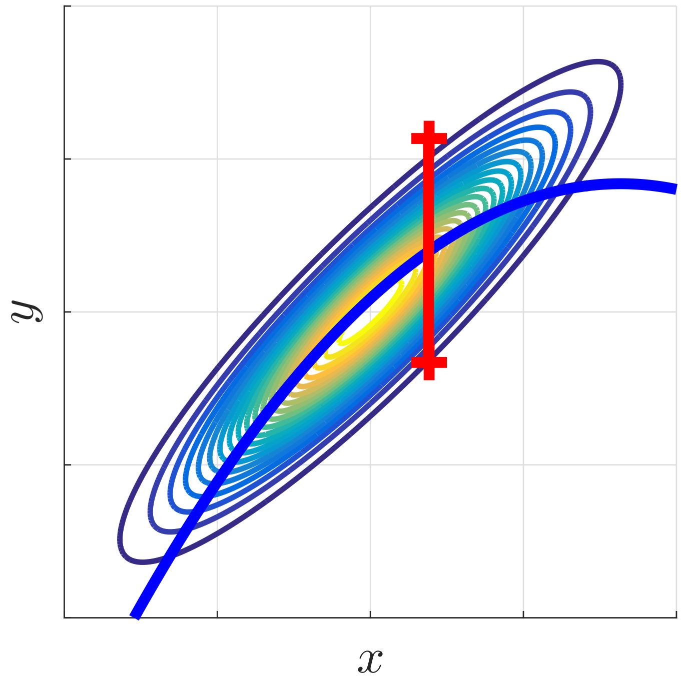
$$\hat{p} = \operatorname*{argmin}_{p\in\mathcal{P}_n} \sum_{i} \big(y_i - p(x_i)\big)^2$$

The **story** of the data and fitted curve is **different**.
But does it matter? **YES**

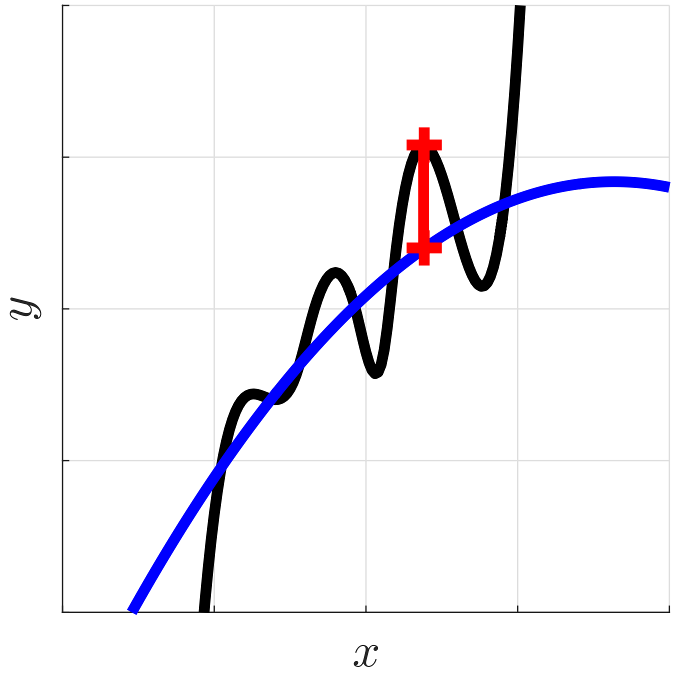
REGRESSION VS. APPROXIMATION

What is *error*?

Regression



Approximation



Confidence interval

$$\hat{p}(x) \pm \underbrace{2\,\widehat{\mathrm{se}}[\,y\,|\,x\,]}_{\text{plug-in estimate of } \textit{standard error}}$$

COMPUTABLE!

Approximation error

$$|\,\hat{p}(x) - f(x)\,|$$

NOT COMPUTABLE!

Error norms

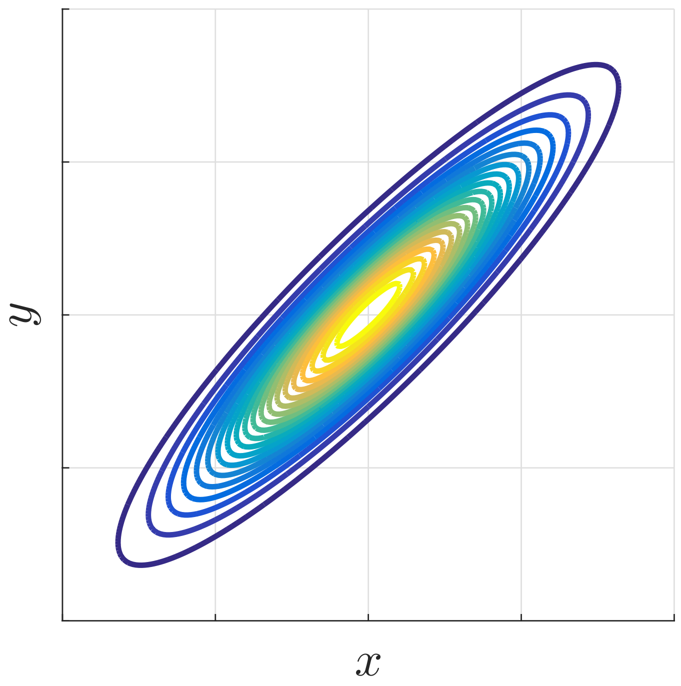
$$\left(\int |\,\hat{p}(x) - f(x)\,|^2 \, \pi(x)\, dx \right)^{1/2}$$

$$\sup_x |\,\hat{p}(x) - f(x)\,|$$

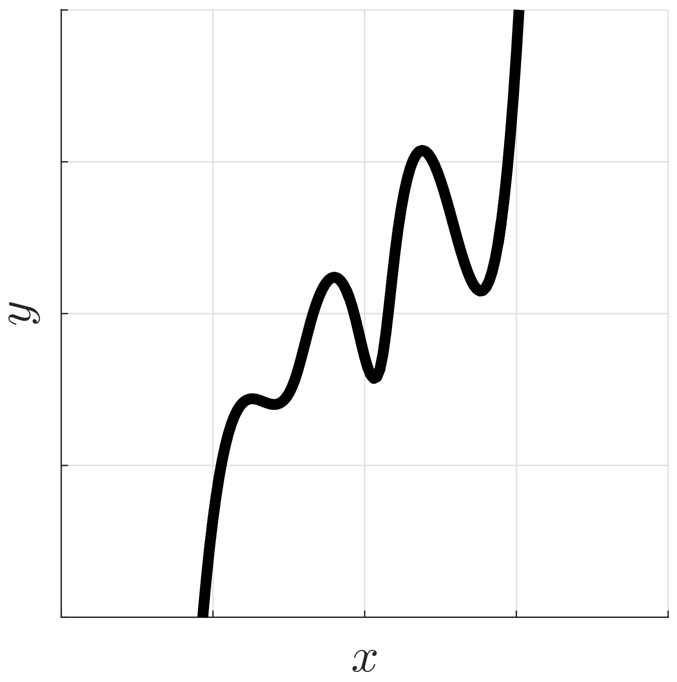
REGRESSION VS. APPROXIMATION

What is *convergence*?

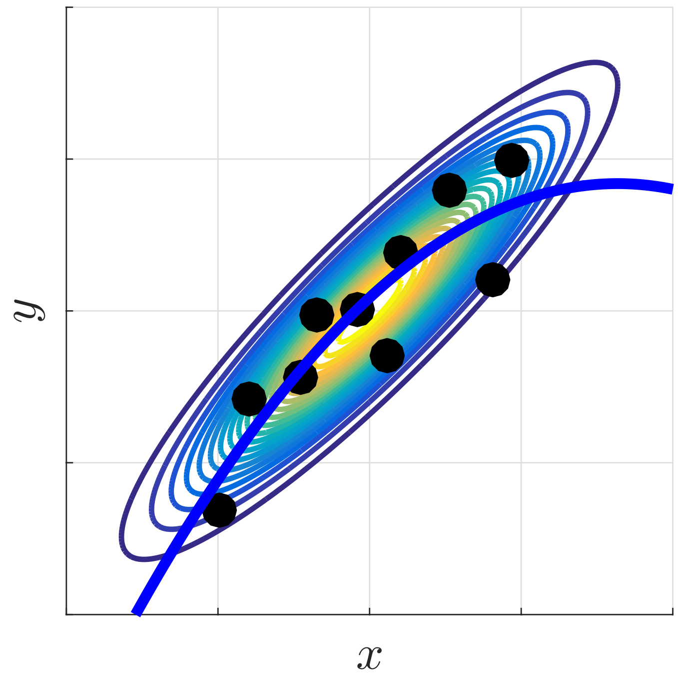
Regression



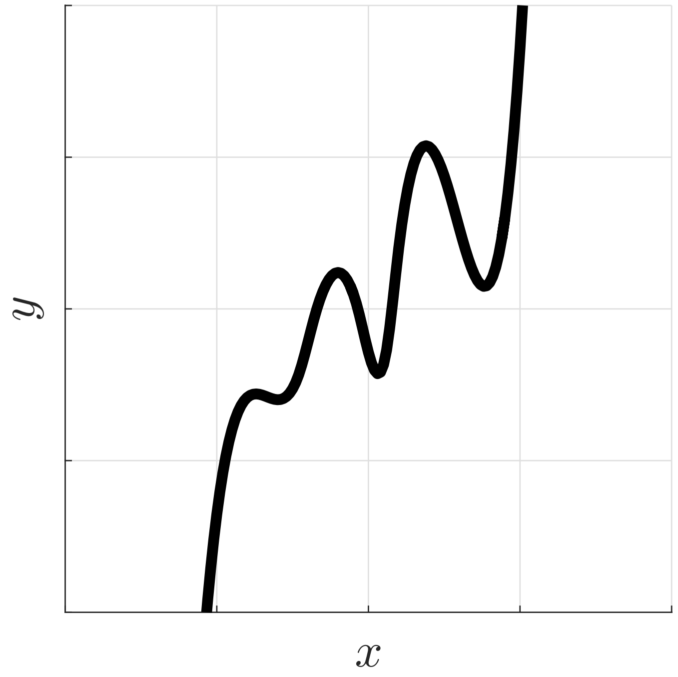
Approximation



Regression



Approximation

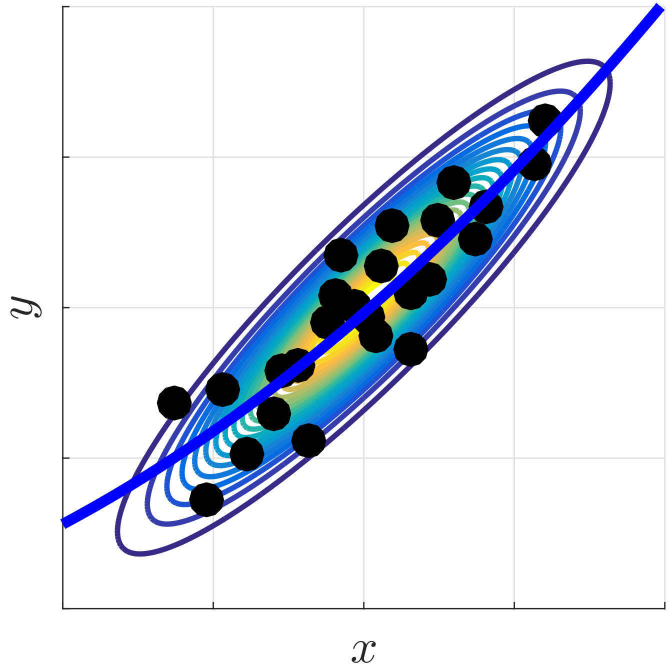


As data increases, *root-n consistency*

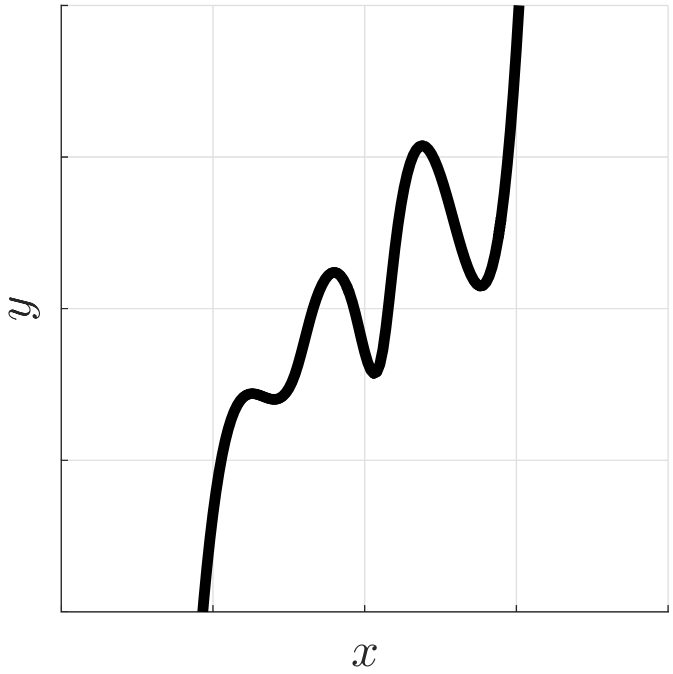
$$\left.\begin{array}{ccc}\hat{\theta} & \rightarrow & \theta \\ \hat{p}(x) & \rightarrow & p(x)\end{array}\right\}$$

"true"
parameters

Regression



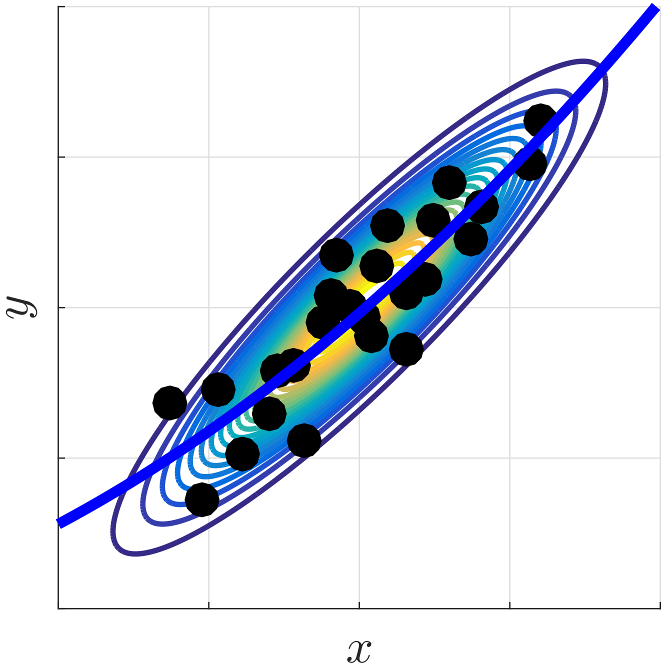
Approximation



As data increases, *root-n consistency*

$$\left.\begin{array}{lcl}\hat{\theta} & \rightarrow & \theta \\ \hat{p}(x) & \rightarrow & p(x)\end{array}\right\} \text{“true” parameters}$$

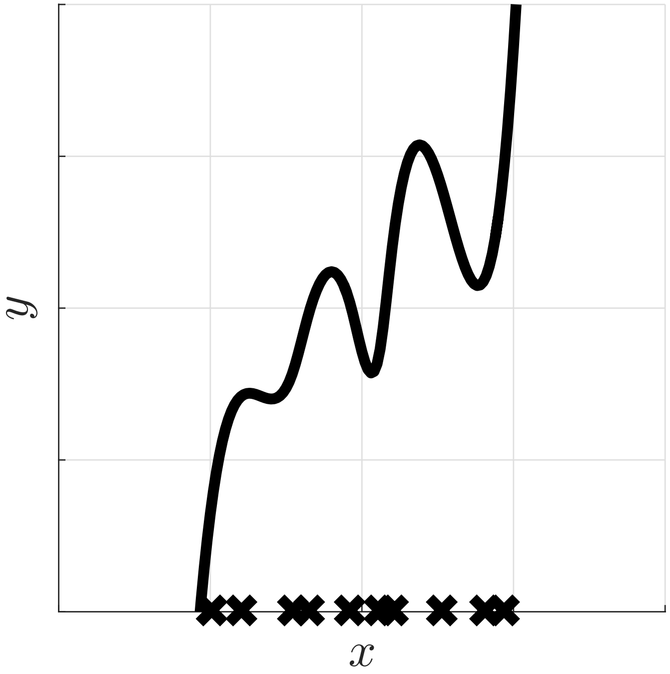
Regression



As data increases, *root-n consistency*

$$\left.\begin{array}{lll} \hat{\theta} & \rightarrow & \theta \\ \hat{p}(x) & \rightarrow & p(x) \end{array}\right\} \text{"true" parameters}$$

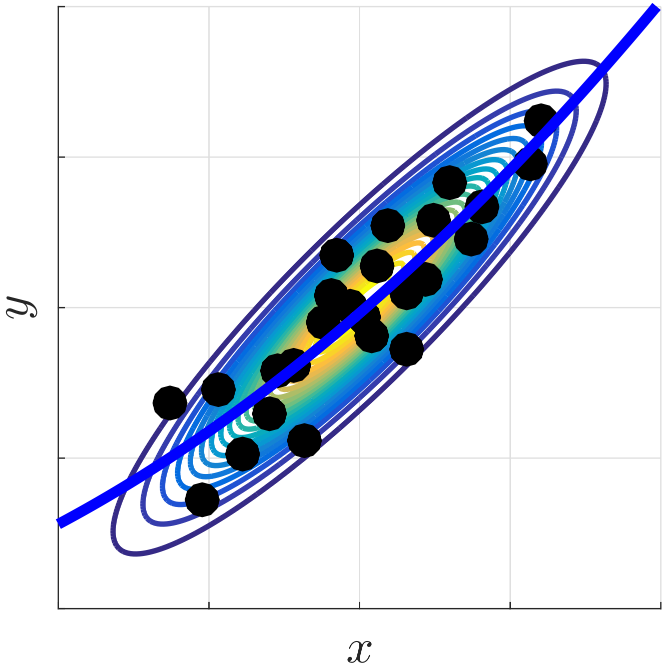
Approximation



As the approximation class grows

$$\|\, \hat{p}(x) - f(x) \,\| \;\rightarrow\; 0$$

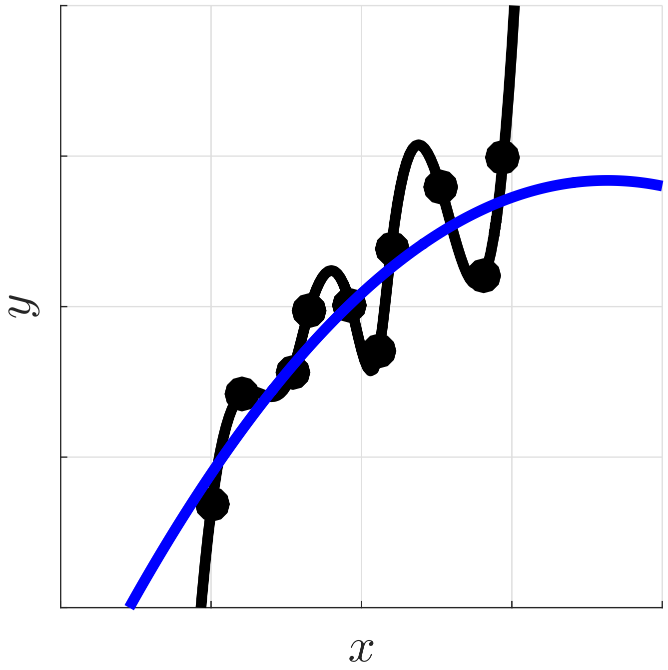
Regression



As data increases, *root-n consistency*

$$\left.\begin{array}{lcl}\hat{\theta} & \to & \theta \\ \hat{p}(x) & \to & p(x)\end{array}\right\} \text{“true” parameters}$$

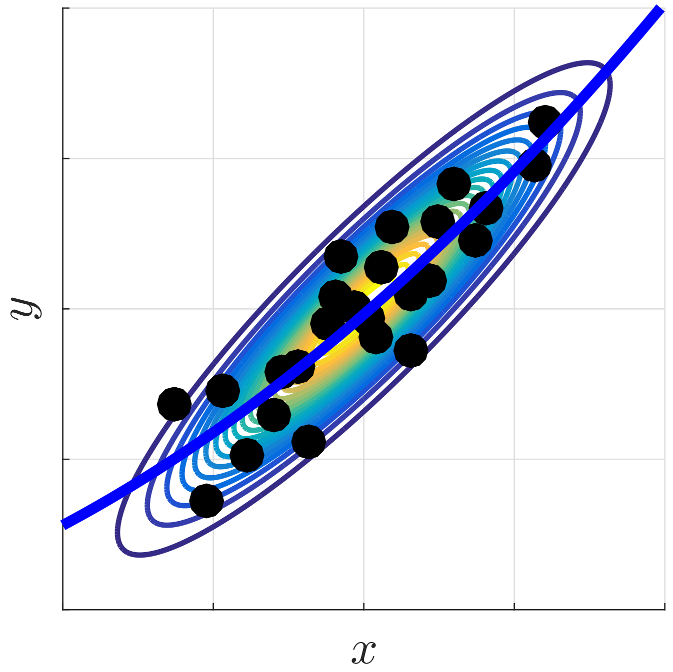
Approximation



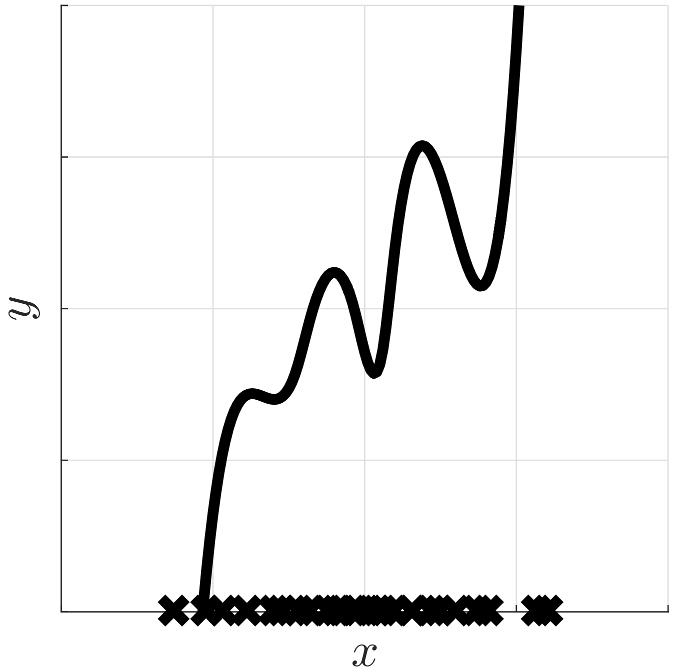
As the approximation class grows

$$\| \hat{p}(x) - f(x) \| \to 0$$

Regression



Approximation



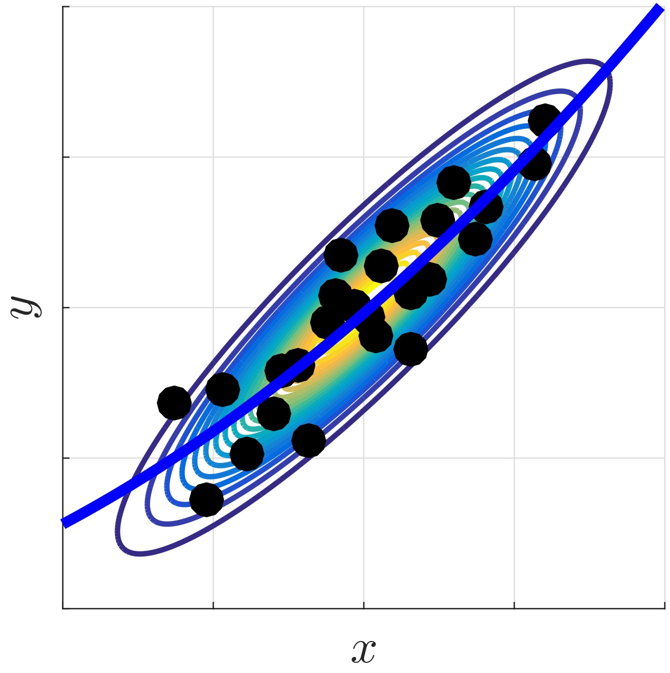
As data increases, *root-n consistency*

$$\left.\begin{array}{lcl} \hat{\theta} & \rightarrow & \theta \\ \hat{p}(x) & \rightarrow & p(x) \end{array}\right\} \text{“true” parameters}$$

As the approximation class grows

$$\left\| \hat{p}(x) - f(x) \right\| \rightarrow 0$$

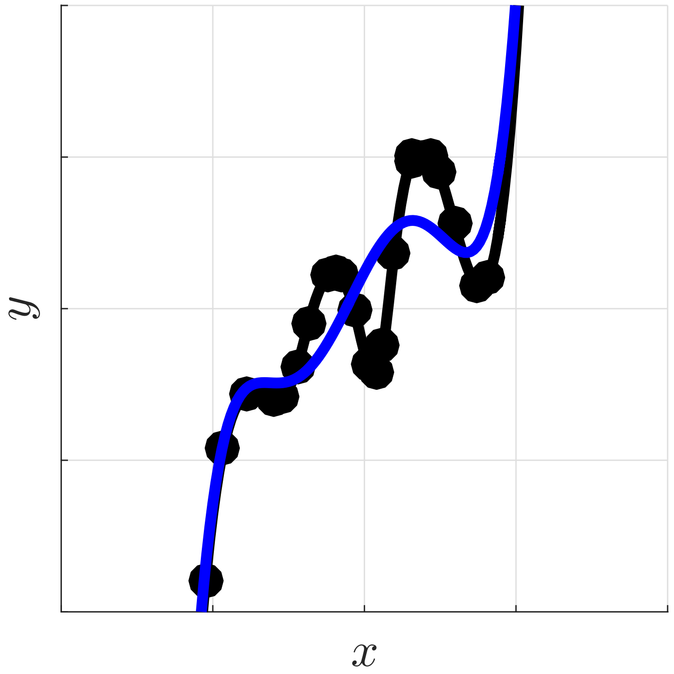
Regression



As data increases, *root-n consistency*

$$\left.\begin{array}{lcl}\hat{\theta} & \to & \theta \\ \hat{p}(x) & \to & p(x)\end{array}\right\} \text{“true” parameters}$$

Approximation

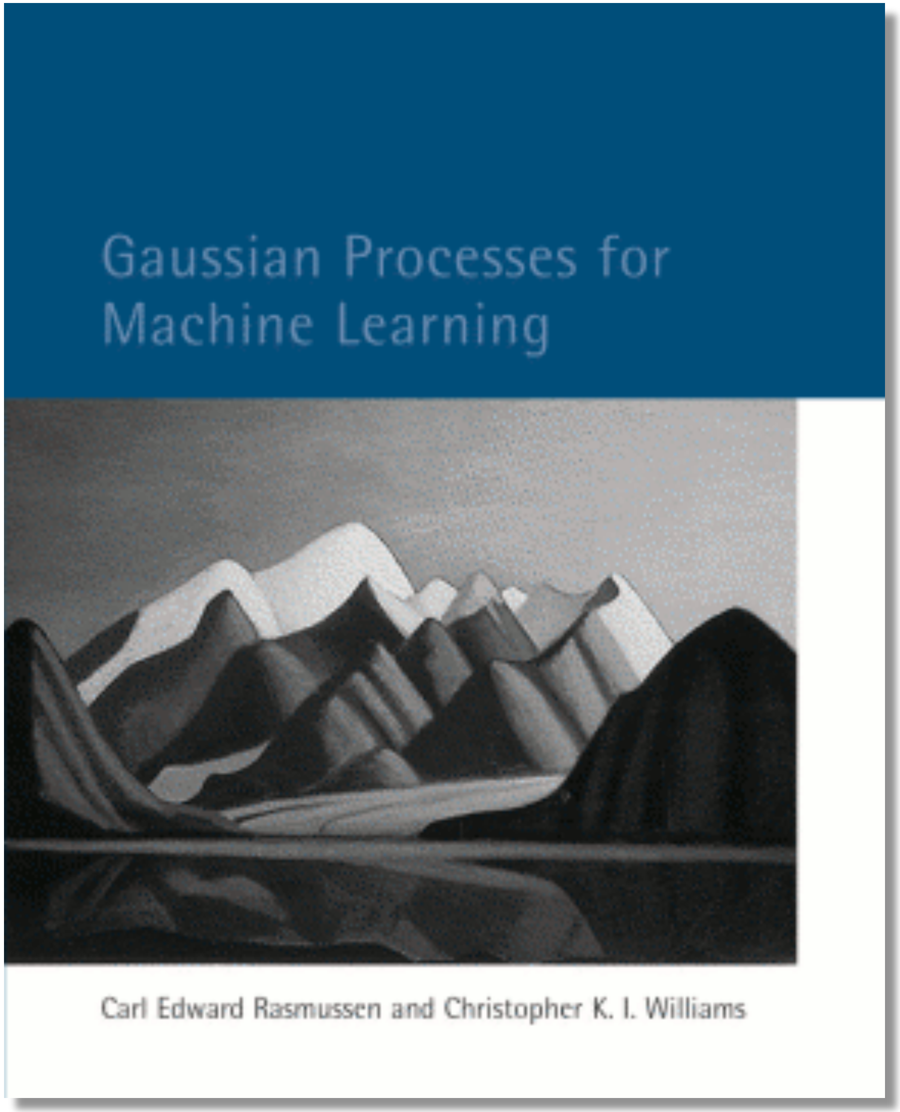


As the approximation class grows

$$\left\| \hat{p}(x) - f(x) \right\| \to 0$$

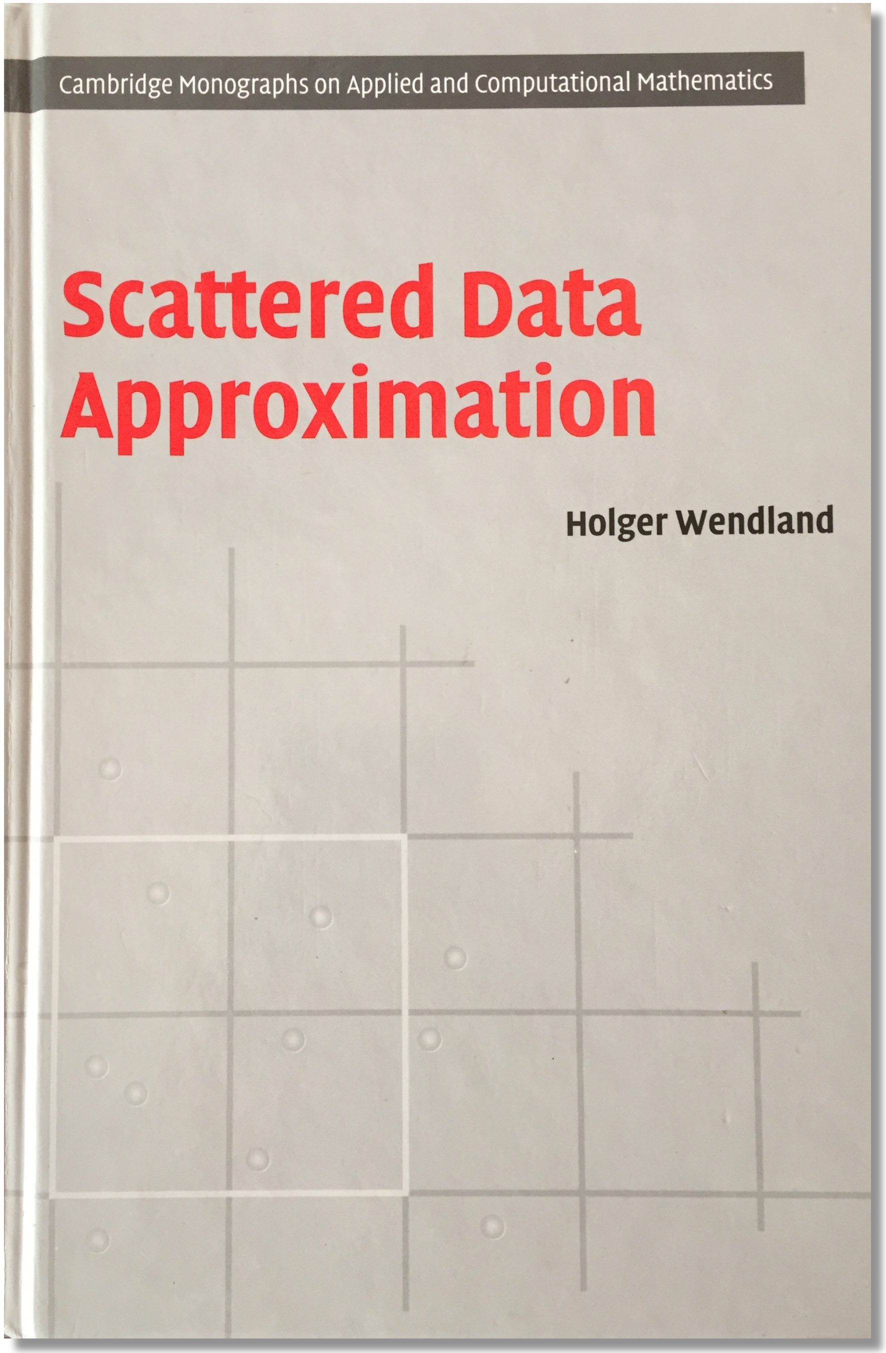
Convergence rate depends on $f(x)$

- high order derivatives
- size of region of analyticity
- Chebyshev coefficients
- …



Gaussian Processes for
Machine Learning

Carl Edward Rasmussen and Christopher K. I. Williams

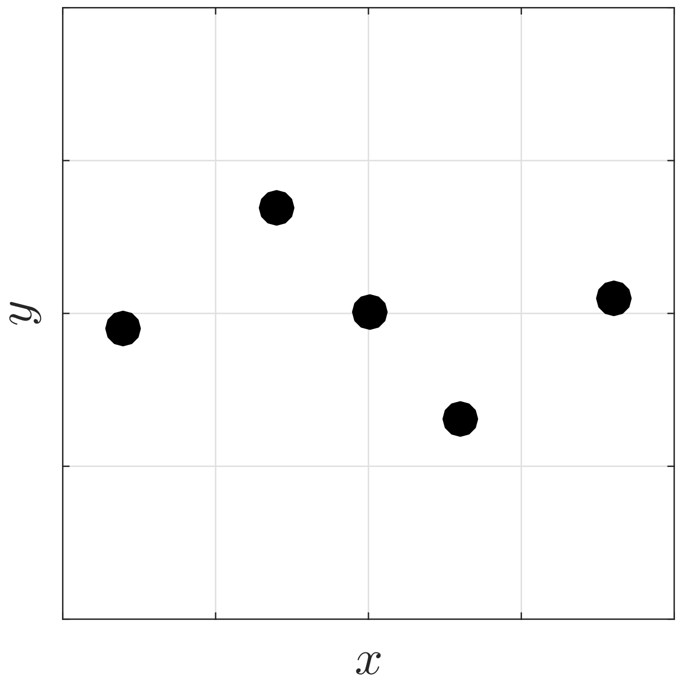


Cambridge Monographs on Applied and Computational Mathematics

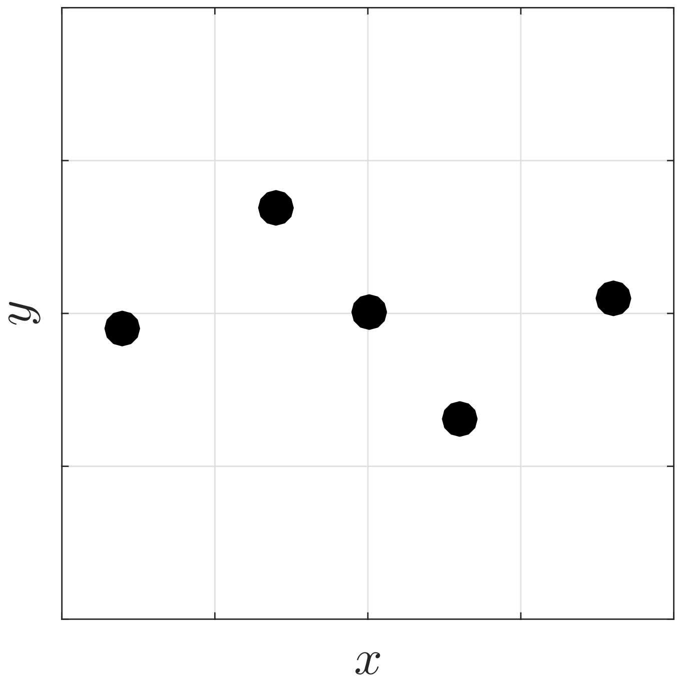
Scattered Data
Approximation

Holger Wendland

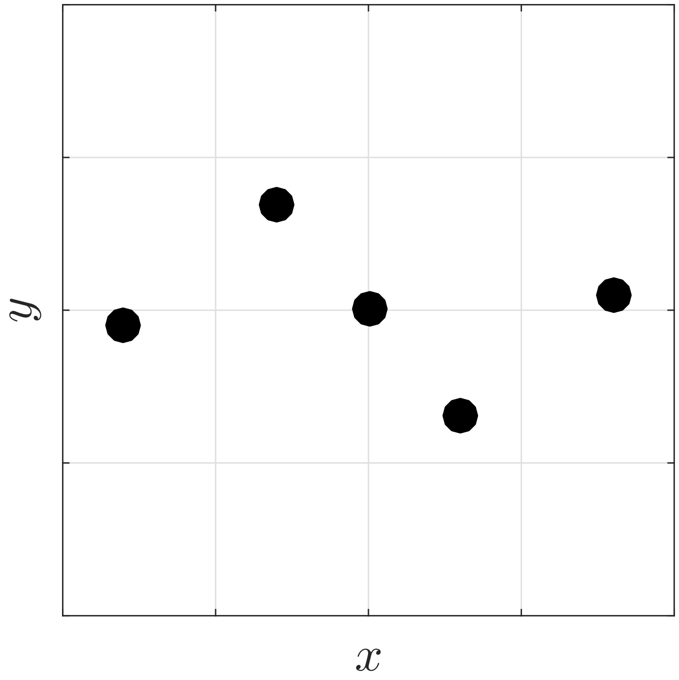
Gaussian process regression



Radial basis approximation



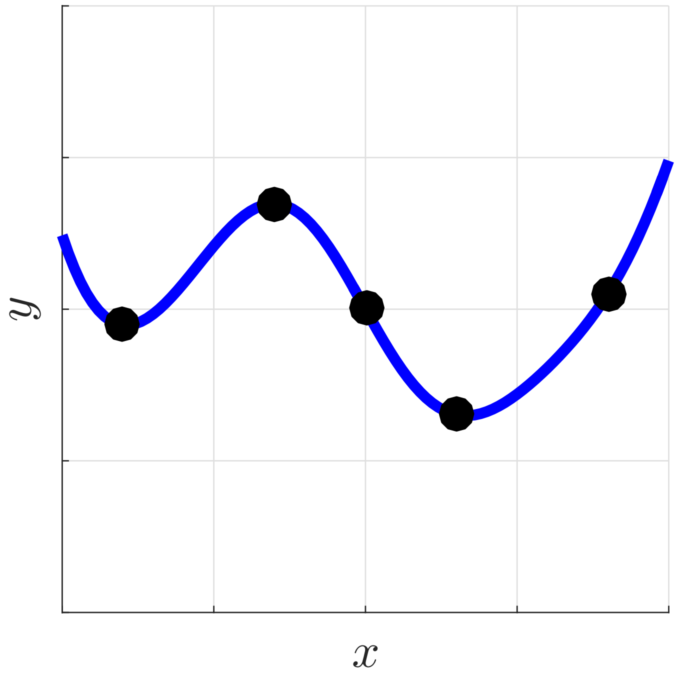
Gaussian process regression



GIVEN

pairs $\{x_i, y_i\}$

Gaussian process regression



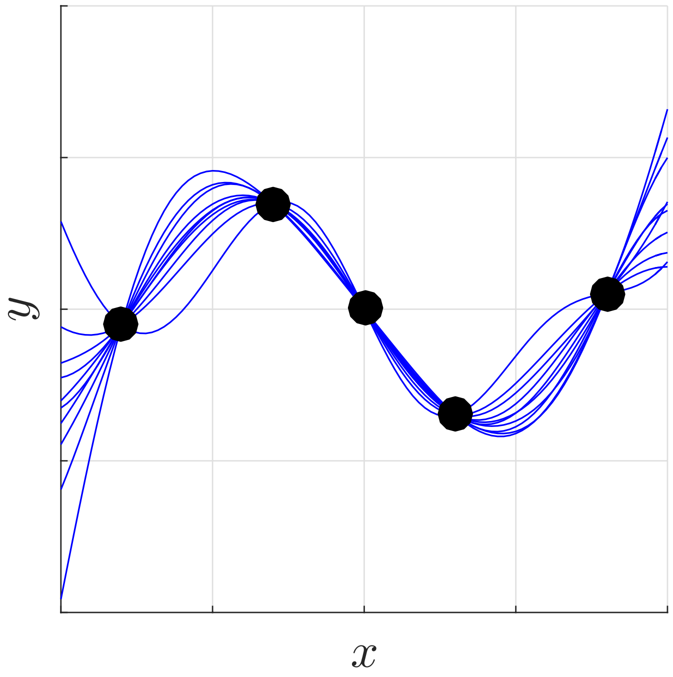
GIVEN

pairs $\{x_i, y_i\}$

ASSUME

$y_i = g(x_i, \omega)$ one realization of a GP

Gaussian process regression



GIVEN

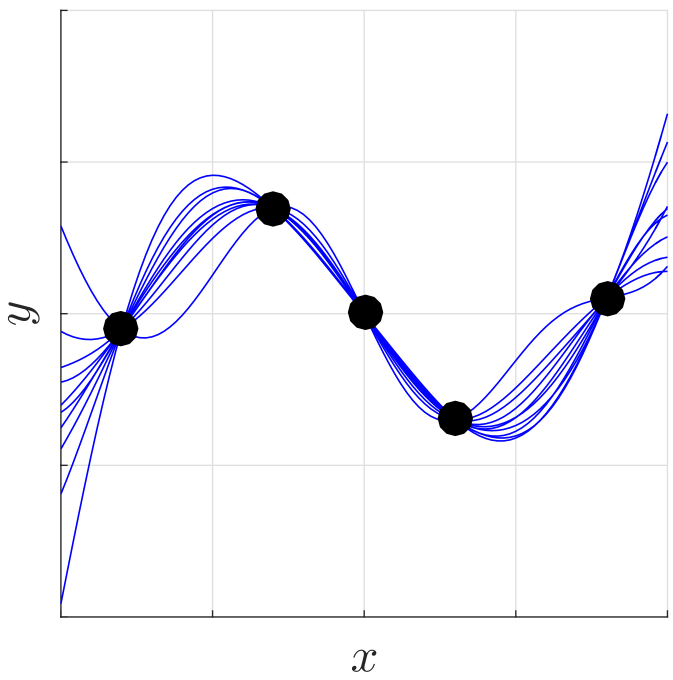
pairs $\{x_i, y_i\}$

ASSUME

$y_i = g(x_i, \omega)$ one realization of a GP

$y_i = g(x_i, \cdot)$ among many possible realizations

Gaussian process regression



CORRELATION MODEL (e.g.)

$$\kappa(x, x';\, \theta) \;=\; \exp(|x - x'|^2/\theta)$$

GIVEN

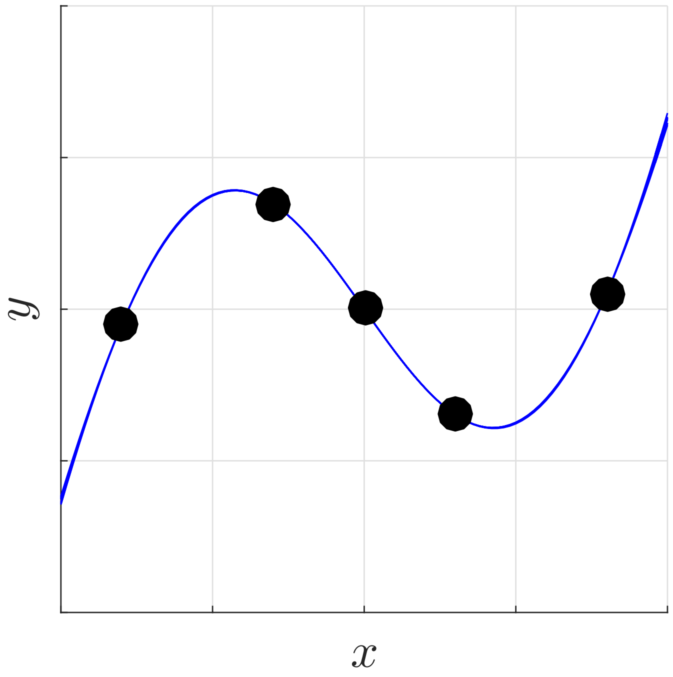
pairs $\{x_i, y_i\}$

ASSUME

$y_i = g(x_i, \omega)$ one realization of a GP

$y_i = g(x_i, \cdot)$ among many possible realizations

Gaussian process regression



CORRELATION MODEL (e.g.)

$$\kappa(x, x'; \theta) \;=\; \exp(|x - x'|^2/\theta)$$

GIVEN

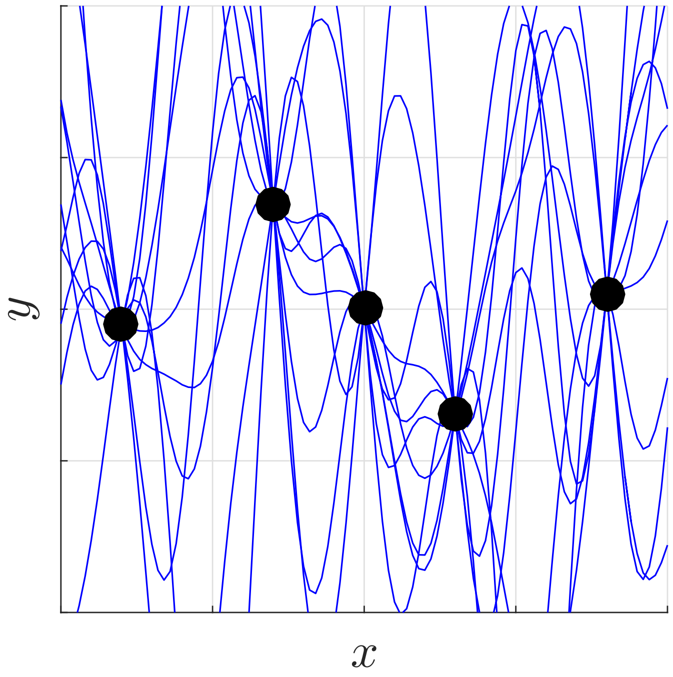
pairs $\{x_i, y_i\}$

ASSUME

$y_i = g(x_i, \omega)$ one realization of a GP

$y_i = g(x_i, \cdot)$ among many possible realizations

Gaussian process regression



CORRELATION MODEL (e.g.)

$$\kappa(x, x'; \theta) \;=\; \exp(|x - x'|^2/\theta)$$

GIVEN

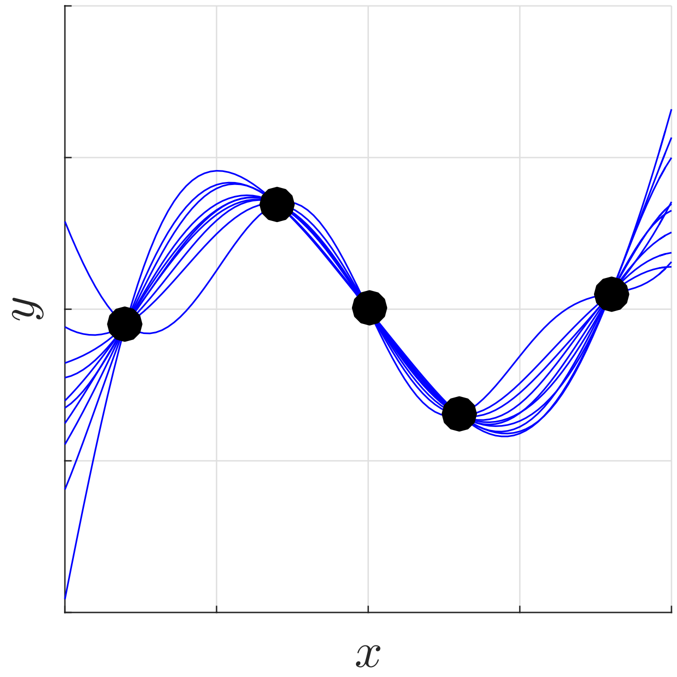
pairs $\{x_i, y_i\}$

ASSUME

$y_i = g(x_i, \omega)$ one realization of a GP

$y_i = g(x_i, \cdot)$ among many possible realizations

Gaussian process regression



CORRELATION MODEL (e.g.)

$$\kappa(x, x'; \theta) \;=\; \exp(|x - x'|^2/\theta)$$

FIT

$$\underset{\theta}{\text{maximize}} \;\; \text{likelihood}(\theta; \{x_i, y_i\})$$

GIVEN

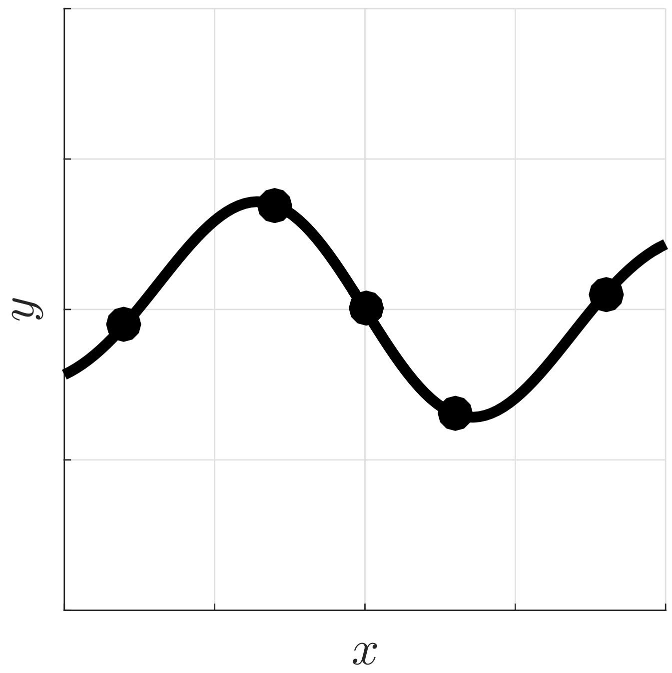
pairs $\{x_i, y_i\}$

ASSUME

$y_i = g(x_i, \omega)$ one realization of a GP

$y_i = g(x_i, \cdot)$ among many possible realizations

Gaussian process regression



GIVEN

pairs $\{x_i, y_i\}$

ASSUME

$y_i = g(x_i, \omega)$ one realization of a GP

$y_i = g(x_i, \cdot)$ among many possible realizations

CORRELATION MODEL (e.g.)

$$\kappa(x, x'; \theta) \;=\; \exp(|x - x'|^2/\theta)$$

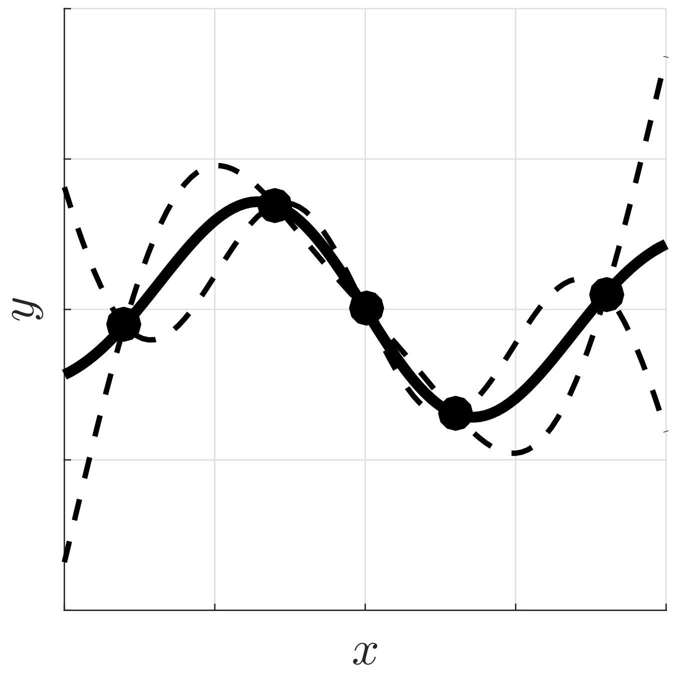
FIT

$$\underset{\theta}{\text{maximize}} \;\; \text{likelihood}(\theta; \{x_i, y_i\})$$

PREDICT (B.L.U.E.)

$$\begin{aligned} y(x) &= \mathbb{E}[\, g(x, \cdot) \,|\, \{x_i, y_i\} \,] \\ &= \sum_i y_i \, w_i(x) \end{aligned}$$

Gaussian process regression



GIVEN

pairs $\{x_i, y_i\}$

ASSUME

$y_i = g(x_i, \omega)$ one realization of a GP

$y_i = g(x_i, \cdot)$ among many possible realizations

CORRELATION MODEL (e.g.)

$$\kappa(x, x'; \theta) \; = \; \exp(|x - x'|^2/\theta)$$

FIT

$$\underset{\theta}{\text{maximize}} \;\; \text{likelihood}(\theta; \{x_i, y_i\})$$

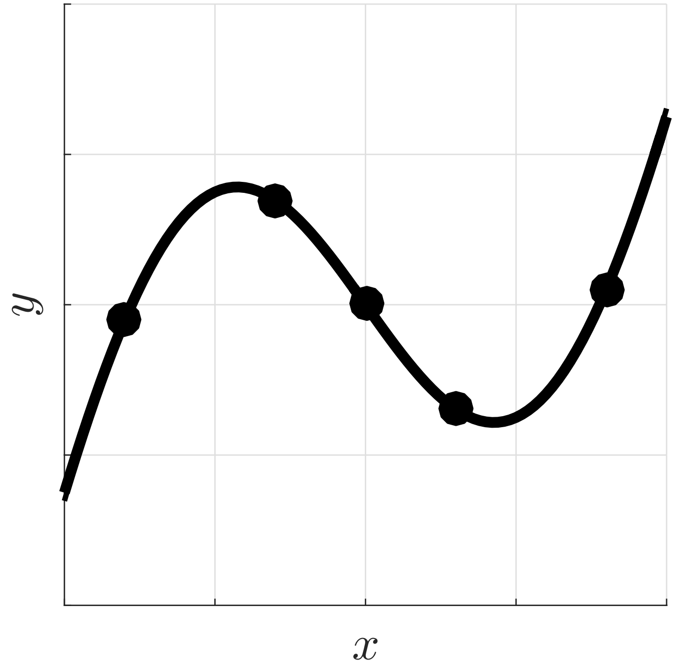
PREDICT (B.L.U.E.)

$$\begin{aligned} y(x) &= \mathbb{E}[\, g(x, \cdot) \,|\, \{x_i, y_i\}\,] \\ &= \sum_i y_i \, w_i(x) \end{aligned}$$

QUANTIFY UNCERTAINTY

$$\mathbb{V}\text{ar}[\, g(x, \cdot) \,|\, \{x_i, y_i\}\,] \; = \; \text{“formula”}$$

Gaussian process regression



GIVEN

pairs $\{x_i, y_i\}$

ASSUME

$y_i = g(x_i, \omega)$ one realization of a GP

$y_i = g(x_i, \cdot)$ among many possible realizations

CORRELATION MODEL (e.g.)

$$\kappa(x, x'; \theta) \;=\; \exp(|x - x'|^2/\theta)$$

FIT

$$\underset{\theta}{\text{maximize}} \;\; \text{likelihood}(\theta; \{x_i, y_i\})$$

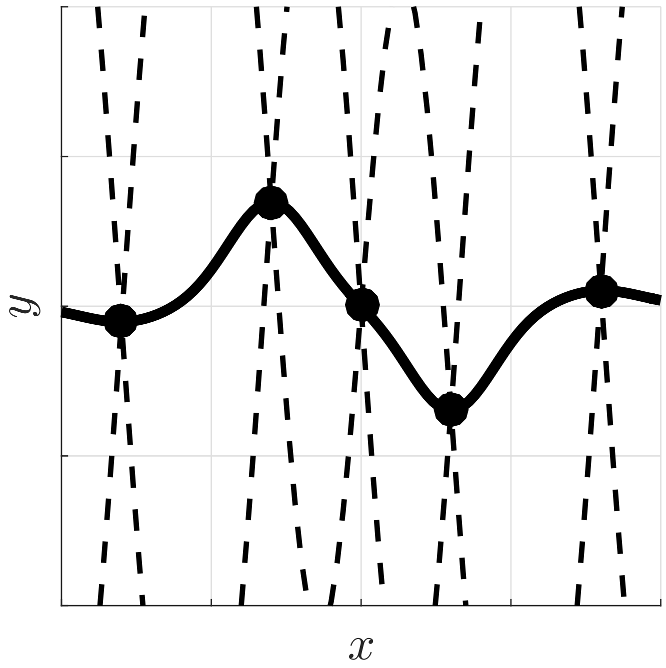
PREDICT (B.L.U.E.)

$$\begin{aligned} y(x) &= \mathbb{E}[\, g(x, \cdot) \,|\, \{x_i, y_i\} \,] \\ &= \sum_i y_i \, w_i(x) \end{aligned}$$

QUANTIFY UNCERTAINTY

$$\mathbb{V}\text{ar}[\, g(x, \cdot) \,|\, \{x_i, y_i\} \,] \;=\; \text{``formula''}$$

Gaussian process regression



GIVEN

pairs $\{x_i, y_i\}$

ASSUME

$y_i = g(x_i, \omega)$ one realization of a GP

$y_i = g(x_i, \cdot)$ among many possible realizations

CORRELATION MODEL (e.g.)

$$\kappa(x, x'; \theta) \;=\; \exp(|x - x'|^2/\theta)$$

FIT

$$\underset{\theta}{\text{maximize}} \;\; \text{likelihood}(\theta; \{x_i, y_i\})$$

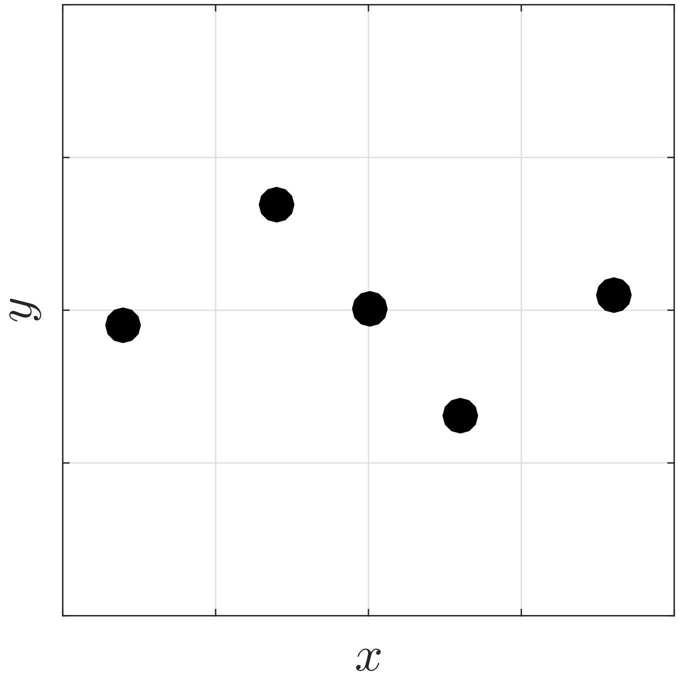
PREDICT (B.L.U.E.)

$$\begin{aligned} y(x) &= \mathbb{E}[\, g(x, \cdot) \,|\, \{x_i, y_i\} \,] \\ &= \sum_i y_i \, w_i(x) \end{aligned}$$

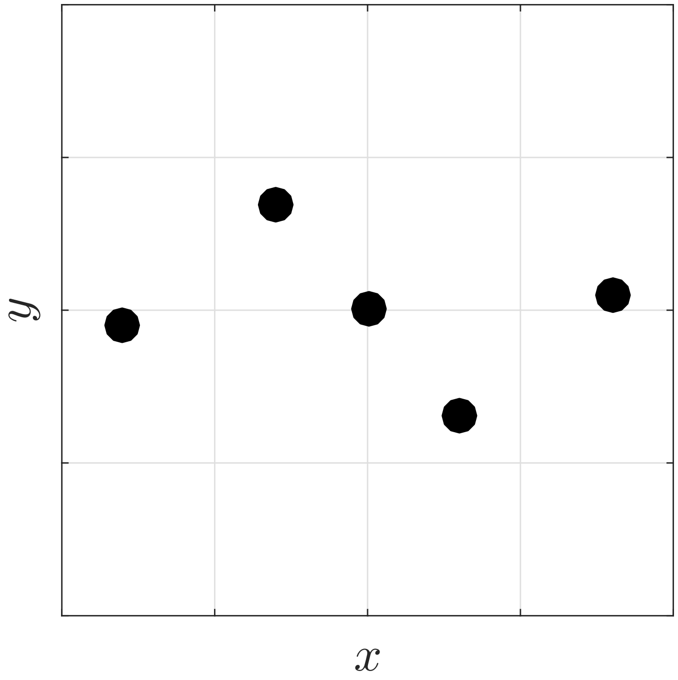
QUANTIFY UNCERTAINTY

$$\mathbb{V}\text{ar}[\, g(x, \cdot) \,|\, \{x_i, y_i\} \,] \;=\; \text{“formula”}$$

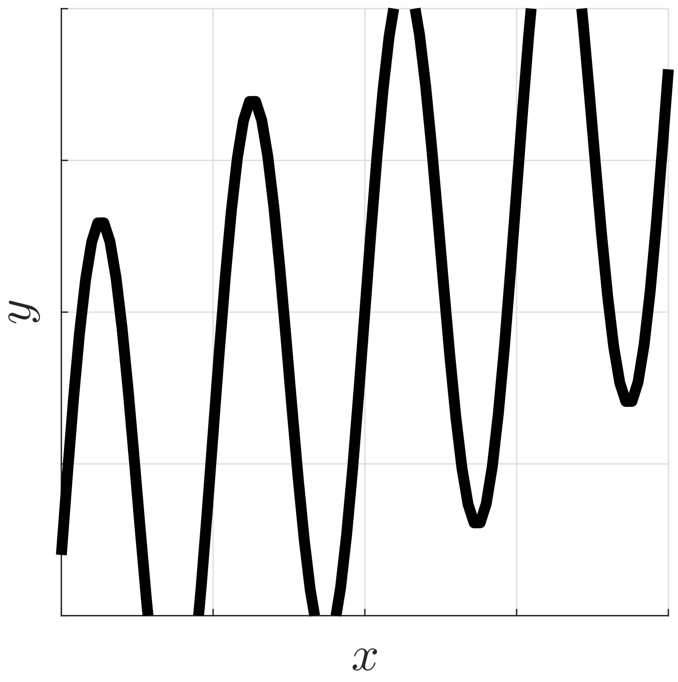
Gaussian process regression



Radial basis approximation



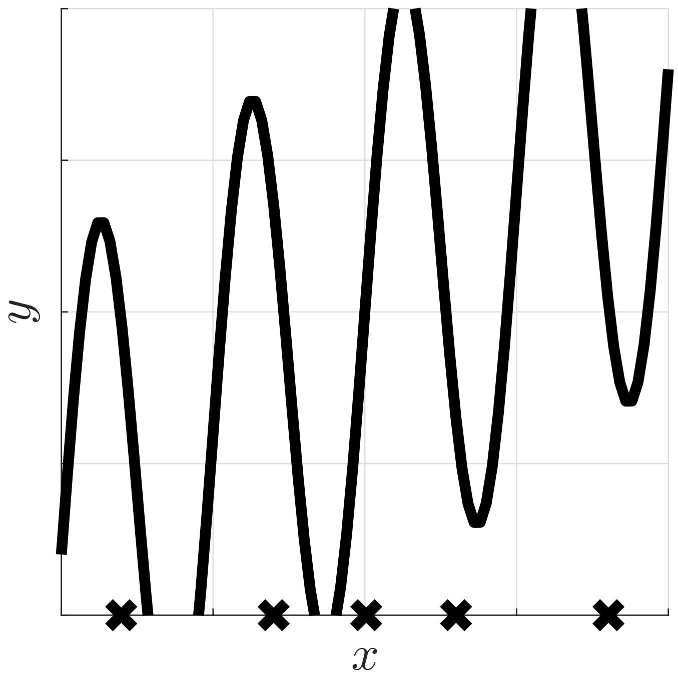
Radial basis approximation



GIVEN

a queryable function $f(x)$

Radial basis approximation

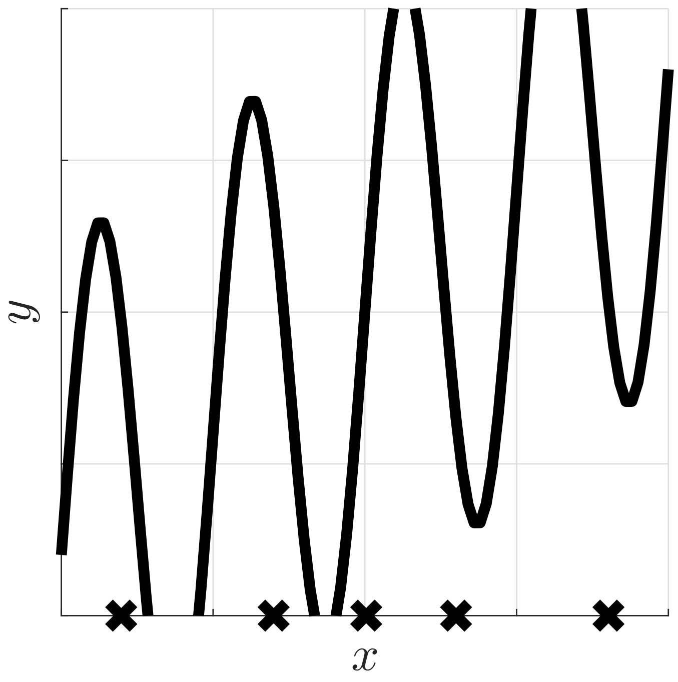


GIVEN

a queryable function $f(x)$

centers $x_1, \ldots, x_n$

Radial basis approximation



GIVEN

a queryable function $f(x)$

centers $x_1, \ldots, x_n$

GOAL

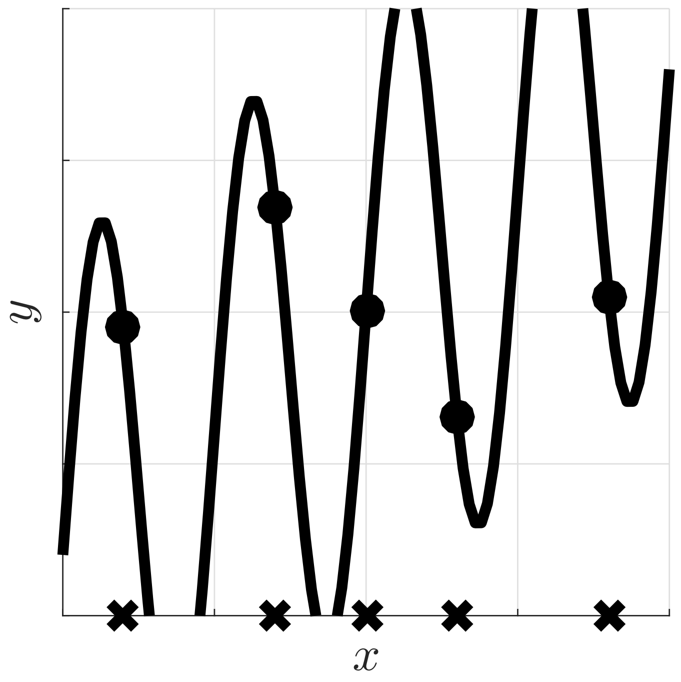
find $s(x)$ such that

the error $\|s - f\|$ is small

Radial basis approximation

QUERY THE FUNCTION

$$y_i = f(x_i)$$



GIVEN

a queryable function $f(x)$

centers $x_1, \ldots, x_n$

GOAL

find $s(x)$ such that

the error $\|s - f\|$ is small

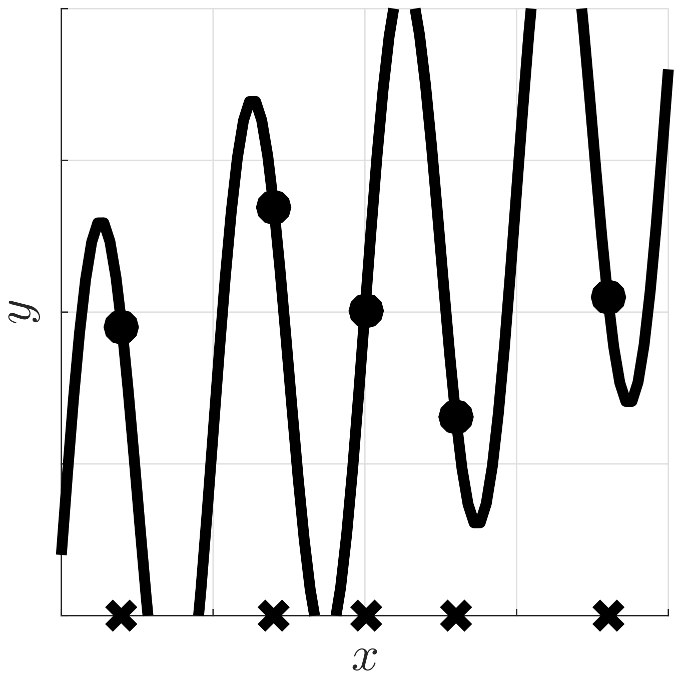
Radial basis approximation

QUERY THE FUNCTION

$$y_i = f(x_i)$$

CHOOSE KERNEL

$$\kappa(x, x';\, \varepsilon) = \exp(\varepsilon\, |x - x'|^2)$$



GIVEN

a queryable function $f(x)$

centers $x_1, \ldots, x_n$

GOAL

find $s(x)$ such that

the error $\|s - f\|$ is small

Radial basis approximation

QUERY THE FUNCTION

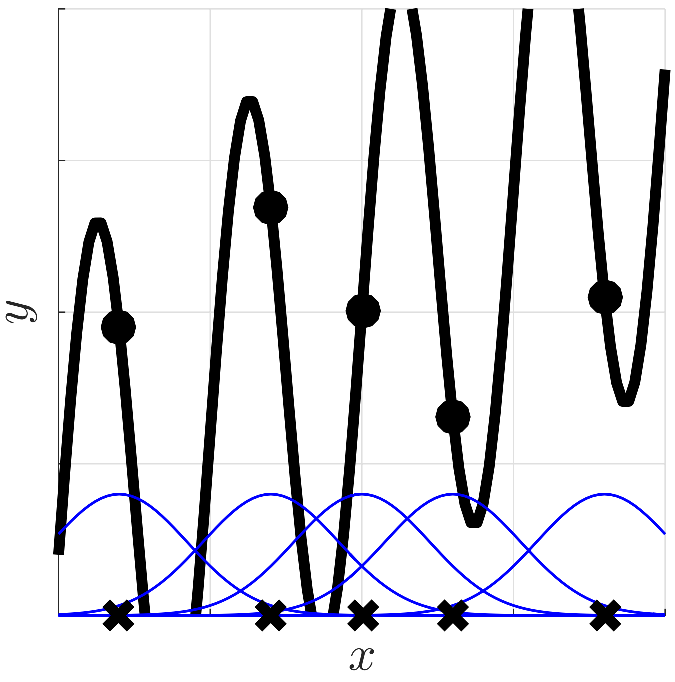
$$y_i = f(x_i)$$

CHOOSE KERNEL

$$\kappa(x, x'; \varepsilon) = \exp(\varepsilon\,|x - x'|^2)$$

DEFINES BASIS

$$\phi_i(x) = \kappa(x, x_i; \varepsilon)$$



GIVEN

a queryable function $f(x)$

centers $x_1, \ldots, x_n$

GOAL

find $s(x)$ such that

the error $\|s - f\|$ is small

Radial basis approximation

QUERY THE FUNCTION

$$y_i = f(x_i)$$

CHOOSE KERNEL

$$\kappa(x, x'; \varepsilon) = \exp(\varepsilon\, |x - x'|^2)$$

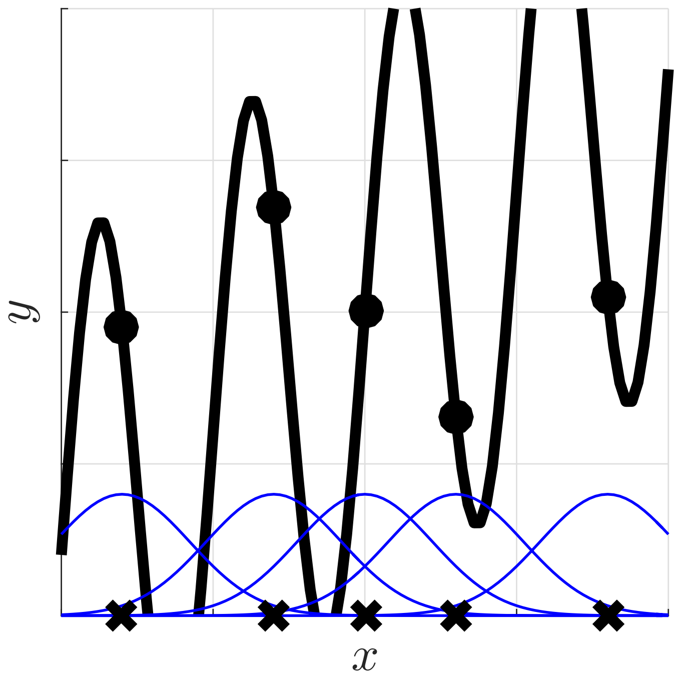
DEFINES BASIS

$$\phi_i(x) = \kappa(x, x_i; \varepsilon)$$

COMPUTE COEFFICIENTS

$$\boldsymbol{K}\mathbf{a} = \mathbf{f}$$

$$\boldsymbol{K}_{ij} = \kappa(x_i, x_j),\ \mathbf{f}_i = f(x_i)$$



GIVEN

a queryable function $f(x)$

centers $x_1, \ldots, x_n$

GOAL

find $s(x)$ such that

the error $\|s - f\|$ is small

Radial basis approximation

QUERY THE FUNCTION

$$y_i = f(x_i)$$

CHOOSE KERNEL

$$\kappa(x, x';\, \varepsilon) = \exp(\varepsilon\, |x - x'|^2)$$

DEFINES BASIS

$$\phi_i(x) = \kappa(x, x_i;\, \varepsilon)$$

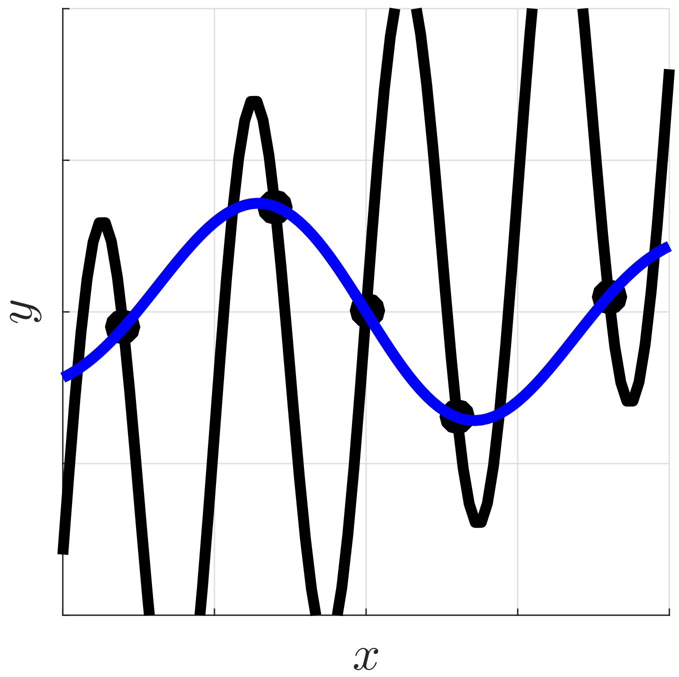
COMPUTE COEFFICIENTS

$$\boldsymbol{K}\mathbf{a} = \mathbf{f}$$

$$\boldsymbol{K}_{ij} = \kappa(x_i, x_j),\ \mathbf{f}_i = f(x_i)$$

PREDICT

$$s(x) = \sum_i a_i\, \phi_i(x)$$



GIVEN

a queryable function $f(x)$

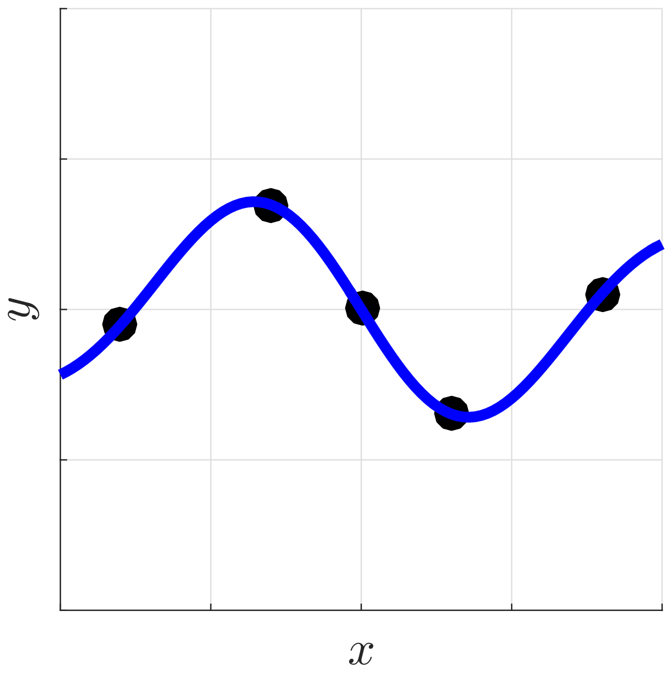
centers $x_1, \ldots, x_n$

GOAL

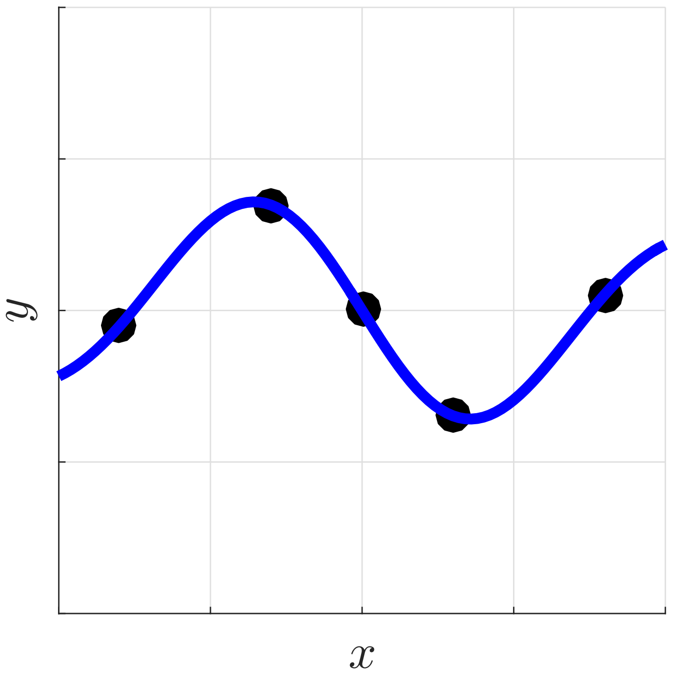
find $s(x)$ such that

the error $\|s - f\|$ is small

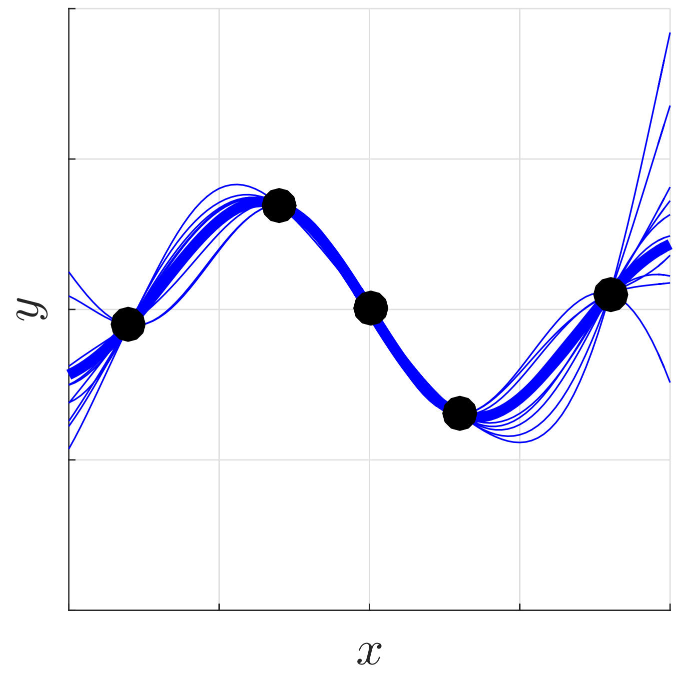
Gaussian process regression



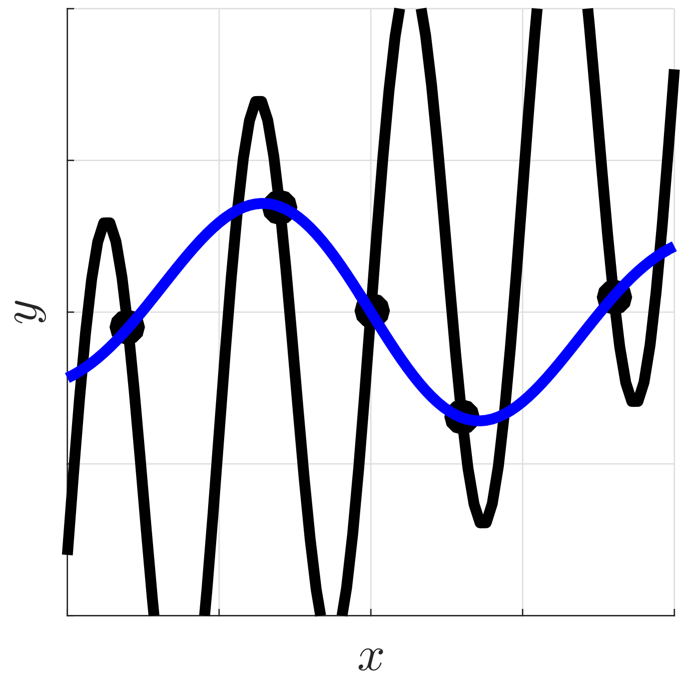
Radial basis approximation



Gaussian process regression



Radial basis approximation



The **story** of the data and fitted curve is **different**.
But does it matter? **YES**

Comments on error and convergence

GP conditional variance is **NOT** error (except possibly under some very specific conditions).

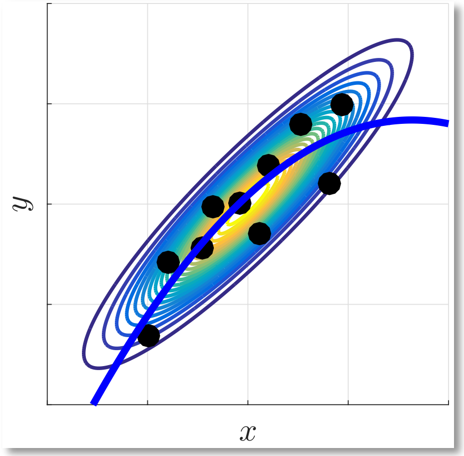
RBF error estimates are *asymptotic* in the *fill distance.*

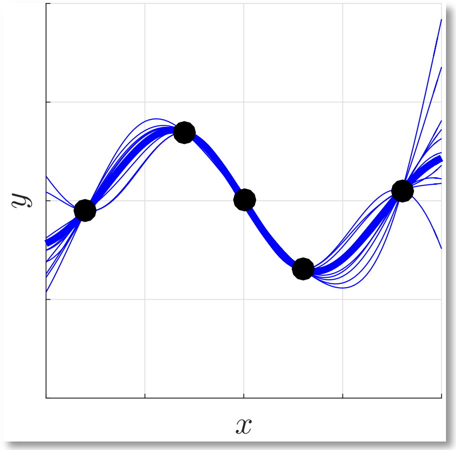
Both approaches make practically unverifiable assumptions about the origin of the data generating function/process.

As a caricature:

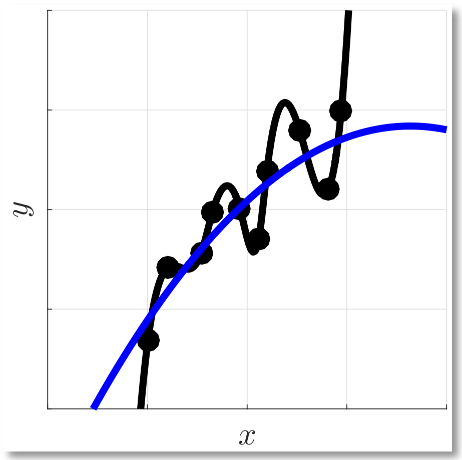
- statisticians try to reduce the error by finding a better model
  - (e.g., solve the fitting problem better)
- mathematicians try to reduce the error with more queries
  - (i.e., sample into asymptopia)

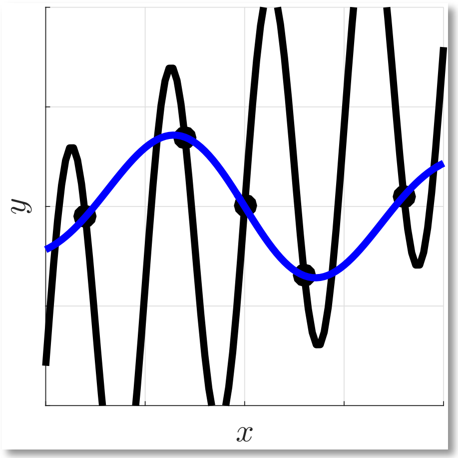
Regression





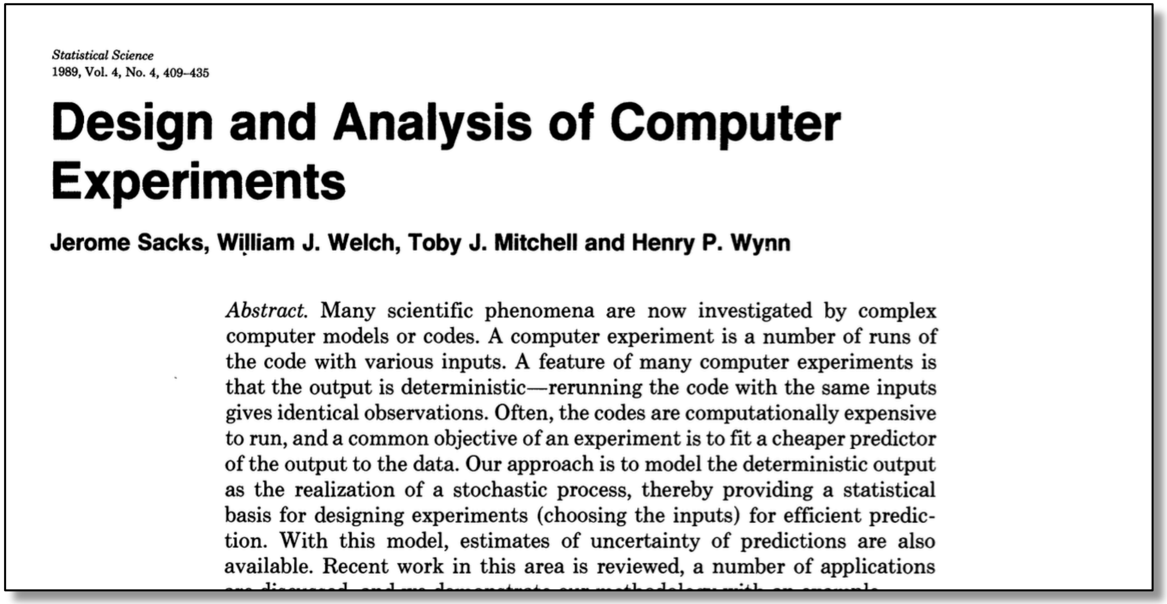
Approximation





Which one is a *computer simulation?*

What is error in a computer simulation?



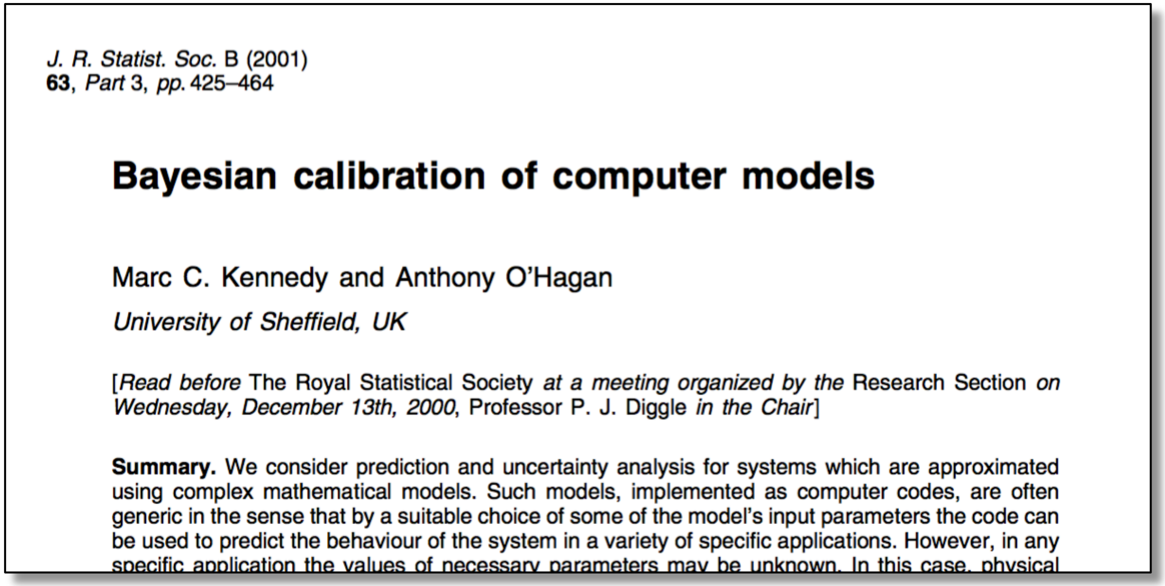
Statistical Science
1989, Vol. 4, No. 4, 409–435

Design and Analysis of Computer Experiments

Jerome Sacks, William J. Welch, Toby J. Mitchell and Henry P. Wynn

Abstract. Many scientific phenomena are now investigated by complex computer models or codes. A computer experiment is a number of runs of the code with various inputs. A feature of many computer experiments is that the output is deterministic—rerunning the code with the same inputs gives identical observations. Often, the codes are computationally expensive to run, and a common objective of an experiment is to fit a cheaper predictor of the output to the data. Our approach is to model the deterministic output as the realization of a stochastic process, thereby providing a statistical basis for designing experiments (choosing the inputs) for efficient prediction. With this model, estimates of uncertainty of predictions are also available. Recent work in this area is reviewed, a number of applications

Sacks et al. (1989)



J. R. Statist. Soc. B (2001)
63, *Part* 3, *pp.* 425–464

Bayesian calibration of computer models

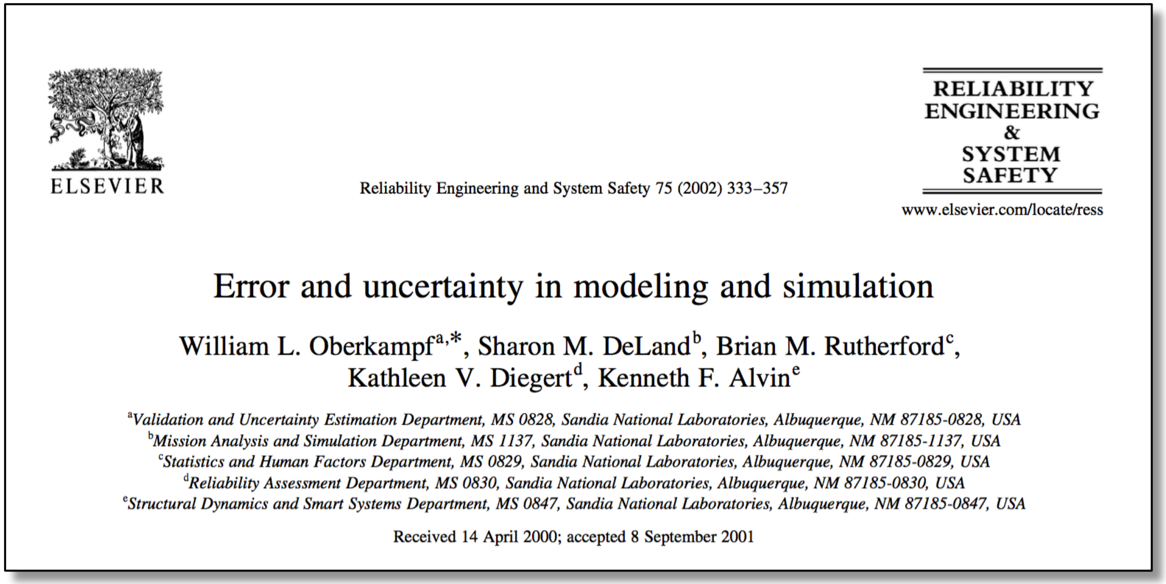
Marc C. Kennedy and Anthony O'Hagan
University of Sheffield, UK

[*Read before* The Royal Statistical Society *at a meeting organized by the* Research Section *on Wednesday, December 13th, 2000,* Professor P. J. Diggle *in the Chair*]

Summary. We consider prediction and uncertainty analysis for systems which are approximated using complex mathematical models. Such models, implemented as computer codes, are often generic in the sense that by a suitable choice of some of the model's input parameters the code can be used to predict the behaviour of the system in a variety of specific applications. However, in any specific application the values of necessary parameters may be unknown. In this case, physical

Kennedy and O'Hagan (2001)

What is error in a computer simulation?



ELSEVIER

Reliability Engineering and System Safety 75 (2002) 333–357

RELIABILITY
ENGINEERING
&
SYSTEM
SAFETY

www.elsevier.com/locate/ress

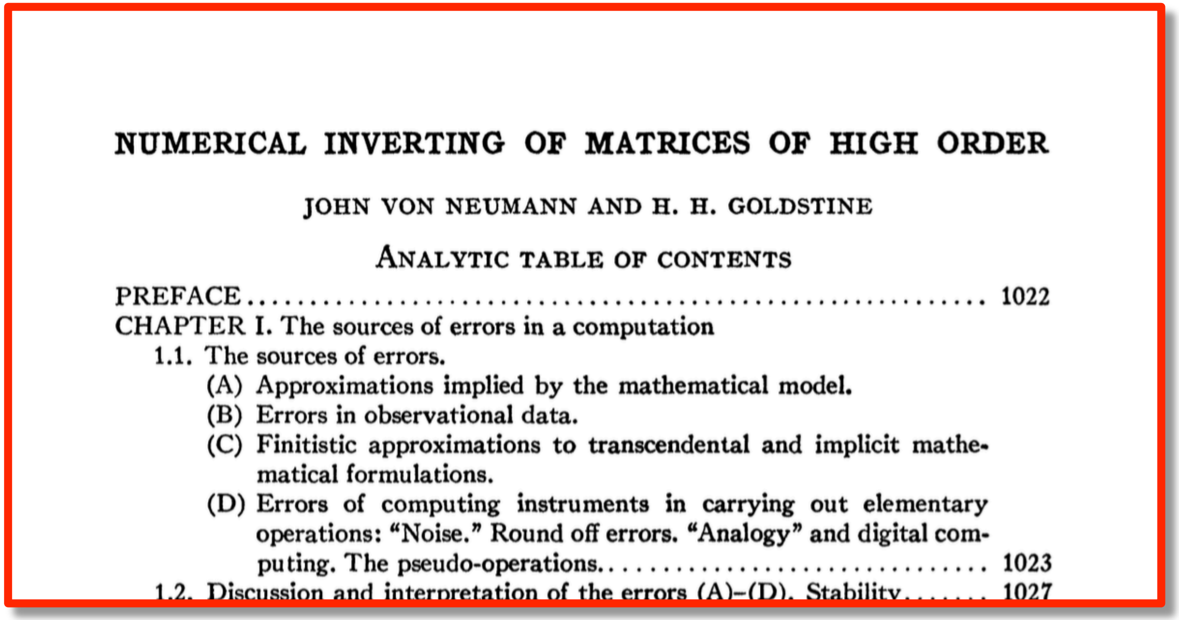
Error and uncertainty in modeling and simulation

William L. Oberkampf[a,*], Sharon M. DeLand[b], Brian M. Rutherford[c],
Kathleen V. Diegert[d], Kenneth F. Alvin[e]

[a]*Validation and Uncertainty Estimation Department, MS 0828, Sandia National Laboratories, Albuquerque, NM 87185-0828, USA*
[b]*Mission Analysis and Simulation Department, MS 1137, Sandia National Laboratories, Albuquerque, NM 87185-1137, USA*
[c]*Statistics and Human Factors Department, MS 0829, Sandia National Laboratories, Albuquerque, NM 87185-0829, USA*
[d]*Reliability Assessment Department, MS 0830, Sandia National Laboratories, Albuquerque, NM 87185-0830, USA*
[e]*Structural Dynamics and Smart Systems Department, MS 0847, Sandia National Laboratories, Albuquerque, NM 87185-0847, USA*

Received 14 April 2000; accepted 8 September 2001

Oberkampf et al. (2002)



NUMERICAL INVERTING OF MATRICES OF HIGH ORDER

JOHN VON NEUMANN AND H. H. GOLDSTINE

ANALYTIC TABLE OF CONTENTS

von Neumann and Goldstine
Bulletin of the AMS (**1947**)

[h/t Joe Grcar]

The von Neumann and Goldstine Catechism

"This analysis of the sources of errors should be objective and strict inasmuch as completeness is concerned, but when it comes to the defining, classifying, and separating of the sources, a certain subjectiveness and arbitrariness is unavoidable. With these reservations, the following enumeration and classification of sources of errors seems to be adequate and reasonable."

Mathematical model

Observations and parameters

Finitistic approximations

Round-off

The von Neumann and Goldstine Catechism

"This analysis of the sources of errors should be objective and strict inasmuch as completeness is concerned, but when it comes to the defining, classifying, and separating of the sources, a certain subjectiveness and arbitrariness is unavoidable. With these reservations, the following enumeration and classification of sources of errors seems to be adequate and reasonable."

Mathematical model

Observations and parameters

Finitistic approximations

Round-off

NOTES

How well math model approximates reality

Model-form error

The von Neumann and Goldstine Catechism

“This analysis of the sources of errors should be objective and strict inasmuch as completeness is concerned, but when it comes to the defining, classifying, and separating of the sources, a certain subjectiveness and arbitrariness is unavoidable. With these reservations, the following enumeration and classification of sources of errors seems to be adequate and reasonable.”

Mathematical model

Observations and parameters

Finitistic approximations

Round-off

NOTES

Forward and inverse UQ

Most of the UQ methods literature

The von Neumann and Goldstine Catechism

"This analysis of the sources of errors should be objective and strict inasmuch as completeness is concerned, but when it comes to the defining, classifying, and separating of the sources, a certain subjectiveness and arbitrariness is unavoidable. With these reservations, the following enumeration and classification of sources of errors seems to be adequate and reasonable."

Mathematical model

Observations and parameters

Finitistic approximations

Round-off

NOTES

Asymptotics from classical numerical analysis

Deterministic numerical **noise**

"Computational noise in deterministic simulations is as ill-defined a concept as can be found in scientific computing."

Moré and Wild (2011)

Summary thoughts

Computer models are deterministic. In my opinion, approximation tools are better suited.

But computational noise is really annoying, if you take it seriously.

LOTS of fundamental research opportunities for applying statistical methods to noise-less data---i.e., the approximation setting.

What does *Bayes* have to do with it?

Practical advice

Everyone, civilized conversation and argumentation!!!

Statisticians, include numerical experiments that demonstrate asymptotic convergence of testing error.

Numerical analysts, convergence analysis of statistical standard error and bootstrap standard error in the context of constructive approximation.

Write three review papers:

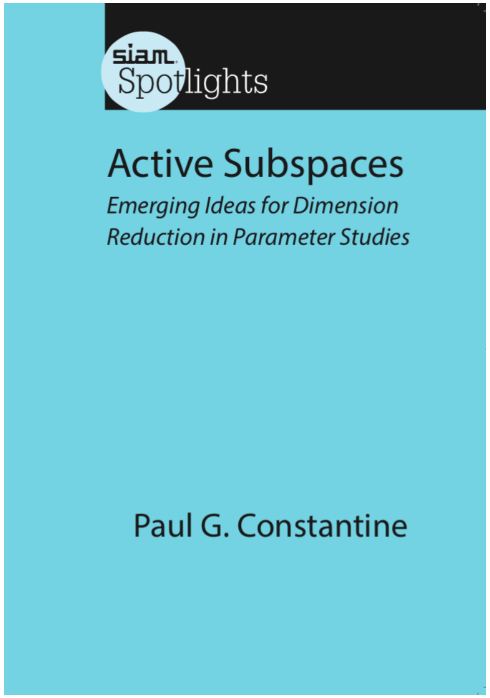
- Regression for numerical analysts
- Approximation for statisticians
- Reconciling perspective with authors from both communities

QUESTIONS?

Why should we care?

What do you do in practice?

Active Subspaces
SIAM (2015)



PAUL CONSTANTINE
Assistant Professor
University of Colorado Boulder
`activesubspaces.org`
`@DrPaulynomial`