60 Years

IAEA

Atoms for Peace and Development

Probabilistic methods for nuclear data

Georg Schnabel
g.schnabel [at] iaea.org

Nuclear Data Section
Division of Physical and Chemical Sciences NAPC
Department for Nuclear Sciences and Applications
IAEA, Vienna

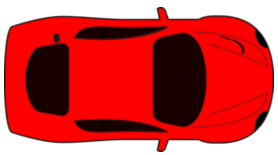
Workshop on Computational Nuclear Science and Engineering

23 May 2022

Short bio

- Studied physics at TU Vienna
- PhD in nuclear data evaluation 2015
- Postdoc at CEA Saclay (2015-2018) and Uppsala University (2018-2019)
- Since 2020 nuclear physicist at Nuclear Data Section of the IAEA dealing with nuclear data library projects, method and code development

Challenge #1: Mars rover



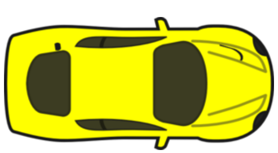
A self-driving vehicle should follow a predefined path on Mars

Assumptions:
Vehicle *approximately* aware of initial position and knows *approximately* speed and direction of movement at any point in time afterwards.

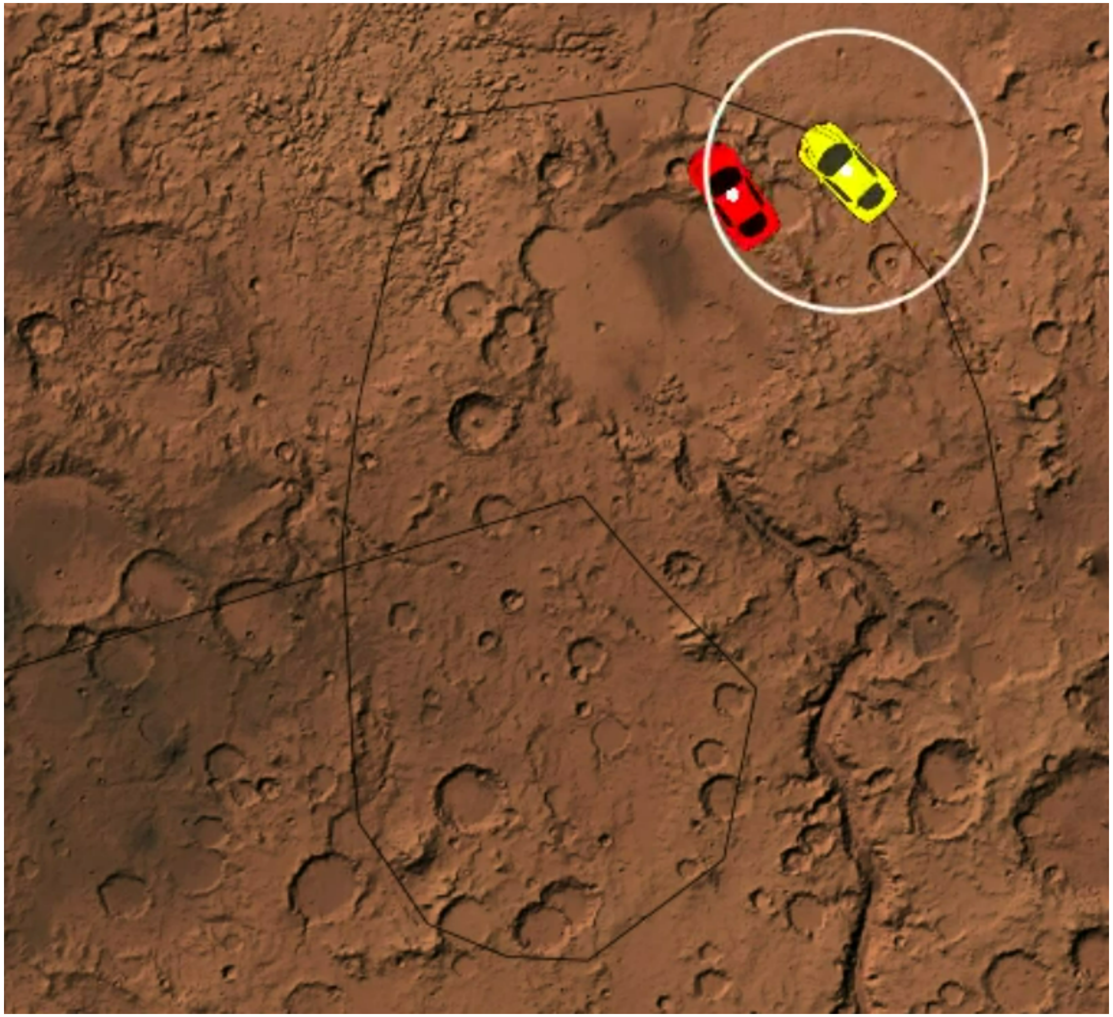
Task:
Determine current position by $x_0 = x_{old} + v*\Delta t$ and adjust speed and direction to stay on desired trajectory

Problem:
Uncertainty about position puts the rover at risk to fall into a deep crater and get damaged

ghost car

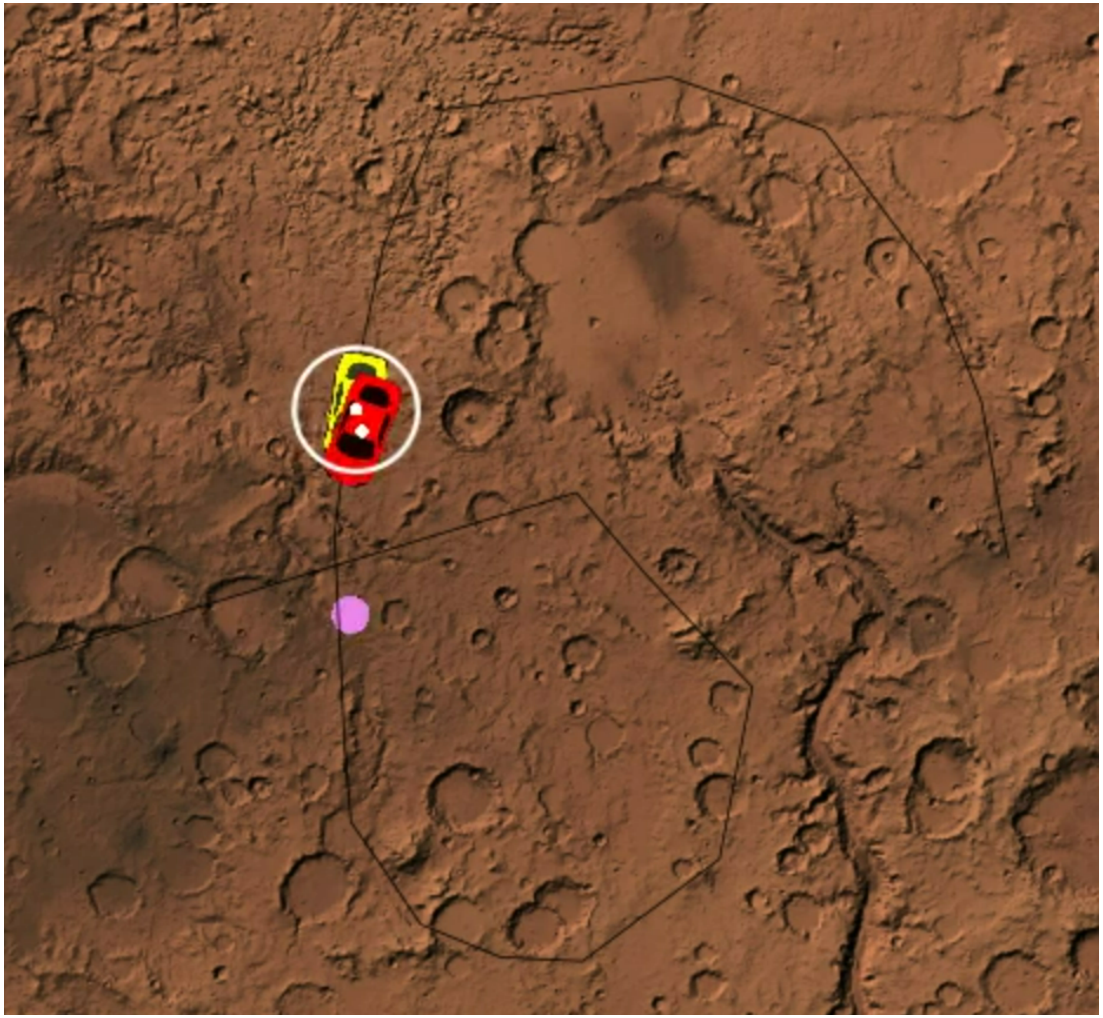


Truth-Estimate divergence



Link to movie

Make use of GPS

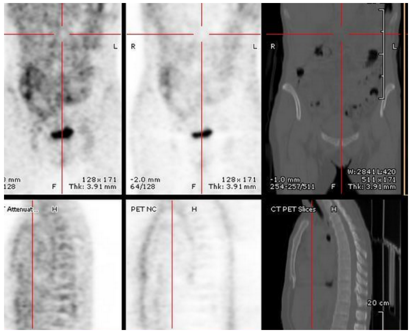


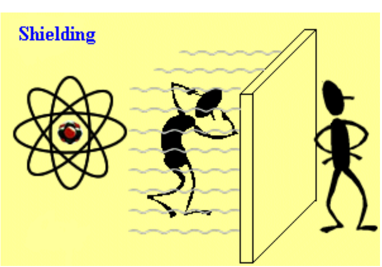
Link to movie

Challenge #2: Face reconstruction

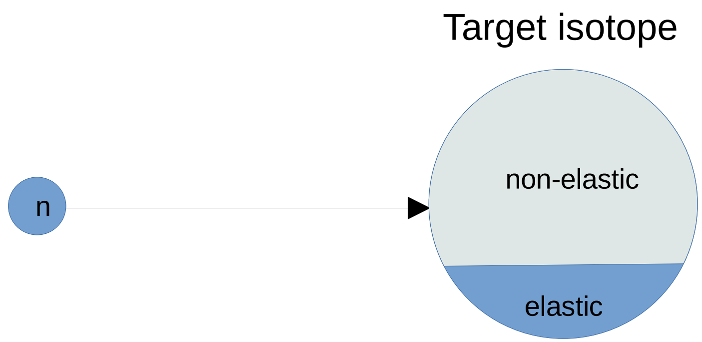
Taken from the A&T face dataset

Nuclear data



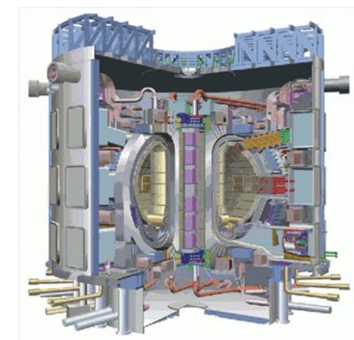


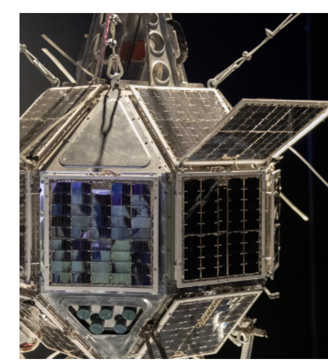
Probabilities of various nuclear interactions involving the atomic nuclei, e.g., cross sections.



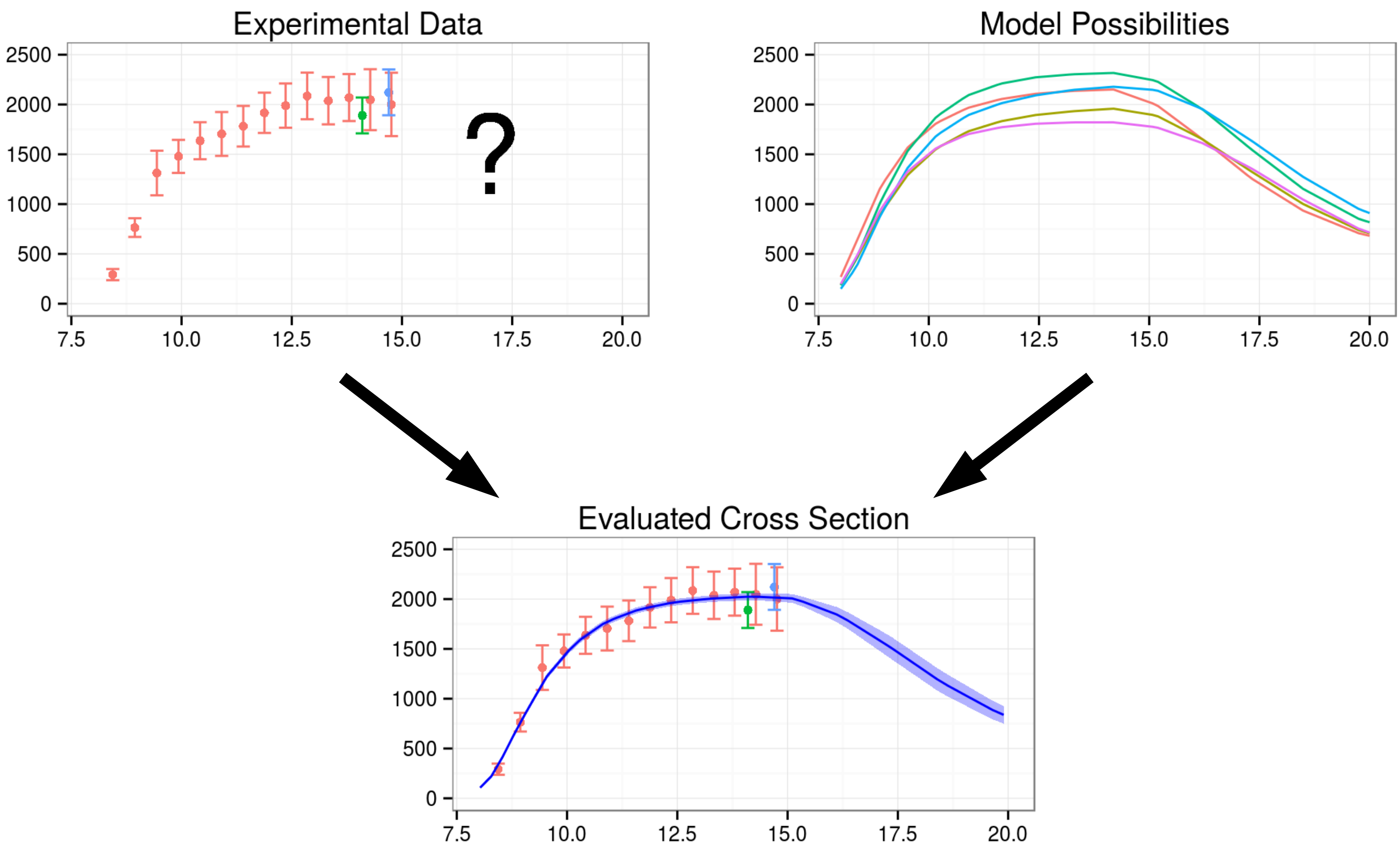
Relevant for:

- Reactor physics
- Radiation dosimetry
- Radiation protection
- Radioactive waste management
- Astrophysics
- Nuclear medicine
- Fusion research
- ...





Challenge #3: Nuclear data evaluation



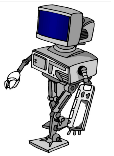
Two formulas for (at least) three applications

$$\vec{p}_1 = \vec{p}_0 + \mathbf{A}_0 \mathbf{S}^T (\mathbf{S}\mathbf{A}_0\mathbf{S} + \mathbf{B})^{-1} (\vec{\sigma}_{\rm exp} - M_{\rm lin}[\,\vec{p}_0]\,)$$

$$\mathbf{A}_1 = \mathbf{A}_0 - \mathbf{A}_0 \mathbf{S}^T (\mathbf{S}\mathbf{A}_0\mathbf{S} + \mathbf{B})^{-1} \mathbf{S}\mathbf{A}_0$$

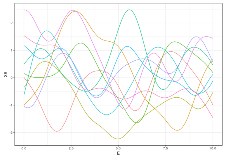
Outline

Why those equations? Where do they come from?



Application of these formulas to different inference problems

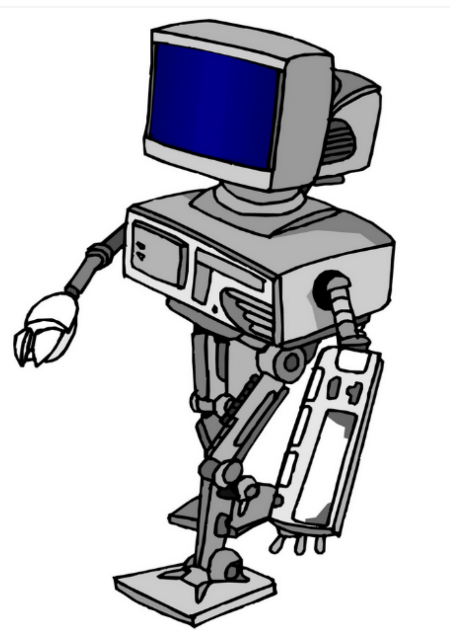
On the way, we will encounter two data science methods



Jaynes Robot

Hypotheses

H1: It rained
H2: It did not rain

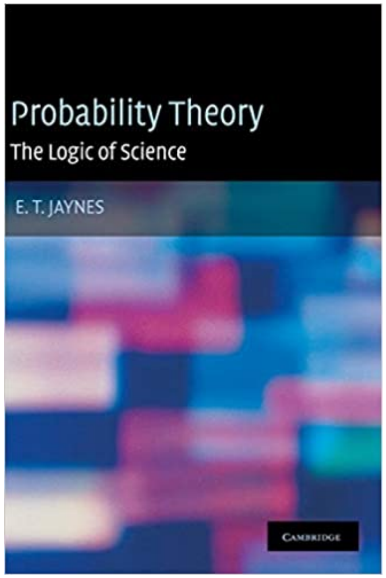


Observations

O1: The ground is wet
O2: The ground is dry

Which hypothesis is true?

Edwin Thompson Jaynes
1922-1998



Consistency with Aristotelian logic

Hypothesis

True

Observation

True

$$\frac{\begin{array}{c}\text{If } A \text{ is true, then } B \text{ is true}\\ A \text{ is true}\end{array}}{\text{Therefore, } B \text{ is true}}$$

Consistency with common sense

Hypothesis

Observation

More plausible

True

$$\frac{\text{If } A \text{ is true, then } B \text{ is true} \qquad B \text{ is true}}{\text{Therefore, } A \text{ becomes more plausible}}$$

Desiderata

(I) Degrees of Plausibility are represented by real numbers

(II) Qualitative Correspondence with common sense

(IIIa) If a conclusion can be reasoned out in more than one way, then every possible way must lead to the same result.

(from E.T. Jaynes, Probability Theory: The Logic of Science)

Richard Threlkeld Cox
1898-1991

Desiderata / Cox theorem

(I) Degrees of Plausibility are represented by real numbers

(II) Qualitative Correspondence with common sense

(IIIa) If a conclusion can be reasoned out in more than one way, then every possible way must lead to the same result.

(from E.T. Jaynes, Probability Theory: The Logic of Science)

Richard Threlkeld Cox
1898-1991

→ **Computation rules of probability theory follow**

e.g., product rule

$$P(H, O) = P(O \mid H) P(H)$$

The statistical split

Frequentist statistics

Ronald Fisher

Egon Pearson

Jerzy Neyman *

Bayesian statistics

Thomas Bayes

Pierre-Simon Laplace

*) MFO - Mathematisches Forschungsinstitut Oberwolfach
https://opc.mfo.de/detail?photo_id=3044

Bayesian update formula

Thomas Bayes
1701-1761

$$P(H \mid O) \propto P(O \mid H)P(H)$$

Pierre Simon Laplace
1749-1827

H hypothesis (e.g., "It rained")
O observation (e.g., "Floor wet")

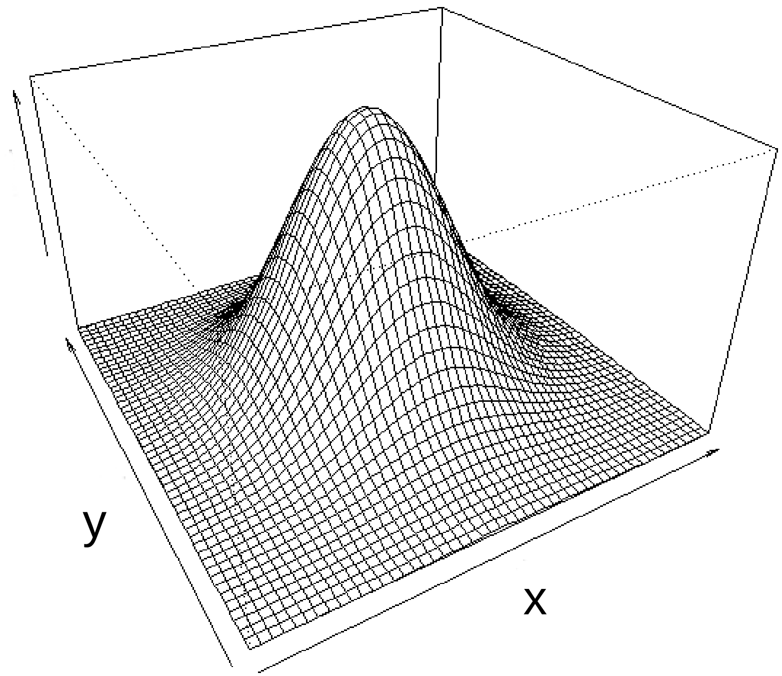
P(H) probability of hypothesis to be true

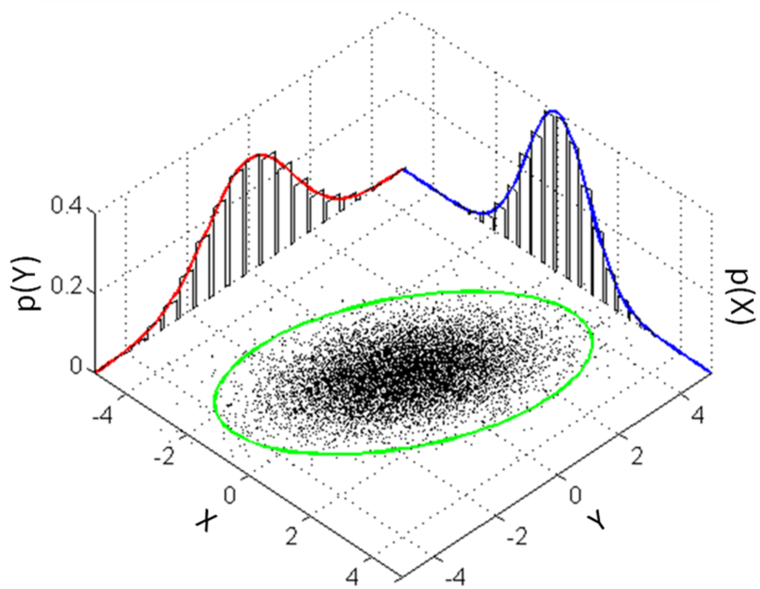
P(O|H) probability to make observation O given hypothesis H is true

P(H|O) probability of hypothesis H given we observed O

Multivariate normal distribution

$$\mathcal{N}(\vec{x}\,|\,\vec{x}_0, \mathbf{A}_0) = \frac{1}{\sqrt{(2\pi)^N |\mathbf{A}_0|}} \exp\left(-\frac{1}{2} (\vec{x} - \vec{x}_0)^T \, \mathbf{A}_0^{-1} \, (\vec{x} - \vec{x}_0) \right)$$





$$\vec{x} = \begin{pmatrix} \text{altitude} \\ \text{temperature} \end{pmatrix}$$

$$\vec{x}_0 = \begin{pmatrix} 1000 \\ 10 \end{pmatrix}$$

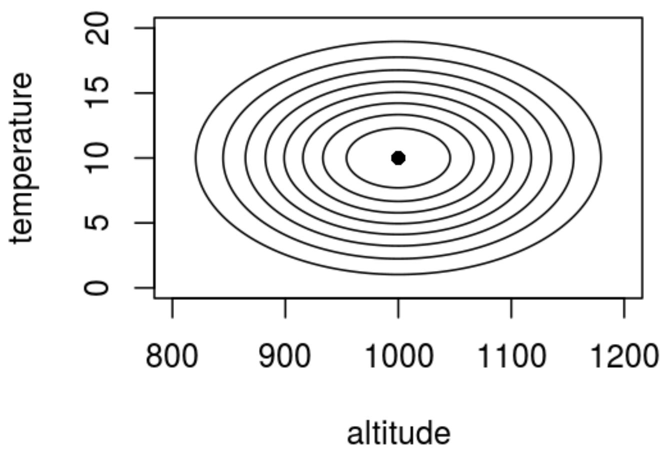
$$\vec{x} = \begin{pmatrix} \text{altitude} \\ \text{temperature} \end{pmatrix}$$

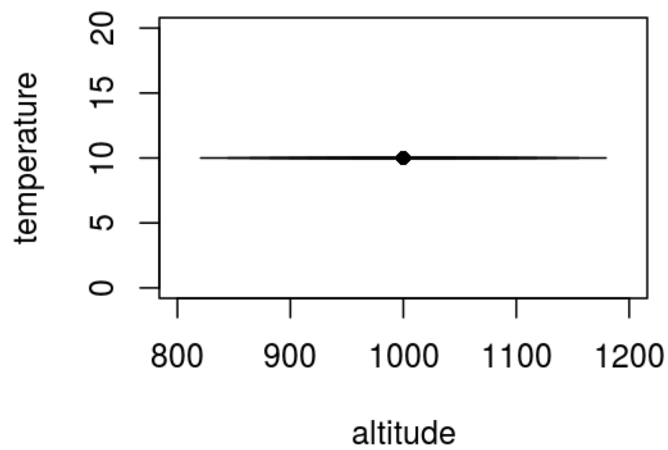
Covariance matrix

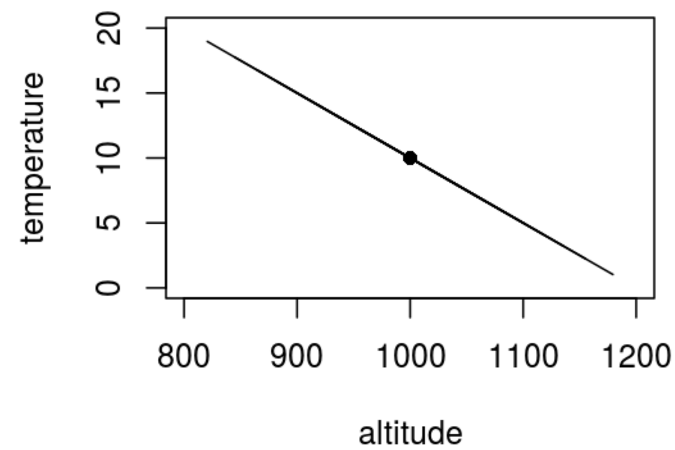
Variance / uncertainty squared

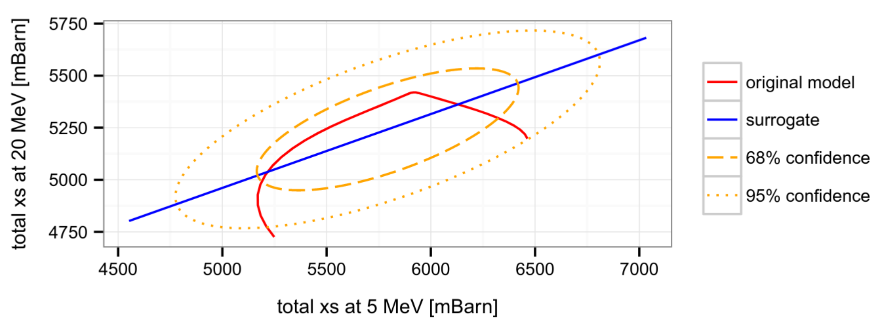
covariance

$$\mathbf{A}_0 = \begin{pmatrix} 100^2 & 0 \\ 0 & 5^2 \end{pmatrix} \qquad \mathbf{A}_0 = \begin{pmatrix} 100^2 & 0 \\ 0 & 0 \end{pmatrix} \qquad \mathbf{A}_0 = \begin{pmatrix} 100^2 & -100 \cdot 5 \\ -100 \cdot 5 & 5^2 \end{pmatrix}$$



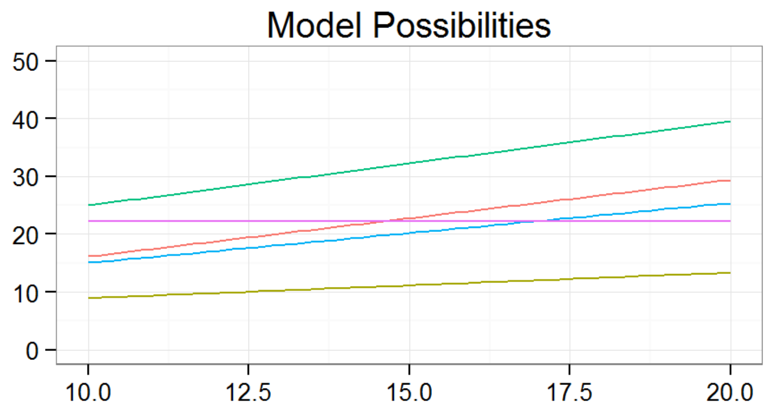


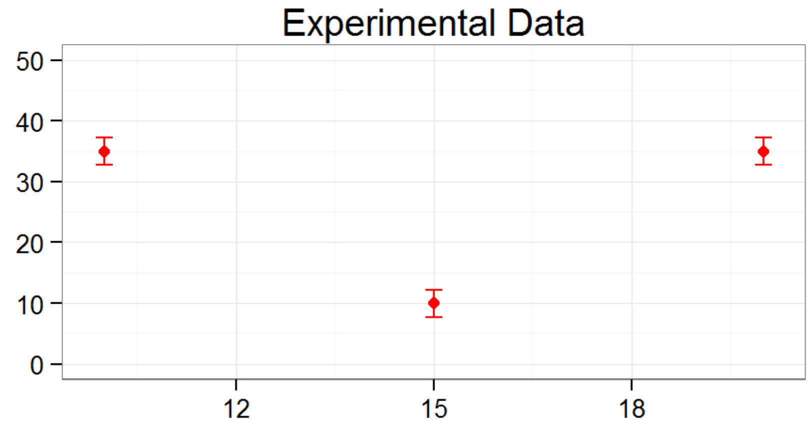




A covariance matrix captures linear relationships and uncertainties

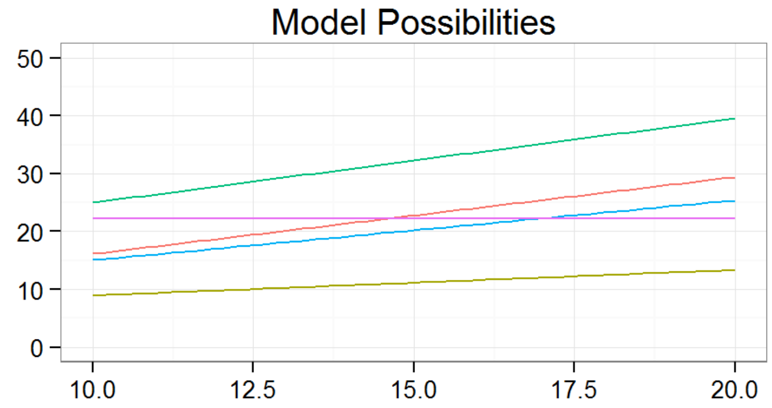
Putting everything together (at the example of linear regression)





$$y(x) = kx + d$$

Putting everything together (prior knowledge "P(H)")



$$y(x) = kx + d$$

$$\pi(\vec{p}) = \mathcal{N}(\vec{p}\,|\,\vec{p}_0, \mathbf{A}_0)$$

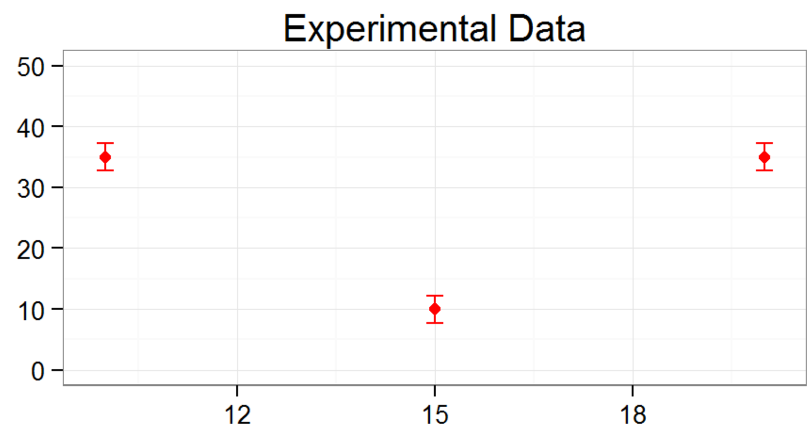
$$k \propto \mathcal{N}(0, \delta_k^2)$$

$$\vec{p}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$d \propto \mathcal{N}(0, \delta_d^2)$$

$$\mathbf{A}_0 = \begin{pmatrix} \delta_k^2 & 0 \\ 0 & \delta_d^2 \end{pmatrix}$$

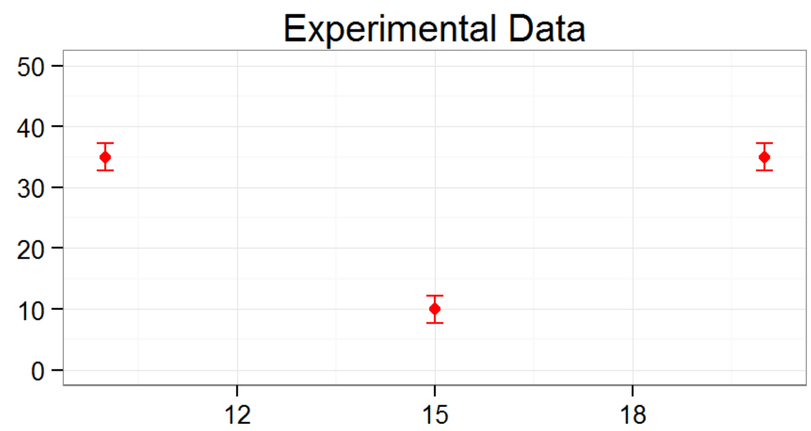
Putting everything together (likelihood "P(O|H)")



$$f(\vec{\sigma}_{\text{exp}} \,|\, \vec{p}) = \mathcal{N}(\vec{\sigma}_{\text{exp}} \,|\, M_{\text{lin}}[\vec{p}], \mathbf{B})$$

$$\vec{\sigma}_{\text{exp}} = \begin{pmatrix} \sigma_{\text{exp},1} \\ \sigma_{\text{exp},2} \\ \vdots \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} \varepsilon_1^2 & 0 & 0 \\ 0 & \varepsilon_2^2 & 0 \\ 0 & 0 & \ddots \end{pmatrix}$$

Putting everything together (likelihood “P(O|H)”)



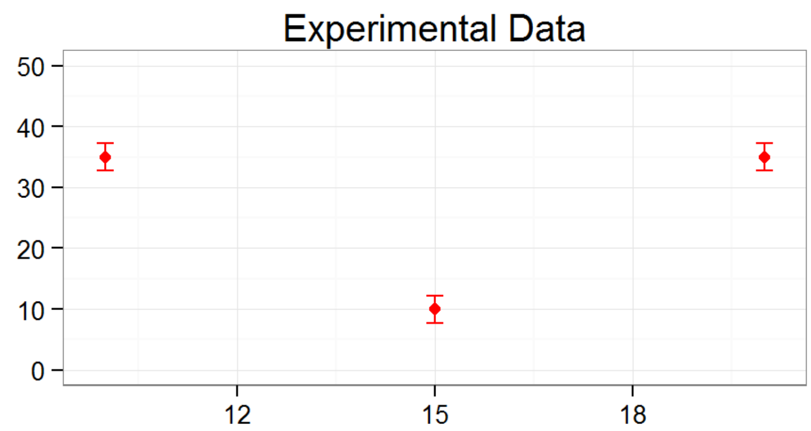
$$f(\vec{\sigma}_{\mathrm{exp}} \,|\, \vec{p}) = \mathcal{N}(\vec{\sigma}_{\mathrm{exp}} \,|\, M_{\mathrm{lin}}[\vec{p}], \mathbf{B})$$

$$\vec{\sigma}_{\mathrm{exp}} = \begin{pmatrix} \sigma_{\mathrm{exp},1} \\ \sigma_{\mathrm{exp},2} \\ \vdots \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} \varepsilon_1^2 & 0 & 0 \\ 0 & \varepsilon_2^2 & 0 \\ 0 & 0 & \ddots \end{pmatrix}$$

Counting statistics
→ Poisson distribution

$$\varepsilon_i = \sqrt{\mathrm{counts}_i}$$

Putting everything together (likelihood "P(O|H)")



$$f(\vec{\sigma}_{\exp} \,|\, \vec{p}) = \mathcal{N}(\vec{\sigma}_{\exp} \,|\, M_{\text{lin}}[\vec{p}], \mathbf{B})$$

$$y(x) = kx + d$$

$$M_{\text{lin}}[\vec{p}] = \mathbf{S}\vec{p} \qquad \mathbf{S} = \begin{pmatrix} E_1 & 1 \\ E_2 & 1 \\ \vdots & \end{pmatrix}$$

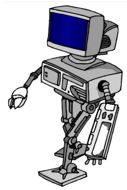
Generalized Least Squares (GLS)

in a nutshell

Prior $\pi(\vec{p}) = \mathcal{N}(\vec{p}\,|\,\vec{p}_0, \mathbf{A}_0)$

Likelihood $f(\vec{\sigma}_{\mathrm{exp}}\,|\,\vec{p}) = \mathcal{N}(\vec{\sigma}_{\mathrm{exp}}\,|\,M_{\mathrm{lin}}[\vec{p}], \mathbf{B})$

$M_{\mathrm{lin}}[\vec{p}] = \vec{\sigma}_{\mathrm{ref}} + \mathbf{S}\vec{p}$



Bayesian update $\pi(\vec{p}\,|\,\vec{\sigma}_{\mathrm{exp}}) \propto f(\vec{\sigma}_{\mathrm{exp}}\,|\,\vec{p})\,\pi(\vec{p})$

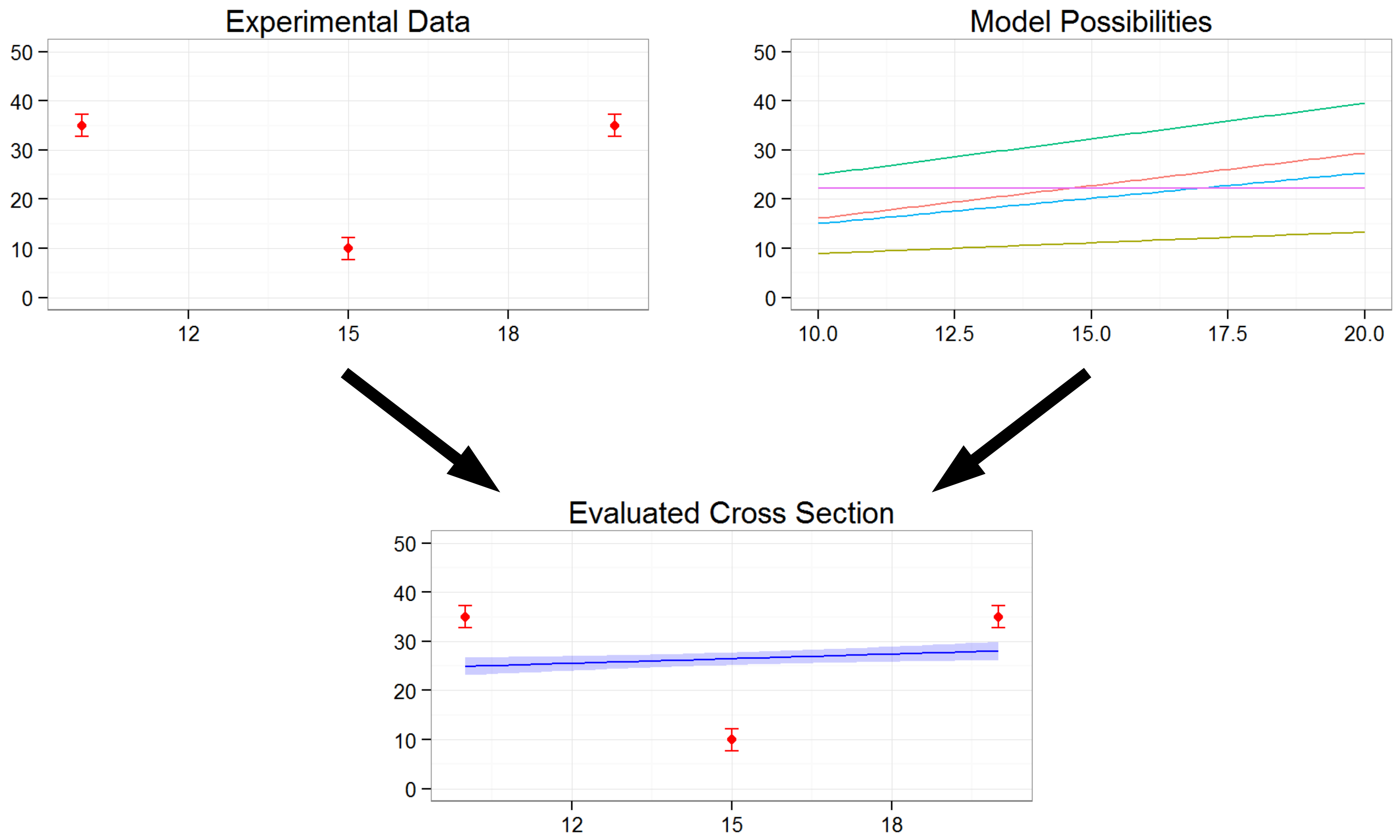
Posterior $\pi(\vec{p}\,|\,\vec{\sigma}_{\mathrm{exp}}) = \mathcal{N}(\vec{p}\,|\,\vec{p}_1, \mathbf{A}_1)$

$$\vec{p}_1 = \vec{p}_0 + \mathbf{A}_0\mathbf{S}^T(\mathbf{S}\mathbf{A}_0\mathbf{S} + \mathbf{B})^{-1}(\vec{\sigma}_{\mathrm{exp}} - M_{\mathrm{lin}}[\,\vec{p}_0]\,)$$

$$\mathbf{A}_1 = \mathbf{A}_0 - \mathbf{A}_0\mathbf{S}^T(\mathbf{S}\mathbf{A}_0\mathbf{S} + \mathbf{B})^{-1}\mathbf{S}\mathbf{A}_0$$

“GLS formulas”

And the result?



Important insight about models

"All models are wrong but some are useful"

- George E. P. Box

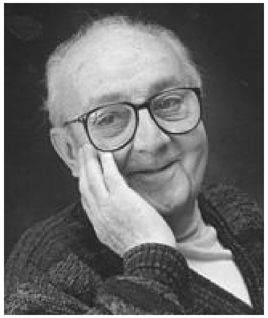
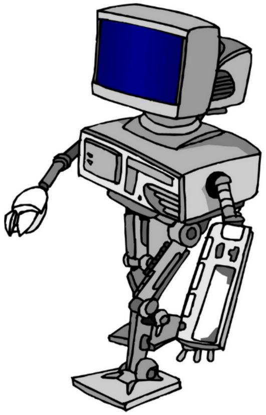


Image credit:
DavidMCEddy at en.wikipedia
CC BY-SA 3.0

Ice cream or pizza?



Reality

How to deal with a deficient model (or to come up with a good one)?

$$y(x) = a_0 + a_1 x + a_2 x^2$$

$$a_0 \propto \mathcal{N}(0, \delta_{a_0}^2) \qquad \vec{p}_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{A}_0 = \begin{pmatrix} \delta_{a_0}^2 & 0 & 0 \\ 0 & \delta_{a_1}^2 & 0 \\ 0 & 0 & \delta_{a_2}^2 \end{pmatrix}$$

$$a_1 \propto \mathcal{N}(0, \delta_{a_1}^2)$$

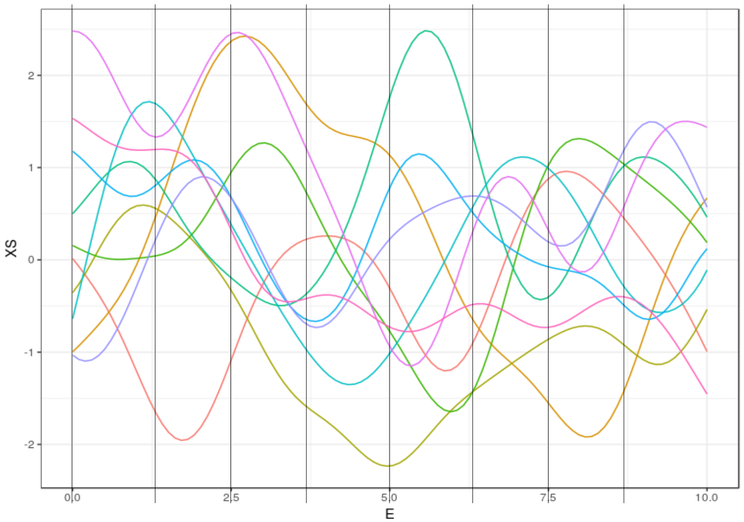
$$a_2 \propto \mathcal{N}(0, \delta_{a_2}^2)$$

But what if it is too difficult to find a good analytic function?

Surrogate approach

Assumptions:

- Access to a good physics model
- Ability to perform hundreds to thousands of model calculations using different parameter values in good time
- Each model calculation yields predictions on a dense mesh and we can use linear interpolation between mesh points



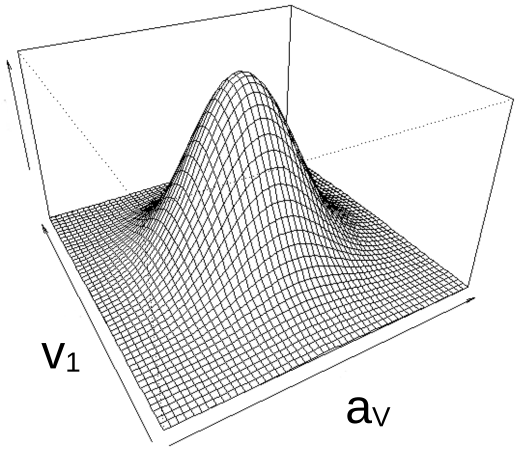
E_1 E_2 E_3 ...

$$\vec{\sigma}_{\mathrm{mod}} = \begin{pmatrix} \sigma(E_1) \\ \sigma(E_2) \\ \sigma(E_3) \\ \sigma(E_4) \\ \vdots \end{pmatrix}$$

Surrogate approach

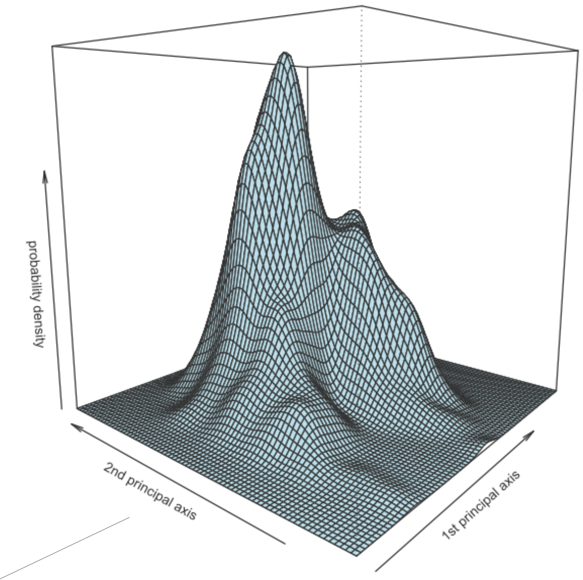
basic idea

$$\pi(\vec{p}_{\mathrm{true}}\,|\,M)$$



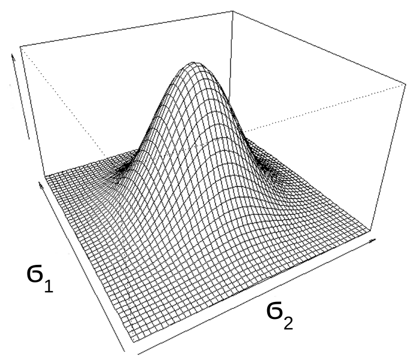
is linked to

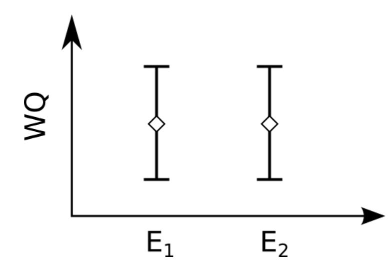
$$\pi(\vec{\sigma}_{\mathrm{true}}\,|\,M)$$



replaced by

$$\pi(\vec{\sigma}_{\mathrm{true}}\,|\,M_{\mathrm{sur}})$$



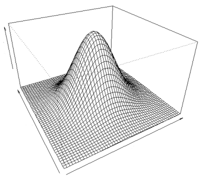


$$\pi(\vec{p}_{\mathrm{true}}\,|\,\vec{\sigma}_{\mathrm{exp}}, M) \propto f(\vec{\sigma}_{\mathrm{exp}}\,|\,\vec{p}_{\mathrm{true}}, M)\,\pi(\vec{p}_{\mathrm{true}}\,|\,M)$$

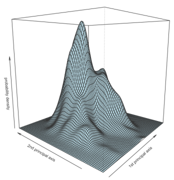
Surrogate approach

construction of M_{sur}

1) draw ensemble $\vec{p}_1\,\vec{p}_2,\vec{p}_3,\ldots,\vec{p}_n$ from prior distribution $\pi(\vec{p}_{\mathrm{true}}\,|\,M)$

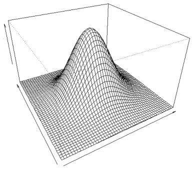


2) use nuclear model to calculate $\vec{\sigma}_1 = M[\,\vec{p}_1\,], \vec{\sigma}_2 = M[\,\vec{p}_2\,],\ldots$



3) estimate multivariate normal distribution in observation space

$$\vec{\sigma}_0 = \frac{1}{n}\sum_{i=1}^{n}\vec{\sigma}_i \qquad \mathbf{A}_0 = \frac{1}{n}\sum_{i=1}^{n}(\vec{\sigma}_i - \vec{\sigma}_0)(\vec{\sigma}_i - \vec{\sigma}_0)^T \qquad M_{\mathrm{sur}}[\,\sigma_{\mathrm{true}}\,] = \mathbf{S}\vec{\sigma}_{\mathrm{true}}$$



Generalized Least Squares formulas

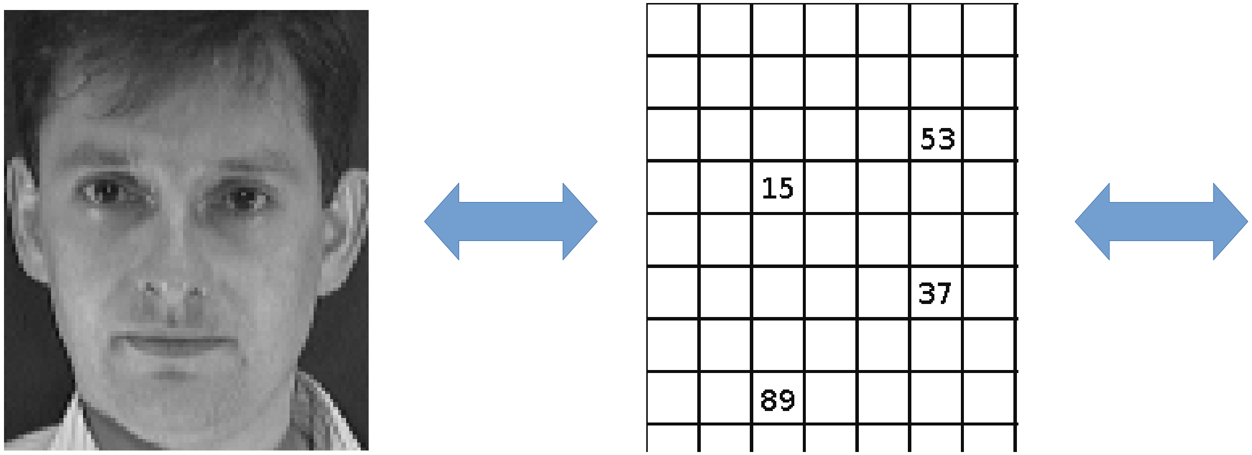
$$\mathbf{A}_1 = \mathbf{A}_0 - \mathbf{A}_0\mathbf{S}^T(\mathbf{S}\mathbf{A}_0\mathbf{S} + \mathbf{B})^{-1}\mathbf{S}\mathbf{A}_0$$

$$\vec{\sigma}_1 = \vec{\sigma}_0 + \mathbf{A}_0\mathbf{S}^T(\mathbf{S}\mathbf{A}_0\mathbf{S} + \mathbf{B})^{-1}(\vec{\sigma}_{\mathrm{exp}} - M_{\mathrm{sur}}[\,\vec{\sigma}_0]\,)$$

Application to ~~nuclear data evaluation~~ face reconstruction

We have a sample of ~~model predictions~~ faces

Representation of faces



$$\vec{x} = \begin{pmatrix} r_{11} \\ r_{12} \\ \vdots \\ r_{21} \\ r_{22} \\ \vdots \end{pmatrix}$$

All the matrices in their glory details

Prior belief

$$\vec{x}_0 = \frac{1}{N}\sum_i^{400} \vec{f}_i \qquad \mathbf{A}_0 = \frac{1}{N}\sum_i^{400}\left(\vec{f}_i - \vec{x}_0\right)\left(\vec{f}_i - \vec{x}_0\right)^T$$

Observation

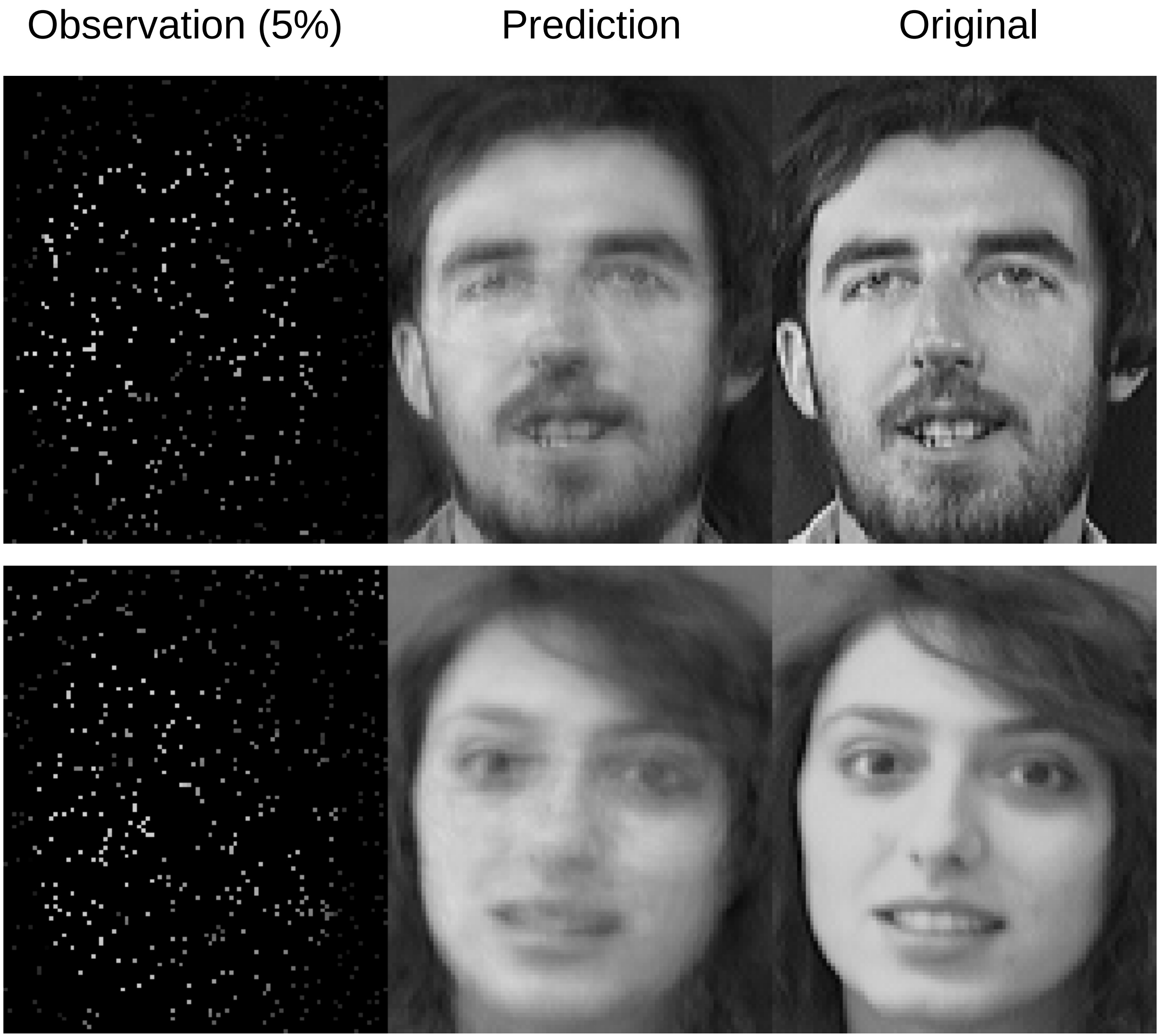
$$\vec{y} = \begin{pmatrix} o_{12} \\ o_{28} \\ \vdots \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 0.1^2 & 0 & \dots \\ 0 & 0.1^2 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad \mathbf{S} = \begin{pmatrix} 0 & 1 & \dots & 0 & \dots \\ 0 & 0 & \dots & 1 & \dots \\ \vdots & \vdots & \ddots & \vdots & \ddots \end{pmatrix} \qquad \vec{x} = \begin{pmatrix} r_{11} \\ r_{12} \\ \vdots \\ r_{21} \\ r_{22} \\ \vdots \end{pmatrix}$$

Improved estimate

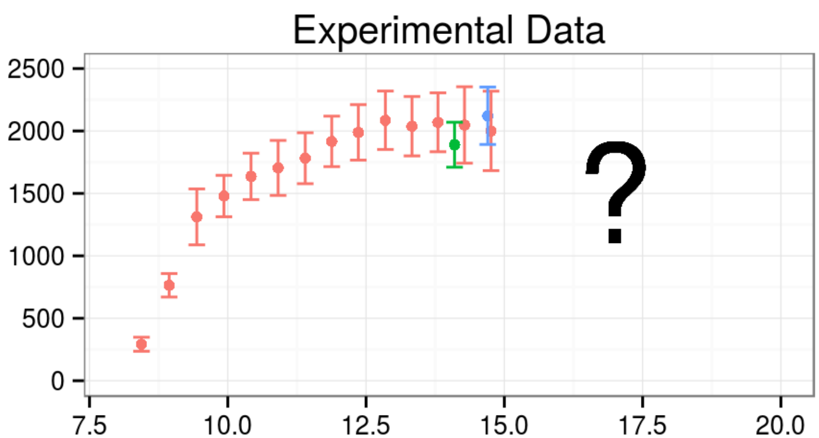
$$\vec{x}_1 = \vec{x}_0 + \mathbf{A}_0\mathbf{S}^T\left(\mathbf{S}\mathbf{A}_0\mathbf{S}^T + \mathbf{B}\right)^{-1}(\vec{y} - \mathbf{S}\vec{x}_0)$$

$$\mathbf{A}_1 = \mathbf{A}_0 - \mathbf{A}_0\mathbf{S}^T\left(\mathbf{S}\mathbf{A}_0\mathbf{S}^T + \mathbf{B}\right)^{-1}\mathbf{S}\mathbf{A}_0$$

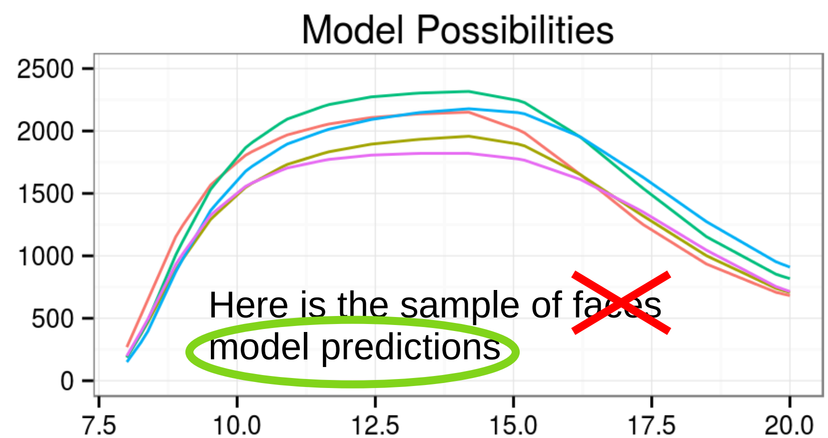
Example result

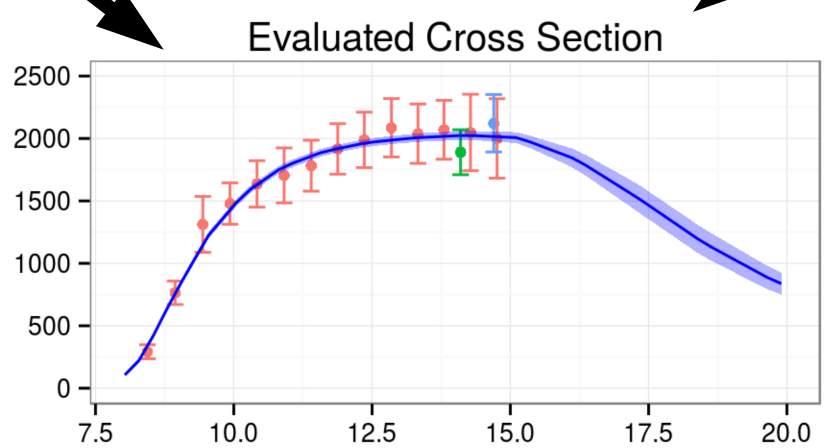


Ok, let's do also nuclear data evaluation with this approach

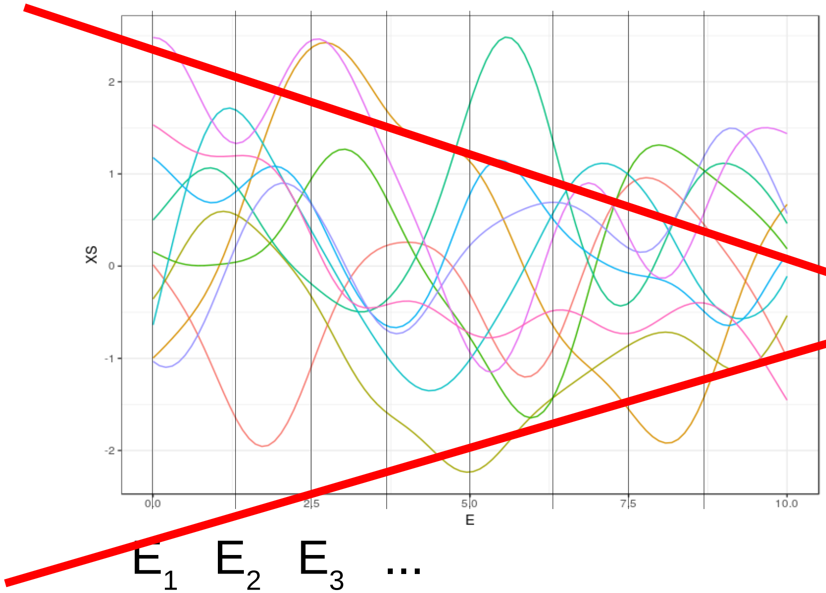


(measured xs = intensities of available pixels)





But what if...

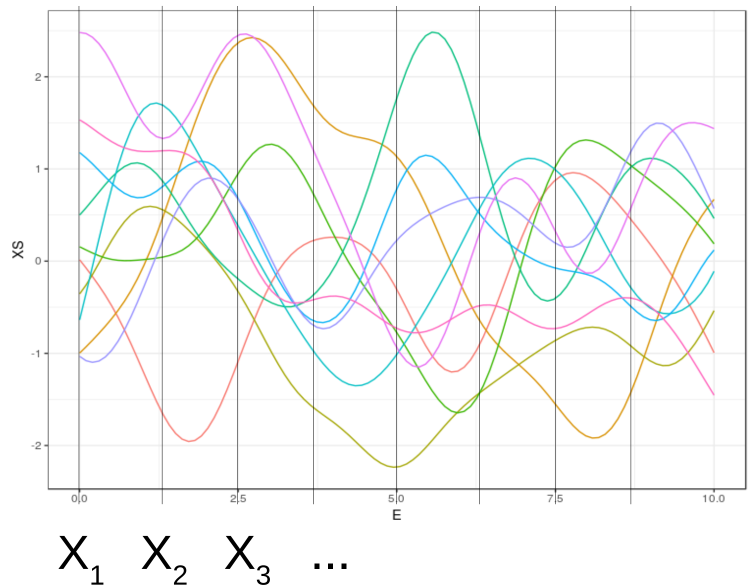


… we don’t have a good physics model nor a comprehensive database

… we don’t know a good analytic function to be used as model

Idea: Use a function to construct covariance matrices

$$\mathbf{y} \sim \mathcal{N}(\vec{m}, \mathbf{K})$$



$$\vec{m} = \begin{pmatrix} \mu(x_1) \\ \mu(x_2) \\ \mu(x_3) \\ \vdots \end{pmatrix}$$

$$\mathbf{K} = \begin{pmatrix} \kappa(x_1, x_1) & \kappa(x_1, x_2) & \kappa(x_1, x_3) & \dots \\ \kappa(x_2, x_1) & \kappa(x_2, x_2) & \kappa(x_2, x_3) & \dots \\ \kappa(x_3, x_1) & \kappa(x_3, x_2) & \kappa(x_3, x_3) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Covariance function

Specify a function that yields the covariance between function values f($\mathbf{x}_1$) and f($\mathbf{x}_2$) for any possible pair $\mathbf{x}_1$ and $\mathbf{x}_2$. This function is called a **covariance function**.

$$\kappa(\boldsymbol{x}_1, \boldsymbol{x}_2) = \delta^2 \exp\left(-\frac{(\boldsymbol{x}_1 - \boldsymbol{x}_2)^2}{2\lambda^2}\right)$$

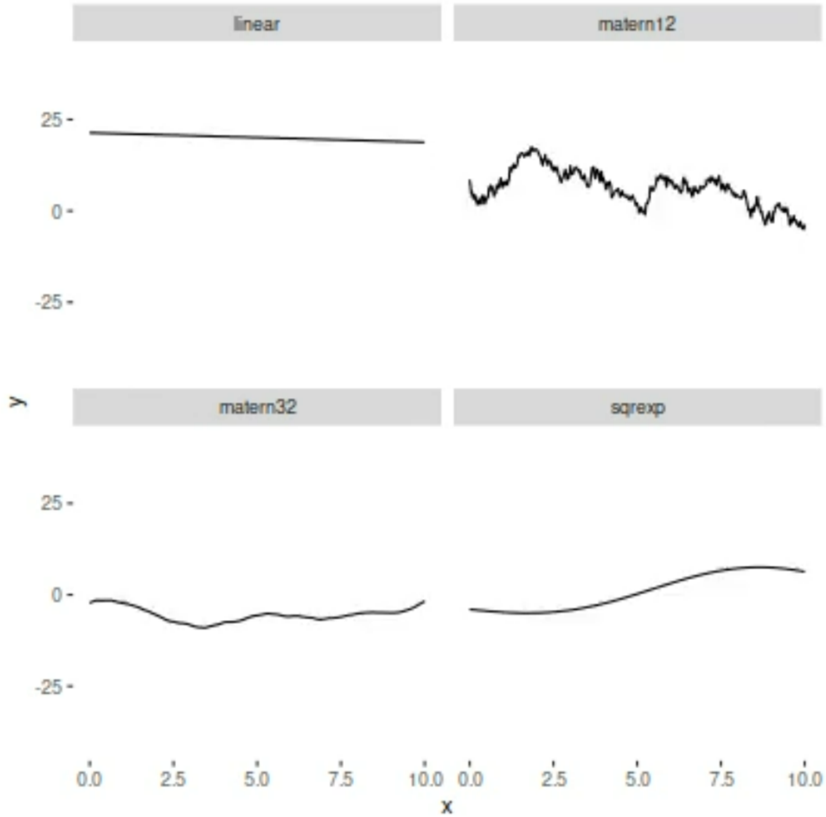
Specify another function μ($\mathbf{x}$) that yields the center value of the process at location $\mathbf{x}$**.** This function is called a **mean function**.

$$\mu(\mathbf{x}) = 0$$

Different covariance functions and the prior knowledge they induce

$$\kappa_{\text{lin}}(x_1, x_2) = x_1 x_2 {\delta_k}^2 + {\delta_d}^2$$

$$\kappa_{\text{mat},1/2}(x_1, x_2) = \delta^2 \exp\left(-\frac{|x_1 - x_2|}{\rho}\right)$$

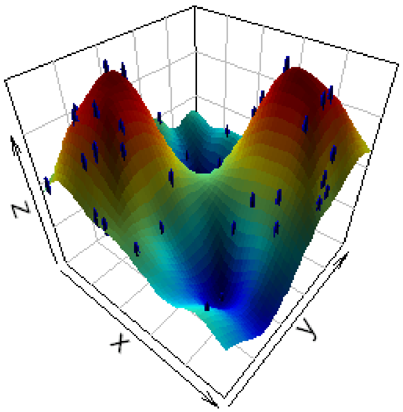


Play video

$$\kappa_{\text{mat},3/2}(x_1, x_2) = \delta^2 \left(1 + \frac{\sqrt{3}|x_1 - x_2|}{\rho}\right) \exp\left(-\frac{\sqrt{3}|x_1 - x_2|}{\rho}\right)$$

$$\kappa_{\text{sqrexp}}(x_1, x_2) = \delta^2 \exp\left(-\frac{1}{2\lambda^2}(x_1 - x_2)^2\right)$$

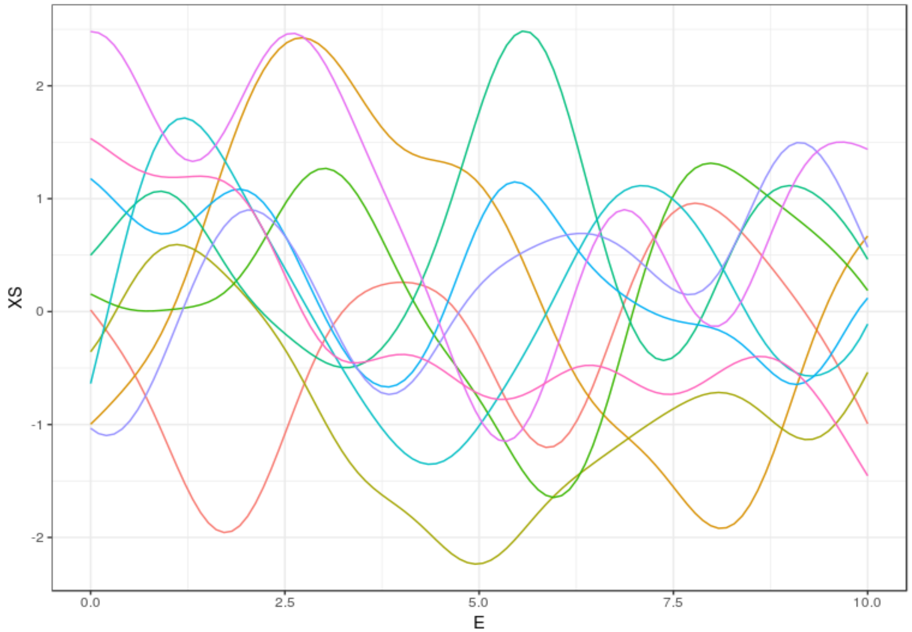
Gaussian process regression



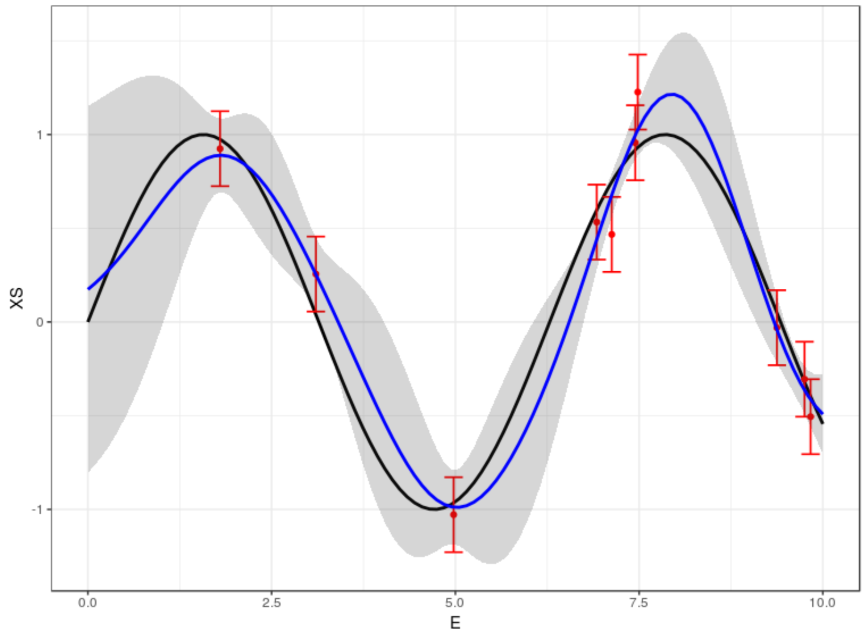
Powerful concept
Directly parametrize covariance matrix and work implicitly with an infinite number of parameters/basis functions!

$$\kappa(\boldsymbol{x}_1, \boldsymbol{x}_2) = \delta^2 \exp\left(-\frac{(\boldsymbol{x}_1 - \boldsymbol{x}_2)^2}{2\lambda^2}\right)$$

Sample from prior (δ=λ=1)

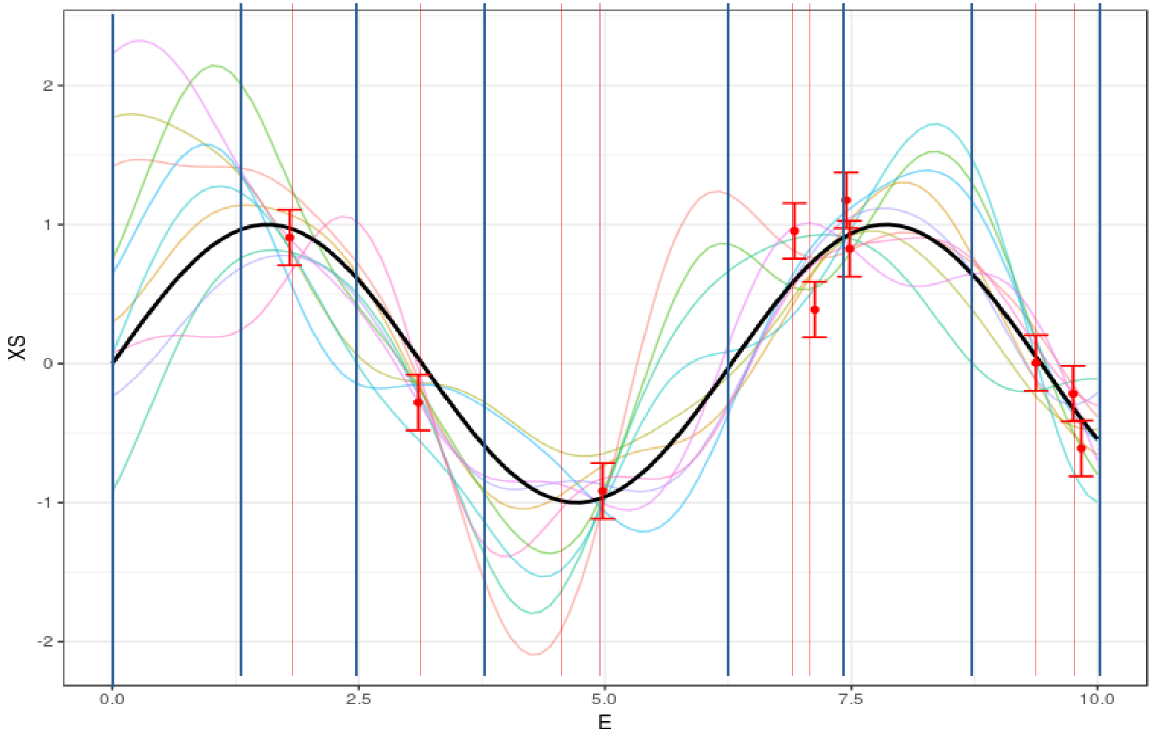


Posterior uncertainty



Computational mesh and experimental mesh

Introduce computational mesh **U** and experimental mesh **V**



A few covariance matrices

$$\mathbf{K}_{UU} = \begin{pmatrix} \kappa(x_{U,1}, x_{U,1}) & \kappa(x_{U,1}, x_{U,2}) & \kappa(x_{U,1}, x_{U,3}) & \cdots \\ \kappa(x_{U,2}, x_{U,1}) & \kappa(x_{U,2}, x_{U,2}) & \kappa(x_{U,2}, x_{U,3}) & \cdots \\ \kappa(x_{U,3}, x_{U,1}) & \kappa(x_{U,3}, x_{U,2}) & \kappa(x_{U,3}, x_{U,3}) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$$\mathbf{K}_{VV} = \begin{pmatrix} \kappa(x_{V,1}, x_{V,1}) & \kappa(x_{V,1}, x_{V,2}) & \kappa(x_{V,1}, x_{V,3}) & \cdots \\ \kappa(x_{V,2}, x_{V,1}) & \kappa(x_{V,2}, x_{V,2}) & \kappa(x_{V,2}, x_{V,3}) & \cdots \\ \kappa(x_{V,3}, x_{V,1}) & \kappa(x_{V,3}, x_{V,2}) & \kappa(x_{V,3}, x_{V,3}) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

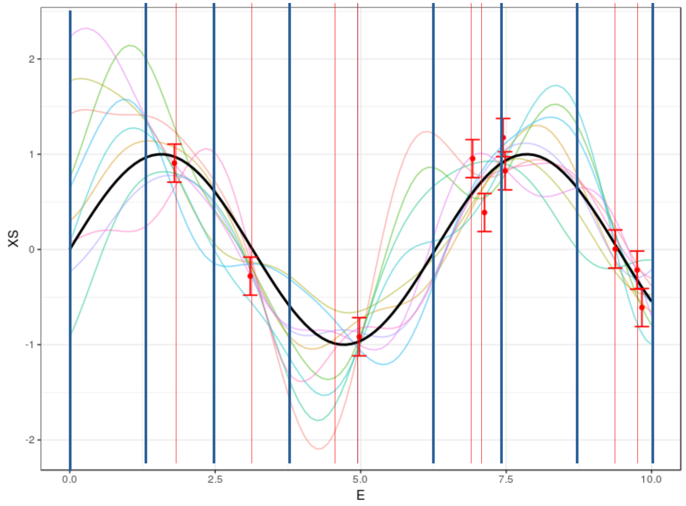
$$\mathbf{K}_{UV} = \begin{pmatrix} \kappa(x_{U,1}, x_{V,1}) & \kappa(x_{U,1}, x_{V,2}) & \kappa(x_{U,1}, x_{V,3}) & \cdots \\ \kappa(x_{U,2}, x_{V,1}) & \kappa(x_{U,2}, x_{V,2}) & \kappa(x_{U,2}, x_{V,3}) & \cdots \\ \kappa(x_{U,3}, x_{V,1}) & \kappa(x_{U,3}, x_{V,2}) & \kappa(x_{U,3}, x_{V,3}) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Applying the GLS formulas

Use GLS formulas:

$$\vec{p}_1 = \vec{p}_0 + \mathbf{A}_0 \mathbf{S}^T (\mathbf{S}\mathbf{A}_0\mathbf{S} + \mathbf{B})^{-1} (\vec{\sigma}_{\text{exp}} - M_{\text{lin}}[\,\vec{p}_0]\,)$$

$$\mathbf{A}_1 = \mathbf{A}_0 - \mathbf{A}_0 \mathbf{S}^T (\mathbf{S}\mathbf{A}_0\mathbf{S} + \mathbf{B})^{-1} \mathbf{S}\mathbf{A}_0$$



Make the following assignments:

$$\mathbf{A}_0 \rightarrow \mathbf{K}_{UU}$$

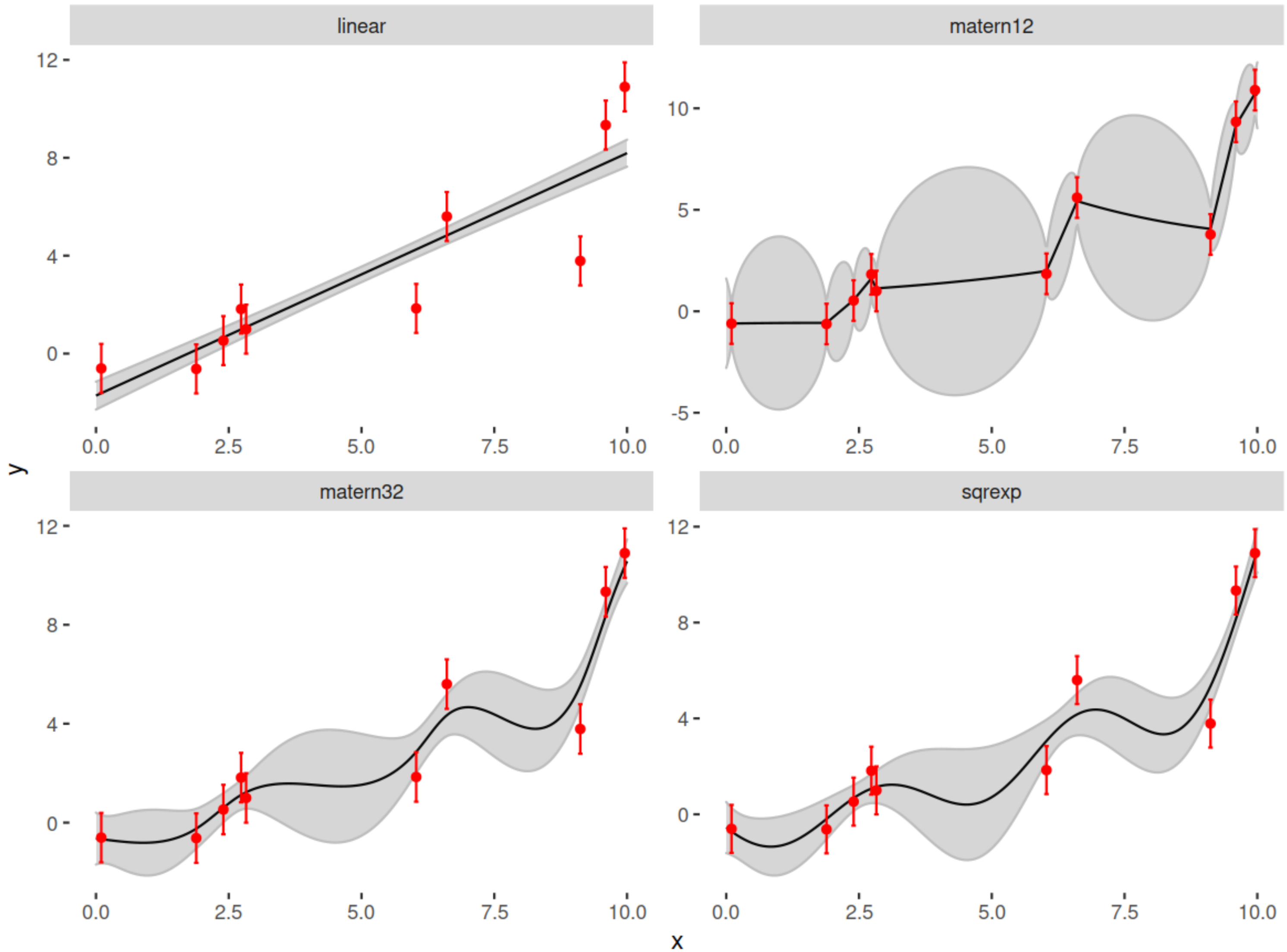
$$\vec{p}_0 \rightarrow \vec{m}_U$$

$$M_{\text{lin}}[\vec{p}_0] \rightarrow \vec{m}_V$$

$$\mathbf{S}\mathbf{A}_0\mathbf{S}^T \rightarrow \mathbf{K}_{VV}$$

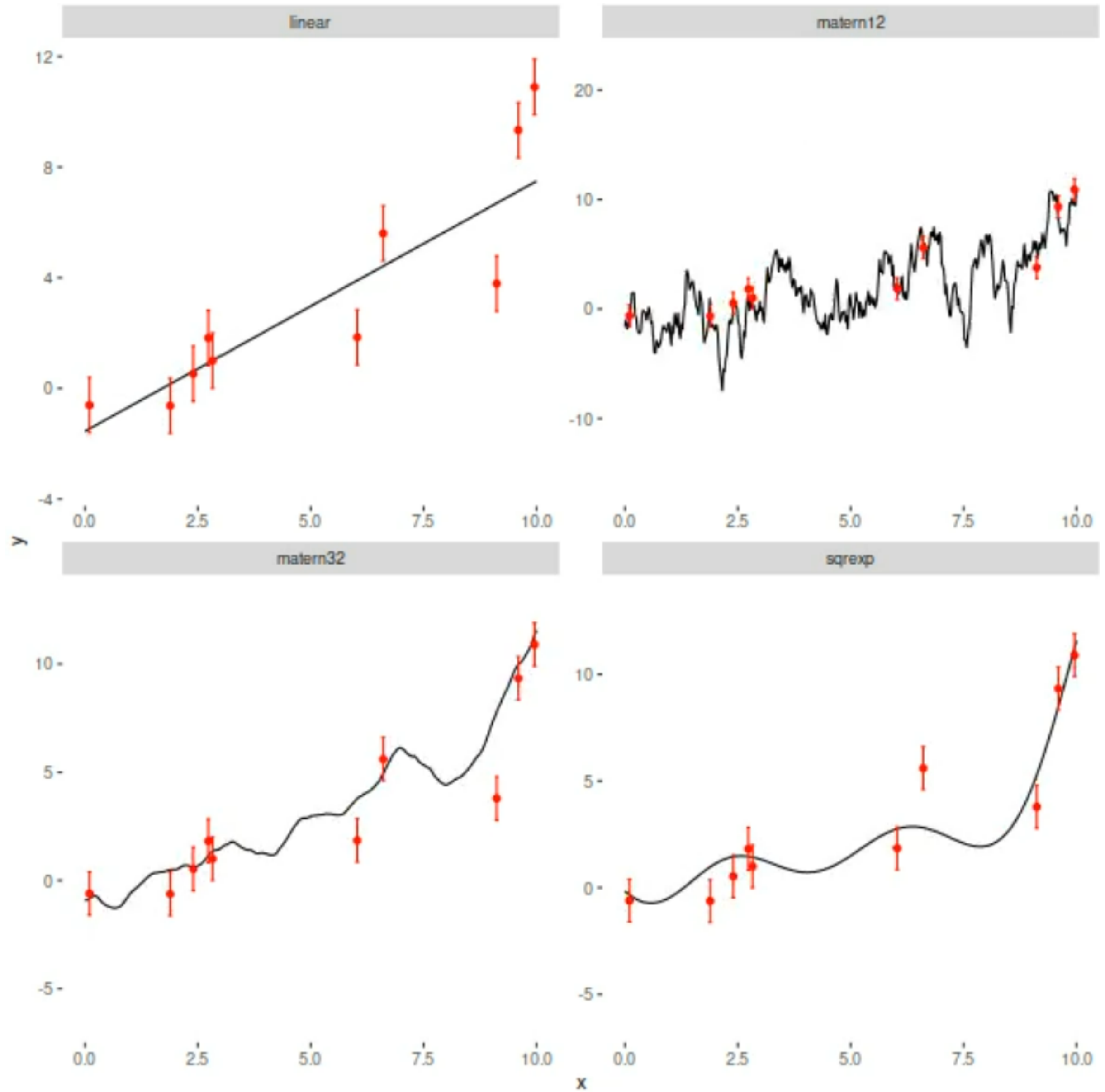
$$\mathbf{S}\mathbf{A}_0 \rightarrow \mathbf{K}_{VU}$$

GPs constrained with observations



The posterior covariance matrix encodes more than just uncertainties!

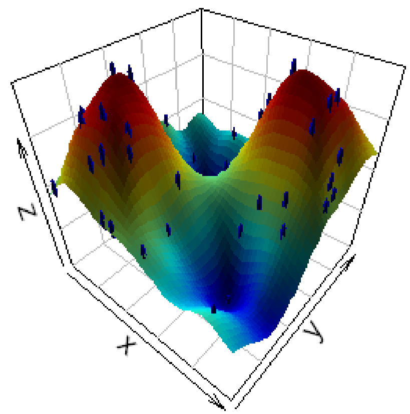
GPs constrained with observations



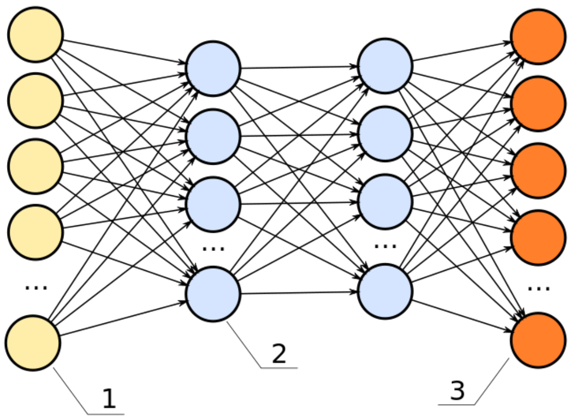
Play video

Comparison to neural networks

GP processes



artificial neural networks



Both approaches …

… are methods for classification and regression

… are universal function approximators

Neural networks …

… scale better to large data sets

… are able to capture non-local features

… are difficult to interpret

GP processes …

… are statistical methods from the ground up (uncertainties)

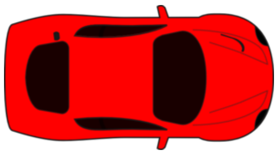
… facilitate the incorporation of prior assumptions

… interface well with existing nuclear data evaluation methods

Summary

- Two formulas for inference for at least three applications (GLS)
- Presented an argument to motivate the use of Bayesian statistics
- Employed the multivariate normal distribution
- Generalized Least Squares method applied in different ways
- Finally, we talked about Gaussian process regression which can be regarded as special case of Bayesian GLS if we discretize the function (or vice-versa)

Bonus: Mars rover



Task: An automated vehicle should follow a predefined path on Mars

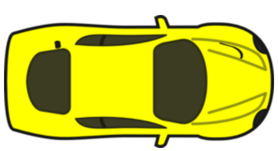
Hypothetical approach:
a) Vehicle is aware of initial position and its speed and direction of movement at any point in time afterwards.
b) Determine current position by $x_0 = x_{old} + v*\Delta t$
c) Adjust speed and direction to stay on desired trajectory

Problems:
Neither initial position nor speed nor direction are perfectly known. This introduces uncertainty about the current position.

State and error propagation:

$$\vec{x}_0 = \vec{x}_{\text{old}} + \vec{v}\Delta t \quad \mathbf{A}_0 = \mathbf{A}_{\text{old}} + \mathbf{V}\Delta t^2$$

ghost car



Update equations

Prior belief

$$\vec{x}_0 = \vec{x}_{\text{old}} + \vec{v}\Delta t \quad \mathbf{A}_0 = \mathbf{A}_{\text{old}} + \mathbf{V}\Delta t^2$$

Observation

$$\vec{y} \ \ldots \ \text{GPS position estimate}$$

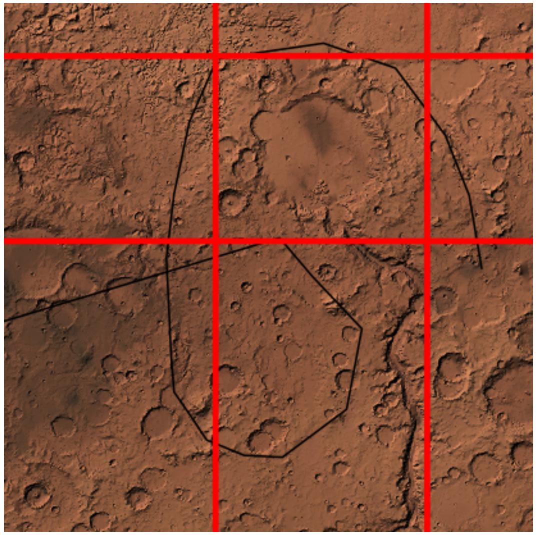
$$\mathbf{B} = \begin{pmatrix} 20^2 & 0 \\ 0 & 20^2 \end{pmatrix} \quad \mathbf{S} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Improved estimate

$$\vec{x}_1 = \vec{x}_0 + \mathbf{A}_0\mathbf{S}^T\left(\mathbf{S}\mathbf{A}_0\mathbf{S}^T + \mathbf{B}\right)^{-1}(\vec{y} - \mathbf{S}\vec{x}_0)$$

$$\mathbf{A}_1 = \mathbf{A}_0 - \mathbf{A}_0\mathbf{S}^T\left(\mathbf{S}\mathbf{A}_0\mathbf{S}^T + \mathbf{B}\right)^{-1}\mathbf{S}\mathbf{A}_0$$

Using a laser guidance system



Prior belief

$$\vec{x}_0 = \vec{x}_{\text{old}} + \vec{v}\Delta t \quad \mathbf{A}_0 = \mathbf{A}_{\text{old}} + \mathbf{V}\Delta t^2$$

Observation

$$\vec{y} \ \ldots \ \text{horizontal position estimate}$$

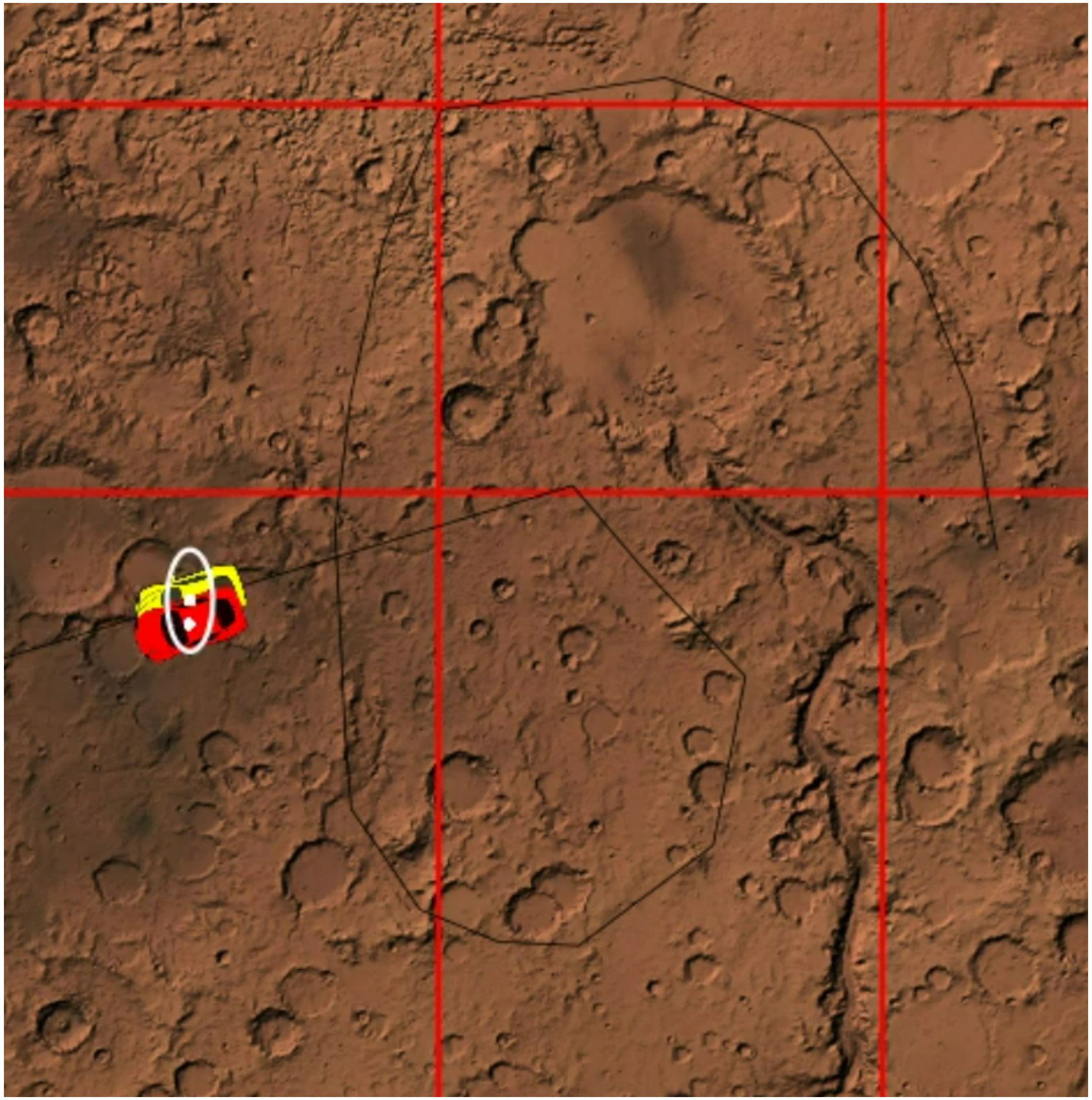
$$\mathbf{B} = \begin{pmatrix} 0.2^2 \end{pmatrix} \quad \mathbf{S} = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

Improved estimate

$$\vec{x}_1 = \vec{x}_0 + \mathbf{A}_0\mathbf{S}^T\left(\mathbf{S}\mathbf{A}_0\mathbf{S}^T + \mathbf{B}\right)^{-1}(\vec{y} - \mathbf{S}\vec{x}_0)$$

$$\mathbf{A}_1 = \mathbf{A}_0 - \mathbf{A}_0\mathbf{S}^T\left(\mathbf{S}\mathbf{A}_0\mathbf{S}^T + \mathbf{B}\right)^{-1}\mathbf{S}\mathbf{A}_0$$

Using a laser guidance system



Link to movie

References

- E. T. Jaynes “Probability Theory: The Logic of Science” (first chapters available online for free)
- C. E. Rasmussen & C. K. I Williams, “Gaussian processes for Machine Learning”, http://www.gaussianprocess.org/gpml/
- GitHub repository with the scripts and videos of this presentation: https://github.com/gschnabel/compnuc-workshop-2022