Parameter Estimation in Epistemologically valid Machine Interval Experiments

Raazesh Sainudiin
Department of Mathematics and Statistics,
University of Canterbury, Private Bag 4800,
Christchurch, New Zealand

joint work with

Alexander Danis and Warwick Tucker
Department of Mathematics,
Uppsala University, Box 480 Uppsala, Sweden

Joint Maths-Philosophy Retreat, Kaikoura, New Zealand

The Dualistic Context ("The Bigger Picture")

The Dualistic Context ("The Bigger Picture")

- **Tradition**: Modern European Empiricism

The Dualistic Context ("The Bigger Picture")

- **Tradition**: Modern European Empiricism
- **Internal Consistency** : Aristotelean Logic

The Dualistic Context (“The Bigger Picture”)

- **Tradition**: Modern European Empiricism
- **Internal Consistency** : Aristotelean Logic
- **Universe of Hypotheses**: Popper’s Falsifiability

The Dualistic Context ("The Bigger Picture")

- **Tradition**: Modern European Empiricism
- **Internal Consistency** : Aristotelean Logic
- **Universe of Hypotheses**: Popper's Falsifiability
- **Empirical Resolution** : Mechatronically Measured Data $\mathcal{D}_o$

The Dualistic Context ("The Bigger Picture")

- **Tradition**: Modern European Empiricism
- **Internal Consistency** : Aristotelean Logic
- **Universe of Hypotheses**: Popper's Falsifiability
- **Empirical Resolution** : Mechatronically Measured Data $\mathcal{D}_o$
- **Model** : Deterministic ODE-IVPs

The Dualistic Context ("The Bigger Picture")

- **Tradition**: Modern European Empiricism
- **Internal Consistency** : Aristotelean Logic
- **Universe of Hypotheses**: Popper's Falsifiability
- **Empirical Resolution** : Mechatronically Measured Data $\mathcal{D}_o$
- **Model** : Deterministic ODE-IVPs
- **Parameter Space** : finite dimensional parameter space

The Dualistic Context ("The Bigger Picture")

- **Tradition**: Modern European Empiricism
- **Internal Consistency** : Aristotelean Logic
- **Universe of Hypotheses**: Popper's Falsifiability
- **Empirical Resolution** : Mechatronically Measured Data $\mathcal{D}_o$
- **Model** : Deterministic ODE-IVPs
- **Parameter Space** : finite dimensional parameter space
- **Approach**: Statistical Decision Theory (set-valued approach)

The Dualistic Context ("The Bigger Picture")

- **Tradition**: Modern European Empiricism
- **Internal Consistency** : Aristotelean Logic
- **Universe of Hypotheses**: Popper's Falsifiability
- **Empirical Resolution** : Mechatronically Measured Data $\mathcal{D}_o$
- **Model** : Deterministic ODE-IVPs
- **Parameter Space** : finite dimensional parameter space
- **Approach**: Statistical Decision Theory (set-valued approach)
- **Engineering Constraints**: Resource-limited Info. Proc.

The Dualistic Context ("The Bigger Picture")

- **Tradition**: Modern European Empiricism
- **Internal Consistency** : Aristotelean Logic
- **Universe of Hypotheses**: Popper's Falsifiability
- **Empirical Resolution** : Mechatronically Measured Data $\mathcal{D}_o$
- **Model** : Deterministic ODE-IVPs
- **Parameter Space** : finite dimensional parameter space
- **Approach**: Statistical Decision Theory (set-valued approach)
- **Engineering Constraints**: Resource-limited Info. Proc.
- **Objective** : Epistemologically Valid Parameter Estimation

The Dualistic Context ("The Bigger Picture")

- **Tradition**: Modern European Empiricism
- **Internal Consistency** : Aristotelean Logic
- **Universe of Hypotheses**: Popper's Falsifiability
- **Empirical Resolution** : Mechatronically Measured Data $\mathcal{D}_o$
- **Model** : Deterministic ODE-IVPs
- **Parameter Space** : finite dimensional parameter space
- **Approach**: Statistical Decision Theory (set-valued approach)
- **Engineering Constraints**: Resource-limited Info. Proc.
- **Objective** : Epistemologically Valid Parameter Estimation
- **Solution**: Computer-aided Proofs & Interval Analysis

Outline

Outline

Statement (Hume, 1777)

"But where different effects have been found to follow from causes, which are to appearance exactly similar, all these various effects must occur to the mind in transferring the past to the future, and enter into our consideration, when we determine the probability of the event."[1]

Outline

Statement (Hume, 1777)

"But where different effects have been found to follow from causes, which are to appearance exactly similar, all these various effects must occur to the mind in transferring the past to the future, and enter into our consideration, when we determine the probability of the event."[1]

- Epistemologically valid experiment

Outline

Statement (Hume, 1777)

"But where different effects have been found to follow from causes, which are to appearance exactly similar, all these various effects must occur to the mind in transferring the past to the future, and enter into our consideration, when we determine the probability of the event."[1]

- Epistemologically valid experiment
- Data from a double pendulum

Outline

Statement (Hume, 1777)

"But where different effects have been found to follow from causes, which are to appearance exactly similar, all these various effects must occur to the mind in transferring the past to the future, and enter into our consideration, when we determine the probability of the event."[1]

- Epistemologically valid experiment
- Data from a double pendulum
- Model with parameter space $\boldsymbol{\Theta} \subsetneq \mathbb{R}^k, k < \infty$

Outline

Statement (Hume, 1777)

"But where different effects have been found to follow from causes, which are to appearance exactly similar, all these various effects must occur to the mind in transferring the past to the future, and enter into our consideration, when we determine the probability of the event."[1]

- Epistemologically valid experiment
- Data from a double pendulum
- Model with parameter space $\mathbf{\Theta} \subsetneq \mathbb{R}^k, k < \infty$
- Action space of point estimation $\mathcal{A} = \mathbf{\Theta}$

Outline

Statement (Hume, 1777)

"But where different effects have been found to follow from causes, which are to appearance exactly similar, all these various effects must occur to the mind in transferring the past to the future, and enter into our consideration, when we determine the probability of the event."[1]

- Epistemologically valid experiment
- Data from a double pendulum
- Model with parameter space $\boldsymbol{\Theta} \subsetneq \mathbb{R}^k, k < \infty$
- Action space of point estimation $\mathcal{A} = \boldsymbol{\Theta}$
- Solution
    - MLE is CSP with epistemologically valid action space $\mathbb{I}\boldsymbol{\Theta}^*$
    - Set-valued integrators, (T,F,?)-based estimators

Outline

Statement (Hume, 1777)

"But where different effects have been found to follow from causes, which are to appearance exactly similar, all these various effects must occur to the mind in transferring the past to the future, and enter into our consideration, when we determine the probability of the event."[1]

- Epistemologically valid experiment
- Data from a double pendulum
- Model with parameter space $\boldsymbol{\Theta} \subsetneq \mathbb{R}^k, k < \infty$
- Action space of point estimation $\mathcal{A} = \boldsymbol{\Theta}$
- Solution
  - MLE is CSP with epistemologically valid action space $\mathbb{I}\boldsymbol{\Theta}^*$
  - Set-valued integrators, (T,F,?)-based estimators
- Blabber on Ongoing Work

Epistemologically valid experiment

Definition

Epistemology is the study of the nature and grounds of knowledge especially with reference to its limits and validity.

Epistemologically valid experiment

Definition

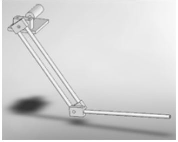
Epistemology is the study of the nature and grounds of knowledge especially with reference to its limits and validity.

Definition

A **statistical experiment** $\mathcal{E}_{\mathcal{P}}$ is the triple $(\mathbf{X}, \mathcal{F}_{\mathbf{X}}, \mathcal{P})$ consisting of a sample space $\mathbf{X}$ of all possible empirically observable realizations of a natural phenomenon $\mathbf{\Phi}$, a sigma-algebra $\mathcal{F}_{\mathbf{X}}$ on $\mathbf{X}$, and a family of probability measures $\mathcal{P} = \{P_\theta, \theta \in \mathbf{\Theta}\}$, where each P_θ is a probability measure on the measurable space $(\mathbf{X}, \mathcal{F}_{\mathbf{X}})$. The θ is an index belonging to the index set $\mathbf{\Theta}$. The index map $d(\theta) = P_\theta : \mathbf{\Theta} \rightarrow \mathcal{P}$ associates every $\theta \in \mathbf{\Theta}$ with $P_\theta \in \mathcal{P}$, in an arbitrary manner that even allows for the index map d to be the identity map with $\mathbf{\Theta} = \mathcal{P}$.

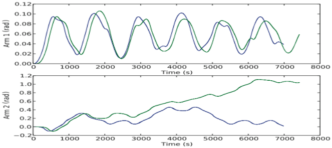
Phenomenon: Damped Double Pendulum Trajectories

A: DP Schematic

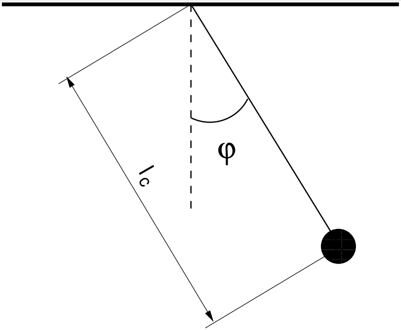


B: Streaming DP data

C: Enclosures of two initially close trajectories



ODE Model: Damped Single Pendulum Trajectories



- Model the arm as a distributed mass with centre of mass located at a distance l_c from the pivot,
- moment of inertia the arm is $\mathcal{I}$ and its mass is m.
- the acceleration due to gravity is $g \approx 9.81\text{ms}^{-2}$,
- φ is the angular position,
- $\dot{\varphi}$ is the angular velocity,

ODE Model: Damped Single Pendulum Trajectories

Kinetic energy of the arm consists of only rotational kinetic energy $T = \frac{1}{2}\mathcal{I}\dot{\varphi}^2$, The potential energy of the pendulum is calculated by considering the geometric position of the centre of mass above the equilibrium position, $V = ml_c g(1 - \cos\varphi)$ Lagrangian of the single pendulum:

$$\mathcal{L} = T - V \tag{1}$$

$$= \frac{1}{2}\mathcal{I}\dot{\varphi}^2 - ml_c g(1 - \cos\varphi). \tag{2}$$

The Euler-Lagrange form:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\right) - \frac{\partial\mathcal{L}}{\partial\varphi} = 0, \tag{3}$$

giving the equation of motion for the single pendulum system,

$$\mathcal{I}\ddot{\varphi} + mgl_c\sin\varphi = 0 \tag{4}$$

or,

$$\ddot{\varphi} = -\xi^2\sin\varphi \tag{5}$$

where $\xi = \sqrt{\frac{ml_c g}{\mathcal{I}}}$.

ODE Model: Damped Single Pendulum Trajectories

To numerically integrate the equation of motion, we convert (5) into a system of first order equations by letting $\dot{\varphi} = \omega$ and differentiating, $\dot{\omega} = \ddot{\varphi}$. Thus we have the system of first order equations,

$$\begin{bmatrix} \dot{\varphi} \\ \dot{\omega} \end{bmatrix} = \begin{bmatrix} \omega \\ -\xi^2 \sin \varphi \end{bmatrix} \quad (6)$$

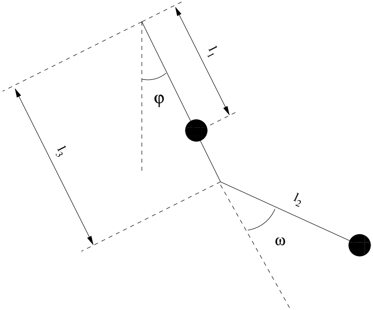
Friction may be added to the system by adding another term to (4). The friction in this case is modeled as proportional to the angular velocity, the torque produced is given by,

$$\tau_b = \mu \dot{\varphi}$$

giving (4) as,

$$\mathcal{I}\ddot{\varphi} + \mu\dot{\varphi} + mgl_c \sin \varphi = 0 \quad (7)$$

ODE Model: Passive Double Pendulum Trajectories



- the centre of mass of the inner arm is distance l_1
- the distance between pivots of the inner arm is l_3
- the centre of mass of the outer arm is distance l_2
- top arm has mass m_1 and moment of inertia of $\mathcal{I}_1$
- similarly for the outer arm they are m_2 and $\mathcal{I}_2$

ODE Model: Passive Double Pendulum Trajectories

After some work...
Derivation of equations via the Euler-Lagrange equations of motion follows in an manner analogous to that presented for the passive single pendulum...
We can this parametric family of vector fields for our statistical experiment with data $\{x(t_i)\}_{t_i \in \mathbb{T}}$ as follows:

$$\boldsymbol{\Theta} \ni \theta, \quad x(t) = \int f(x * * * * * * * * * * * *; \theta)$$

Here, x_i is a sample time and $y_i = (\varphi_1, \varphi_2)$ gives the angular positions of each arm at time x_i

Epistemologically valid experiment

Definition

Epistemology is the study of the nature and grounds of knowledge especially with reference to its limits and validity.

Epistemologically valid experiment

Definition

Epistemology is the study of the nature and grounds of knowledge especially with reference to its limits and validity.

Epistemological Considerations:

Epistemologically valid experiment

Definition

Epistemology is the study of the nature and grounds of knowledge especially with reference to its limits and validity.

Epistemological Considerations:

- Limits on Numerical Resolution

Epistemologically valid experiment

Definition

Epistemology is the study of the nature and grounds of knowledge especially with reference to its limits and validity.

Epistemological Considerations:

- Limits on Numerical Resolution
- Limits on Empirical Resolution

Epistemologically valid experiment

Definition

Epistemology is the study of the nature and grounds of knowledge especially with reference to its limits and validity.

Epistemological Considerations:

- Limits on Numerical Resolution
- Limits on Empirical Resolution
- Limits on Linguistic Resolution (future work!)

Epistemologically valid experiment

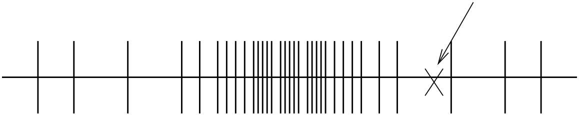
Definition

Epistemology is the study of the nature and grounds of knowledge especially with reference to its limits and validity.

Epistemological Considerations:

- Limits on Numerical Resolution
- Limits on Empirical Resolution
- Limits on Linguistic Resolution (future work!)
- Limits on ...

Limits on Numerical resolution (LNR)



Computers support a finite set of fixed length floating-point numbers of the form

$$x = \pm m \cdot b^e = \pm 0.m_1 m_2 \cdots m_p \cdot b^e$$

where, m is the signed mantissa of precision p, b is the base (usually 2) and e, bounded between $\underline{e}$ and $\overline{e}$, is the exponent. When $b = 2$, the digits of the mantissa $m_1 = 1$ and $m_i \in \{0, 1\}, \forall\, i, 1 < i \leq p$ [3].

Numerical Errors due to LNR

Numerical Errors due to LNR

- Overflow Error $(12! = 479001600, 13! \neq 1932053504)$

Numerical Errors due to LNR

- Overflow Error $(12! = 479001600, 13! \neq 1932053504)$
- Rounding Error (actual result - computed result)

Numerical Errors due to LNR

- Overflow Error $(12! = 479001600, 13! \neq 1932053504)$
- Rounding Error (actual result - computed result)
- Cancellation Error (accumulated round-off error)

Numerical Errors due to LNR

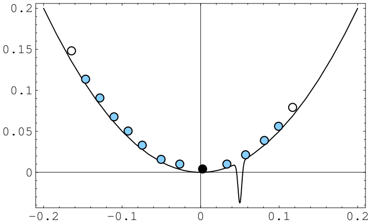
- Overflow Error ($12! = 479001600, 13! \neq 1932053504$)
- Rounding Error (actual result - computed result)
- Cancellation Error (accumulated round-off error)
- Truncation Error (from doing only finitely many operations)

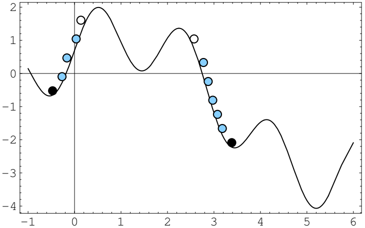
Numerical Errors due to LNR

- Overflow Error ($12! = 479001600, 13! \neq 1932053504$)
- Rounding Error (actual result - computed result)
- Cancellation Error (accumulated round-off error)
- Truncation Error (from doing only finitely many operations)
- Conversion Error (decimal to finite set of binary numbers)

Numerical Errors due to LNR

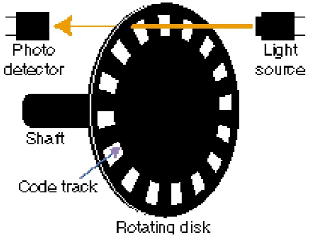
- Overflow Error $(12! = 479001600, 13! \neq 1932053504)$
- Rounding Error (actual result - computed result)
- Cancellation Error (accumulated round-off error)
- Truncation Error (from doing only finitely many operations)
- Conversion Error (decimal to finite set of binary numbers)
- Heuristic punctual local optimization is not rigorous!





Limits on Empirical Resolution

In order to make the ground of knowledge about $\mathbf{\Phi}$ with LER epistemologically sound, the empirically indiscernible sets must be allowed to enter the statistical experiment as data.



Data

- lossless compression (minimal sufficient statistic) of the trajectory
- the measurable discrete state transitions along with the transition time
- time stamps, arm-position states are integers representing intervals

sample_number, encoder1, encoder2
26 0 0
1042 -1 0
1578 -1 -1
6752 -2 -1
.
.
.
1222243 -2 20480
1229330 -1 20480

Limits on Empirical Resolution

Words of Vladik Kreinovich (two recent Los Alamos Reports on Meaurement Errors)

In many such situations, the only thing we know is the upper bound d on the measurement error. Thus, after we get the measured value X, the only information that we have about the actual (unknown) value x is that x belongs to the interval $[X - d, X + d]$.

Limits on Empirical Resolution

Words of Vladik Kreinovich (two recent Los Alamos Reports on Meaurement Errors)

In many such situations, the only thing we know is the upper bound d on the measurement error. Thus, after we get the measured value X, the only information that we have about the actual (unknown) value x is that x belongs to the interval $[X - d, X + d]$.
Here, we have two choices:

Limits on Empirical Resolution

Words of Vladik Kreinovich (two recent Los Alamos Reports on Meaurement Errors)

In many such situations, the only thing we know is the upper bound d on the measurement error. Thus, after we get the measured value X, the only information that we have about the actual (unknown) value x is that x belongs to the interval $[X - d, X + d]$.
Here, we have two choices:

(a) we can ask an expert and come up with a subjective probability distribution on this interval. However, there is no guarantee that this distribution is correct, and that the recommendations based on this subjective expert distribution are valid for the actual (unknown) distribution of the measurement error.

Limits on Empirical Resolution

Words of Vladik Kreinovich (two recent Los Alamos Reports on Meaurement Errors)

In many such situations, the only thing we know is the upper bound d on the measurement error. Thus, after we get the measured value X, the only information that we have about the actual (unknown) value x is that x belongs to the interval $[X - d, X + d]$.
Here, we have two choices:

(a) we can ask an expert and come up with a subjective probability distribution on this interval. However, there is no guarantee that this distribution is correct, and that the recommendations based on this subjective expert distribution are valid for the actual (unknown) distribution of the measurement error.

(b) Another approach is to use robust statistics – a special type called interval computations. We do not know the exact distribution, we only know that this distribution is located on the interval. So, we want to make conclusions which are valid no matter what this distribution is.

Epistemologically Valid Experiment

We want an **epistemologically valid experiment** that accounts for the physical limits on

- empirical resolution ("show what you can actually see")
- numerical resolution ("compute what you actually can")

Solution

Solution

- Action Space $\mathcal{A}$ of the classical estimation problem is merely the parameter space $\mathbf{\Theta}$

Solution

- Action Space $\mathcal{A}$ of the classical estimation problem is merely the parameter space $\boldsymbol{\Theta}$
- Epistemologically valid action space $\boldsymbol{\mathcal{A}}$ is a machine-representable Hausdorff-extension the Parameter Space

$$\boldsymbol{\Theta} \xrightarrow{\text{EN}} \mathbb{I}\boldsymbol{\Theta}^* := \mathbb{I}\boldsymbol{\Theta} \cup \emptyset$$

Solution

- Action Space $\mathcal{A}$ of the classical estimation problem is merely the parameter space $\boldsymbol{\Theta}$
- Epistemologically valid action space $\boldsymbol{\mathcal{A}}$ is a machine-representable Hausdorff-extension the Parameter Space

$$\boldsymbol{\Theta} \xrightarrow{\text{EN}} \mathbb{I}\boldsymbol{\Theta}^* := \mathbb{I}\boldsymbol{\Theta} \cup \emptyset$$

- $\mathbb{I}\boldsymbol{\Theta}$ is the set of all compact boxes in $\boldsymbol{\Theta}$.

Solution

- Action Space $\mathcal{A}$ of the classical estimation problem is merely the parameter space $\boldsymbol{\Theta}$
- Epistemologically valid action space $\boldsymbol{\mathcal{A}}$ is a machine-representable Hausdorff-extension the Parameter Space

$$\boldsymbol{\Theta} \xrightarrow{\text{EN}} \mathbb{I}\boldsymbol{\Theta}^* := \mathbb{I}\boldsymbol{\Theta} \cup \emptyset$$

- $\mathbb{I}\boldsymbol{\Theta}$ is the set of all compact boxes in $\boldsymbol{\Theta}$.
- $\emptyset$ has to be added to our epistemologically valid $\boldsymbol{\mathcal{A}}$

Solution

- Action Space $\mathcal{A}$ of the classical estimation problem is merely the parameter space $\mathbf{\Theta}$
- Epistemologically valid action space $\boldsymbol{\mathcal{A}}$ is a machine-representable Hausdorff-extension the Parameter Space

$$\mathbf{\Theta} \xrightarrow{\text{EN}} \mathbb{I}\mathbf{\Theta}^* := \mathbb{I}\mathbf{\Theta} \cup \emptyset$$

- $\mathbb{I}\mathbf{\Theta}$ is the set of all compact boxes in $\mathbf{\Theta}$.
- $\emptyset$ has to be added to our epistemologically valid $\boldsymbol{\mathcal{A}}$
- identifiability of the extended experiment indexed by $\mathbb{I}\mathbf{\Theta}$ in terms of symmetric set difference follows from identifiability of the original experiment indexed by $\mathbf{\Theta}$ and inclusion monotony of the index map (likelihood or conditional probability of data given parameter)

Thanks

- Many thanks to:
    - Piers Lawrence for completing the physical double pendulum in Civil Engg Dept.'s Lathe (Alan Nicholson), Richard Brown coordinated Electronic design and Mike Stuart did it
    - UCDMS for supporting the double pendulum project (especially)
        - Bob Broughton (logistics, parts order, etc)
        - David Wall ($ and kind words)
    - Douglas Bridges et al's ConstruMath Grant for UppsalaCAPA-CanterburyUCDMS air-traffic

Bibliography

D Hume.
An enquiry concerning human understanding: Section VI - OF PROBABILITY.
In CW Eliot, editor, *The Harvard classics:English philosophers of the seventeenth and eighteenth centuries, 1910 edition*, volume 37. The Collier Press, 1777.

A Neumaier.
Interval methods for systems of equations.
Cambridge university press, 1990.

IEEE Task P754.
ANSI/IEEE 754-1985, Standard for Binary Floating-Point Arithmetic.
IEEE, New York, 1985.

AN Shiryaev.
Probability.
Springer-Verlag, 1989.

Piers Lawrence, Michael Stuart, Richard Brown, Warwick Tucker and Raazesh Sainudiin.
A mechatronically measurable double pendulum for machine interval experiments.
Indian Statistical Institute Technical Report, isibang/ms/2010/11, October 25, 2010