Promises and Limits of Machine Learning

Thomas J. Sargent
SUSTech Frontier Science Symposium
Shenzhen, China

November 11, 2023

Theme

"Parameter" means different things to different scientists, and in different applications

Richard Feynman's Metaphor

- Ignorant of the rules of chess, an observer watches two people play one game of chess
- From those data, the observer is asked to infer the rules of the game of chess
- Likewise, a physicist infers laws of nature from incomplete observations of a physical system

Relevance of Richard Feynman's Metaphor

- Feynman's metaphor literally describes what a structural economic statistician tries to do
- From incomplete observations of prices and quantities, an economic statistician wants to infer a game that generated those prices and quantities

A game

- Lists of
  - players
  - available choices
  - payoffs
  - information sets
- a timing protocol
- an equilibrium concept

Physics Metaphors and Economics

von Neumann and Morgenstern (chapter 1, 1944) and Koopmans (1946) described two types of statistical models:

- Kepler stage: descriptive model with curve-fitting parameters
- Newton stage: structural model with parameters characterizing physical invariants

Physics Metaphors and Economics

Parameters two types of models:

- Kepler's third law: $T^2 = Kr^3$; invariant parameters: $2, 3, K$
- Newton's interpretation of Kepler's third law:
  - $K = \frac{2\pi}{GM}$; G invariant parameter
  - $F = mr\omega^2 = G\frac{mM}{r^2}$ (centripetal force);

Imitating Feynman in Economics

On a little reflection, it is difficult to feel any general optimism as to the solubility of this problem . . . One would need to know a great deal about pigeons in order to construct a robot which would serve as a good imitator of an actual pigeon in Skinner's tests. As time-series econometricians, we are in the position of attempting to do exactly this, using observations only on pigeons in their natural environment. If any success is to be possible, it will clearly involve some boldness in the use of economic theory.

Robert E. Lucas, Jr. From the introduction to Rational Expectations and Econometric Practice, *University of Minnesota Press, 1981.*

Machine Learning

Components:

- data and labels: $\{x_i, y_i\}_{i=0}^{N}$
- a function to learn: $y_i = f(x_i)$
- a parameterized set of possible functions: $f \in \mathcal{F} = \{f_{\theta \in \Theta}\}$
- curse of dimensionality

Curse of Dimensionality

... simple arguments ... reveal the impossibility of learning from generic high dimensional data as a result of the curse of dimensionality ...

Geometric Deep Learning: Grids, Groups Graphs, Geodesics, and Gauges, M. Bronstein et al. (2021)

Curse of Dimensionality

While simple arguments . . . reveal the impossibility of learning from generic high dimensional data as a result of the curse of dimensionality, there is hope for physically-structured data, where we can employ two fundamental principles: symmetry and scale invariance. . . .

Geometric Deep Learning: Grids, Groups Graphs, Geodesics, and Gauges, M. Bronstein et al. (2021)

Machine Learning

Components:

- data and labels: $\{x_i, y_i\}_{i=0}^{N}$
- function to learn: $y_i = f(x_i)$
- function class: $f \in \mathcal{F} = \{f_{\theta \in \Theta}\}$
- coping with curse of dimensionality
  - add 'fake data' in form of a prior
  - capitalize on associated inductive bias
  - justify inductive bias as "regularization"

Two Types of Priors

- geometric
  - composition of simple functions: $f = h \circ g \circ k \circ h \circ g \circ k \circ \cdots$
  - design 'architecture' of component functions h, g, k to build in symmetries (invariances)
  - use chain rule to get FONC's
- economic
  - posit a dynamic game with small number of parameters (invariants) that describe players' payoffs and physical constraints
  - use restrictions across players' equilibrium strategies to infer parameters
  - interpret those parameters as "invariants" that allow the economist to study consequences of historically unprecedented economic policies

Two Types of Priors

Common desiderata:

- Invariants
- Regularizers

Geometric Priors

Symmetries and architectures, e.g., CNN's (convolutional neural networks)

- translational invariance
- scale invariance (pooling)

Geometric Priors

Symmetries and architectures, e.g., GNN's (graph neural networks)

- translational invariance
- scale invariance (pooling)
- permutation invariance

Economic example: inflation dynamics

- Data: $\{\mu_t, \pi_t\}_{t=0}^{T}$
- Inflation sequences generated by "nature" and observed by statistician (oops, I mean a data scientist):

$$\pi_t = d_0 \left(\frac{1 - d_1^t}{1 - d_1} \right) + d_1^t \pi_0$$

$$\mu_t = g_0 + g_1 \left(\frac{1 - g_2^t}{1 - g_2} \right) + g_3 g_2^t \mu_0$$

Descriptive Model of Inflation dynamics

- Data: $\{\mu_t, \pi_t\}_{t=0}^{T} \equiv (\vec{\mu}, \vec{\pi})$
- Inflation sequences generated by "nature":

$$\pi_t = d_0 \left(\frac{1 - d_1^t}{1 - d_1} \right) + d_1^t \pi_0$$

$$\mu_t = g_0 + g_1 \left(\frac{1 - g_2^t}{1 - g_2} \right) + g_3 g_2^t \mu_0$$

- Function to learn: $(\mu_t, \pi_t) = f(t)$
- Functional form
- Parameters: $\theta = (d_0, d_1, g_0, g_1, g_2, g_3) \in \Theta$

Descriptive Model of Inflation dynamics

- Kepler stage aim of acquiring concise description of observed inflation dynamics
- No reverse engineering of “rules of game”

Economic Prior for Inflation Dynamics

- Reverse engineer “rules of a game”
- Want parameters that are **invariant** to alterations in parameters of money growth process $\vec{\mu}$.

Economic Prior for Inflation Dynamics

(based on Structural Model of Calvo (1978) and Chang (1999))
(see `https://python-advanced.quantecon.org/calvo.html`)

- Lists of
  - people – private agents, central bank and treasury officials, ...
  - their preferences
  - their constraints
- timing protocol
- Nash equilibrium

Economic Prior

- Function to estimate: $(\mu_t, \pi_t) = h(t), h \in \mathcal{H} = \{h_{\delta \in \Delta}\}$
- $\delta = a_0, a_1, a_2, \beta; \gamma; \beta$
- Parameters to estimate: $\delta \in \Delta$:
  - Monetary-fiscal authority's purposes: a_0, a_1, a_2, β
  - Tax distortion costs: γ
  - Private agents' money demand interest elasticity: α

Implications of Economic Prior

$\delta = (a_0, a_1, a_2, \beta; \gamma; \alpha) \in \Delta$ and the $g(t)$ function is

$$\pi_t = d_0 \left(\frac{1 - d_1^t}{1 - d_1} \right) + d_1^t \pi_0$$

$$\mu_t = b_0 + b_1 d_0 \left(\frac{1 - d_1^t}{1 - d_1} \right) + b_1 d_1^t \pi_0$$

where

$$(d_0, d_1, b_0, b_1) = \phi(\delta)$$
$$(\pi_0, \mu_0) = \psi(\delta)$$

These are instances of the hallmark "cross-equation restrictions" of rational-expectations econometrics.

Two Classes of Priors

Geometric:

- Descriptive
- Measurement without Theory (T.C. Koopmans, 1946)
- Auxiliary (Gallant and Tauchen, 1996)

Economic:

- Structural (Koopmans, 1946-1950; Lucas (1973), Rust, Wolpin, . . .,)

Deploying ML for Structural Models

- As input to solving "direct problems"
  - Hamilton-Jacobi-Bellman equations take form $f_\theta = T(f_\theta), \quad \theta \in \Theta$
  - T is an operator, Θ a manifold of parameters describing game
  - f_θ is a *value* or *policy* function
  - Approximate f_θ with composition of functions parameterized by auxiliary parameters
- "Deep learning" approaches to solving partial differential equations associated with HJB equations

Concluding Remarks

To answer "what is a parameter?" modern AI and ML distinguishes two types of statistical model

- Descriptive (Kepler stage)
- Structural (Newton stage)

In this way, it respects the "Lucas Critique".