Optimal Control and Inverse Problems for Data-Driven Decision-Making

Nina Chafee, Micah Chandler, Benjamin Lebdaoui, Meredith Orton

Haverford College,University of North Georgia, University of Montevallo, LaGrange College

Emory Models Meet Data REU 2025

EMORY
UNIVERSITY

Mentored by Dr. John Darges, Emory University
This work was supported by NSF Grant DMS-2349534

We're farmers. . .

. . . and we want to efficiently grow crops!

The Problem:

- What's good for crops is bad for local environment
- Rabbits are particularly vulnerable to change
- Wolves are endangered

What measures can we take that will maximize crops without decreasing wolves?

Our Project

GOAL:

- Focus on ODE, dynamical systems

$$\frac{dy}{dt} = F(y(t), t)$$

- Control some outcome for a real world system

Problems:

- Only have noisy measurements
- Unknown parameters

Solution: Parameter estimation

Square One

- What is a model?

$$\overset{\text{input}}{\overbrace{}}\quad \underset{\text{output}}{y} = F(\overset{\text{input}}{t}; \underset{\text{parameter}}{\theta})$$

- What is the data?

$$d_i = F(t; \theta) + \eta$$

Definition: Inverse Problems

Based on observed data, reconstruct model with parameter estimation

The ODE Models

Lotka-Volterra:

$$\frac{dp}{dt} = \theta_1\, p - \theta_2\, p\, r, \qquad p(0) = p_0$$
$$\frac{dr}{dt} = \theta_4\, p\, r - \theta_3\, r, \qquad r(0) = r_0$$

Lorenz System[1]:

$$\frac{dx}{dt} = \alpha(y - x), \qquad x(0) = x_0$$
$$\frac{dy}{dt} = x(\rho - z) - y, \qquad y(0) = y_0$$
$$\frac{dz}{dt} = xy - \beta z, \qquad z(0) = z_0$$

[1] Fowler, et al., *The complex Lorenz equations*, 1982.

Estimating Parameters With a Deterministic Approach

The Deterministic Approach

- Parameters $\hat{\theta}$ are **fixed unknowns**
- Data is pairings: (t_i, d_i)
- Measure difference with **objective function**

$$\text{Ordinary Least Squares (OLS): } \operatorname{argmin}_{\theta} \sum_{i=1}^{N} (F(t_i;\theta) - d_i)^2$$

Analyzing the data:

1. Guess initial conditions
2. Optimize objective
3. Plot true and estimated trajectories

Deterministic Results: Lotka-Volterra

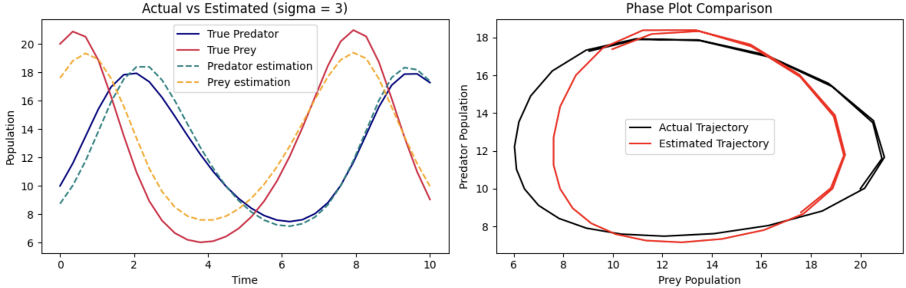


Figure: Deterministic estimation on synthetic noisy data ($\sigma = 3.0$) with initial condition $r_0 = 10$, $p_0 = 10$, parameter guess $\theta_0 = \begin{bmatrix} 1.0 & 0.1 & 1.0 & 0.1 \end{bmatrix}$.

Deterministic Results: Lorenz Model

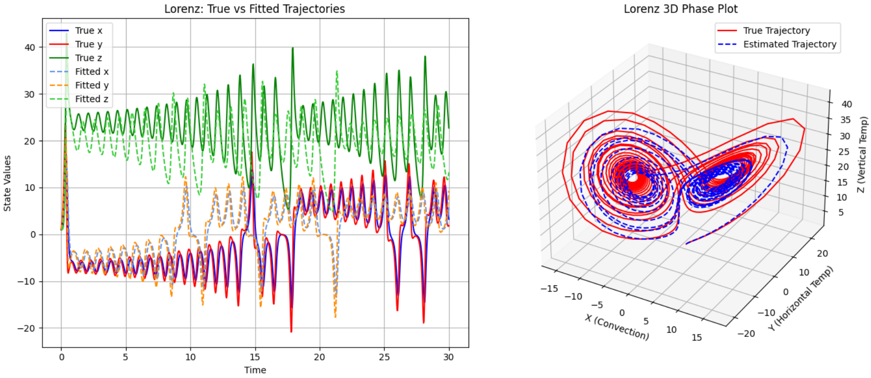


Figure: Frequentist estimation on synthetic noisy data ($\sigma = 0.2$) with initial condition $x_0 = 1$, $y_0 = 1$, $z_0 = 1$ parameter guess $\theta_0 = \begin{bmatrix} 9 & 25 & 2 \end{bmatrix}$.

What Would Bayes Do?

Problems with deterministic approach:

1. Can't incorporate uncertainty
2. No inclusion of prior knowledge

Definition: Bayesian Parameter Estimation

Treat parameters as random variables and incorporate prior beliefs to estimate parameters as distributions

GOAL: Based on the data, capture the distribution of likely parameters.

Bayesian Approach!

Prior

Choose $\Pi_{\text{prior}}(\theta)$ based on what we know about the parameters

Probability of data given parameters assuming Gaussian noise

$$\Pi_{\text{like}}(d \mid \theta) = \exp(-1/2 * \Sigma(F(x_i;\theta) - d_i)^2/\sigma^2)$$

Probability of parameters given data (Bayes' Rule)

$$\Pi_{\text{post}}(\theta \mid d) = \frac{\Pi_{\text{like}}(d \mid \theta) * \Pi_{\text{prior}}(\theta)}{\int \Pi_{\text{like}}(d \mid \theta) * \Pi_{\text{prior}}(\theta)\, d\theta}$$

Traits of a Distribution We Want to Capture

$$\mathbb{E}(\theta) = \int \theta * \Pi_{post}(\theta \mid d)\, d\theta \qquad \text{Mean}$$

$$\mathrm{Var}(\theta) = \int \theta^2 * \Pi_{post}(\theta \mid d)\, d\theta - \mathbb{E}(\theta)^2 \qquad \text{Variance}$$

$$\mathrm{Cov}(\theta) = \int (\theta - \theta_{\text{mean}})(\theta - \theta_{\text{mean}})^T * \Pi_{\text{post}}(\theta \mid d)\, d\theta \qquad \text{Covariance}$$

Problem: integration is costly!

How Do Get a Sample?

Definition: Markov Chain Monte Carlo (MCMC) methods

MCMC methods generate approximate samples of a target distribution using a chain of random variables.

METROPOLIS-HASTINGS ALGORITHM:

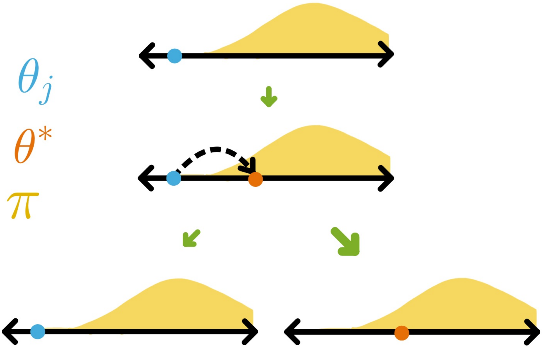
Ingredients:

- Proposal distribution: $q(\theta_j, \theta_{j-1})$
- Target distribution: $\Pi_{\text{post}}(\theta \mid d)$

Steps:

1. θ_j, sample $\theta^* \sim q(\theta_j)$
2. Evaluate quality of θ^* compared to θ_j using Π_{post}
3. Based on evaluation accept $(\theta_{j+1} = \theta^*)$ or reject $(\theta_{j+1} = \theta_j)$

How Do Get a Sample?



Steps:

1. θ_j, sample $\theta^* \sim q(\theta_j)$
2. Evaluate quality of θ^* compared to θ_j using Π_{post}
3. Based on evaluation accept ($\theta_{j+1} = \theta^*$) or reject ($\theta_{j+1} = \theta_j$)

What Are We Looking For?

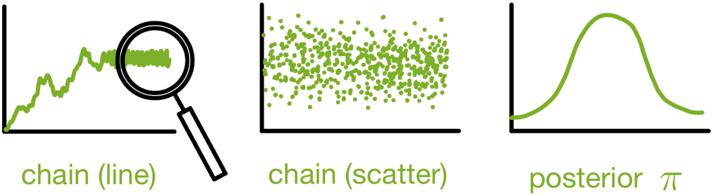


Figure: Heuristic depiction of a “good” chain and possible posterior distribution.

Note:

“Nice” Markov chains approach a unique **stationary distribution**.

Good news: success is guaranteed!

What Are We Looking For?

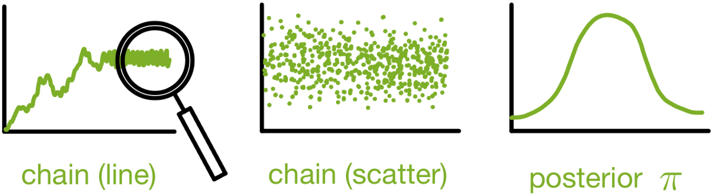


Figure: Heuristic depiction of a "good" chain and possible posterior distribution.

Note:

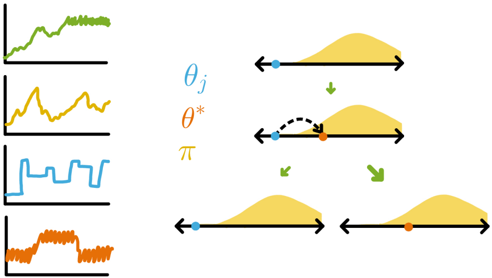
"Nice" Markov chains approach a unique **stationary distribution**.

Good news: success is guaranteed!

Bad news: this could take forever!

Crisis Aversion

$q(\theta)$	**Problem**	**Solution**
Too narrow	Limited exploration	Increase proposal scale
Too wide	Low acceptance rate	Decrease proposal scale
Wrong shape	Poor mixing	Match posterior geometry



MCMC Results: Lotka-Volterra

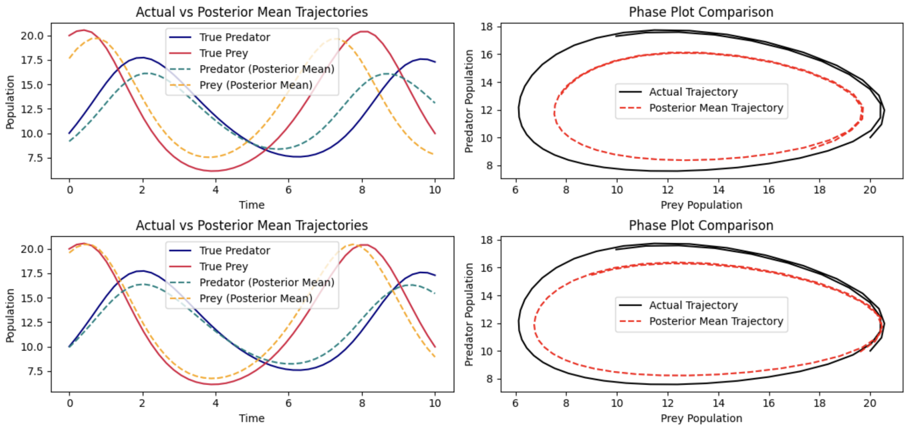


Figure: Bayesian ensemble sampling (above) and subsequent Metropolis-Hastings (below) estimation on synthetic noisy data ($\sigma = 3.0$) with initial condition $r_0 = 10$, $p_0 = 10$, parameter guess $\theta_0 = \begin{bmatrix} 1.0 & 0.1 & 1.0 & 0.1 \end{bmatrix}$. Inputs for Metropolis-Hastings recovered from ensemble estimates.

MCMC Results: Lorenz Model

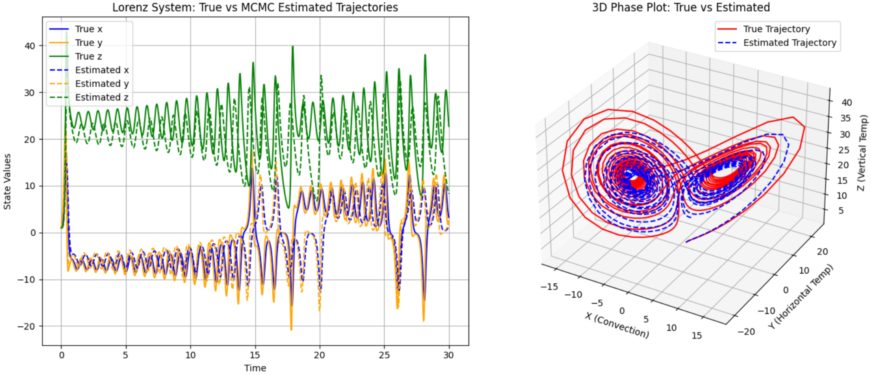


Figure: Bayesian ensemble sampling estimation on synthetic noisy data (σ=0.2) with initial condition x_0=1, y_0=1, z_0=1, parameter guess $\theta_0 = \begin{bmatrix} 9 & 25 & 2 \end{bmatrix}$.

Back to the Farm...

Rabbits have become a pest. How do we intervene without hurting wolf population?

Example: Introduce limited-time rabbit hunting to control the population.

- When should we intervene?
- How strongly should we act?

We must balance this with natural predation.

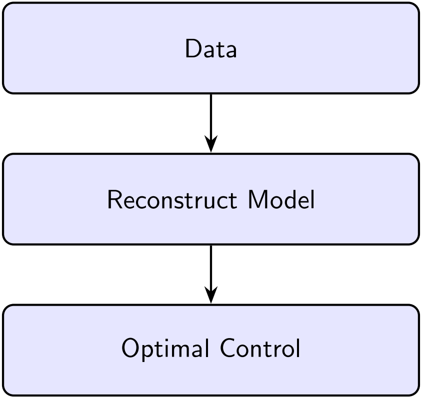
Optimal Control allows us to maximize crop yields, without hurting wolves.

- Start with a basic model:
$$\frac{dx}{dt} = f(x(t), t)$$

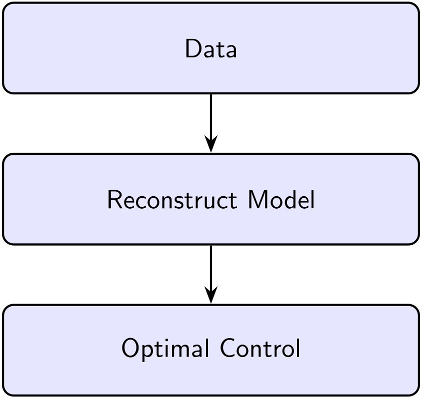
- Now introduce a **control** input $\alpha(t)$:
$$\frac{dx}{dt} = f(x(t), \alpha(t), t)$$

GOAL: $\min_{\alpha \in A} c_{x,t}(\alpha)$

What's Next?



What's Next?



Thank you!

Murdered Darlings Below

Simplifying the Transition Kernel

Transition Kernel: $\alpha(\theta_j, \theta^*) = \min\left\{1,\ \dfrac{\Pi_{\text{post}}(\theta^*)\, q(\theta^*, \theta_j)}{\Pi_{\text{post}}(\theta_j)\, q(\theta_j, \theta^*)}\right\}$

Problem: This is in terms of our posterior distribution

Definition: Bayes' Formula

$$\Pi_{post}(\theta|d) = \Pi_{like}(d|\theta)\Pi_{prior}(\theta)$$

Note:

$$\text{Assuming } \varepsilon \sim \triangleright(\mu, \sigma^{\wedge}) \implies \frac{q(\theta^*,\theta_j)}{q(\theta_j,\theta^*)} = 1$$

$$\alpha(\theta_j, \theta^*) = \min\left\{1,\ \frac{\Pi_{\text{like}}(\theta^*)\, \Pi_{\text{post}}(\theta^*)}{\Pi_{\text{like}}(\theta_j)\, \Pi_{\text{post}}(\theta_j)}\right\}$$

Convert to Log-Distributions

Likelihood:

- $\Pi_{loglike}(\theta) = \dfrac{-\Sigma(d_i - y_i)^2}{2\sigma^2}$

Prior:

- $\log(\Pi_{prior}(\theta)) = \Pi_{logprior}(\theta)$

Posterior:

- $\Pi_{logpost}(\theta) = \Pi_{loglike}(\theta) + \Pi_{logprior}(\theta)$

Pros:

- Reduces complexity and computational time
- Exponentials become constants
- Multiplication becomes addition

Ensemble Sampling Pseudo-Algorithm

For time $j = 1$ to M and each walker ${}^k\theta$:

1. Pick ${}^j\theta \neq^k \theta$
2. Propose $\theta^* = {}^j\theta + Z * ({}^k\theta) - {}^j\theta$, where $Z \sim \text{Uniform}(\sqrt{1/a}, \sqrt{a})^2$
3. Calculate $\alpha = \left(1, Z^{d-1} * \frac{\Pi_{\text{post}}(\theta^*)}{\Pi_{\text{post}}({}^k\theta)}\right)$
4. Accept θ^* as the next in sequence ${}^k\theta$ with probability α

Pros:

- Parallelizable
- Efficient with correlation

Metropolis-Hasting from Scratch

For this experiment, we defined our own likelihood, prior, posterior, transition kernel, proposal mechanism, and Markov Chain construction. The proposal used the Cholesky decomposition of the covariance matrix V by defining

$$q_{proposal} = q_{current} + Rz$$

where R is Chol(V), and $z \sim \triangleright(0, 1)$ the initial guess was the result of the OLS estimator

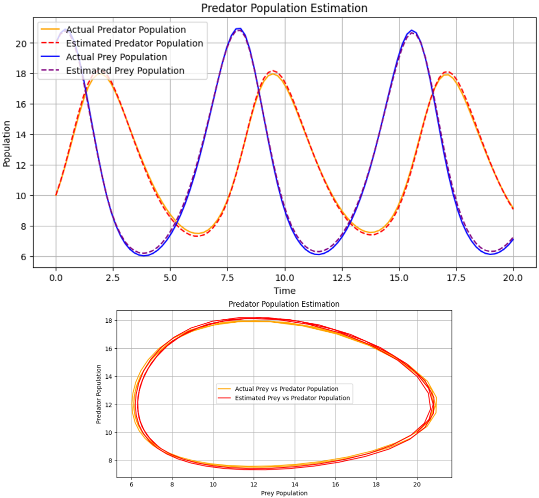


Figure: Metropolis-Hastings result with initial guess $\theta = [0.75, 0.09052, 0.9645, 0.07896]$ (result from OLS) and $\sigma = 1.0$

Optimal Control Deep Dive

Ingredients:

- Admissible controls $\mathbb{A} = \{\alpha : [O, T] \to C\}$ where C is compact in $\mathbb{R}^n$
- Value function $u(x, t) = \inf_{\alpha \in \mathbb{A}} c_{x,t}(\alpha)$
- Cost function $c_{x,t} : \mathbb{A} \to \mathbb{R}$, $c(\alpha) = g(\mathbf{x}(t)) + \int_0^T h(\mathbf{x}(s), \alpha(s))ds$

Steps:

1.

Problems:

- Defining cost function
- Defining admissible controls
- Optimizing over uncountable spaces
- Time of introduction

Optimal Control for the Lorenz System

The Lorenz system is well known for its chaotic behavior. Our goal in the following weeks is to find control inputs $V_1(t)$, $V_2(t)$, $V_3(t)$ that, when added to the system equations, guide the system toward an equilibrium state and stabilize its dynamics.

Controlled Lorenz System:

$$\frac{dx}{dt} = \alpha(y - x) + V_1(t)$$
$$\frac{dy}{dt} = x(\rho - z) - y + V_2(t)$$
$$\frac{dz}{dt} = xy - \beta z + V_3(t)$$

With the right control functions, the state variables $x(t), y(t), z(t)$ can converge to a desired equilibrium point, suppressing the system's chaotic behavior.

Source: El-Gohary, A., & Bukhari, F. (2021). *Optimal Control of Lorenz System During Different Time Intervals.*

Optimal Control of Predator-Prey System

One example of optimal control in dynamic systems is its application to the Lotka–Volterra predator-prey model. In this case, we introduce a control input to influence the prey population:

- $u(t)$: the rate of pesticide application

When pesticides are applied, the prey equation becomes:

$$\frac{dp}{dt} = \theta_1 p - \theta_2 pr - b_1 u(t) p$$

Source: Goh, B. S., Leitmann, G., & Vincent, T. L. (1974). *Optimal Control of a Prey-Predator System*. *J. Opt. Theory Appl.*, 14(4), 265–280.