Convergence and Equilibrium for Stochastic Models of Ecological Disturbances

James Broda
jbroda@bowdoin.edu

Nicollet Island-Northeast Fire, 1893

California wildfires, 2019

California wildfires, 2019

An ecological problem

How can one model the carbon content of an ecosystem with randomly occurring disturbances of random severity?

Today's plan

1. Discrete time and state introduction

2. Continuous time and state Semistochastic model

Questions to keep in mind

1. What information can we extract from equilibrium distributions?
2. Why do convergence rates matter?

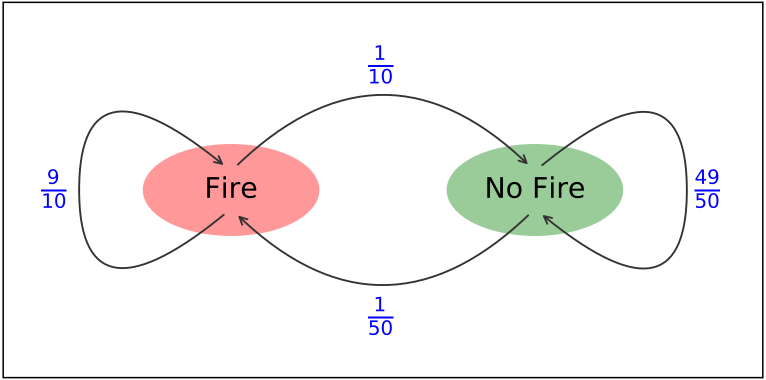
Markov chains

1. $\mathcal{X}$: State space
2. μ_0 : Initial distribution
3. Q : Transition matrix

Two-state Markov chain

1. $\mathcal{X}$: $\{$Fire, No Fire$\}$
2. μ_0 : Is there a fire now?
3. Q : Environment, beliefs

Notebook example



Evolved distributions

- Probability vectors

$$\mu_n = \begin{pmatrix} \text{Prob. of fire on } n^{th} \text{ day} \\ \text{Prob. of no fire on } n^{th} \text{ day} \end{pmatrix}^T$$

- Initial probability vector: μ_0

$$\mu_n = \mu_0 \begin{pmatrix} \frac{9}{10} & \frac{1}{10} \\ \frac{1}{50} & \frac{49}{50} \end{pmatrix}^n = \mu_0 Q^n$$

Equilibrium distributions

- Limiting:

$$\lim_{n\to\infty} \mu_n = \pi \quad \text{for any } \mu_0$$

- Invariant:

$$\mu_0 = \pi \Rightarrow \mu_n = \pi \quad \text{for all } n > 0$$

Note: These two characterizations are not always equivalent!

An invariant approach

$$\pi = \pi Q$$
$$= \pi \begin{pmatrix} \frac{9}{10} & \frac{1}{10} \\ \frac{1}{50} & \frac{49}{50} \end{pmatrix}$$

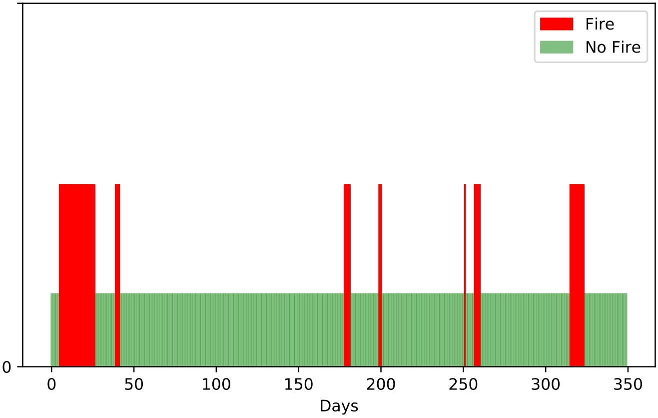
π is a *left*-eigenvector of the transition matrix, Q, with eigenvalue 1.

$$\pi = \left(\frac{1}{6}, \quad \frac{5}{6}\right)$$

Question:
What information can we extract from this equilibrium distribution?

Answer:
In the long run, there will be fires on 1 out of 6 days.

Realization of Markov chain



What about the other eigenvalue?

The eigenvalues of Q satisfy $\lambda_1 = 1$ and $|\lambda_2| = 1 - \beta$ with $\beta = |1 - \frac{1}{10} - \frac{1}{50}| = \frac{6}{50}$.

$$Q = \begin{pmatrix} \frac{9}{10} & \frac{1}{10} \\ \frac{1}{50} & \frac{49}{50} \end{pmatrix}$$

The other eigenvalue is related to the rate at which an arbitrary initial distribution, μ_0, converges to π. One can show

$$\mathrm{d}_{TV}(\mu_n, \pi) \leq (1 - \beta)^n .$$

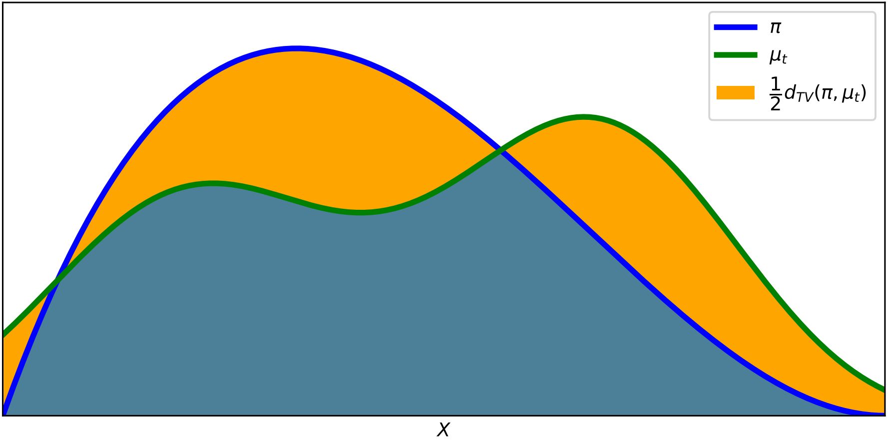
Total variation distance

Given probability distributions, μ and ν:

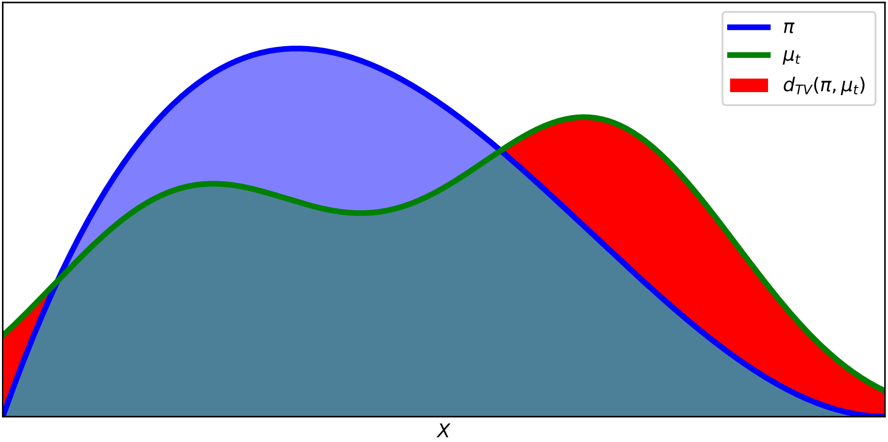
$$d_{TV}(\mu, \nu) = \sup_A |\mu(A) - \nu(A)|$$

$$d_{TV}(\mu, \nu) = \sup_{0 \le f \le 1} |\mathbb{E}_\mu(f) - \mathbb{E}_\nu(f)|$$

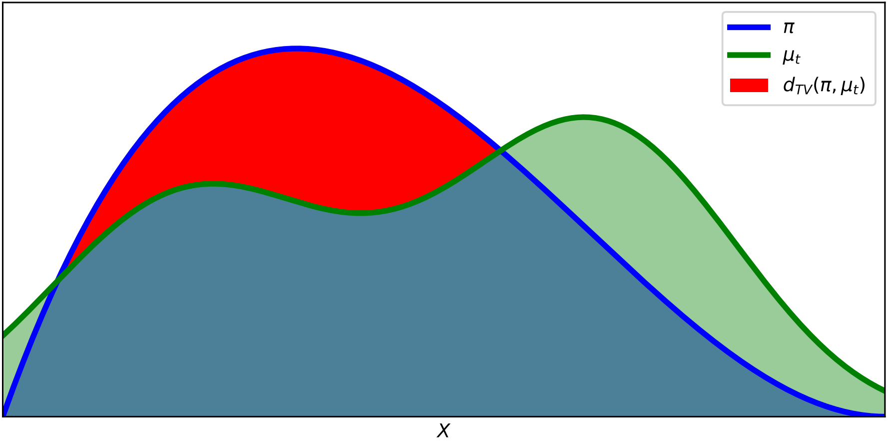
Total variation distance



Total variation distance



Total variation distance



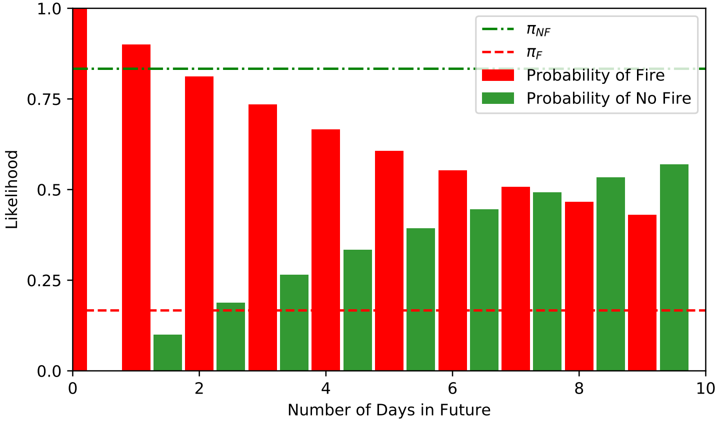
Convergence rates

Question:
Why do convergence rates matter?

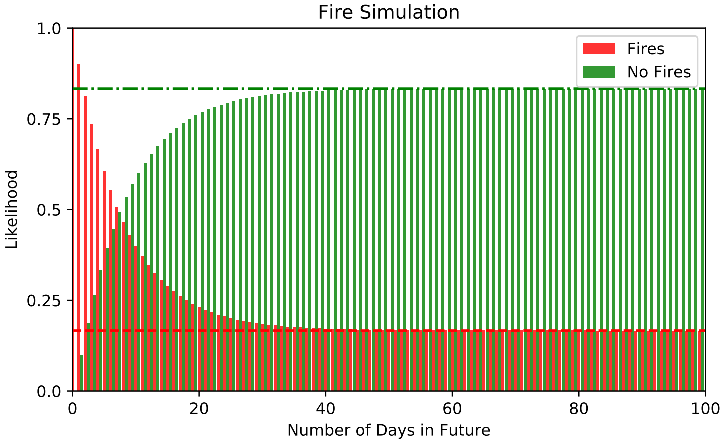
Answer:
The “long run” may be a long time coming.

Markov chains are “memoryless”, but need time to forget.

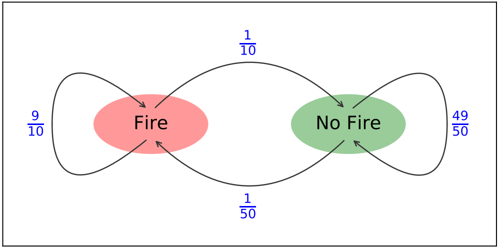
Evolution of probabilities



Evolution of probabilities

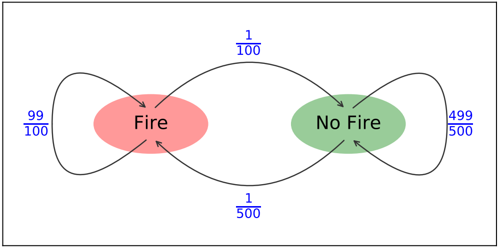


FAST:



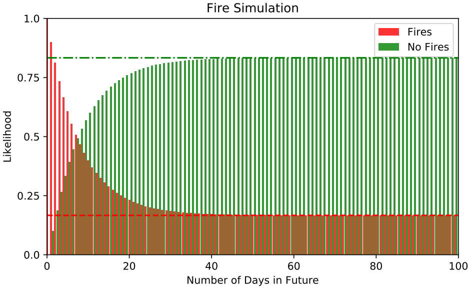
Fire
No Fire
$\frac{9}{10}$
$\frac{1}{10}$
$\frac{49}{50}$
$\frac{1}{50}$

SLOW:

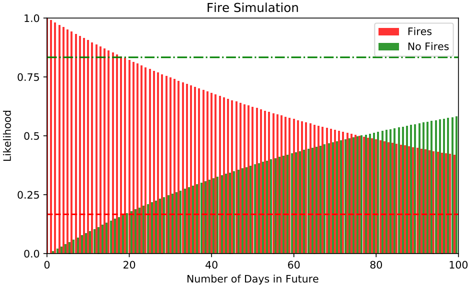


Fire
No Fire
$\frac{99}{100}$
$\frac{1}{100}$
$\frac{499}{500}$
$\frac{1}{500}$

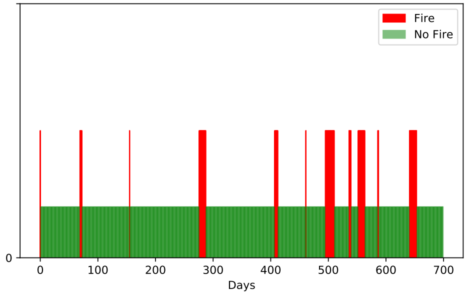
FAST:


Fire Simulation
Fires
No Fires
Likelihood
Number of Days in Future
1.0
0.75
0.5
0.25
0.0
0
20
40
60
80
100

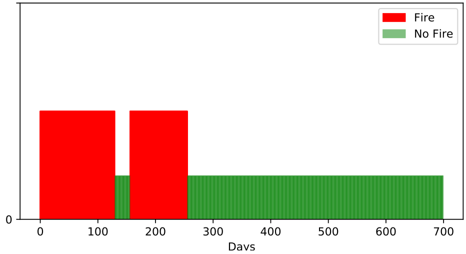
SLOW:


Fire Simulation
Fires
No Fires
Likelihood
Number of Days in Future
1.0
0.75
0.5
0.25
0.0
0
20
40
60
80
100

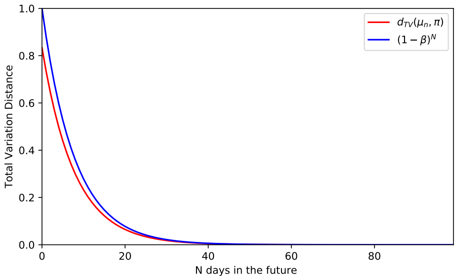
FAST:


Fire
No Fire
0
0
100
200
300
400
500
600
700
Days

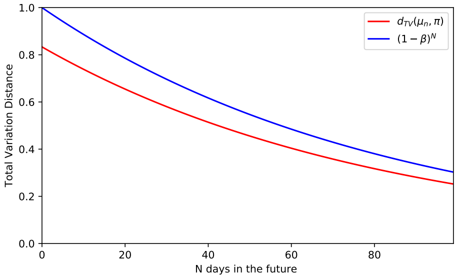
SLOW:


Fire
No Fire
0
0
100
200
300
400
500
600
700
Days

FAST:


$d_{TV}(\mu_n, \pi)$
$(1-\beta)^N$
Total Variation Distance
N days in the future
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80

SLOW:


$d_{TV}(\mu_n, \pi)$
$(1-\beta)^N$
Total Variation Distance
N days in the future
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80

Rewind

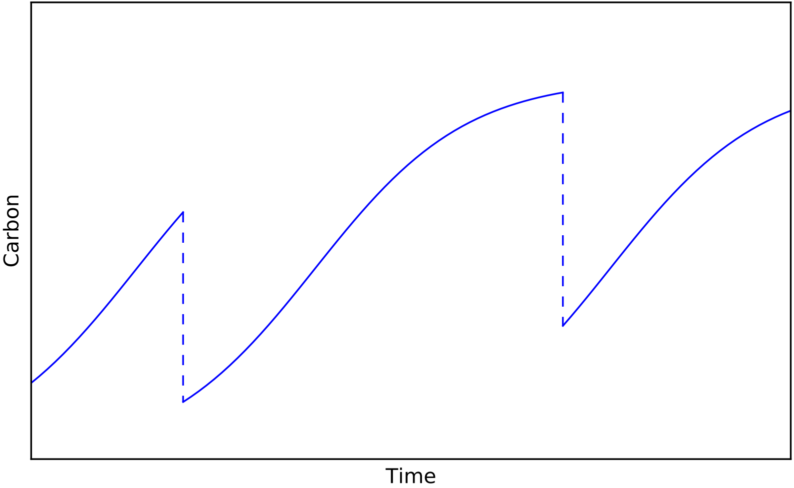
How can one model the carbon content of an ecosystem and account for randomly occurring disturbances of random severity?

Semistochastic model for the carbon content of an ecosystem

Model design:

- Carbon content increases deterministically between disturbances.
- Fires occur at random times and release carbon.
- Severity of fires is random.

Semistochastic model



Components of semistochastic model

Growth Rate – Deterministic evolution

$$\frac{\mathrm{d}x}{\mathrm{d}t} = g(x)\,, \qquad x(t) = \phi^t(x_0)$$

Disturbance Rate – Probability per unit time

$$\Lambda(x)$$

Disturbance Kernel – Severity of disturbances

$$P(x, A)$$

Components of semistochastic model

Growth Rate – Deterministic evolution

$$\frac{\mathrm{d}x}{\mathrm{d}t} = g(x)\,, \qquad x(t) = \phi^t(x_0)$$

Disturbance Rate – Probability per unit time

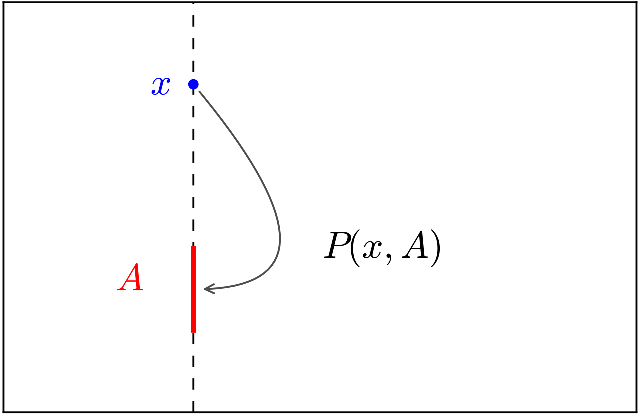
$$\Lambda(x)$$

Disturbance Kernel – Severity of disturbances

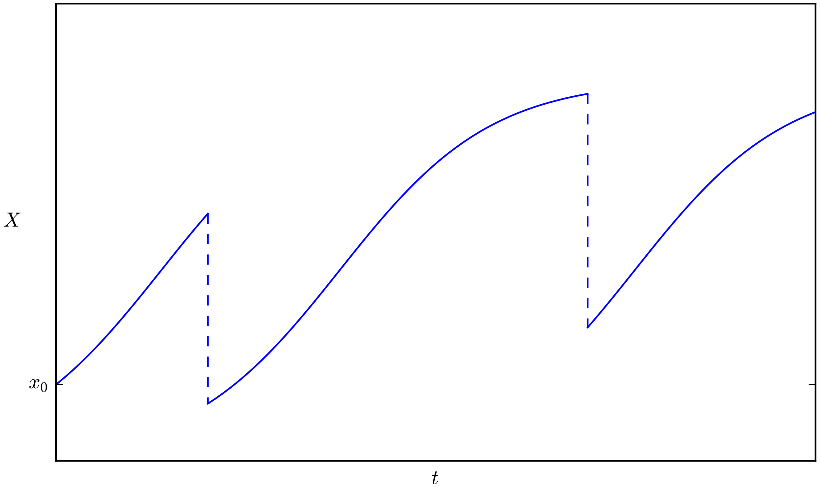
$$P(x, A)$$

This state dependence is important!!

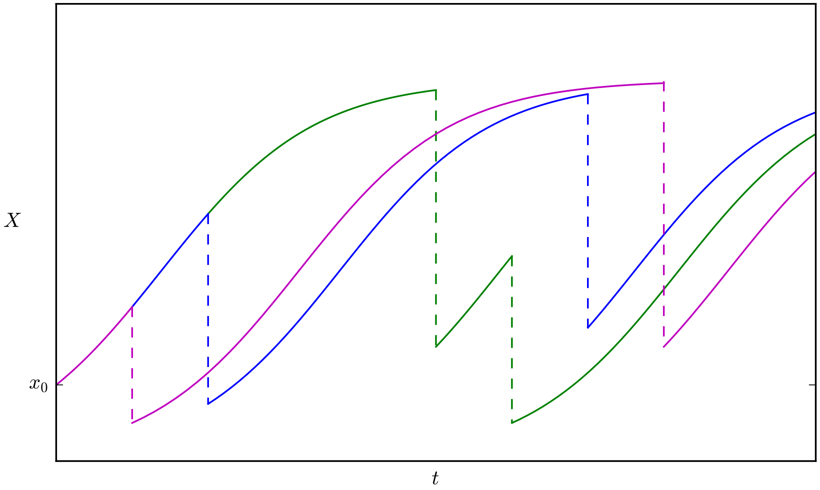
Disturbance kernel



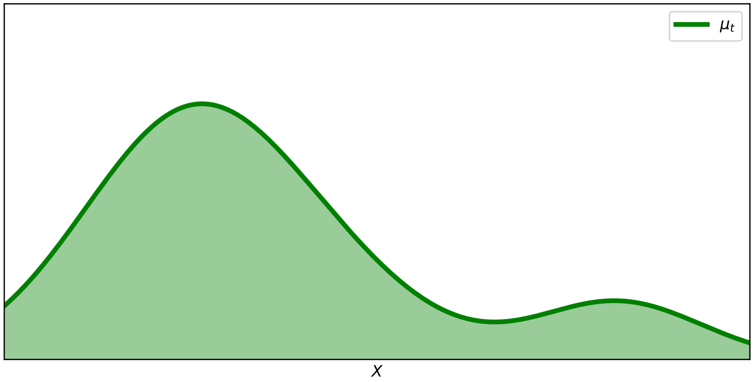
Stochastic flow-kick model



Many paths



Distribution of X_t



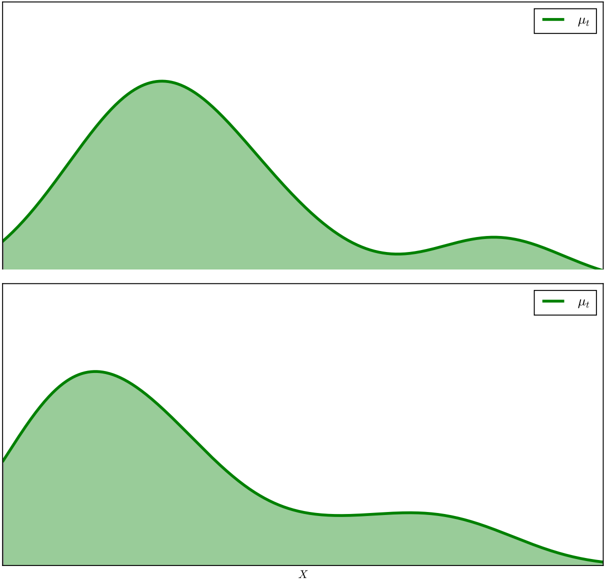
Evolution operator

- Initial Distribution:

$$\mu_0$$

- Evolved Distributions:

$$\mu_0 \to \mu_t =: \mu_0 \mathcal{U}^t$$


μ_t
μ_t
X

A direct approach: Differential forward equation for time-dependent density

For absolutely continuous distributions and disturbance kernel,

$$\frac{\mathrm{d}\mu_t}{\mathrm{d}x} = \rho(x,t)\,, \quad P(x,\mathrm{d}y) = p(x,y)\mathrm{d}y$$

$$\partial_t\rho(x,t) = -\partial_x\left(g(x)\rho(x,t)\right) - \Lambda(x)\rho(x,t) + \int p(y,x)\Lambda(y)\rho(y,t)\,\mathrm{d}y\,.$$

A direct approach: Differential forward equation for time-dependent density

For absolutely continuous distributions and disturbance kernel,

$$\frac{\mathrm{d}\mu_t}{\mathrm{d}x} = \rho(x,t)\,, \quad P(x,\mathrm{d}y) = p(x,y)\mathrm{d}y$$

$$\partial_t \rho(x,t) = -\partial_x\left(g(x)\rho(x,t)\right) - \Lambda(x)\rho(x,t) + \int p(y,x)\Lambda(y)\rho(y,t)\,\mathrm{d}y\,.$$

Deterministic evolution

A direct approach: Differential forward equation for time-dependent density

For absolutely continuous distributions and jump kernel,

$$\frac{\mathrm{d}\mu_t}{\mathrm{d}x} = \rho(x,t)\,, \quad P(x,\mathrm{d}y) = p(x,y)\mathrm{d}y$$

$$\partial_t \rho(x,t) = -\partial_x\left(g(x)\rho(x,t)\right) - \Lambda(x)\rho(x,t) + \int p(y,x)\Lambda(y)\rho(y,t)\,\mathrm{d}y\,.$$

Disturbance occurrence

A direct approach: Differential forward equation for time-dependent density

For absolutely continuous distributions and jump kernel,

$$\frac{\mathrm{d}\mu_t}{\mathrm{d}x} = \rho(x,t)\,, \quad P(x,\mathrm{d}y) = p(x,y)\mathrm{d}y$$

$$\partial_t \rho(x,t) = -\partial_x\left(g(x)\rho(x,t)\right) - \Lambda(x)\rho(x,t) + \int p(y,x)\Lambda(y)\rho(y,t)\,\mathrm{d}y\,.$$

Disturbance severity

Fundamental questions

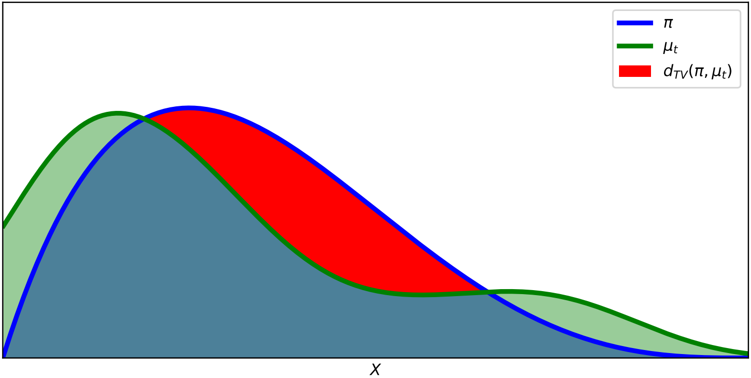
- Does there exist a distribution π on $\mathcal{X}$ with

$$\mathrm{d}_{TV}(\mu_t, \pi) \to 0 \ \text{ as } \ t \to \infty\,?$$

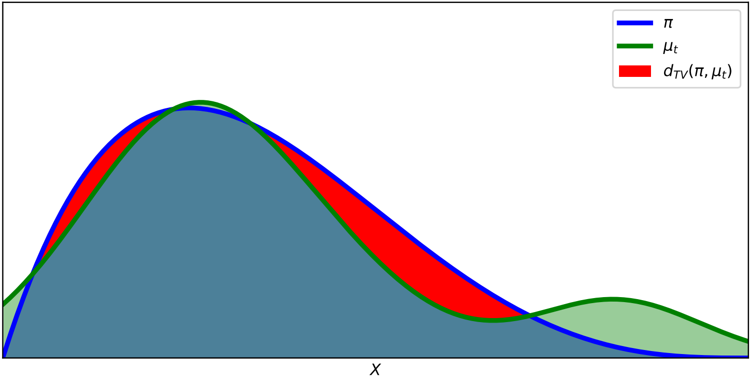
- Given $\delta > 0$, how large must t be so that

$$\mathrm{d}_{TV}(\mu_t, \pi) < \delta\,?$$

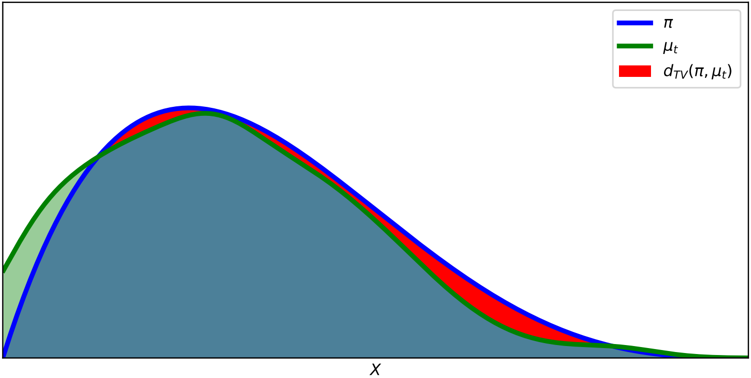
Evolved distributions



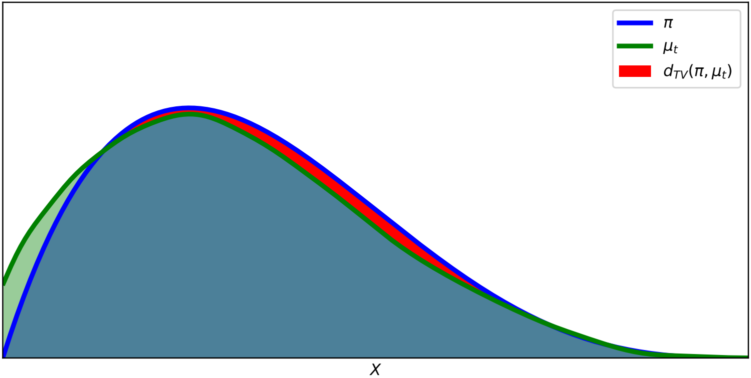
Evolved distributions



Evolved distributions



Evolved distributions



Equilibrium distributions

Question:
What information can we extract from this equilibrium distribution?

Answer:
We can compute the fraction of time (in the long run) that the process X_t spends in any measurable subset of our state space.

Theorem (B.)

Under appropriate assumptions, X_t is uniformly ergodic with a unique stationary distribution, π, and

$$\mathrm{d}_{TV}(\mu_t, \pi) \leq (1 - \beta)^t$$

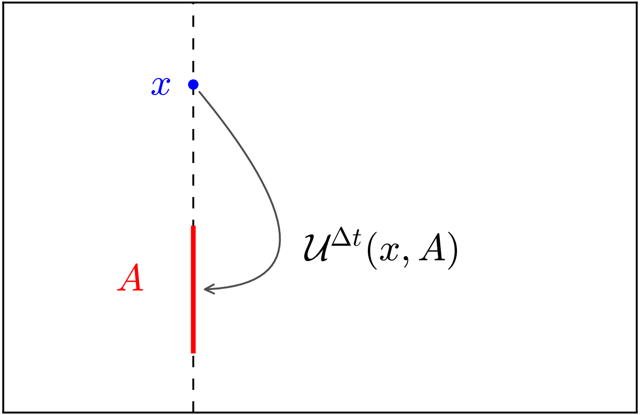
for any initial distribution μ_0.

Note: The value of β is explicitly constructed.

How is β computed?

1. Discretize the process (in time).
2. Develop minorization for the discretization.
3. Deduce bounds for the original continuous-time process.

Discrete time transition kernel



Uniform minorization

Ingredients:

1. Probability measure η on $\mathcal{X}$
2. $\beta > 0$

With

$$\mathcal{U}^{\Delta t}(x, A) \geq \beta\eta(A)$$

for any measurable set A and all $x \in \mathcal{X}$.

Weak formulation

Let $f : \mathcal{X} \mapsto \mathbb{R}$ be an observable, then

$$\begin{aligned} \langle \mu_{\Delta t}, f \rangle &= \langle \mu_0 \mathcal{U}^{\Delta t}, f \rangle \\ &= \langle \mu_0, \mathcal{U}^{\Delta t} f \rangle \end{aligned}$$

with

$$[\mathcal{U}^{\Delta t} f](x) := \mathbb{E}[f(X_{\Delta t}) \,|\, X_0 = x]$$

and

$$\langle \mu_0, \mathcal{U}^{\Delta t} f \rangle = \int [\mathcal{U}^{\Delta t} f](x) \mathrm{d}\mu_0(x)$$

Alternative characterization of minorization

The minorization condition

$$\mathcal{U}^{\Delta t}(x, A) \geq \beta\eta(A)$$

is equivalent to requiring for any nonnegative observable f, and for all $x \in A$,

$$[\mathcal{U}^{\Delta t} f](x) \geq \beta \int f(y) \mathrm{d}\eta(y) \ .$$

Infinitesimal generator

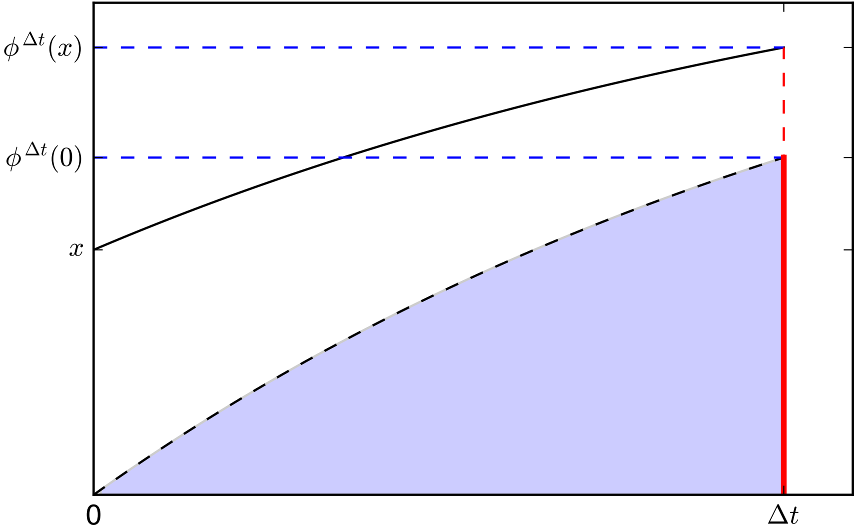
To control the discrete-time evolution operator, $\mathcal{U}^{\Delta t}$, we can study the infinitesimal generator $\mathcal{L}$ of $\mathcal{U}^t$.

$\mathcal{L}$ acts on observables, f, according to

$$[\mathcal{L}f](x) = \lim_{t \searrow 0} \frac{\mathcal{U}^t f(x) - f(x)}{t} .$$

In our case,

$$[\mathcal{L}f](x) = f'(x)g(x) + \Lambda(x) \int P(x, \mathrm{d}y)[f(y) - f(x)] .$$


$\phi^{\Delta t}(x)$
$\phi^{\Delta t}(0)$
x
0
Δt

Uniform minorization for compact one-dimensional state spaces

1. $\frac{\mathrm{d}\eta}{\mathrm{d}x} = \frac{\Delta t - \psi(0,x)}{\mathrm{C}_1} \mathbb{1}\{0 \leq x \leq \phi^{\Delta t}(0)\}$

2. $\beta_{\Delta t} = e^{-\lambda \Delta t} \mathrm{C}_2 \int_0^{\phi^{\Delta t}(0)} [\Delta t - \psi(0,x)] \, \mathrm{d}x$

with ψ defined by

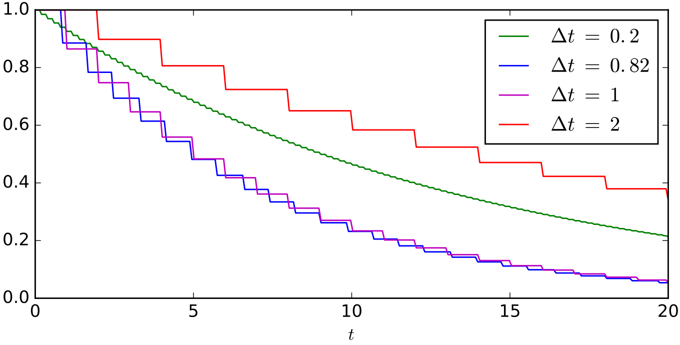
$$x_1 = \phi^t(x_0) \quad \iff \quad t = \psi(x_0, x_1)$$

and

$$\lambda \geq \Lambda(x) \text{ for all } x \in \mathcal{X}$$

Optimal Δt

Plots of $(1 - \beta_{\Delta t})^{\lfloor t/\Delta t \rfloor}$ vs. t for various Δt



Question:
Why do convergence rates matter?

Answers:

- Determine how long before “long-run” averages are realized.
- Provide guidance for numerical methods of approximating stationary distributions.
- Relevant for sub-sampling techniques used with Monte Carlo methods in likelihood-based inference.

Applications of semistochastic / piecewise-deterministic models

- Ecological disturbances
- Precipitation models
- Growth-fragmentation processes
- Human behaviour
- Viral reproduction

Select references

1. J. Broda, A. Grigo, and N.P. Petrov. Convergence rates for semistochastic processes. *Discrete and Continuous Dynamical Systems - B*, **24(1)** (2019), 109-125.
2. Benaim, Michel, et al. Qualitative properties of certain piecewise deterministic Markov processes. *Annales de l'IHP Probabilites et statistiques.***51(3)** (2015), 1040-1075.
3. M. C. A. Leite, N. P. Petrov and E. Weng, Stationary distributions of semistochastic processes with disturbances at random times and with random severity, *Nonlinear Analysis: Real World Applications*, **13** (2012), 497–512.
4. B. Beckage, W. J. Platt and L. J. Gross, Vegetation, fire, and feedbacks: a disturbance-mediated model of savannas, *The American Naturalist*, **174** (2009), 805–818.
5. B. J. Cairns, Evaluating the expected time to population extinction with semi-stochastic models, *Mathematical Population Studies*, **16** (2009), 199–220.