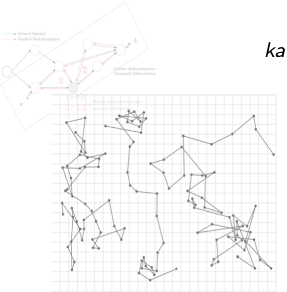
Calibrating Multivariate Lévy Processes with Neural Networks

Kailai Xu, Eric Darve

ICME, Stanford University

kailaix@stanford.edu *darve@stanford.edu*

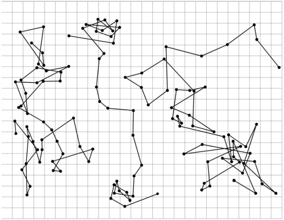
July 2020



Outline

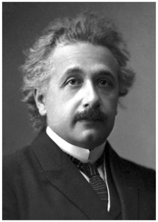
1. Introduction to Lévy Processes
2. Inverse Modeling Methodology
3. Application to Calibrating Lévy Processes
4. Open Source Software: ADCME

Gaussian Processes



Random Walk of Pollen Seeds

Robert Brown



Albert Einstein

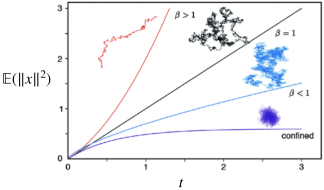
- Brownian Motion: $\mathbb{E}(\|x\|^2) \propto t$
- Gaussian Process: Brownian motion with drifts

$$\mathbf{W}_t = \mathbf{b}t + \Sigma \mathbf{B}_t \quad \mathbf{b} \in \mathbb{R}^d, \Sigma \in \mathbb{R}^{d\times d}$$

From Gaussian Proccesses to Lévy Processes

- However, an increasing number of natural phenomena do not fit into the relatively simple Brownian motion framework.

$$\mathbb{E}(\|x\|^2) \propto t^{\beta}$$



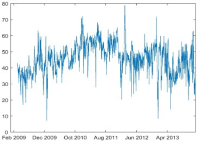
- Lévy processes generalize Gaussian processes by allowing a heavy-tailed step-length distribution in the random walk. They have been found successful in describing anomalous diffusion.

Modeling with Lévy Processes

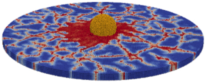
Bird flocking

Predators Movement Patterns

Lévy Processes



Stocks



Solid Mechanics

Outline

Mathematical Formulation

Definition (Lévy Processes)

The paths of Lévy processes $\mathbf{X}_t$ can be described by

$$\mathbf{X}_t = \mathbf{b}t + \Sigma\mathbf{W}_t + \sum_{k=1}^{N_t} J_k$$

where $\mathbf{b} \in \mathbb{R}^d$, $\Sigma \in \mathbb{R}^{d\times d}$, $\mathbf{W}_t$ is a standard Brownian motion, N_t is a Poisson process, and J_k is an i.i.d. sequence of random variables, which describes the jump.

The characteristic function of $\mathbf{X}_t$ is described by

$$\phi(\boldsymbol{\xi}) = \mathbb{E}[e^{\mathrm{i}\langle\boldsymbol{\xi},\mathbf{X}_t\rangle}] =$$

$$\exp\left[t\left(\mathrm{i}\langle\mathbf{b},\boldsymbol{\xi}\rangle - \frac{1}{2}\langle\boldsymbol{\xi},\mathbf{A}\boldsymbol{\xi}\rangle + \int_{\mathbb{R}^d}\left(e^{\mathrm{i}\langle\boldsymbol{\xi},\mathbf{x}\rangle} - 1 - \mathrm{i}\langle\boldsymbol{\xi},\mathbf{x}\rangle\mathbf{1}_{\|\mathbf{x}\|\leq 1}\right)\nu(d\mathbf{x})\right)\right]$$

Inverse Modeling

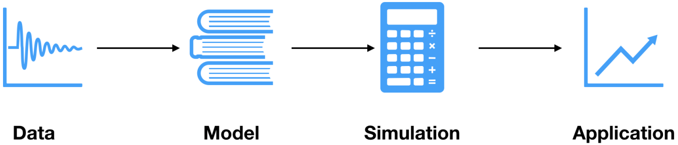
The ultimate goal of inverse modeling is to make predictions of future behaviors, which in turn requires us to find the causes of behaviors.

- Make Predictions $\Rightarrow$ Forward Problem

$$(\mathbf{b}, \mathbf{A}, \nu) \longrightarrow \mathbf{X}_{\Delta t}, \mathbf{X}_{2\Delta t}, \mathbf{X}_{3\Delta t}, \ldots$$

- Find Causes $\Rightarrow$ Inverse Problem

$$\mathbf{X}_{\Delta t}, \mathbf{X}_{2\Delta t}, \mathbf{X}_{3\Delta t}, \ldots \longrightarrow (\mathbf{b}, \mathbf{A}, \nu)$$

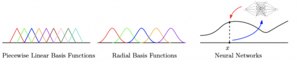


Our Approach

Match the characteristic function!

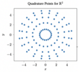
Model

$$\nu(\mathbf{x}) \approx \nu_{\boldsymbol{\theta}}(\mathbf{x})$$



$$\downarrow$$

$$\phi(\boldsymbol{\xi}) \approx \phi_{\boldsymbol{\theta}}(\boldsymbol{\xi}) := \exp\left[\Delta t \sum_{i=1}^{n_q} \left(e^{\mathrm{i}\langle \boldsymbol{\xi}, \mathbf{x}_i\rangle} - 1 - \mathrm{i}\langle \boldsymbol{\xi}, \mathbf{x}_i\rangle \mathbf{1}_{\|\mathbf{x}_i\|\le 1}\right) \nu_{\boldsymbol{\theta}}(\mathbf{x}_i) w_i\right]$$



$$\downarrow$$

$$\boxed{\min_{\boldsymbol{\theta}} \frac{1}{m} \sum_{i=1}^{m} \|\hat\phi_n(\boldsymbol{\xi}_i) - \phi_{\boldsymbol{\theta}}(\boldsymbol{\xi}_i)\|^2}$$

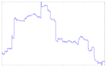
$$\uparrow$$

Data

$$\hat\phi_n(\boldsymbol{\xi}) := \frac{1}{n} \sum_{i=1}^{n} \exp(\mathrm{i}\langle \boldsymbol{\xi}, \mathbf{X}_{i\Delta t} - \mathbf{X}_{(i-1)\Delta t}\rangle),\ \boldsymbol{\xi} \in \mathbb{R}^d$$

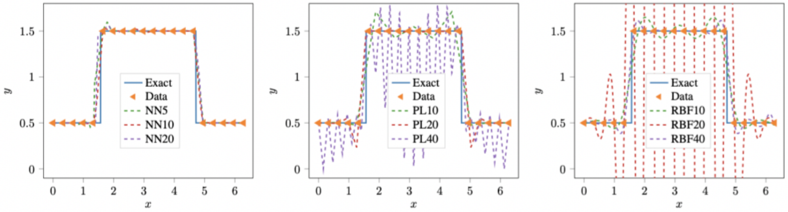
$$\uparrow$$

$$\{\mathbf{X}_{i\Delta t}\}_i$$



Neural Networks

- Unique challenge for estimating $\nu(\mathbf{x})$: non-smoothness and data-insufficiency (requires regularization).
- Neural network is adaptive to discontinuities and acts as a regularizer.



Neural Network **Piecewise Linear** **Radial Basis Functions**

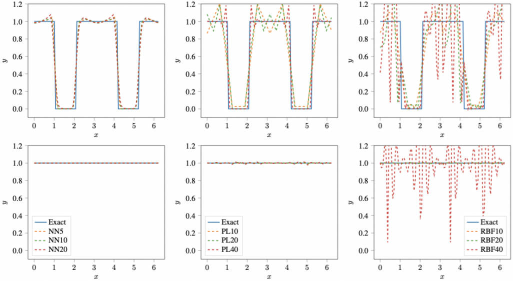
Outline

Example 1: Multivariate α-stable Process

- The multivariate α-stable process is a special Lévy process with the characteristic function

$$\phi(\boldsymbol{\xi}) = \mathbb{E}\left(\exp(\mathrm{i}\Delta t\langle \mathbf{X}, \boldsymbol{\xi}\rangle)\right) = \exp\left(-\Delta t \int_{\mathbb{S}^d} |\langle \mathbf{s}, \boldsymbol{\xi}\rangle|^{\alpha} \Gamma(\mathbf{s}) d\mathbf{s}\right)$$

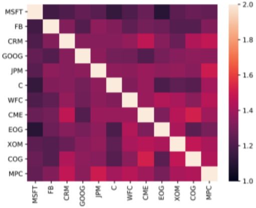
$$\alpha = 0.75, \Gamma(\mathbf{s}) = \mathbf{1}_{|s_1|>0.5}(\mathbf{s}), \text{ and } \Gamma(\mathbf{s}) = 1, \ \mathbf{s} = (s_1, s_2), \ \mathbf{s} \in \mathbb{S}^2$$

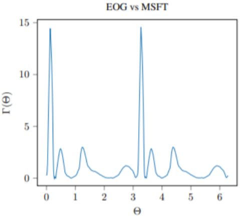


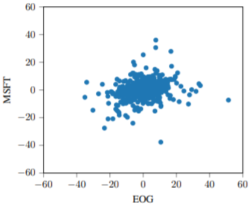
Function	NN5	NN10	NN20	PL10	PL20	PL40	RBF10	RBF20	RBF40
Step	0.7500	0.7499	0.7498	0.7493	0.7494	0.7500	0.7482	0.7483	0.7504
Constant	0.7499	0.7500	0.7500	0.7500	0.7500	0.7499	0.7500	0.7500	0.7499

Example 2: Application to the Stock Market

- Identify the jump diffusion intensity and heteroscedasticity movement of stock prices.







Outline

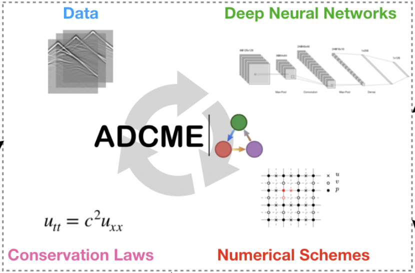
1 Introduction to Lévy Processes

2 Inverse Modeling Methodology

3 Application to Calibrating Lévy Processes

4 Open Source Software: ADCME

ADCME: A Powerful Inverse Modeling Library for Scientific Machine Learning



PoreFlow.jl
Geomechanics
Viscoelasticity
Multiphase Flow
Multiphysics
* coming soon

FwiFlow.jl
Multiphase Flow
Nonlocal Operators
https://github.com/lidongzh/FwiFlow.jl

ADSeismic.jl
General Seismic Inversion
https://github.com/kailaix/ADSeismic.jl

NNFEM.jl

NNFEM.jl
Constitutive Law Modeling
Hyperelasticity
Elasto-Plasticity
https://github.com/kailaix/NNFEM.jl