Automatic Differentiation of Stochastic Equations

August 30, 2025

Abstract

This document provides an introduction to automatic differentiation (AD) techniques applied to stochastic differential equations (SDEs). We discuss theoretical foundations, computational methods, and applications. Exercises with solutions are provided to reinforce key concepts.

1 Introduction

Stochastic differential equations (SDEs) play a crucial role in various fields, including finance, physics, and engineering. Differentiating functionals of SDEs is essential for optimization and sensitivity analysis. Automatic differentiation (AD) provides a powerful tool for computing derivatives with high precision and efficiency.

This book covers:

- Fundamentals of stochastic calculus and SDEs,
- Principles of automatic differentiation,
- Application of AD to stochastic processes,
- Computational techniques and implementation.

2 Background on Stochastic Calculus

A stochastic differential equation takes the form:

$$dX_t = f(X_t, t)dt + g(X_t, t)dW_t, \tag{1}$$

where W_t is a Wiener process, and f, g are suitable functions. The Itô lemma is fundamental for differentiating functionals of stochastic processes:

$$dF(X_t, t) = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial X}dX_t + \frac{1}{2}\frac{\partial^2 F}{\partial X^2}g^2 dt. \tag{2}$$

3 Principles of Automatic Differentiation

Automatic differentiation decomposes a function into elementary operations and applies the chain rule systematically. AD methods can be categorized into:

- Forward mode AD: Efficient for functions with small input dimensions,
- Reverse mode AD: Efficient for functions with large input dimensions, common in machine learning and sensitivity analysis.

For a function $y = f(x)$, forward mode computes:

$$\dot{y} = f'(x)\dot{x}, \tag{3}$$

where $\dot{x}$ represents the derivative seed. Reverse mode propagates sensitivities backwards:

$$\bar{x} = \bar{y} f'(x). \tag{4}$$

4 Applying AD to Stochastic Equations

Differentiating functionals of SDEs requires combining AD with stochastic calculus. Given an expectation:

$$J(\theta) = \mathbb{E}[F(X_T; \theta)], \tag{5}$$

where X_T follows an SDE parameterized by θ, we compute $\nabla_\theta J$ using:

1. Pathwise differentiation (if the expectation is smooth),
2. Score function method (likelihood ratio method),
3. Adjoint methods via reverse mode AD.

5 Exercises

5.1 Exercise 1: Differentiation of Expectations

Given the functional:

$$J(\theta) = \mathbb{E}[e^{-\theta X_T}], \tag{6}$$

where X_T satisfies $dX_t = -X_t dt + dW_t$, compute $\frac{dJ}{d\theta}$ using pathwise differentiation.

5.2 Exercise 2: Reverse Mode AD for SDEs

Implement a reverse mode AD algorithm to compute the gradient of:

$$J(\theta) = \mathbb{E}[X_T^2] \tag{7}$$

for the SDE $dX_t = \theta X_t dt + dW_t$.

6 Solutions

6.1 Solution to Exercise 1

We first compute the expectation:

$$J(\theta) = \mathbb{E}[e^{-\theta X_T}]. \tag{8}$$

Taking the derivative under expectation:

$$\begin{aligned} \frac{dJ}{d\theta} &= \mathbb{E}\left[\frac{d}{d\theta} e^{-\theta X_T}\right] \\ &= \mathbb{E}\left[-X_T e^{-\theta X_T}\right]. \end{aligned}$$

Thus, the solution is:

$$\frac{dJ}{d\theta} = -\mathbb{E}[X_T e^{-\theta X_T}]. \tag{9}$$

6.2 Solution to Exercise 2

We define the adjoint variable:

$$\bar{X}_T = \frac{dJ}{dX_T} = 2X_T. \tag{10}$$

Using backpropagation through discretized dynamics, we compute:

$$\frac{dJ}{d\theta} = \mathbb{E}\left[\int_0^T 2X_t \frac{\partial X_T}{\partial \theta} dt\right]. \tag{11}$$

Expanding X_T explicitly:

$$X_T = X_0 e^{\theta T} + \int_0^T e^{\theta(T-s)} dW_s. \tag{12}$$

Differentiating with respect to θ:

$$\frac{\partial X_T}{\partial \theta} = T X_0 e^{\theta T} + \int_0^T (T-s) e^{\theta(T-s)} dW_s. \tag{13}$$

Substituting into the integral:

$$\frac{dJ}{d\theta} = \mathbb{E}\left[\int_0^T 2X_t \left(T X_0 e^{\theta T} + \int_0^T (T-s) e^{\theta(T-s)} dW_s\right) dt\right]. \tag{14}$$

This final expression provides an explicit form for computing the gradient using reverse mode AD.

7 Conclusion

We introduced AD techniques for SDEs, including forward and reverse mode methods. These techniques enable efficient sensitivity analysis in stochastic models.