Neural SDEs Made Easy: SDEs are Infinite-Dimensional GANs

Patrick Kidger, James Foster, Xuechen Li, Harald Oberhauser, Terry Lyons

UNIVERSITY OF OXFORD

Summary

SDEs are a map from noise (Brownian motion) to a solution distribution. They may be sampled from via numerical SDE solvers, but they do not admit a notion of probability density.

We observe that this is exactly the same as the generator of a GAN. By adding a discriminator, we train arbitrary neural SDEs as continuous-time generative models for time series; e.g. to model financial stocks.

Background: SDEs

Consider an SDE of the form

$$X_0 \sim \mu, \quad \mathrm{d}X_t = f(t, X_t)\,\mathrm{d}t + g(t, X_t) \circ \mathrm{d}W_t.$$

Here:

- μ is the initial distribution.
- f is the drift, g is the diffusion.
- W is Brownian motion.

The (strong) solution to an SDE is the unique map F such that $F(\mu, W) = X$. It is a map from noise distributions to a target distribution. It can be sampled from via SDE solvers; it does not have a probability density.

Background: GANs

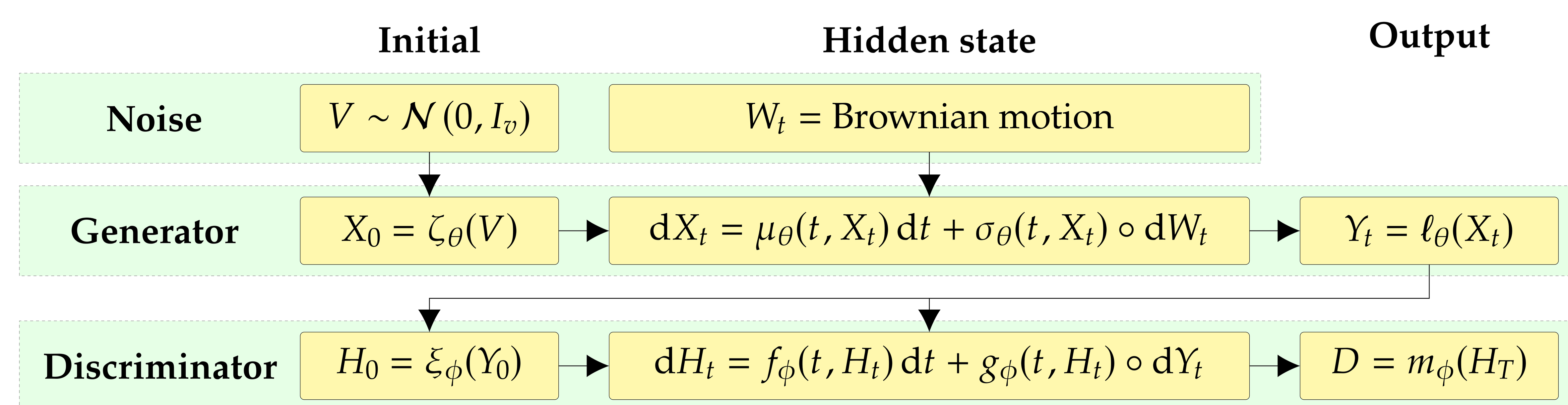
Given some noise distribution μ, and a target distribution ν, then the generator G_θ of a GAN is trained such that $G_\theta(\mu) = \nu$.

It is a map from a noise distribution to a target distribution. It can be sampled from, but its probability density is impossible to compute.

Method

We see that SDEs and GANs have a lot in common. It goes further: SDEs are usually trained by matching specific statistics (e.g. option prices). Meanwhile GANs are trained by matching a learnt scalar statistic: the discriminator.

Thus by using a neural SDE as a generator—and a neural CDE as the discriminator—we can train a general neural SDE as a model for time series.



Results

Dataset	Performance Metric	Neural SDE	CTFP	Latent ODE
Financial Stocks	Classification	**0.357 ± 0.045**	0.165 ± 0.087	0.000239 ± 0.000086
	Prediction	**0.144 ± 0.045**	0.725 ± 0.233	46.2 ± 12.3
	MMD	**1.92 ± 0.09**	2.70 ± 0.47	60.4 ± 35.8
CNN Training Weights	Classification	0.507 ± 0.019	**0.676 ± 0.014**	0.0112 ± 0.0025
	Prediction	**0.00843 ± 0.00759**	0.0808 ± 0.0514	0.127 ± 0.152
	MMD	**5.28 ± 1.27**	12.0 ± 0.5	23.2 ± 11.8
Beijing Air Quality	Classification	0.589 ± 0.051	**0.764 ± 0.064**	0.392 ± 0.011
	Prediction	**0.395 ± 0.056**	0.810 ± 0.083	0.456 ± 0.095
	MMD	**0.000160 ± 0.000029**	0.00198 ± 0.00001	0.000242 ± 0.000002

Find out more:

Full paper: `https://openreview.net/pdf?id=padYzanQNbg`
Library: `https://github.com/google-research/torchsde`
Email: `kidger@maths.ox.ac.uk`