

Numerical time evolution of probabilities

Concept. To extend the formalism that we developed in previous chapters to... For the mathematically-inclined, this chapter will take a detailed look at how our formalism can be extended to focus on probabilistic modelling. For the programmers, the software described in this chapter lives in the public Git repository: <https://github.com/umbralcalc/denmm-torch>.

2.1 Probabilistic formalism

In this section we will return to the stochadex formalism that we introduced in the first chapter of this book. As we discussed at that point; this formalism is appropriate for sampling from nearly every stochastic phenomenon that one can think of. However, when trying robustly assess how far a model is from accurately describing a set of real-world data, trying to use only generated samples of the model process can be difficult. Instead, in this section, we are going to extend this formalism to look at how probability theory can help with this data comparison problem in a systematic way.

So, how do we begin? In the first chapter, we defined the general stochastic process with the formula $X_{t+1}^i = F_{t+1}^i(X_{0:t}, z, \mathbf{t})$. This equation also has an implicit *master equation* associated to it that fully describes the time evolution of the *probability density function* $P_{t+1}(X|z)$ of $X_{0:t+1} = X$ given that the parameters of the process are z . This can be written as

$$P_{t+1}(X|z) = P_t(X'|z)P_{(t+1)t}(x|X', z), \quad (2.1)$$

where for the time being we are assuming the state space is continuous in each of the matrix elements and $P_{(t+1)t}(x|X', z)$ is the conditional probability that $X_{t+1} = x$ given that $X_{0:t} = X'$ at time t and the parameters of the process are z .

If we wanted to just look at the distribution over the latest row $X_{t+1} = x$, we could achieve this

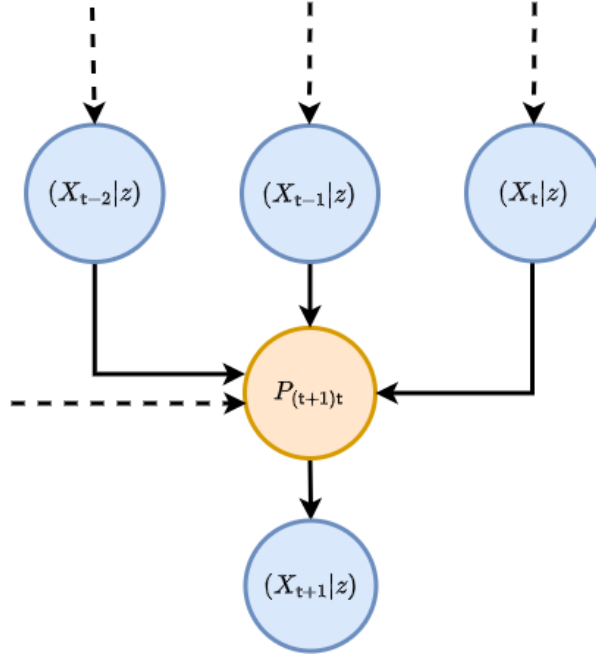


Figure 2.1: Graph representation of Eqs. (2.1) and (2.2).

through marginalisation over all of the previous matrix rows in Eq. (2.1) like this

$$P_{t+1}(x|z) = \int_{\Omega_t} dX' P_{t+1}(X|z) = \int_{\Omega_t} dX' P_t(X'|z) P_{(t+1)t}(x|X', z). \quad (2.2)$$

But what is Ω_t ? You can think of this as just the domain of possible matrix X' inputs into the integral which will depend on the specific stochastic process we are looking at.

The symbol dX' in Eq. (2.2) is our shorthand notation throughout the book for taking a sum of sub-domain integrals over each matrix row; where each row measure is a Cartesian product of n elements (a Lebesgue measure), i.e.,

$$\int_{\Omega_t} dX' = \frac{1}{t} \sum_{t'=0}^t \int_{\omega_{t'}} d^n x' = \frac{1}{t} \sum_{t'=0}^t \int_{\omega_{t'}} \prod_{i=0}^n d(x')^i, \quad (2.3)$$

where lowercase x, x', \dots values will always refer to individual rows within the state matrices. Note that $1/t$ here is a normalisation factor — this just normalises the sum of all probabilities to 1 given that there is a sum over t' . Note also that, if the process is defined over continuous time, we would need to replace

$$\frac{1}{t} \sum_{t'=0}^t \rightarrow \frac{1}{t(t)} \sum_{t'=0}^t \delta t(t'). \quad (2.4)$$

To try and understand what Eqs. (2.1) and (2.2) are saying, we find it's helpful to think of an iterative relationship between probabilities; each of which is connected by their relative conditional probabilities. This kind of thinking is also illustrated in Fig. 2.1.

Without loss of generality, we can relate the latest probabilities to those from deeper into the past by chaining conditional probabilities together in a non-Markovian equivalent of the Chapman-Kolmogorov equation

$$\begin{aligned}
P_{t+1}(x|z) &= \int_{\Omega_{t-1}} dX'' P_{t-1}(X''|z) \int_{\omega_t} d^n x' P_{t(t-1)}(x'|X'', z) P_{(t+1)t}(x|X', z) \\
&= \int_{\Omega_{t-2}} dX''' P_{t-2}(X'''|z) \int_{\omega_{t-1}} d^n x'' P_{(t-1)(t-2)}(x''|X''', z) \\
&\quad \times \int_{\omega_t} d^n x' P_{t(t-1)}(x'|X'', z) P_{(t+1)t}(x|X', z) \\
&= \dots \\
&= \int_{\Omega_{t-s}} dX''' P_{t-s}(X'''|z) \prod_{s'=0}^{s-1} \left\{ \int_{\omega_{t-s'}} d^n x' P_{(t-s')(t-s'-1)}(x'|X'', z) \right\} P_{(t+1)t}(x|X', z).
\end{aligned} \tag{2.5}$$

Depending on the temporal correlation structure of the process, the conditional probabilities can be factorised. For example, processes with second or third-order temporal correlations would be described by the following expressions

$$P_{(t+1)t}(x|X', z) = \frac{1}{t} \sum_{t'=0}^t \int_{\omega_{t'}} d^n x' P_{t'}(x'|z) P_{(t+1)t'}(x|x', z) \tag{2.6}$$

$$P_{(t+1)t}(x|X', z) = \frac{1}{t} \sum_{t'=0}^t \frac{1}{t'} \sum_{t''=0}^{t'} \int_{\omega_{t'}} d^n x' \int_{\omega_{t''}} d^n x'' P_{t''}(x''|z) P_{t't''}(x'|x'', z) P_{(t+1)t'}(x|x', x'', z). \tag{2.7}$$

Let's imagine that x is just a scalar (as opposed to a row vector) for simplicity in the expressions. We can then discretise the 1D space over x into separate i -labelled regions such that $[P]_{t+1}^i - [P]_t^i = \mathcal{J}_{t+1}^i$, where the probability current \mathcal{J}_{t+1}^i for the factorised processes above would be defined as

$$\mathcal{J}_{t+1}^i = -[P]_t^i + \frac{1}{t} \sum_{t'=0}^t \sum_{i'=0}^N \Delta x [P]_{t'}^{i'} [P]_{(t+1)t'}^{i'i'} \tag{2.8}$$

$$\mathcal{J}_{t+1}^i = -[P]_t^i + \frac{1}{t} \sum_{t'=1}^t \frac{1}{t'-1} \sum_{t''=0}^{t'-1} \sum_{i'=0}^N \sum_{i''=0}^N \Delta x^2 [P]_{t''}^{i''} [P]_{t't''}^{i'i''} [P]_{(t+1)t't''}^{ii'i''}. \tag{2.9}$$

The $[P]_{(t+1)t't''}^{ii'i''}$ tensor, in particular, will have $N^3 t(t^2 - 1)$ elements. Note that the third-order temporal correlations can be evolved by identifying the pairwise conditional probabilities as time-dependent state variables and evolving them according to the following relation

$$[P]_{(t+1)t't''}^{ii'i''} = \frac{1}{t} \sum_{t'=1}^t \sum_{i'=0}^N \Delta x [P]_{t't''}^{i'i''} [P]_{(t+1)t't''}^{ii'i''}. \tag{2.10}$$

What other classes of process can be described by Eqs. (2.1) and (2.2)? For Markovian phenomena, the equations no longer depend on timesteps older than the immediately previous one, hence Eq. (2.2) reduces to just

$$P_{t+1}(x|z) = \int_{\omega_t} d^n x' P_t(x'|z) P_{(t+1)t}(x|x', z). \quad (2.11)$$

An analog of Eq. (2.2) exists for discrete state spaces as well. We just need to replace the integral with a sum and the schematic would look something like this

$$P_{t+1}(x|z) = \sum_{\Omega_t} P_t(X'|z) P_{(t+1)t}(x|X', z), \quad (2.12)$$

where we note that the P 's in the expression above all now refer to *probability mass functions*.

In this section we looked into how the mathematical formalism used in the stochadex could be extended with probability theory.

Bibliography