Direct ensemble methods

We now turn to the question of how to compute probabilistic information for stochastic dynamical systems, especially multiscale systems. We consider here the most direct method using N-sample ensemble. This is a very useful method for a variety of problems.

The basis of direct ensemble methods is the law of large numbers for an independent sequence of random variables. Let us discuss this briefly.

A sequence $\{\tilde{X}_n, n \in \mathbb{N}\}$ of random variables is said to be independent if their joint probabilityes factorize. More precisely, if we choose any subset $\{\tilde{X}_{n_1} \dots \tilde{X}_{n_k}\}$ and any events $A_k \subseteq \mathbb{R}^{m_k} = \text{range}(\tilde{X}_{n_k})$, then

$$(*) \qquad P(\tilde{X}_{n_1} \in A_1 \;\&\; \tilde{X}_{n_2} \in A_2 \;\&\; \dots \;\&\; \tilde{X}_{n_k} \in A_k) = \prod_{p=1}^{k} P(\tilde{X}_{n_p} \in A_p).$$

The property $(*)$ is the necessary and sufficient condition for the random variables to be independent. An equivalent condition is the factorization of the joint PDF of any finite subset of the random variables

$$p_{X_{n_1} \dots X_{n_k}}(x_1, \dots, x_k) = \prod_{p=1}^{k} p_{X_{n_p}}(x_p).$$

We furthermore say that the sequence of random variables is independent and identically distributed (i.i.d.) if the distributions P_{X_n} (or, equivalently, the pdf's p_{X_n}) of all the members of the sequence are the same!

$$P_{X_n} = P_X \quad \text{or} \quad p_{X_n} = p \qquad \forall \; n \in \mathbb{N}.$$

The law of large numbers (LLN) is the following basic theorem of probability theory: If $\{\tilde{X}_n, n \in \mathbb{N}\}$ is an iid sequence of random variables such that

$$\mathbb{E}(|\tilde{X}|) < +\infty$$

and if

$$\overline{X}_N = \frac{1}{N} \sum_{n=1}^{N} \tilde{X}_n$$

is the N-sample empirical average, then

$$\lim_{N \to \infty} \overline{X}_N = \mathbb{E}(\tilde{X}) \quad \text{a.s.}$$

Here "a.s." abbreviates "almost surely", i.e. with probability one!

This theorem provides a very simple method to compute statistical information for dynamical systems based on N-sample ensembles of solutions. In this approach, one solves the underlying dynamical equations, for example

$$d\tilde{x}_n = f(\tilde{X}_n, t)\,dt + g(\tilde{X}_n, t)\, d\tilde{W}_n(t)$$
$$\tilde{X}_n(t_0) = \tilde{X}_n^{(0)}$$

for $n = 1, \ldots, N$. Here, $\{\tilde{W}_n(t), n \in \mathbb{N}\}$ is an i.i.d sequence of Brownian motions and $\{X_n^{(0)}, n \in \mathbb{N}\}$ is an i.i.d set of initial conds. drawn from the distribution $P_{X^{(0)}}$.

Any statistical quantity that can be represented as an expectation $\mathbb{E}(\cdot)$ can be approximated from such an N-sample ensemble.

For example, pth-order multitime moments can be approximated by

$$\mu^{(N)}_{i_1 \cdots i_p}(t_1, \ldots, t_p) = \frac{1}{N} \sum_{n=1}^{N} \tilde{X}_{n i_1}(t_1) \cdots \tilde{X}_{n i_p}(t_p)$$

$$\xrightarrow{N \to \infty} \mathbb{E}\left(\tilde{X}_{i_1}(t_1) \cdots \tilde{X}_{i_p}(t_p) \right)$$

$$= \mu_{i_1 \cdots i_p}(t_1, \ldots, t_p).$$

Likewise, the single-time PDF can be obtained

$$P_X^{(N)}(x,t) = \frac{1}{N} \sum_{n=1}^{N} \delta^d(x - \tilde{X}_n(t)) \xrightarrow{N \to \infty} \mathbb{E}\left(\delta^d(x - \tilde{X}(t))\right) = P_X(x,t)$$

or multitime PDF's:

$$P_X^{(N)}(x_1, t_1; \ldots; x_p, t_p) \equiv \frac{1}{N} \sum_{n=1}^{N} \delta^d(x_1 - \tilde{X}_n(t_1)) \cdots \delta^d(x_p - \tilde{X}_n(t_p))$$

$$\xrightarrow{N \to \infty} \mathbb{E}\left(\delta^d(x_1 - \tilde{X}(t_1)) \cdots \delta^d(x_p - \tilde{X}(t_p)) \right)$$

$$= P_X(x_1, t_1; \ldots; x_p, t_p)$$

etc! This approach of using N-sample ensembles is called the Monte Carlo method.

We consider the Pros and Cons of the direct ensemble methods for computing statistical-dynamical ~~systems~~ quantities with N-sample ensembles.

$\boxed{\text{PROS}}$ (1) One only needs to use the original dynamics (plus a method for generating ensembles of initial conditions & random noise)

~~This is an important advantage in practical computing.~~

(2) The method is convergent in the limit $N \to \infty$.

No approximation is involved, other than the restriction of finite N. By computing a large enough ensemble with enough members N, the resulting approximations to statistical quantities will be as accurate as desired.

The disadvantages are exactly the same!

$\boxed{\text{CONS}}$ (1) One must use the original dynamics

If the dynamical system is multiscale with many degrees of freedom, then the underlying equations are usually expensive to solve. One can then often to compute small numbers of individual ensemble members.

(2) The convergence rate of direct ensemble methods is very slow as $N \to \infty$.

To see this, we must consider another fundamental result of probability theory, the <u>central limit theorem</u> (CLT). This result may be understood as describing the correction term to the law of large numbers (LLN). Suppose that $\tilde{X}$ is a random vector which not only satisfies

$$\mathbb{E}\left(|\tilde{X}|\right) < +\infty$$

so that the mean value $\mu = \mathbb{E}(\tilde{X})$ exists, but also

$$\mathbb{E}\left(|\tilde{X}-\mu|^2\right) < +\infty$$

so that the <u>covariance matrix</u> C

$$C = \mathbb{E}\left((\tilde{X}-\mu)(\tilde{X}-\mu)^T\right)$$

also exists. Let S be any square root of C, so that $C = SS^T$. Then if $\{\tilde{X}_n, n \in \mathbb{N}\}$ is any iid sequence of r.v.'s with the distribution of $\tilde{X}$, the CLT states that

$$\frac{1}{N}\sum_{n=1}^{N} \tilde{X}_n \underset{N\to\infty}{\sim} \mu + \frac{1}{\sqrt{N}} S\tilde{N} + o\left(N^{-1/2}\right),$$

where $\tilde{N} = (\tilde{N}_1, \ldots \tilde{N}_d)$ is a vector of $N(0,1)$ r.v.'s.
The above result means, more precisely, that

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} \sum_{n=1}^{N} (\tilde{X}_n - \mu) = S\tilde{N} \equiv \tilde{Y}$$

where the convergence is in the sense of probability distributions and $\tilde{Y}$ is an $N(0,C)$ random ~~variable~~ vector:

$$P_Y(x) = \frac{1}{\sqrt{(2\pi)^d \det C}} \exp\left[-\frac{1}{2}(x-\mu)^T C^{-1}(x-\mu)\right] \leftarrow \mathcal{N}(\underset{\mathbb{R}^d}{\mu}, \underset{\mathbb{R}^{d\times d}}{C})$$