Report: Monte Carlo Simulations

Thibeau Wouters

December 3, 2020

Contents

1 Non-uniform random numbers

Suppose the random variable X satisfies $X \sim \mathcal{U}(0,1)$, f is a continuous probability distribution with corresponding cumulative distribution F and $Y = F^{-1}(X)$. Then Y is distributed with probability density f.

Proof. We will show that $P(Y \leq y) = F(y)$. For this, note that

$$\begin{aligned} P(Y \leq y) &= P(F^{-1}(X) \leq y) \\ &= P\left(F(F^{-1}(X)) \leq F(y)\right) \\ &= P(X \leq F(y)) = F(y) \end{aligned}$$

since F is by definition a non-increasing function, and for all $x \in [0,1]$, we have $P(X \leq x) = x$, since $X \sim \mathcal{U}(0,1)$. □

1.1 First Question

We apply Proposition 1 on three given probability densities by determining and inverting F.

First function: Consider the probability density function f_1 of a $\mathcal{U}(-2,1)$ distribution, i.e.:

$$f_1 : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} \frac{1}{3} & x \in [-2,1] \\ 0 & \text{elsewhere} \end{cases} . \tag{1}$$

Then the cumulative distribution function F_1 is

$$F_1 : \mathbb{R} \to [0,1] : x \mapsto \begin{cases} 0 & x < -2 \\ \frac{x+2}{3} & x \in [-2,1] \\ 1 & x \geq 1 \end{cases} . \tag{2}$$

The inverse function can be obtained, and is

$$F_1^{-1} : [0,1] \to \mathbb{R} : y \mapsto \begin{cases} 0 & y = 0 \\ 3y - 2 & y \in (0,1) \\ 1 & y = 1 \end{cases} . \tag{3}$$

TODO

Second function: Consider the probability density function f_2 of an exponentially distributed random variable:

$$f_2 : \mathbb{R} \to \mathbb{R}^+ : x \mapsto \begin{cases} e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases} . \tag{4}$$

2 Gaussian RNG

3 Monte Carlo integration (I)

We will estimate the following integral via Monte Carlo methods seen in the course notes:

$$\mathcal{I} = \int_{\mathbb{R}^3} \exp\left(-\frac{r^2}{2}\right)(x+y+z)^2 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,, \tag{5}$$

where $r^2 = x^2 + y^2 + z^2$. We first compute this result analytically. We expand the square:

$$\mathcal{I} = \int_{\mathbb{R}^3} \exp\left(-\frac{r^2}{2}\right)(x^2+y^2+z^2+2xy+2yz+2xz) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,. \tag{6}$$

We see that there are only 2 different sorts of integrandi: the first three terms give the same result, and the last three terms also yield the same result. It is sufficient to calculate the following two integrals:

$$I_1 = \int_{\mathbb{R}^3} x^2 \exp\left(\frac{x^2+y^2+z^2}{2}\right)\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \quad I_2 = 2\int_{\mathbb{R}^3} xy \exp\left(\frac{x^2+y^2+z^2}{2}\right)\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \quad . \tag{7}$$

These integrals are straightforward, relying on the following well-known results:

$$\begin{aligned}
\int_{\mathbb{R}} \exp(-\frac{x^2}{2})\,\mathrm{d}x &= \sqrt{2\pi} \\
\int_{\mathbb{R}} x\exp(-\frac{x^2}{2})\,\mathrm{d}x &= 0 \\
\int_{\mathbb{R}} x^2\exp(-\frac{x^2}{2})\,\mathrm{d}x &= \sqrt{2\pi}\,.
\end{aligned}$$

The first result is the Gaussian integral, the second result is simply because the integrand is an odd function, and the third result can easily be derived from the Gaussian integral.

Therefore, we have that

$$I_1 = \int_{\mathbb{R}} x^2 \exp\left(-\frac{x^2}{2}\right)\mathrm{d}x \int_{\mathbb{R}} \exp\left(-\frac{y^2}{2}\right)\mathrm{d}y \int_{R} \exp\left(-\frac{z^2}{2}\right)\mathrm{d}z = (2\pi)^{3/2} \tag{8}$$

and that

$$I_2 = 2\int_{\mathbb{R}} x \exp\left(-\frac{x^2}{2}\right)\mathrm{d}x \int_{\mathbb{R}} y \exp\left(-\frac{y^2}{2}\right)\mathrm{d}y \int_{\mathbb{R}} \exp\left(-\frac{z^2}{2}\right)\mathrm{d}z = 0\,. \tag{9}$$

So we end up with the final result that

$$\mathcal{I} = 3(2\pi)^{3/2}\,. \tag{10}$$

We now approximate the integral via Monte Carlo methods, using the factor $\exp\left(-\frac{r^2}{2}\right)$ as a density function. We have to impose normalisation to this function, i.e. we have to choose N such that the function $W(x,y,z) = N\exp\left(-\frac{x^2+y^2+z^2}{2}\right)$ is normalised to one. We calculate, using the results mentioned above, that

$$N\int_R \exp\left(-\frac{x^2+y^2+z^2}{2}\right)\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = N\left(\exp\left(-\frac{x^2}{2}\right)\right)^3 = N(2\pi)^{3/2}\,, \tag{11}$$

and hence the normalised density function is

$$W(x,y,z) = \frac{1}{(2\pi)^{3/2}}\exp\left(-\frac{x^2+y^2+z^2}{2}\right)\,, \tag{12}$$

which is the density function for a normal distribution in three variables, with mean zero and variance 1, and no correlations between the variables.

To approximate the integral, let f denote the integrand, i.e. $\mathcal{I} = \int_{\mathbb{R}^3} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ and let $g(x,y,z) := (x+y+z)^2$. Since f can be factorised in a part proportional to W and the function g, we have that

$$\mathcal{I} = \int_{\mathbb{R}^3} f(\mathbf{x})\,\mathrm{d}^3\mathbf{x} \approx \frac{1}{N}\sum_{n=1}^{N}\frac{f(\mathbf{x}_n)}{W(\mathbf{x}_n)} \approx \frac{1}{N}(2\pi)^{3/2}\sum_{n=1}^{N} g(\mathbf{x}_n)\,, \tag{13}$$

where we used $\mathbf{x} := (x,y,z)$ for notational simplicity, and $\mathbf{x}_n, n = 1,\ldots,N$ are generated according to the distribution $W(\mathbf{x}_n)$. Comparing this with equation (10), we see that in order to have a good approximation, we would like $\sum_{n=1}^{N} g(\mathbf{x}_n) \approx 3$.

We compute...

We look at the variances ...

As a final remark, we note that we cannot change the roles of the factors of the integrand f. With this, we mean that we cannot use the factor $(x+y+z)^2$ as density, and $\exp\left(-\frac{r^2}{2}\right)$ as the integrand (or which we called g in the solution above). This is because the factor $(x+y+z)^2$ cannot be normalised in the domain of integration, which is $\mathbb{R}^3$, which we require for the density function. Indeed, it is clear that we cannot choose a $N \in \mathbb{R}$ such that

$$N\int_{\mathbb{R}^3}(x+y+z)^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = 1\,, \tag{14}$$

and hence this factor cannot be used to define a density function.

4 Monte Carlo integration (II)

We again try to approximate the integral from the previous section, which was

$$\mathcal{I} = \int_{\mathbb{R}^3} \exp\left(-\frac{r^2}{2}\right)(x+y+z)^2 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,, \tag{15}$$

but now we use the Metropolis algorithm. Therefore, we need to define a Markov chain, characterised by the transition probabilities $P(\vec{r} \to \vec{r}')$, of which the dynamics converge to the distribution $w(\vec{r}) = \frac{1}{(2\pi)^{3/2}} \exp(-r^2/2)$. This will be the case if we require detailed balance in the following way:

$$\frac{P(\vec{r} \to \vec{r}')}{P(\vec{r}' \to \vec{r})} = \exp\left(-\frac{r'^2 - r^2}{2}\right). \tag{16}$$

We split the transition probability in two parts: the selection probability g and acceptance ratio A:

$$P(\vec{r} \to \vec{r}') = g(\vec{r} \to \vec{r}')A(\vec{r} \to \vec{r}')\,. \tag{17}$$

In our case, g is generated by going from $\vec{r}$ to $\vec{r}' = \vec{r} + \vec{\delta}$, where $\vec{\delta} = (\delta_x, \delta_y, \delta_z)$ and the components δ_i are chosen independently from a uniform distribution centered at the origin and of width h, called $W_h(\delta)$. The acceptance ratio is given by

$$A(\delta) = \begin{cases} 1 & \text{if } \; r'^2 < r^2, \\ \exp\left(-\frac{r'^2 - r^2}{2}\right) & \text{otherwise}\,. \end{cases} \tag{18}$$

Using the Metropolis algorithm, we plot the trajectory followed by the x, y coordinates of the points, where we use $N = 10^4$ steps, starting from $\vec{r} = (0,0,0)$ and $\vec{r} = (10,10,10)$. We do this for both $h = 1$ and $h = 0.1$. ...

We also verify, by plotting a histogram of the positions for a run with $N = 10^6$ that the dynamics generates an equilibruim distribution of points proportional to $\exp(-r^2/2)$.

5 Stochastic Matrices

In this problem, we will look at stochastic matrices. More intro...

Suppose we are given a non-negative initial vector w. Then the stochastic matrix generates non-negative vectors at all later times. Indeed, the stochastic matrix lets the system evolve according to the prescription that if the system is at time t described by $w(t)$, then at a later time step $t + \Delta t$, it is described by the vector

$$w(t + \Delta t) = Pw(t)\,, \tag{19}$$

where

$$P : (P)_{\mu\nu} = P(\nu \to \mu) \tag{20}$$

is the stochastic matrix. By definition, all the entries of the stochastic matrix are non-negative. Therefore, if we assume that the vector $w(t)$ has only non-negative entries and we expand the matrix product in equation (19), we end up with a sum of products of non-negative quantities, hence these are again non-negative. So we conclude that $w(t + \Delta t)$ is again a non-negative vector.

This immediately implies by an inductive argument that if we start with an initial non-negative vector w, the stochastic matrix generates non-negative vectors at all later times.

Assume that P and Q are two $N \times N$ stochastic matrices. We now claim that PQ is again a stochastic matrix. To prove this, we have to show that *i)* all the elements of PQ are non-negative, and *ii)* that the sum of the columns of PQ are equal one.

To prove *i)*, we note that the an element of the matrix PQ is simply

$$(PQ)_{ij} = \sum_{k=1}^{N} P_{ik} Q_{kj} \geq 0 \,, \tag{21}$$

since both P_{ik}, Q_{kj} are non-negative for all $i \in \{1, \ldots, N\}$ by the assumption that P and Q are stochastic matrices.

To show *ii)*, we have to show that for all $j \in \{1, \ldots, N\}$

$$\sum_{i=1}^{N} (PQ)_{ij} = 1 \,. \tag{22}$$

We verify this by expanding the matrix product: take $j \in \{1, \ldots, N\}$ arbitrarily, then

$$\begin{aligned} \sum_{i=1}^{N} (PQ)_{ij} &= \sum_{i=1}^{N} \left(\sum_{k=1}^{N} P_{ik} Q_{kj} \right) = \sum_{i,k=1}^{N} P_{ik} Q_{kj} \\ &= \sum_{k=1}^{N} \left(Q_{kj} \sum_{i=1}^{N} P_{ik} \right) = 1 \,, \end{aligned}$$

since for all $k \in \{1, \ldots, N\}$, we have that $\sum_i P_{ik} = 1$ since P is a stochastic matrix, and also since for all $j \in \{1, \ldots, N\}$, we have $\sum_k Q_{kj} = 1$ by the same assumption on Q.

We now introduce the norm $||w||_1 = \sum_j |w_j|$. We consider an arbitrary vector y, not necessarily non-negative, and show that

$$||Py||_1 \leq ||y||_1 \,. \tag{23}$$

We can verify this directly by substituting the expansion of the matrix product AND WHAT ELSE?

$$||Py||_1 = \sum_{j=1}^{N} |(Py)_j| = \sum_{j=1}^{N} \left| \sum_{i=1}^{N} P_{ji} y_i \right| \leq \sum_{j,i=1}^{N} P_{ji} |y_i|$$
$$= \sum_{i=1}^{N} \left(|y_i| \sum_{j=1}^{N} P_{ji} \right) = \sum_{i=1}^{N} |y_i| = ||y||_1 \,,$$

where we used both properties of the definition of a stochastic matrix.

An immediate consequence from equation (23) is that an eigenvector y of a stochastic matrix P, then its eigenvalue λ satisfies $|\lambda| \leq 1$. Indeed, we find that

$$||Py||_1 = ||\lambda y||_1 = |\lambda| \sum_{j=1}^{n} |y_j| \leq \sum_{j=1}^{n} |y_j| \,,$$

from which we immediately find $|\lambda| \leq 1$.

We now show that a stochastic matrix P has at least one eigenvalue $\lambda = 1$. Take any stochastic $N \times N$ matrix P and define the $1 \times N$ row-vector $z = (1, 1, \ldots, 1)$. We claim that z is a left eigenvector of P with eigenvalue $\lambda = 1$. Indeed, we note that for all $i \in \{1, \ldots, N\}$, we have

$$(zP)_i = \sum_{j=1} z_i P_{ij} = \sum_{j=1} P_{ij} = 1 = z_i \,,$$

such that we indeed have that $zP = z$. Note that we only had to make use of the property that the sum of columns of a stochastic matrix is equal to one. Since the stochastic matrix P was arbitrary, this shows that any stochastic matrix P has at least one eigenvalue equal to one.

6 Detailed balance

Consider an system with three states with energies $E_1 < E_2 < E_3$. The dynamics are such that the system can make the transitions $1 \to 2$ with probability a, $2 \to 3$ with probability b, $3 \to 1$ with probability c or stay in the current state, where a, b, c are arbitrary rates for these transitions. Note that we require $0 < a, b, c < 1$.

The stochastic matrix governing the evolution of this system is

$$P = \begin{pmatrix} 1-a & 0 & c \\ a & 1-b & 0 \\ 0 & b & 1-c \end{pmatrix} . \tag{24}$$

The dynamics of this system is ergodic: starting from 1, we have a non-zero probability to directly go to 2, and there exists a path from 1 to 3, which is $1 \to 2 \to 3$. From 2, there is a non-zero probabilty to go to 3, and there exists a path from 2 to 1, which is $2 \to 3 \to 1$. Starting from 3, there is a non-zero probability to go to 1, and there exists a path from 3 to 2, which is $3 \to 1 \to 2$. Hence any state in the system can be reached from any other state.

To verify if detailed balance is satisfied or not, we first compute the stationary state ω corresponding to this dynamics. This is the state that satisfied $P\omega = \omega$, and hence is an eigenvector of the matrix P with eigenvalue equal to one. In the previous section, we have proven that for any stochastic matrix P, such a vector always exists. Hence we can directly look for its components by solving the eigenvalue equation, which yields the system of equations

$$\begin{cases} (1-a)\omega_1 + c\omega_3 &= \omega_1 \\ a\omega_1 + (1-b)\omega_2 &= \omega_2 \\ b\omega_2 + (1-c)\omega_3 &= \omega_3 \end{cases} \iff \begin{cases} -a\omega_1 + c\omega_3 &= 0 \\ a\omega_1 - b\omega_2 &= 0 \\ b\omega_2 - c\omega_3 &= 0 \end{cases} . \tag{25}$$

It can be directly verified that $\omega = (c, \frac{ac}{b}, a)$ is a solution to this set of equations (remember that $b > 0$). The actual form of the components is irrelevant; the crucial point is that all components of the stationary state vector are different from zero, by the assumptions made on a, b, c. This information is sufficient to show that detailed balance is **not** satisfied for this dynamics. The problem here is that if we look again at the stochastic matrix responsible for the development of the dymanics, given in equation (24), we see that the zeroes in the matrix are not present in a symmetric manner. Indeed, take $\mu = 2$ and $\nu = 1$; then $P(2 \to 1) = 0$, while $P(1 \to 2) \neq 0$, and hence the condition for detailed balance

$$\omega_\nu P(\nu \to \mu) = \omega_\mu P(\mu \to \nu) \,, \tag{26}$$

is not satisfied for this choice of μ and ν, since the left hand side is different from zero, while the right hand side is equal to zero.

In the lecture notes, we saw that detailed balance is a sufficient condition to get a Boltzmann distribution for the equilibrium of the dynamics. Here we show that even though detailed balance is not satisfied, we can still choose the rates a, b, c in such a manner that the equilibrium distribution is a Boltzmann distribution, proving that detailed balance is in fact not a necessary condition. In order to have a Boltzmann distribution, we require that the probability of finding the system in state i is given by

$$p_i = \frac{e^{-\beta E_i}}{Z} \,, \tag{27}$$

where $Z = \sum_{i=1}^{3} e^{-\beta E_i}$ is the partition function. Looking back at our result for the stationary state ω, we find that we can achieve this by setting

$$a = \frac{e^{-\beta E_3}}{Z}, \quad b = \frac{e^{-\beta(E_1+E_3-E_2)}}{Z}, \quad c = \frac{e^{-\beta E_1}}{Z} \,. \tag{28}$$

Since then we indeed have

$$\omega_1 = c = \frac{e^{-\beta E_1}}{Z}$$
$$\omega_2 = \frac{ac}{b} = \frac{e^{-\beta E_2}}{Z}$$
$$\omega_3 = a = \frac{e^{-\beta E_3}}{Z},$$

such that the equilibrium distribution of the dynamics is a Boltzmann distribution.

7 Ising Model: Uniform sampling

8 Metropolis algorithm for the Ising model