When does a system admit an equilibrium probability distribution

Sabarno Saha [1]

November 22, 2024

[1]Department of Physical Sciences IISERK

1. Stochastic Description
   - Stochastic Processes
   - Markov Chains
   - CTMC
   - Transition Rate Matrix
2. Steady State and Equilibrium Distribution
   - Conditions
   - Detailed Balance Condition
3. Perron Frobenius Theorem
   - Irreducibility
   - Strongly connected graphs
   - Stationary State distribution

5. Graph/ Network Theory
   - Graphs
   - Trees
   - Handshaking lemma
   - An "obvious" theorem

6. Equilibrium Distribution

A Stochastic Description

Stochastic Process

A stochastic process is a sequence of random variables where the indexing of the variables often carries the notion of time.

An example would be Brownian Motion, which is described by the Wiener process.

$$P(\hat{W}(t + \Delta t) = x | \hat{W}(t) = x') = \frac{1}{\sqrt{2\pi\Delta t}} \exp\left(-\frac{(x - x')^2}{2\Delta t} \right)$$

Most of the physical processes that we study in classical statistical physics is modelled as Markov Chains.

Markov Property

Let $\{X_n\}$ be a stochastic process. The markov property is defined as

$$\mathbb{P}(X_{n+1}|X_n, X_{n-1}, ..., X_0) = \mathbb{P}(X_{n+1}|X_n)$$

Any stochastic process satisfying the Markov property is called a Markov Chain.

Essentially the future of the system is independent of the past given the present state.

Continuous Time Markov Chains(CTMC)

These are markov chains with the index as time $\{X(t)\}_t$. Let us define a state space as the set of all values S a random variable can assume. Then we use the notation

$$P(X(t) = s) \equiv P(s; t)$$

We can rewrite the markov property as

$$P(s_n; t_n | s_{n-1}; t_{n-1}, s_{n-2}; t_{n-2}, ..., s_0; t_0) = P(s_n; t_n | s_{n-1}; t_{n-1})$$

where $s, s_i \in S\ \forall i$.

Now using the theorem of total probability we can write the Chapman Kolmogorov equation as

Chapman Kolmogorov Equation

$$P(s; t | s_0; t_0) = \sum_{s' \in S} P(s; t | s'; t') P(s'; t' | s_0; t_0)$$

where $t > t' > t_0$

We now choose a time interval dt and rewrite the Chapman Kolmogorov equation, dropping the explicit dependence on x_0, t_0

$$P(s; t + dt) = \sum_{s' \in S} P(s; t + dt | s'; t) P(s'; t)$$

Since this is an infinitesimal time interval, we can write the above equation as

$$P(s; t + dt | s'; t) = \delta_{ss'} + Q_{ss'} dt + o(dt)$$

where Q is the transition rate matrix.

The transition rate matrix should satisfy the following conditions

1. $Q_{ss'} \geq 0$ for $s \neq s'$
2. $Q_{ss} = -\sum_{s' \neq s} Q_{ss'}$

The first condition ensures that the transition rate is non-negative and the second condition essentially ensures that the probability of transitioning from a state s to any state is 1.

Thus we can write

$$\mathbb{P}(t + dt) = \mathbb{I} + Qdt$$

where $\mathbb{I}$ is the identity matrix.

Transition Rate Matrix

We define $k_{xx'}$ as the jump rate from state x' to x. We thus define the elements of the transition rate matrix as

$$p(x; t + dt | x'; t') = Q_{xx'} = k_{xx'} \qquad x \neq x'$$

And the diagonal elements are thus given as

$$Q_{xx} = -\sum_{x' \neq x} k_{xx'}$$

The change in probability of being in state x at time t is the **inflow** of probability from all other states which is $k_{xx'} p(x'; t)$ minus the **outflow** of probability $k_{x'x} p(x; t)$. We can now write the master equation as

$$\frac{\mathrm{d}p(x; t)}{\mathrm{d}t} = \sum_{x' \in S} k_{xx'} p(x'; t) - k_{x'x} p(x; t)$$

This becomes clearer when we define the probability current.

These equations describe the time evolution of the system that are modelled as being in the a probabilistic combination of a set of states at any given time. These are phenomenologically modelled first order differential equations.

These are used in a variety of fields in physics, and other related fields. These are used for studying birth and death processes, non-equilibrium statistical mechanics, chemical reactions, and a whole lot more.
The most common one that we see is the one used here is of the form

$$\frac{d\mathbf{p}}{dt} = \mathbf{p}$$

where is the matrix of connections and $\mathbf{p}$ is the probability distribution column vector.

Probability Current

The probability current is defined as the flow of probability from state x' to state x. We define the probability current as

$$J_{xx'} = k_{xx'}p(x';t) - k_{x'x}p(x;t)$$

The master equation can now be written as

$$\frac{\mathrm{d}p(x;t)}{\mathrm{d}t} = \sum_{x' \in S} J_{xx'}$$

A bit of physics coming up ahead :)

Microscopic Reversibility

If for any allowed jump $x \to x'$ (i.e. $k_{xx'} > 0$), the reverse jump $x' \to x$ ($k_{x'x} > 0$) is also allowed, then the system is said to be microscopically reversible.

These systems are quite often represented as jump networks, where the nodes of the graph represent the states x and the arrows(edges) $x' \rightarrow x$ represent the allowed jumps i.e. the jumps with non-zero transition rates($k_{xx'} > 0$).

Some Physics

The **jump networks** dealt with in physics are **strongly connected**. That is, given any two states x and x', there is a non-zero probability of transitioning from x to x' in a finite number of steps.

This property has some physical justification. Coupled together with microscopic reversibility, this property ensures that the graphs aren't disconnected. Disconnected graphs often lead to systems with non-interacting components, which can be studied independently reducing to the strongly connected graph.

Strongly connected graphs also posses a property called Irreducibility. Irreducibility is a necessary condition for the Perron Frobenius theorem to hold.
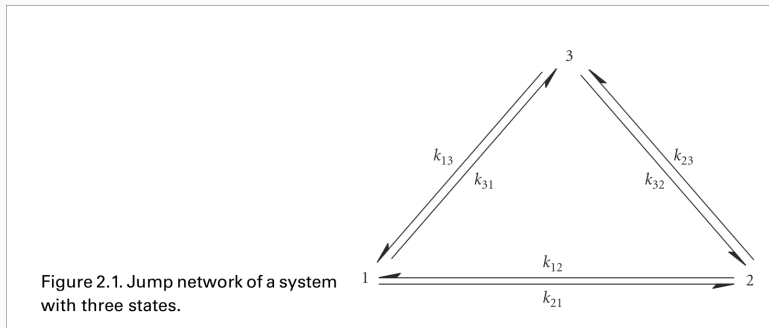
Here are is an example of a jump network :

Figure 2.1. Jump network of a system with three states.

Figure 1: Source : [1]

Steady State & Equilibrium Distribution

Steady state and equilibrium distribution are often used interchangeably. However, they are defined quite differently. The condition for a system to be in equilibrium is a bit more constrained than that of a steady state.

The necessary condition for both to hold is that the probability of being in a state x at time t is independent of time. Thus, essentially our master equation equates to 0.

$$\frac{\mathrm{d}p(x;t)}{\mathrm{d}t} = \sum_{x' \neq x} J_{xx'} = 0$$

We will show that under certain conditions, the system relaxes to a stationary state.

$$\lim_{t \to \infty} p(x;t) = p^{st}(x)$$

Detailed Balance Condition

The independence of $p(x)$ on time can be ensured by the following conditions:

1. $\sum_{x' \neq x} J_{xx'} = 0$
2. $J_{xx'} = 0$ for all states x, x'

Condition 1 is the condition for a steady state, while condition 2 is the condition for equilibrium.

Detailed Balance Condition

The detailed balance condition is defined as

$$k_{xx'} p^{st}(x') = k_{x'x} p^{st}(x) \quad \forall \, x, x' \in S$$

$$J_{xx'} = 0 \quad \forall x, x' \in S$$

When the jump networks satisfy the detailed balance condition, the system is said to be in equilibrium.

Perron Frobenius Theorem

Reducibility

A matrix $\mathbb{A}$ is said to be reducible when there exists a Permutation matrix P such that

$$P^T A P = \begin{pmatrix} \mathbb{X} & \mathbb{Y} \\ 0 & \mathbb{Z} \end{pmatrix}$$

where $\mathbb{X}$ and $\mathbb{Z}$ are square matrices.

We can easily see that if all the elements of $\mathbb{A}$ are positive, then the matrix is irreducible. *This is a necessary condition for the Perron Frobenius theorem to hold.*

We can reframe the Irreducibility problem in terms of graphs.

Irreducibility in terms of Graphs

- The graph $G(\mathbb{A})$ of a matrix $\mathbb{A}_{n \times n}$ is a directed graph on n nodes $N_1, ..., N_n$ in which there is a directed edge from N_i to N_j if and only if $A_{ij} \neq 0$.

- A graph $G(\mathbb{A})$ is called strongly connected if for every pair of nodes N_i and N_j there is a directed path from N_i to N_j.

- A matrix $\mathbb{A}$ is irreducible iff the graph of the matrix is strongly connected.

We have already assumed that the jump network is strongly connected.

We can also think of Irreducibility from the point of view of the markov chains. Two states are said to be **communicating** if there is a non-zero probability of transitioning from one state to the other in a finite number of steps. This is an equivalence relation and thus partitions the state space into equivalence classes.

The markov chain is said to be irreducible if there is only one equivalence class, which is the state space of the Markov chain.

Here in our jump networks any state can be reached from any other state in a finite number of steps. Thus our markov chain is irreducible.
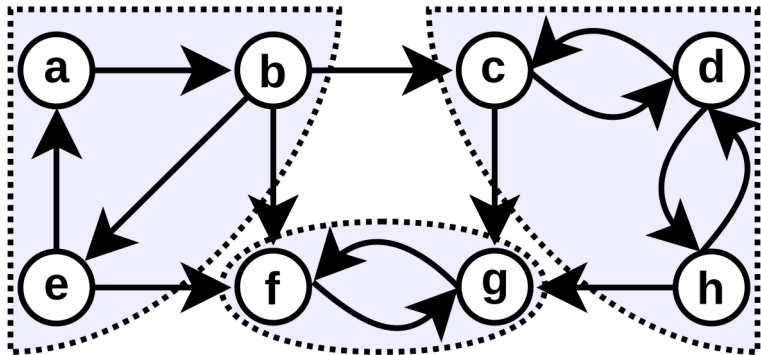
Figure 2: Source : Wikipedia

Primitive

- A matrix $\mathbb{A}$ is called primitive if there exists a positive integer k such that $\mathbb{A}^k$ is a positive matrix (i.e. all elements are positive).

- A matrix $\mathbb{A}$ is primitive if it has only one eigenvalue on the spectral circle.

- A non-negative irreducible matrix $\mathbb{A}$ is primitive with $r = \rho(\mathbb{A})$ iff $\lim_{k \to \infty} (\mathbb{A}/r)^k$ exists, in which case

$$\lim_{k \to \infty} \left(\frac{\mathbb{A}}{r} \right)^k = \frac{pq^T}{q^T p} > 0$$

where p and q are the Perron vectors for $\mathbb{A}$ and $\mathbb{A}^T$ respectively.

The Perron Frobenius Theorem

Perron Frobenius Theorem

Let $\mathbb{A}$ be an irreducible matrix.

- There exists a positive real number λ such that

$$\mathbb{A}\pi = \lambda\pi$$

  called the Perron-Frobenius eigenvalue. Moreover, the eigenvalue is the spectral radius of the matrix.

- The eigenspace corresponding to the Perron-Frobenius eigenvalue is one-dimensional, that is, the corresponding eigenvector is non-degenerate.

- Then there exists a unique probability vector π such that

$$\mathbb{A}\pi = \lambda\pi$$

  called the Perron vector. Moreover, the vector π is strictly positive.

- Furthermore, there are no non-negative eigenvectors of $\mathbb{A}$ except for all positive multiples of π.

Stationary State Dist.

Let P be the transition probability matrix for an irreducible Markov chain and let π be the perron vector for the matrix P.

- The kth step probability matrix is given by P^k since the (i, j)th entry of P^k is the probability of transitioning from state j to state i in exactly k steps.
- The $k$the step distribution is given by $p(k) = P^k p(0)$
- If $\mathbf{P}$ is primitive and if e denotes the column of all ones, then the limit

$$\lim_{k \to \infty} \mathbf{P}^k = \pi e^T \qquad \text{and} \qquad \lim_{k \to \infty} p(k) = \pi$$

Proof:
All stochastic matrices P have a spectral radius $\rho(P) = 1$. All column sums
equal to 1. Thus, e is an eigenvector of P^T with eigenvalue 1, where e is the
column vector with all ones in its entries. We also have $\|P\|_\infty = 1$. Using the
fact that $\rho(\star) \leq \|\star\|$ for every matrix norm(See [2]), we get that

$$1 \leq \rho(P) \leq \|P\|_\infty = 1 \quad \Rightarrow \rho(P) = 1$$

We know that P is primitive with $\rho(P) = 1$, we know that

$$\lim_{k \to \infty} P^k = \frac{pq^T}{q^T p}$$

where p and q are the perron vectors for P and P^T respectively. The perron
vector for P^T is e and let the perron vector for P be π.

We thus have,

$$\lim_{k \to \infty} \mathbf{P}^k = \frac{\pi e^T}{e^T \pi}$$

From the conservation of probability we have that $\sum_i \pi_i = 1$, we have that $e^T \pi = \sum_i \pi_i = 1$. Thus we have

$$\lim_{k \to \infty} \mathbf{P}^k = \pi e^T$$

Now we have

$$\lim_{k \to \infty} p(k) = (\lim_{k \to \infty} \mathbf{P}^k) p(0) = \pi e^T p(0)$$

Again $e^T p(0) = \sum_i p_i(0) = 1$. This leads to

$$\lim_{k \to \infty} p(k) = \pi$$

Note that the stationary limit is independent of the initial distribution $p(0)$ here.

We now actually verify that π is indeed the stationary distribution. Let us recast the master equation into matrix form.

$$\frac{dp(t)}{dt} = Qp(t) = \frac{1}{dt}(\mathbf{P} - \mathbb{I})p(t)$$

where Q is the transition rate matrix. Now we have

$$(\mathbf{P} - \mathbb{I})\pi = \mathbf{P}\pi - \pi = \pi - \pi = 0$$

Thus the infinite time limit of the probability distribution is indeed the stationary distribution.

$$\Rightarrow \frac{dp(t)}{dt} = 0$$

Now the stationary distribution must be unique since the Perron root of any irreducible matrix has a one dimensional eigenspace. Thus the eigenvector must be unique due to the conservation of probability. $\qquad\square$

SIDE NOTE: After completing the presentation, I found a proof that doesn't use Perron Frobenius theorem. However, that only proves it for a transition rate matrix with all positive entries.[1]

Stationary State Dist. for Imprimitive matrices

Let P be the transition probability matrix for an irreducible Markov chain and let π be the perron vector for the matrix P.

- If $\mathbf{P}$ is imprimitive and if e denotes the column of all ones, then the limit

$$\lim_{k \to \infty} \frac{\mathbb{I} + \mathbf{P} + ... + \mathbf{P}^{k-1}}{k} = \pi e^{T}$$

and

$$\lim_{k \to \infty} \frac{p(0) + p(1) + ... + p(k-1)}{k} = \pi$$

We use Cesaro sums to define the stationary state probability distribution. However, we won't deal with it here.

Graph Theory

- A graph is a collection of vertices and edges. The edges connect the vertices.
- The degree of a vertex is the number of edges connected to it.
- A walk is a sequence of vertices connected by edges.
- A trail is a walk with no repeated edges.
- A cycle is a non-empty trail that starts and ends at the same vertex.
- A connected graph is a graph where there is a path between every pair of vertices.

Trees

A tree is a connected graph with no cycles.

Here is one property of a tree that we plan to use later.
A tree with n vertices has $n - 1$ edges.

Handshaking Lemma

The sum of the degrees of all the vertices of a graph is equal to twice the number of edges.

$$\sum_{i=1}^{n} d_i = 2|E|$$

where d_i is the degree of the ith vertex and $|E|$ is the number of edges.

An "obvious" Theorem

A tree has at least two vertices of degree 1.

Proof

Every tree has $n - 1$ edges, so the sum of the degrees of all vertices of any tree has to be $2(n - 1)$. But if there are fewer than two vertices of degree one, then the sum of the degrees of all vertices must be at least $2(n - 1) + 1$, which is a contradiction.

$\square$

Equilibrium Distribution

Equilibrium dist. for trees

Let the jump network be a tree. Then the stationary state distribution is the equilibrium distribution.

Proof

We know that there exists at least 2 vertices with degree 1. Let us label one of them as x. Thus, for this vertex $\sum_{x' \neq x} J_{xx'} = 0$ for all states x' connected to it, under the stationary distribution. Since this is only connected to another vertex say a, We have $J_{xa} = 0$. We can then remove the vertex x and the edge $\{x, a\}$. This is still a tree, so we can do this iteratively. We can reach a point where we have only one vertex left. For the lone vertex, the detailed balance condition is trivially satisfied. Thus, the equilibrium distribution is the stationary distribution.

□

Cyclic Network

A cyclic network is a graph with a cycle in it.

Cyclic Networks

For cyclic networks, the system admits an equilibrium distribution if for any sequence states $(x_0, ..., x_n)$ all different from each other

$$k_{x_0 x_1} k_{x_1 x_2} ... k_{x_{n-1} x_n} k_{x_n x_0} = k_{x_0 x_n} k_{x_n x_{n-1}} ... k_{x_1 x_0}$$

We can see that non-vanishing stationary currents can only survive in loops. If we allow for all cycles the forward and back jump rates to be equal, we can then allow for the balance of probability currents.

Admission of Equilibrium Distribution

For cyclic networks, the system admits an equilibrium distribution if for any sequence states $(x_0, ..., x_n)$ all different from each other

$$k_{x_0 x_1} k_{x_1 x_2} ... k_{x_{n-1} x_n} k_{x_n x_0} = k_{x_0 x_n} k_{x_n x_{n-1}} ... k_{x_1 x_0}$$

Note that this also includes the acyclic cases as well. For all sequences $(x_0, ..., x_n)$ that is not a cycle with x_0 we get $k_{x_0 x_n} = k_{x_n x_0} = 0$. Thus, this condition is trivially satisfied for acyclic networks.

Questions?

Danke Schön

Luca Peliti and Simone Pigolotti, *Stochastic Thermodynamics*, Princeton University Press, Princeton, NJ, 2023.

Carl D. Meyer, *Applied Linear Algebra and Matrix Analysis*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.