Appendix A4. Discrete Markov Stochastic Processes: A Few Remarks

Consider a finite number N of discrete states labelled by indices a, b ….
The set Γ of these states is called the space of states.

Let $\Pi_{ab} \geq 0$ be the probability of transition from state b to state a.
Since state b must transit to some state, even if only to itself, we have

$$\sum_{a=1}^{N} \Pi_{ab} = 1 \qquad \text{(A4.1)}$$

or $\quad \boldsymbol{U}^T \boldsymbol{\Pi} = \boldsymbol{U}^T \qquad$ (A4.1m)

where $\boldsymbol{U} = (1, \ ..., \ 1)^T$ is the right eigenvector for eigenvalue 1 of the transfer matrix $\boldsymbol{\Pi} = \{ \Pi_{ab} \}$.

Let

$P_n(a)$ = Probability of being in state a at time n

then

$$0 \leq P_n(a) \leq 1 \qquad \& \qquad \sum_a P_n(a) = 1 \qquad \text{(A4.2a)}$$

and

$$P_{n+1}(a) = \sum_b \Pi_{ab} P_n(b) \qquad \text{(A4.2)}$$

which is known as the evolution (or master) equation.

In matrix form, we have

$$\begin{aligned} \boldsymbol{P}_{n+1} &= \boldsymbol{\Pi} \, \boldsymbol{P}_n \qquad \text{(A4.2m)} \\ &= \boldsymbol{\Pi}^{n+1} \boldsymbol{P}_0 \end{aligned}$$

where the vector $\boldsymbol{P} = \{ P(a) \}$, with eq(A4.2a), is called a distribution.

Eq(A4.1 or A4.1m) means that every column of $\boldsymbol{\Pi}$ has a sum of 1.

Eq(A4.2) implies

$$\begin{aligned} \sum_a P_{n+1}(a) &= \sum_b \sum_a \Pi_{ab} P_n(b) \\ &= \sum_b P_n(b) = 1 \qquad \text{[Eq(A4.1) used]} \end{aligned}$$

so that probability is conserved.

In matrix form,

$$\boldsymbol{U}^T \boldsymbol{P}_{n+1} = \boldsymbol{U}^T \boldsymbol{\Pi} \, \boldsymbol{P}_n = \boldsymbol{U}^T \boldsymbol{P}_n = 1$$

A4.1. The Spectrum of the Transition Matrix

As already mentioned, the vector $\boldsymbol{U}$, with

$$U_a = 1 \qquad \forall \, a$$

is a left eigenvector of $\boldsymbol{\Pi}$ with eigenvalue 1, since

$$\sum_a U_a \Pi_{ab} = \sum_{a=1}^{N} \Pi_{ab} = 1 = U_b$$

or $\quad \boldsymbol{U}^T \boldsymbol{\Pi} = \boldsymbol{U}^T$

The corresponding right eigenvector $\boldsymbol{V}$ (of eigenvalue 1) satisfies

$$\boldsymbol{\Pi} \boldsymbol{V} = \boldsymbol{V} \qquad \text{(A4.3m)}$$

or $\sum_b \Pi_{ab} V_b = V_a$ (A4.3)

$\boldsymbol{V}$ is a stationary solution to eq(A.4.2) since

$$\boldsymbol{\Pi}^n \boldsymbol{V} = \boldsymbol{V} \qquad \forall\, n$$

If

$$V_a \geq 0 \;\; \forall\, a$$

then

$$P_n(a) = \frac{V_a}{\sum_a V_a} \qquad \text{or} \qquad \boldsymbol{P}_n = \frac{\boldsymbol{V}}{\boldsymbol{U}^T \boldsymbol{V}}$$

is a equilibrium distribution.

More generally, let $\boldsymbol{V}$ be a right eigenvector of $\boldsymbol{\Pi}$ with eigenvalue ν , i.e.,

$$\boldsymbol{\Pi}\, \boldsymbol{V} = \nu \boldsymbol{V} \qquad \text{(A4.4m)}$$

$$\sum_b \Pi_{ab} V_b = \nu\, V_a \qquad \text{(A4.4)}$$

then

$$\sum_{a,b} \Pi_{ab} V_b = \sum_b V_b \qquad \text{[Eq(A4.1) used]}$$

$$= \nu \sum_a V_a \qquad \text{[From eq(A4.4)]}$$

Or, in matrix form,

$$\boldsymbol{U}^T \boldsymbol{\Pi}\, \boldsymbol{V} = \boldsymbol{U}^T \boldsymbol{V} = \nu\, \boldsymbol{U}^T \boldsymbol{V}$$

$$\rightarrow \qquad \nu = 1 \qquad \text{or} \qquad \sum_a V_a = 0 \qquad \text{(A4.4a)}$$

Taking the absolute value of eq(A4.4) gives

$$|\nu|\;|V_a| = \left|\sum_b \Pi_{ab} V_b\right|$$

$$\leq \sum_b \Pi_{ab}\, |V_b| \qquad \text{(A4.5)}$$

$$\rightarrow \qquad |\nu| \sum_a |V_a| \leq \sum_{a,b} \Pi_{ab}\, |V_b| \quad = \sum_b |V_b|$$

$$\therefore \qquad |\,\nu\,| \;\leq 1 \qquad \text{(A4.6)}$$

where the equality holds only if

$$|V_a| = \sum_b \Pi_{ab}\, |V_b| \qquad \text{(A4.7)}$$

which, by eq(A4.3), means that $\boldsymbol{P} = \left\{\frac{|V_a|}{\sum_a |V_a|}\right\}$ is a stationary probability distribution.

Thus, eqs(A4.6-7) seems to imply that, if $\boldsymbol{V}$ is the eigenvector of $|\,\nu\,| = 1$, then $\boldsymbol{P} = \left|\frac{\boldsymbol{V}}{\boldsymbol{U}^T \boldsymbol{V}}\right|$ is a stationary distribution. This is misleading since the statement actually holds only for the special case $\nu = 1$. The proof is as follows.

If $\boldsymbol{V}$ is the eigenvector for $|\,\nu\,| = 1$, then

$$\boldsymbol{\Pi}\, \boldsymbol{V} = \nu\, \boldsymbol{V} \qquad \text{(a)}$$

$$\rightarrow \qquad |\,\boldsymbol{\Pi} \boldsymbol{V}\,| = |\,\boldsymbol{V}\,|$$

If eq(A4.7) also holds, then

$$|\,\boldsymbol{\Pi} \boldsymbol{V}\,| = \boldsymbol{\Pi}\,|\,\boldsymbol{V}\,|$$

which is possible only if $\boldsymbol{V}$ is real & non-negative. Since $\boldsymbol{\Pi}$ is real, this in turn means $\boldsymbol{V}$ cannot satisfy eq(a) unless ν is real so that $\nu = 1$. QED

If for some subset I of $\{1, \ldots, N\}$,

$$V_a = 0 \qquad \forall\, a \in I$$

then eq(A4.7) implies

$$0 = \sum_{b \notin I} \Pi_{a\,b}\, |V_b| \qquad \forall\, a \in I \qquad \text{(A4.8)}$$

Since both $\Pi_{a\,b}$ & $|V_b|$ are non-negative, this means

$$\Pi_{a\,b} = 0 \qquad \forall\, a \in I \;\&\; b \notin I \qquad \text{(A4.8a)}$$

Note that interchanging rows i & j and then columns i & j amounts to exchanging the labels of the states i & j. Such an operation therefore has no effect on the physics of the system.

Hence, by suitably renaming the relevant states, we can write eq(A4.8a) in block matrix form as

$$\boldsymbol{\Pi} = (\Pi_{I\,J}) = \left(\begin{array}{c|c} * & 0 \\ \hline * & * \end{array}\right) \qquad \text{(A4.8b)}$$

where the subscript J denotes the subset $\{b \notin I\}$.

Since,

$$\left(\begin{array}{c|c} * & 0 \\ \hline * & * \end{array}\right)\begin{pmatrix} 0 \\ \boldsymbol{b} \end{pmatrix} = \begin{pmatrix} 0 \\ * \end{pmatrix}$$

any state with only components in $\{b \notin I\}$ will never evolve into states in I.

Thus, if there exists a vector $\boldsymbol{V}$ such that the a^{th} component of $\boldsymbol{\Pi}\,\boldsymbol{V}$ vanishes, then state a is unreachable by any state that has no finite a^{th} component, i.e.,

$$\begin{pmatrix} \boldsymbol{a} \\ 0 \end{pmatrix} \neq \left(\begin{array}{c|c} * & 0 \\ \hline * & * \end{array}\right)^n \begin{pmatrix} 0 \\ \boldsymbol{b} \end{pmatrix} \qquad \forall\, n \qquad \text{(A4.8c)}$$

Connectivity Assumption

A subspace Γ of $\mathcal{V}$ is connected (with respect to a given $\boldsymbol{\Pi}$) if, for any pair of states $a, b \in \Gamma$,

$$\exists\, \{c_1, \ldots, c_r\} \text{ such that } \qquad \Pi_{b\,c_r}\, \Pi_{r\,c_{r-1}} \ldots \Pi_{1\,a} > 0 \quad \text{(A4.9a)}$$

This means any state has a non-zero probability to reach any other state.

Let $\boldsymbol{E}(a)$ be the basis vector representing state a, i.e.,

$$\boldsymbol{E}(a) = \{\, E_b(a) = \delta_{b\,a} \}$$

then eq(A4.9a) means that for any pair of $\boldsymbol{E}(a)$ & $\boldsymbol{E}(b)$, there exists an n such that

$$\boldsymbol{E}(a)^T\, \boldsymbol{\Pi}^n\, \boldsymbol{E}(b) \neq 0$$

Hereafter, Γ is assumed to be connected.

Note that, by the arguments of the last section, Γ is disconnected even if there is only one off-diagonal $\Pi_{a\,b} = 0$.

In view of eqs(A4.2a & A4.8), the stationary distribution must has no vanishing component. Otherwise, any distribution on its way to equilibrium will eventually satisfy eq(4.8), which in turn implies Γ is disconnected.

Moreover, all eigenvalues of $\boldsymbol{\Pi}$ are non-degenerate, i.e., the eigenspace of every eigenvalue has dimension $n = 1$. (A4.9b)

Proof is by contradiction as follows.

If $n > 1$, then there exists at least 2 vectors $\boldsymbol{V}$ & $\boldsymbol{W}$ such that

$$\boldsymbol{\Pi}\,\boldsymbol{V} = \nu\,\boldsymbol{V} \qquad \boldsymbol{\Pi}\,\boldsymbol{W} = \nu\,\boldsymbol{W}$$

Let

$$\boldsymbol{U} = \alpha \boldsymbol{V} + \beta \boldsymbol{W} \qquad \alpha, \beta = \text{constants}$$
$$\to \quad \Pi \boldsymbol{U} = \nu \boldsymbol{U}$$

Since α & β are arbitrary, we can always choose them so that one component of $\boldsymbol{U}$, and hence $\Pi \boldsymbol{U}$, vanishes. Γ is therefore disconnected. QED.

The stationary distribution given by the eigenspace of $\nu = 1$ [see eq(A4.3)] is therefore a single vector with components

$$P(a) = \frac{1}{\mathcal{Z}} e^{-E(a)} \qquad \text{(A4.9)}$$

with

$$E(a) > 0 \qquad \& \qquad \mathcal{Z} = \sum_a e^{-E(a)}$$

Note that eq(A4.9) merely ensures

$$1 > P(a) > 0 \qquad \& \qquad \sum_a P(a) = 1$$

as befits a probability distribution.

Let $\boldsymbol{S}$ be the similarity transform that diagonalize Π, i.e.,

$$\boldsymbol{S} \Pi \boldsymbol{S}^{-1} = \text{diag}(\nu_1, \ldots, \nu_N) \qquad \{\nu_j\} = \text{eigenvalues of } \Pi$$
$$\to \quad \boldsymbol{S} \Pi^n \boldsymbol{S}^{-1} = \boldsymbol{S} \Pi \boldsymbol{S}^{-1} \boldsymbol{S} \Pi \boldsymbol{S}^{-1} \ldots \boldsymbol{S} \Pi \boldsymbol{S}^{-1}$$
$$= \text{diag}(\nu_1^n, \ldots, \nu_N^n)$$

i.e, the eigenvalues of Π^n are simply $\{\nu_1^n, \ldots, \nu_N^n\}$.

If $|\nu_j| < 1$, then

$$\lim_{n\to\infty} |\nu_j^n| = \lim_{n\to\infty} |\nu_j|^n = 0 \to \qquad \lim_{n\to\infty} \nu_j^n = 0$$

This means, as $n \to \infty$, all states in the eigenspace of $|\nu| < 1$ evolves to the zero vector. The stationary distribution thus resides in the eigenspace of $|\nu| = 1$.

If Π has one eigenvalue $\nu = 1$, while all other eigenvalues have $|\nu| < 1$, then any distribution will evolve eventually to the stationary distribution $\boldsymbol{P}$ given by eq(A4.9), which is thus the equilibrium distribution.

Other Eigenvalues with Modulus One

Since all elements of Π is real, its trace & its determinant are also real.
I.e., the sum & product of its eigenvalues are real.
Which means its eigenvalues are either real, or appear in complex conjugate pairs.

If there exists another eigenvalue with $|\nu| = 1$, i.e.,

$$\nu = e^{i\theta} \qquad \text{with} \qquad \theta \neq 0 \quad (\text{mod}\, 2\pi)$$
$$\Pi \boldsymbol{V} = e^{i\theta} \boldsymbol{V}$$

then $\boldsymbol{V}$ satisfies

$$\Pi^n \boldsymbol{V} = e^{i n \theta} \boldsymbol{V}$$

Writing $\boldsymbol{V} = \boldsymbol{V}_R + i \boldsymbol{V}_I$, then $\boldsymbol{V}$ rotates counterclockwise by an angle θ in the complex plane after each time-step. Since Π does not change the magnitude $|\boldsymbol{V}|$ of $\boldsymbol{V}$, the distribution $\boldsymbol{P} = |\boldsymbol{V}|$ is a stationary distribution, though not in the sense of eq(A4.7).

If θ is of the form

$$\theta = 2\pi \frac{n}{m} \qquad \text{where } m, n \;\&\; \frac{N}{m} \text{ are natural numbers}$$

then the motion is periodic with period m.

The components of $\boldsymbol{V}$ can then be sorted by their phases, which take only m values:

$$\theta_j = j\,\theta \qquad\qquad j = 0,\ 1,\ \ldots,\ m-1$$

For convenience, we re-label the states so that the 1st $\frac{N}{m}$ components of $\boldsymbol{V}$ all have phase 0, the 2nd $\frac{N}{m}$ components all have phase θ, & so on .... The eigen-equation can be written in block form as

$$\boldsymbol{\Pi}\,\boldsymbol{V} = \begin{pmatrix} \boldsymbol{\Pi}_{00} & \boldsymbol{\Pi}_{01} & \cdots & \boldsymbol{\Pi}_{0,m-1} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Pi}_{k\,0} & \boldsymbol{\Pi}_{k\,1} & \cdots & \boldsymbol{\Pi}_{k,m-1} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Pi}_{m-1,\,0} & \boldsymbol{\Pi}_{m-1,\,1} & \cdots & \boldsymbol{\Pi}_{m-1,\,m-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{V}_0 \\ \vdots \\ e^{i\,k\,\theta}\,\boldsymbol{V}_k \\ \vdots \\ e^{i\,(m-1)\,\theta}\,\boldsymbol{V}_{m-1} \end{pmatrix} = e^{i\,\theta} \begin{pmatrix} \boldsymbol{V}_0 \\ \vdots \\ e^{i\,k\,\theta}\,\boldsymbol{V}_k \\ \vdots \\ e^{i\,(m-1)\,\theta}\,\boldsymbol{V}_{m-1} \end{pmatrix}$$

where $\boldsymbol{V}_j$ are real vectors of dimension N/m.

Consider the typical k^{th} row. Since $\boldsymbol{\Pi}$ is real & non-negative, the only way to obtain the right hand side term $e^{i\,(k+1)\,\theta}\,\boldsymbol{V}_k$ is to set

$$\boldsymbol{\Pi}_{k\,j} = 0 \qquad \forall\ j \neq k+1$$

& $\quad \boldsymbol{\Pi}_{k,\,k+1}\,\boldsymbol{V}_{k+1} = \boldsymbol{V}_k$

During the evolution of the system, the $\boldsymbol{V}_k$ segments thus undergo cyclic permutations of period m.

Infinite Number of States

See Zinn-Justin’s text.