Limit Distributions for Finite State, Homogeneous Discrete Time, Irreducible Markov Chains

In the long run what fraction of the time is the system (MC) in a particular state?

To prepare, need:

**Defn** The period \( d(i) \) of a state \( i \in S \) is given by

\[
d(i) = \min \{ n \geq 0 : (P^n)_{ii} > 0 \}
\]

If \( d(i) = 1 \), then \( i \) is said to be a periodic.

Examples: Random walk on graph (reflecting BC)

![Graph](image)

\( d(A) = 2 \)

In fact the period of any state in a given communication class is the same as the period of any other state in the class.

So \( d(j) = 2 \) for all \( j \in S \) in this example.
General idea: A period $d$ Markov chain can be decomposed into $d$ subsets, the MC dynamics proceed from one subset to the next.

We'll focus on aperiodic irreducible MC's for the most part.

Theorem: If $\{X_n\}_{n=0}^{\infty}$ is an aperiodic, irreducible FSDT homogeneous MC w/ irreducible transition probability matrix $P$, then it has a unique stationary distribution $\pi^*$ and
$$\lim_{n \to \infty} \text{Prob}(X_n = j) = \pi^*_j. \quad (*)$$

That is, $\pi^*$ is the limit distribution for the MC.
Idea of Proof:

Proof A): Finite dimensional linear algebra
(Calculer Ch. 1)

Thm says \[ \lim_{n \to \infty} \langle \varnothing, p^n \rangle = \prod_{i \in \text{dist,}} \]

Why would this be true?

If limit exists, it has to be stat. dist.

because \[ \lim_{n \to \infty} \langle \varnothing, p^n \rangle = \lim_{n \to \infty} \langle \varnothing, p^{n+1} \rangle = \lim_{n \to \infty} \langle \varnothing, p^n \rangle \]

(check that limit exists, normalizable)

Let's write \( p \) in Jordan form:

\[ p = A J A^{-1} \]

where \( J \) has Jordan block form

If diagonalizable \[ J = \begin{pmatrix}
\lambda_1 & 0 & 0 & \cdots & 0 \\
0 & \lambda_2 & 0 & \cdots & 0 \\
0 & 0 & \ddots & \ddots & 0 \\
0 & 0 & \cdots & \lambda_{n-1} & 0 \\
0 & 0 & \cdots & 0 & \lambda_n
\end{pmatrix} \]

eigenvalues \( \lambda_1, \ldots, \lambda_n \).
\[ p^n = Q J^n Q^{-1} \]

For large \( n \), Jordan blocks with:

- \( |\lambda| > 1 \): amplified
- \( |\lambda| = 1 \): persists
- \( |\lambda| < 1 \): decay

If \( p^n \) is to project all vectors \( \vec{v} \) onto the vector \( \vec{v}_1 \) (which is an eigenvector with eigenvalue 1), then need all other eigenvalues satisfy \( |\lambda| < 1 \).

What do we need to guarantee this?

Irreducibility (otherwise multiple eigenvalues = 1)

A periodicity avoids eigenvalues \( \lambda \) with

\( |\lambda| = 1 \) but \( \lambda \neq 1 \).
So given $\text{periodicity} + \text{irreducibility}$, how prove them?

Perron-Frobenius Theorem (Karlin + Taylor App. 2) (Lawler Ex1.15)

If matrix $A$ has all entries $A_{ij} > 0$
then there exists an eigenvalue $\lambda$
of maximal amplitude such that:
1) $\lambda$ is real and positive
2) $\lambda$ is a simple eigenvalue
3) The unique eigenvector associated
to $\lambda$ has all entries $\geq 0$.

Use this for the proof of the limit that:
Irreducible + aperiodic $\Rightarrow$ There is an $N \to \infty$
such that for $n \geq N$, $(P^n)_{ij} > 0$ for all $i, j \in S$.

$P$ may have some $0$ entries (so can't use $P^T$ on $P$)
but $P^n$ has all positive entries for $n \geq N$. 
Apply PPT to $p^n$

$p^n$ has a simple real eigenvalue $\lambda$ which has maximal amplitude.

I also know that $\pi$ is an eigenvector of $p^n$ w/ eigenvalue 1.

Is $\lambda = 1$ or is $\lambda > 1$?

Observe that $\sum_{j \in S} (p^n)_{ij} = 1$.

Since the eigenvector $\pi$ corresponding to $\lambda$ has all nonnegative entries,

$$(p^n \pi)_i = \sum_{j \in S} (p^n)_{ij} \pi_j \leq \|\pi\|_\infty \leq (p^n)_{ij} \max_{j \in S} \pi_j$$

$\pi^n \pi_i \leq \|\pi\|_\infty$

This implies $\lambda^n \leq 1 \Rightarrow \lambda \leq 1 \Rightarrow \lambda > 1$.

So therefore, $\pi$ is the unique eigenvector of eigenvalue $\lambda = 1$, it has all nonnegative entries, and all other eigenvalues $\mu$ satisfy $|\mu| < 1$. This shows $\pi$ is the limit distribution.
Comment: General principle is that amount of time you have to wait for transient effects from initial data to die out and for the MC to reach its limit distribution is $n$ such that

$$|\lambda_2|^n < 1$$

where $\lambda_2$ is next biggest eigenvalue (after 1).

If $\lambda_2$ is near 1 $\Rightarrow$ metastable states (C. Schütte, Berlin)

Another remark: One can show that the approach to stat dist. can be described in terms of information theory/entropy.

Define the relative entropy of a Prob. dist. $\bar{p}$ with respect to stat dist. $\Pi$ is

$$I(\bar{p}, \Pi) = \sum_j p_j \ln \left( \frac{p_j}{\Pi_j} \right)$$

This is a Lyapunov function for MC, meaning

$$I(\bar{p}^{(n+1)}, \Pi) < I(\bar{p}^{(n)}, \Pi)$$

and

$$\lim_{n \to \infty} I(\bar{p}^{(n)}, \Pi) = 0$$

Haken, Synergetics
Proof B): Coupling: Trendy and generalizes to $\infty$-state MC

We know that we have stat. dist. $\mathbb{P}$ with $\pi_j = \pi_j$. If we have a MC $\{X_n\}_{n=0}^{\infty}$ which starts w/ dist. $\mathbb{P}$: $\Pr(X_0 = j) = \pi_j$. and evolves w/ $P$, then obviously $\lim_{n \to \infty} \Pr(X_n = j) = \pi_j$.

Consider a MC $\{X_n\}_{n=0}^{\infty}$ with same trans matrix $P$ but arbitrary initial dist. $\mathbb{Q}_j = \Pr(X_0 = j)$.

If I can show that $\Pr(X_n = Y_n$ for some $n \geq \infty)$ then the Markov property will show that $\lim_{n \to \infty} \Pr(X_n = j) = \lim_{n \to \infty} \Pr(Y_n = j) = \pi_j$. 

[Resnick 2.13]
How show $\text{Prob}(X_n = Y_n \text{ for some } n \to \infty)$?
- aperiodicity + irreducibility

: reducible

: $X_n$ and $Y_n$ can be out-of-phase or periodic.

Resnick: View $\Xi_n = (X_n, Y_n)$

This is a powerful idea;
Goes from existence of a stat. dist.
To prove it is the unique, long-time limit of system.

Malliavin: Two-dim. Navier-Stokes equ
Examples Quality Inspection Protocols

Production line

Inspection station

Product ships if not inspected
or if inspected but not
defective.

Regular adaptive sampling scheme:
Start by inspecting every product
until you get M good inspections
in a row. Then I switch to sampling
only 1 in every r products, with
regular spacing. Whenever a defect
is found, revert to inspecting every
product, until M good consecutive
products are found again.

M = 3  r = 4

defect
Questions (for any inspection protocol):

1) How often is an inspected product defective?

2) What fraction of shipped products are defective?

3) What fraction of products are inspected? (effort)

With answers, could try to optimize

Write $M$ and $n$ or compare schemes.

Use Markov chain ideas to make predictions
about these questions.

Have to model unpredictability in the
status of products.

Simplest version: Each product is defective
with prob. $p$, independent of the status
of any other product.
Markov Chain Model

We will choose state space $S = \{0, \ldots, M\}$ as describing the state of the machine. We will let $X_n$ be the state of machine after the $n$th inspection.

$X_n = j$ means after $n$th inspection, the machine has seen $j$ consecutive good products.

(Don't need to include state of product since it's iid and act as input noise to machine state).

$$P_{ij} = \text{Prob}(X_{n+1} = j \mid X_n = i)$$

$$P = \begin{pmatrix}
0 & 1-p & 0 & 0 & 0 & \cdots & 0 & 0 \\
p & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & p & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & p & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & p & 0 \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & p \\
\end{pmatrix}$$