$$P^{m,m+n+r} = P^{m,m+n} P^{m+n,m+n+r}$$

$$P^{m,m+n} = P^{n}$$

Proof

$$P_{ij}^{m,m+n+r} = P(X_{m+n+r} = j \mid X_m = i)$$

$$= \underbrace{\sum_k P(X_{m+n+r} = j, X_{m+n} = k \mid X_m = i)}_{\text{The chain must pass through some intermediate step } k.}$$

Now for events A, B, C

$$P(A \cap B \mid C) = P(A \mid B \cap C) P(B \mid C)$$

A: $X_{m+n+r} = j$

B: $X_{m+n} = k$

C: $X_m = i$

$$P_{ij}^{m,m+n+r} = \sum_{k \in S} P(X_{m+n+r} = j \mid X_{m+n} = k, X_m = i) P(X_{m+n} = k \mid X_m = i)$$

Markov property

$$= \sum_k P(X_{m+n+r} = j \mid X_{m+n} = k) P(X_{m+n} = k \mid X_m = i)$$

$\boxed{\#6 \quad 2/7}$

As a consequence

$$(2.2.2) \qquad P^{m,m+n} = P^{0,0+n}.$$

We introduce the notation

$$P^{m,m+n} \to P^n, \quad P_{ij}^{m,m+n} \to P_{ij}^n$$

<u>Theorem 2.2.2</u>

Let $\{\mu_i^n\}$,

$$\mu_i^n = P(X_n = i), \qquad i = 0, 1, 2, \ldots$$

be the pmf for state X_n. Set

$$\vec{\mu}^n = (\mu_0^n, \mu_1^n, \ldots).$$

Then,

$$\vec{\mu}^{m+n} = \vec{\mu}^m P^n$$

$$\vec{\mu}^n = \vec{\mu}^0 P^n$$

<u>Proof</u>

$$\mu_j^{m+n} = P(X_{m+n} = j) = \sum_{i=0}^{\infty} P(X_{m+n} = j \mid X_m = i) P(X_m = i)$$

$$= \sum_{i=0}^{\infty} P_{ij}^n \mu_i^m = (\vec{\mu}^m P^n)_{ij}$$

<u>by</u> Thm 2.2.1.

Note: In standard linear algebra courses, we represent transformations using matrices on column vectors, $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix}$, $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & \cdots & \\ \vdots & & \end{pmatrix}$

$\vec{v} \to A\vec{v}$ = new column vector (matrix-vector product)

We can also represent transformations using row vectors:

$$\vec{w} = (w_1, w_2, w_3, \ldots)$$

$\vec{w} \to \vec{w}A$ = new row vector (vector-matrix product)

(Recall the rule "row into column")

We conclude that the intermediate time evolution of a Markov chain is determined by $\vec{\mu}^0$ and P.

<u>Example 2.2.1</u>

Consider Ex. 2.1.7 with

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 0 & 1 \end{pmatrix}$$

$$P^2 = \begin{pmatrix} 1\times 1 & 0 & 0 \\ 1/4 + \frac{1}{2}\frac{1}{4} & \frac{1}{2}\cdot\frac{1}{2} & \frac{1}{2}\cdot\frac{1}{4} + 1 \\ 0 & 0 & 1\times 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 3/8 & 1/4 & 3/8 \\ 0 & 0 & 1 \end{pmatrix}$$

symmetric

We see that the 0's and 1's will be present in any power of P. Likewise, the middle entry is $\left(\frac{1}{2}\right)^n$. By symmetry, $P_{21}^n = P_{23}^n$ for all n, and this can be worked out

$$P^n = \begin{pmatrix} 1 & 0 & 0 \\ (1-(1/2)^n)/2 & \left(\frac{1}{2}\right)^n & (1-(1/2)^n)/2 \\ 0 & 0 & 1 \end{pmatrix}$$

We can draw some interesting conclusions. For example, the probability of observing genotype Aa in subsequent generations goes to zero rapidly. Anticipating future discussion,

$$\lim_{n\to\infty} P^n = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 1 \end{pmatrix}$$

If we start with a plant with genotype AA or aa, we stay there forever with probability 1. If we start with a plant with genotype Aa, we have equal probability of ending up with a plant with genotype AA or aa, and we reach such a plant in very few generations with large probability.

In this example, there is some large time limiting behavior.

Example 2.2.2

For the random walk example 2.1.2, recall (2.1.2)

$$X_n = X_0 + \sum_{k=1}^{n} B_k$$

$\{B_k\}$ = sequence of Bernoulli variables with values $\{-1, 1\}$, prob. $\{1-p, p\}$

X_n = position of the particle after n steps starting at X_0.

The set of realizations of the random walk can be described as the set of vectors

$$\{ \vec{s} = (s_0, s_1, s_2, \ldots) \}$$

where

$$s_0 = X_0$$

$$s_{i+1} - s_i = \pm 1, \quad i \geq 1$$

Each such vector describes a sample path.

The probability that the first n steps of a walk follow a given path $\vec{s} = (s_0, s_1, \dots, s_n)$ is

$$p^r (1-p)^\ell$$

$r =$ # steps of s to the right

$\ell =$ # steps of s to the left

i.e.

$$r = \text{number in } \{ i : s_{i+1} - s_i = +1 \}$$

$$\ell = \text{number in } \{ i : s_{i+1} - s_i = -1 \}$$

Any event can be expressed as a collection of suitable paths and the probability will be the sum of the component probabilities.

For example,

$$P(X_n = j \mid X_0 = i) = \sum_{r=0}^{n} N_n^r(i, j)\, p^r (1-p)^{n-r}$$

$N_n^r(i, j) =$ number of paths $(s_0, s_1, \dots, s_n)$ with $s_0 = i$, $s_n = j$ and exactly r rightward steps

Clearly, .

$$r + l = n$$

$$r - l = j - i = \text{rightward displacement}$$

$$\Rightarrow \quad r = \frac{1}{2}(n+j-i)$$

$$l = \frac{1}{2}(n-j+i)$$

Thus,

$$P(X_n = j \mid X_0 = i) = \binom{n}{\frac{1}{2}(n+j-i)} p^{\frac{1}{2}(n+j-i)} (1-p)^{\frac{1}{2}(n-j+i)}$$

when

$$\frac{1}{2}(n+j-i) \text{ is an integer in } [0, 1, \ldots, n]$$

and

$$P(X_n = j \mid X_0 = i) = 0$$

otherwise.

This follows because there are exactly $\binom{n}{r}$ paths of length n with r rightward steps and $n-r$ leftward steps. Summarizing

$$(2.2.2) \quad P(X_n = j \mid X_0 = i) = \begin{cases} \binom{n}{\frac{1}{2}(n+j-i)} p^{\frac{1}{2}(n+j-i)} (1-p)^{\frac{1}{2}(n-j+i)} & n+j-i \text{ even} \\ 0 & \text{otherwise} \end{cases}$$

Example 2.2.3

Consider a system that fluctuates between off (0) and on (1) states at discrete times. If the system is off at some time, the probability that it is on at the next time is p, $0 \le p \le 1$. If it is on at some time, the probability it is off at the next time is q, $0 \le q \le 1$.

X_n = state at time n
State space is $\{0,1\}$

X_{n+1} is determined only by X_n.

$$P_{00} = P(X_{n+1} = 0 \mid X_n = 0) = 1-p$$

$$P_{01} = P(X_{n+1} = 1 \mid X_n = 0) = p$$

$$P_{11} = P(X_{n+1} = 1 \mid X_n = 1) = 1-q$$

$$P_{10} = P(X_{n+1} = 0 \mid X_n = 1) = q$$

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

To compute powers of P, we use

eigenvalue decompositions. If P is a diagonalizable matrix, there are matrices Q, Λ, where

$$\Lambda = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_m \end{pmatrix}$$

is a diagonal matrix of eigenvalues and Q is an invertible matrix, such that

$$P = Q^{-1} \Lambda Q$$

In this case,

$$P^n = Q^{-1} \Lambda^n Q$$

with

$$\Lambda^n = \begin{pmatrix} \lambda_1^n & & & 0 \\ & \lambda_2^n & & \\ & & \ddots & \\ 0 & & & \lambda_m^n \end{pmatrix}$$

We compute Λ, Q:

$$\begin{vmatrix} 1-p-\lambda & p \\ q & 1-q-\lambda \end{vmatrix} = (1-p-\lambda)(1-q-\lambda) - qp$$

$$= 1-q-\lambda-p+pq+p\lambda-\lambda+\lambda q+\lambda^2-qp$$

$$= 1-q-p+(p+q-2)\lambda+\lambda^2$$

$$= (\lambda-1)\left(\lambda-(1-p-q)\right).$$

This means $\lambda_1 = 1$, and $\lambda_2 = 1-p-q$ are the eigenvalues.

For $\lambda_1 = 1$,

$$\begin{pmatrix} 1-p-\lambda & p \\ q & 1-q-\lambda \end{pmatrix} = \begin{pmatrix} -p & p \\ q & -q \end{pmatrix}$$

so the eigenvector satisfies

$$\begin{pmatrix} -p & p \\ q & -q \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow v_1 = v_2$$

We choose $\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

For $\lambda_2 = 1-p-q$,

$$\begin{pmatrix} 1-p-\lambda & p \\ q & 1-q-\lambda \end{pmatrix} = \begin{pmatrix} 1-p-(1-p-q) & p \\ q & 1-q-(1-p-q) \end{pmatrix}$$

$$= \begin{pmatrix} q & p \\ q & p \end{pmatrix}$$

so

$$\begin{pmatrix} q & p \\ q & p \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \quad q v_1 + p v_2 = 0$$

A solution is $v_1 = p$, $v_2 = -q$, and we set

$$\vec{v}_2 = \begin{pmatrix} p \\ -q \end{pmatrix}$$

Recall that Q is the matrix whose columns are the eigenvectors taken in the order corresponding to Λ. We have

$$Q = \begin{pmatrix} 1 & p \\ 1 & -q \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 1-p-q \end{pmatrix}$$

We compute Q^{-1}: recall

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$$Q^{-1} = \frac{1}{-q-p}\begin{pmatrix} -q & -p \\ -1 & 1 \end{pmatrix} = \frac{1}{p+q}\begin{pmatrix} q & p \\ 1 & -1 \end{pmatrix}$$

Hence,

$$P^n = \frac{1}{p+q}\begin{pmatrix} q & p \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1^n & 0 \\ 0 & (1-p-q)^n \end{pmatrix}\begin{pmatrix} 1 & p \\ 1 & -q \end{pmatrix}$$

$$= \frac{1}{p+q}\begin{pmatrix} q & p \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & p \\ (1-p-q)^n & -q(1-p-q)^n \end{pmatrix}$$

and

$$P^n = \frac{1}{p+q}\begin{pmatrix} q & p \\ q & p \end{pmatrix} + \frac{(1-p-q)^n}{p+q}\begin{pmatrix} p & -p \\ -q & q \end{pmatrix}$$

This is is the finite time result.

While we are here, we consider what happen as $n \to \infty$.

If $p = q = 0$, $P = I$, and the initial state never changes ($\lim_{n\to\infty} P^n = P$). $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, the identity matrix

If $p = q = 1$, then

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad P^2 = I, \; P^3 = P, \; P^4 = I, \ldots$$

The states simply fluctuate back and forth with each time, and there is no limiting behavior.

Otherwise, $|\lambda_2| = |1-p-q| < 1$, hence

$$\lim_{n\to\infty} P^n = \frac{1}{p+q}\begin{pmatrix} q & p \\ q & p \end{pmatrix}$$

Now, the rows of the limit are identical. Hence, for large times, the probabilities that the system is off and on are approximately

$$\frac{q}{p+q}, \quad \frac{p}{p+q}$$

respectively, regardless of initial state.

In this situation, we call the distribution $\left\{\frac{q}{p+q}, \frac{p}{p+q}\right\}$ on the state space $\{\text{off}, \text{on}\}$ a <u>limiting</u> <u>distribution</u>.

We would like to know when this holds.

§2.3 An Intermediate Time Analysis.

We consider some intermediate time questions that generalizes Ex. 2.2.1.

<u>Example 2.3.1</u>

Consider a Markov chain $\{X_n\}$ whose

transition probability matrix between states $\{0,1,2\}$ is

$$P = \begin{pmatrix} 1 & 0 & 0 \\ \alpha & \beta & \gamma \\ 0 & 0 & 1 \end{pmatrix}, \quad \alpha,\beta,\gamma > 0, \quad \alpha+\beta+\gamma = 1$$

If the Markov chain starts in state 0 or 2, it remains in 0 or 2 respectively (with probability 1). If it starts in state 1, it remains in state 1 for some (random) time then moves to state 0 or 2, where it is trapped or absorbed, i.e. remains for ever.

—start # 7 2/12

Questions:

1) In which state is the process absorbed into?

2) How long does it take to reach an absorbed state?

Let

$$T = \min\{n \geq 0, X_n = 0 \text{ or } X_n = 2\}$$

be the time of absorption.