Introduction / Definitions

Note Title 8/24/2005

ODE's $\dot{x}_t = f(x_t)$ ← no memory; once you know initial state, future is determined

Real world has noise. However, memory is often "short"; once you know where you start, the past does not help predict the future

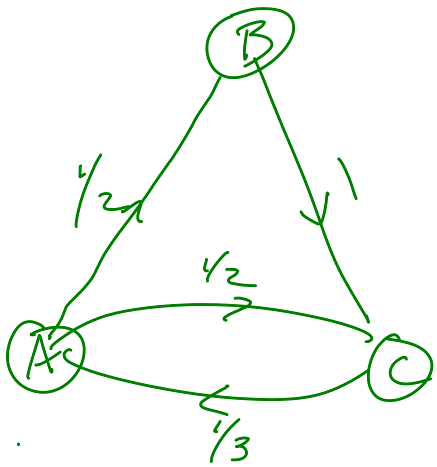
Markov property

Course subject: Markov processes with

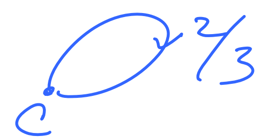
- discrete state space
- $\underbrace{\text{discrete}}_{\mathbb{Z}^+ = \{0,1,\ldots\}}$ or $\underbrace{\text{continuous time}}_{\mathbb{R}_+ = [0,\infty)}$

Diagrams to think about:

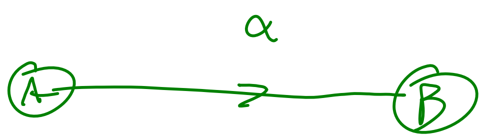
a) Discrete time



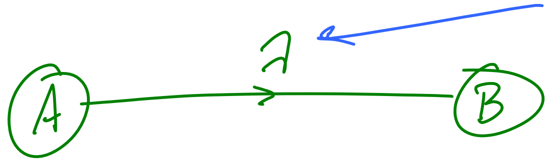
implicit loop



b) Discrete time



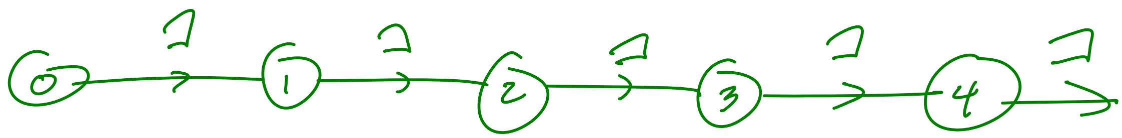
c) Continuous time



jumps to B after exponential clock (parameter λ) rings

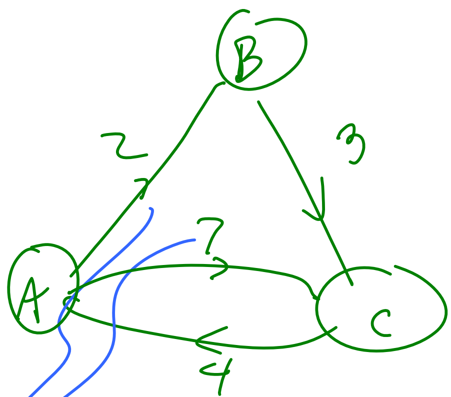
$E(\lambda)$

d) Continuous time



(Poisson process)

e) Continuous time



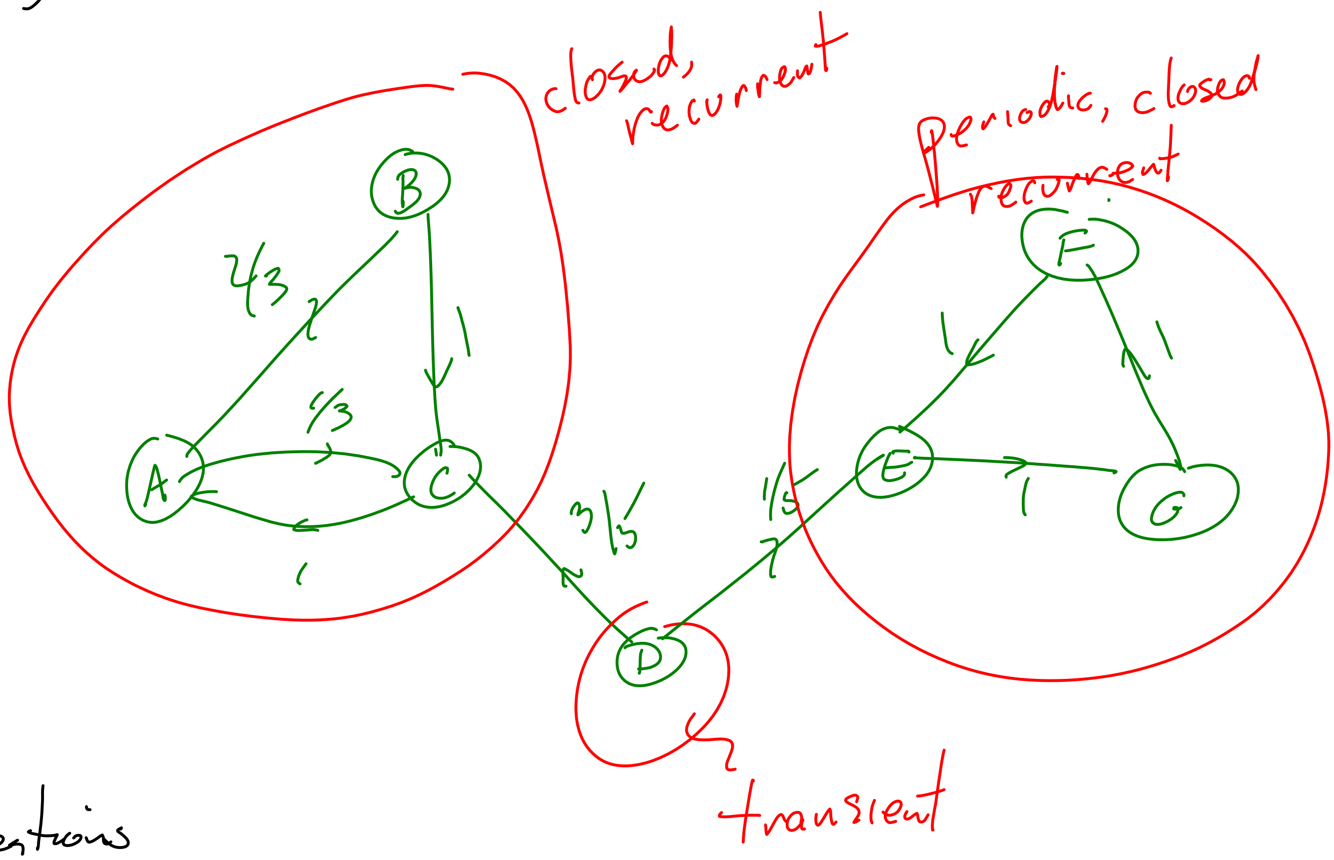
if starting at A: $E(2)$ & $E(7)$

if $E(2)$ rings first, go to B

if $E(7)$ rings first, go to C

More later

f) Discrete time



Questions

start from D: what is the probability of hitting F?

start from C: what is probability of hitting B?

start from C: how long (on average) until you hit B?

start from C: how much time (on average) is spent in state A?

Formal Setup:

Background probability triple $(\Omega, \mathcal{F}, \mathbb{P})$

Ω: event space

$\mathcal{F}$: σ-algebra of subsets of Ω

$\left.\begin{array}{l} \phi \in \mathcal{F} \\ A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F} \\ \{A_n\}_1^\infty \subset \mathcal{F} \Rightarrow \bigcup_{n=1}^\infty A_n \in \mathcal{F} \end{array}\right\}$ measurable space

$\mathbb{P}$: probability measure on $(\Omega, \mathcal{F})$

$\mathbb{P}(\Omega) = 1$

$0 \leq \mathbb{P}(A) \leq 1$ all $A \in \mathcal{F}$

$\{A_n\}_1^\infty \subset \mathcal{F}$ disjoint $\Rightarrow \mathbb{P}\left(\bigcup_{n=1}^\infty A_n\right) = \sum_{n=1}^\infty \mathbb{P}(A_n)$

Recall: - If $(E, \mathcal{E})$ is a second measurable space,

$X : \Omega \to E$ is a random variable if

$\{\omega \in \Omega : X(\omega) \in S\} \in \mathcal{F}$ for all $S \in \mathcal{E}$

- $\{A_n\}_1^n \subset \mathcal{F}$, $\{\alpha_n\}_1^n \subset \mathbb{R}$;

$$\mathbb{E}\Big(\sum_n \alpha_n \chi_{A_n}\Big) \overset{def}{=} \sum_n \alpha_n \mathbb{P}(A_n)$$

$$\chi_A(\omega) \overset{def}{=} \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{else} \end{cases}$$

$A \subset \Omega$

expectation operator defined by taking limits of simple functions.

State Space S (S countable)

Recall: An I-valued stochastic process is a sequence $(X_n)_0^\infty$ of I-valued random variables.

Let's now define our basic objects of study

Definition An I-valued stochastic process is **Markovian** if

$$\mathbb{P}\{X_{n+1}=i_{n+1} / X_0=i_0, X_1=i_1 \cdots X_n=i_n\}$$
$$=\mathbb{P}\{X_{n+1}=i_{n+1} / X_n=i_n\}$$

for all $i_0, i_1 \sim i_n$ such that

$$\mathbb{P}\{X_0=i_0, X_1=i_1 \sim X_n=i_n\}>0 \ \& \ \mathbb{P}\{X_n=i_n\}>0.$$

Definition A Markov process is **time-homogeneous** if $\mathbb{P}\{X_{n+1}=j / X_n=i\}=\mathbb{P}\{X_1=j / X_0=i\}$

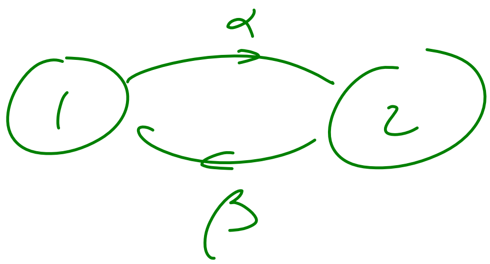
for all j such that $\mathbb{P}\{X_n=i\}>0$ & $\mathbb{P}\{X_0=i\}>0$.

NOTE: We will assume that all Markov processes are time-homogeneous

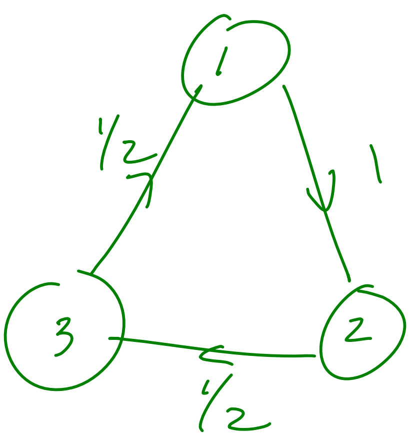
Thus each Markov process defines a stochastic matrix P which satisfies

$$P_{ij} = \mathbb{P}\{X_{n+1}=j \mid X_n = i\} \quad \text{all } i,j\in I,\ n\geq 0$$

such that $\mathbb{P}\{X_n=i\}>0$.



$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$



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

Definition $P = (p_{ij} : i, j \in I)$ is a stochastic matrix if

1. $p_{ij} \geq 0$ all i & j
2. $\sum_j p_{ij} = 1$ all i.

Returning to the inspiration of ODE's, to uniquely solve an ODE, we need the dynamics (a vector field) and an initial condition. For a Markov process, P gives the dynamics. We also need an initial distribution, given by

Definition $\lambda = (\lambda_i ; i \in I)$ is a distribution if

1. $\lambda_i \geq 0$ all $i \in S$
2. $\sum_i \lambda_i = 1$

Definition Let λ be a distribution and let P be a stochastic matrix. We say that a stochastic process X is Markov (λ, P) if

1. $\mathbb{P}\{X_0 = i\} = \lambda_i$ all $i \in I$
2. $\mathbb{P}\{X_{n+1} = i_{n+1} / X_0 = i_0 \cdots X_n = i_n\} = p_{i_n i_{n+1}}$

   all $i_0, i_1 \sim i_{n+1}$ such that

   $\mathbb{P}\{X_0 = i_0, X_1 = i_1 \sim X_n = i_n\} > 0.$

Claim: If X is Markov (λ, P) then X is Markov, in the sense that

$$\mathbb{P}\{X_{n+1} = i_{n+1} / X_0 = i_0 \sim X_n = i_n\} = \mathbb{P}\{X_{n+1} = i_{n+1} / X_n = i_n\}.$$

Pf: We have that

$$\mathbb{P}\{X_{n+1}=i_{n+1} \mid X_n=i_n\} = \frac{\mathbb{P}\{X_{n+1}=i_{n+1}, X_n=i_n\}}{\mathbb{P}\{X_n=i_n\}}$$

$$= \sum_{i_0, i_1 \sim i_{n-1}} \frac{\mathbb{P}\{X_{n+1}=i_{n+1}, X_n=i_n, X_{n-1}=i_{n-1} \cdots X_0=i_0\}}{\mathbb{P}\{X_n=i_n\}}$$

$$= \sum_{i_0, i_1 \sim i_{n-1}} \mathbb{P}\{X_{n+1}=i_{n+1} \mid X_0=i_0 \sim X_n=i_n\} \times \frac{\mathbb{P}\{X_0=i_0 \sim X_n=i_n\}}{\mathbb{P}\{X_n=i_n\}}$$

partition of $\{X_n=i\}$

$$= P_{i_n i_{n+1}} \sum_{i_0 \sim i_{n-1}} \frac{\mathbb{P}\{X_0=i_0 \sim X_n=i_n\}}{\mathbb{P}\{X_n=i_n\}}$$

$$= P_{i_n i_{n+1}} = \mathbb{P}\{X_{n+1}=i_{n+1} \mid X_0=i_0 \sim X_n=i_n\} \qquad \blacksquare$$

Now let's confirm that λ & P uniquely specify X

Claim If X is Markov (λ, P),

$$\mathbb{P}\{X_0 = i_0, X_1 = i_1 \sim X_n = i_n\}$$

$$= \lambda_{i_0} P_{i_0 i_1} \sim P_{i_{n-1} i_n}.$$

Pf Clearly this is true for $n=0$. If it is true for n, then

$$\mathbb{P}\{X_0 = i_0, X_1 = i_1 \sim X_{n+1} = i_{n+1}\}$$

$$= \mathbb{P}\{X_{n+1} = i_{n+1} / X_0 = i_0 \sim X_n = i_n\}$$
$$\times \mathbb{P}\{X_0 = i_0 \sim X_n = i_n\}$$

$$= p_{i_n i_{n+1}} \lambda_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n} \qquad \blacksquare$$

This easily gives us the law of each X_n.

<u>Claim</u> $\quad \mathbb{P}\{X_n = i\} = (\underbrace{\lambda}_{\text{row vector}} P^n)_i$

<u>Pf</u>

$$\mathbb{P}\{X_n = i\} = \sum_{i_0, i_1, \dots, i_{n-1}} \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i}$$

$$= (\lambda P^n)_i \qquad \blacksquare$$

A useful initial distribution is one which puts all its mass at a single point.

Definition δ_i is given by $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for all $j \neq i$ (Dirac distribution at i)

Definition $\mathbb{P}_i$ is the law of an (δ_i, P) Markov process

Let's next show that, like ODE's, Markov processes can be restarted.

Definition If X is a stochastic process and $n \geq 0$, we define $\hat{X}^n_k \stackrel{def}{=} X_{n+k}$ all $k \geq 0$

Claim If $(X_n)_0^\infty$ is Markov(λ, P), then, conditional on $\{X_m = i\}$, $(\hat{X}^m_n)_{n=0}^\infty$ is Markov(δ_i, P) and independent of $(X_0, X_1 \sim X_m)$

Pf

$$\mathbb{P}\{\underbrace{X^m_0 = i_0, X^m_1 = i_1 \sim X^m_n = i_n}_{\text{future}}, \underbrace{X_0 = j_0, X_1 = j_1 \sim X_m = j_m}_{\text{past}} \mid X_m = i\}$$

$$= \mathbb{P}\{X_m = i_0 = i, X_{m+1} = i_1 \sim X_{m+n} = i_n, X_0 = j_0, X_1 = j_1 \sim X_m = j_m\} / \mathbb{P}\{X_m = i\}$$

$$= \frac{\lambda_{j_0} P_{j_0 j_1} \cdots P_{j_{m-1}, j_m} \delta_{j_m, i} \, \delta_{i, i_0} P_{i_0, i_1} P_{i_2 i_3} \sim P_{i_{m-1}, i_m}}{\mathbb{P}\{X_m = i\}}$$

past;

$$= \mathbb{P}\{X_0 = j_0 \sim X_m = j_m \mid X_m = i\}$$

$$= \mathbb{P}\{X_0 = j_0 \sim X_m = j_m \mid X_m = i\}$$

$$\underbrace{\delta_{i,i_0} p_{i_0,i_1} \text{———} p_{i_{m-1},i_m}}_{\text{Markov}(\delta_i, P)};$$

ie

$$\mathbb{P}\{X^m_0 = i_0 \sim X^m_n = i_n, X_0 = j_0 \sim X_m = j_m \mid X_m = i\}$$

$$= \mathbb{P}\{X_0 = j_0 \sim X_{m-1} = j_m \mid X_m = i\} \delta_{i i_0} p_{i_0 i_1} \sim p_{i_{n-1} i_n} \quad ✩$$

Thus,

$$\mathbb{P}\{X^m_0 = i_0, X^m_1 = i_1 \sim X^m_{n-1} = i_{n-1} \mid X_m = i\}$$

$$= \sum_{j_0, j_1 \sim j_m \in I} \mathbb{P}\{X^m_0 = i_0 \sim X^m_{n-1} = i_{n-1}, X_0 = j_0$$

$$\sim X_n = j_m \mid X_m = i\}$$

$$= \sum_{j_0 \sim j_{m-1} \in I} \mathbb{P}\{X_0 = j_0 \sim X_m = j_m \mid X_m = i\}$$

$$\delta_{i i_0} p_{i_0 i_1} \text{———} p_{i_{m-1} i_m}$$

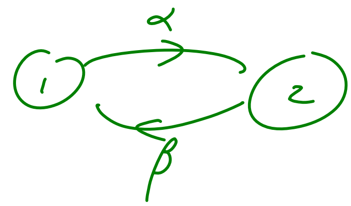
$$= \delta_{i i_0} p_{i_0 i_1} \text{———} p_{i_{m-1} i_m}.$$

Thus, conditional on $\{X_m = i\}$, X^m is Markov (δ_i, P).

Use this in $\star$ to get conditional independence ☒

Two-state Markov chain

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$



Compute $P_{11}^{(n)}$

$$P_{11}^{(n+1)} = \left(\begin{pmatrix} P_{11}^{(n)} & P_{12}^{(n)} \\ P_{21}^{(n)} & P_{22}^{(n)} \end{pmatrix} \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \right)_{11}$$

$$= (1-\alpha) P_{11}^{(n)} + \beta P_{12}^{(n)}$$

Since

$$P_{11}^{(n)} + P_{12}^{(n)} = \mathbb{P}_1\{X_n = 1\} + \mathbb{P}_1\{X_n = 2\} = 1,$$

$$P_{11}^{(n+1)} = (1-\alpha)P_{11}^{(n)} + \beta(1-P_{11}^{(n)})$$

$$\begin{cases} P_{11}^{(n+1)} = (1-\alpha-\beta)P_{11}^{(n)} + \beta \\ P_{11}^{(0)} = \delta_{1,1} \end{cases}$$

If $\alpha+\beta \neq 0$,

$$P_{11}^{(n)} = (1-\alpha-\beta)^n P_{11}^{(0)} + \sum_{j=1}^{n} (1-\alpha-\beta)^{n-j}\beta$$

$$= (1-\alpha-\beta)^n P_{11}^{(0)} + \beta \sum_{j=0}^{n-1} (1-\alpha-\beta)^j$$

$$= (1-\alpha-\beta)^n P_{1,1}^{(0)} + \frac{\beta}{\alpha+\beta}\{1-(1-\alpha-\beta)^n\}$$

$$= \frac{\beta}{\alpha+\beta} + \left(P_{11}^{(0)} - \frac{\beta}{\alpha+\beta}\right)(1-\alpha-\beta)^n$$

$$P_{11}^{(n)} = \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n$$

$$P_{21}^{(n)} = \frac{\beta}{\alpha+\beta} - \frac{\beta}{\alpha+\beta}(1-\alpha-\beta)^n$$

$$P^{(n)}_{12} = \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n$$

$$P^{(n)}_{22} = \frac{\alpha}{\alpha+\beta} + \frac{\beta}{\alpha+\beta}(1-\alpha-\beta)^n$$

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

If $\alpha+\beta=0$, then

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad P^n = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Note If $\alpha+\beta\neq 0$, $P^n \longrightarrow \dfrac{\begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix}}{(\alpha+\beta)}$.

Thus, for any distribution λ,

$$\pi = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right)$$

$$\underbrace{\lambda P^n}_{\text{distribution of } X_n} \xrightarrow{n\to\infty} (\lambda_1\ \lambda_2)\frac{\begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix}}{(\alpha+\beta)} = \frac{(\beta\ \alpha)}{(\alpha+\beta)}$$

$$\lambda P^{n+1} = (\lambda P^n) P$$

$$\pi = \pi P$$

stationary distribution (more later); long-time behavior of distribution of X_n

$$\pi P = 1 \pi$$

π is <u>left</u> eigenvector of P associated with eigenvalue 1; recall that since P is stochastic, $\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}$ is <u>right</u> eigenvector of P associated with eigenvector 1

Alternate Calculation

Eigenvalues of P

$$0 = \det(\lambda I - P) = \det\begin{pmatrix} \lambda-(1-\alpha) & -\alpha \\ -\beta & \lambda-(1-\beta) \end{pmatrix}$$

$$= (\lambda-(1-\alpha))(\lambda-(1-\beta)) - \alpha\beta$$

$$= \lambda^2 - \lambda(2-\alpha-\beta) + (1-\alpha)(1-\beta) - \alpha\beta$$

$$= \lambda^2 - \lambda(2-\alpha-\beta) + 1-\alpha-\beta$$

$$\Rightarrow \lambda = \frac{2-\alpha-\beta \pm \sqrt{(2-\alpha-\beta)^2 - 4(1-\alpha-\beta)}}{2}$$

Note:

$$(2-\alpha-\beta)^2 - 4(1-\alpha-\beta) = (1+1-\alpha-\beta)^2 - 4(1-\alpha-\beta)$$

$$= 1 + 2(1-\alpha-\beta) + (1-\alpha-\beta)^2 - 4(1-\alpha-\beta)$$

$$= 1 - 2(1-\alpha-\beta) + (1-\alpha-\beta)^2$$

$$= (1-(1-\alpha-\beta))^2 = (\alpha+\beta)^2$$

$$\lambda = \frac{2-\alpha-\beta \pm (\alpha+\beta)}{2} = \frac{2-(\alpha+\beta) \pm (\alpha+\beta)}{2}$$

$$= \{1,\ 1-\alpha-\beta\}$$

$$Pe_1 = 1 e_1 \qquad \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \begin{pmatrix} e_{11} \\ e_{21} \end{pmatrix} = \begin{pmatrix} e_{11} \\ e_{21} \end{pmatrix}$$

$$(1-\alpha)e_{11} + \alpha e_{21} = e_{11}$$

$$-\alpha e_{11} + \alpha e_{21} = 0$$

$$e_{11} = e_{21} \ ; \quad e_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$Pe_2 = (1-\alpha-\beta) e_2 \qquad \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \begin{pmatrix} e_{12} \\ e_{22} \end{pmatrix} = (1-\alpha-\beta) \begin{pmatrix} e_{12} \\ e_{22} \end{pmatrix}$$

$$(1-\alpha)e_{12} + \alpha e_{22} = (1-\alpha-\beta) e_{12}$$

$$\alpha e_{12} = -\beta e_{22} \qquad e_{12} = -\frac{\alpha}{\beta} e_{22}$$

$$e_2 = \begin{pmatrix} -\alpha \\ \beta \end{pmatrix}$$

$$E = [e_1 : e_2] = \begin{pmatrix} 1 & -\alpha \\ 1 & \beta \end{pmatrix}$$

$$P[e_1 : e_2] = [e_1 : e_2] \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1-\alpha-\beta \end{pmatrix}}_{\Lambda}$$

$$PE = E\Lambda$$

E invertible (eigenvectors for different eigenvalues are orthogonal)

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

$$E^{-1} = \frac{\begin{pmatrix} \beta & \alpha \\ -1 & 1 \end{pmatrix}}{(\alpha+\beta)}$$

$$P^n = E \Lambda^n E^{-1} = \begin{pmatrix} 1 & -\alpha \\ 1 & \beta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (1-\alpha-\beta)^n \end{pmatrix} \frac{\begin{pmatrix} \beta & \alpha \\ -1 & 1 \end{pmatrix}}{\alpha+\beta}$$

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

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

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

Alternate Calculation

Suppose we want to compute only $P_{11}^{(n)}$.

From eigenvalue analysis, we know that

$$P_{11}^{(n)} = c_1 + c_2(1-\alpha-\beta)^n$$

We also know $$1 = P_{11}^{(0)} = c_1 + c_2$$

$$1-\alpha = P_{11}^{(1)} = c_1 + c_2(1-\alpha-\beta)$$

Solve simultaneously

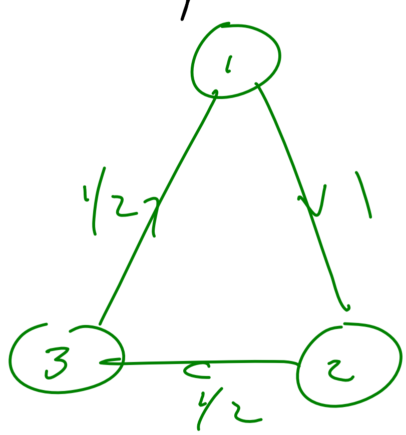
$$\alpha = c_2 - c_2(1-\alpha-\beta) = c_2(\alpha+\beta)$$

$$c_2 = \frac{\alpha}{\alpha+\beta}$$

$$c_1 = \frac{\beta}{\alpha+\beta}$$

$$P_{11}^{(n)} = \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n$$

Another example



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

Problem: find $P_{11}^{(n)}$.

① eigenvalues of P

$$0 = \det(\lambda I - P) = \det\begin{pmatrix} \lambda & -1 & 0 \\ 0 & \lambda - 1/2 & -1/2 \\ -1/2 & 0 & \lambda - 1/2 \end{pmatrix}$$

$$= \lambda(\lambda - 1/2)^2 + 1(-1/4)$$

$$= \lambda^3 - \lambda^2 + \lambda/4 - 1/4$$

Note: Since P is a stochastic matrix,

$$P\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}1\\1\\1\end{pmatrix}$$

$\lambda = 1$ is an eigenvalue $\longleftarrow$ Always true

Synthetic division

$$\begin{array}{r|l} & \lambda^2 + 0\lambda + 1/4 \\ \hline \lambda - 1 & \lambda^3 - \lambda^2 + \lambda/4 - 1/4 \\ & \underline{\lambda^3 - \lambda^2} \quad \underline{\lambda/4 - 1/4} \\ & \quad\ 0 \qquad\quad\ \ 0 \end{array}$$

$$0 = \lambda^3 - \lambda^2 + \lambda/4 - 1/4 = (\lambda - 1)(\lambda^2 + 1/4)$$

eigenvalues are: $\lambda = 1$, $\lambda = i/2$, $\lambda = -i/2$.

$i = \sqrt{-1}$

general form of $P_{11}^{(n)}$ is

$$P_{11}^{(n)} = a + b\left(\frac{i}{2}\right)^n + c\left(\frac{-i}{2}\right)^n$$

complex eigenvalues must appear in pairs of conjugates

Note: $(i/2)^n = \left(\frac{1}{2}\right)^n i^n = \left(\frac{1}{2}\right)^n$

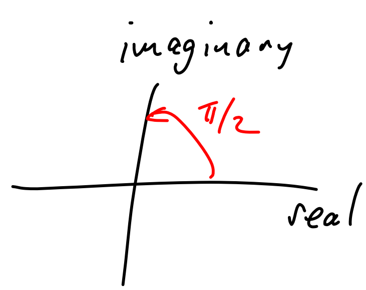
$(-i/2)^n = \left(\frac{1}{2}\right)^n (-i)^n$

$$P_{11}^{(n)} = a + b\left(\frac{1}{2}\right)^n i^n + c\left(\frac{1}{2}\right)^n (-i)^n$$

$$= a + \left(\frac{1}{2}\right)^n \{ b i^n + c(-i)^n \}$$

Note: $e^{i\theta} = \cos\theta + i\sin\theta$

$$i = e^{i\pi/2} = \underbrace{\cos \pi/2}_{0} + i\underbrace{\sin \pi/2}_{1}$$



$$-i = e^{-i\pi/2} = \underbrace{\cos(-\pi/2)}_{0} + i\underbrace{\sin(-\pi/2)}_{-1}$$

$$i^n = e^{n\pi i/2} = \cos n\pi/2 + i \sin n\pi/2$$

$$(-i)^n = e^{-n\pi i/2} = \cos(-n\pi/2) + i \sin(-n\pi/2)$$

$$= \cos n\pi/2 - i \sin n\pi/2$$

$$P_{11}^{(n)} = a + \left(\tfrac{1}{2}\right)^n \{ b \cos n\pi/2 + b i \sin n\pi/2$$

$$+ c \cos n\pi/2 - c i \sin n\pi/2 \}$$

b, c complex

$$= a + \left(\tfrac{1}{2}\right)^n \{ \underbrace{(b+c)}_{\hat{b}} \cos n\pi/2 + \underbrace{(b-c) i}_{\hat{c}} \sin n\pi/2 \}$$

$$= \hat{a} + \left(\tfrac{1}{2}\right)^n \{ \hat{b} \cos n\pi/2 + \hat{c} \sin n\pi/2 \}$$

We know:

$$P_{11}^{(0)} = 1$$

$$P_{11}^{(1)} = 0$$

$$P_{11}^{(2)} = \left(\begin{pmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix} \right)_{11}$$

$$= 0$$

$$1 = P_{11}^{(0)} = \hat{a} + \hat{b} \qquad (1)$$

$$0 = P_{11}^{(1)} = \hat{a} + \frac{1}{2}\hat{c} \qquad (2)$$

$$0 = P_{11}^{(2)} = \hat{a} - \frac{1}{4}\hat{b} \qquad (3)$$

Simultaneously solve:

$$\hat{a} = 1 - \hat{b} \qquad (1)$$

$$0 = 1 - \hat{b} - \frac{1}{4}\hat{b} \qquad (3)$$

$$= 1 - \frac{5}{4}\hat{b}$$

$$\hat{b} = \frac{4}{5}$$

$$\hat{a} = 1 - \frac{4}{5}$$

$$\hat{a} = \frac{1}{5}$$

$$\hat{c} = -2\hat{a} = -2/5$$

$$\hat{c} = -2/5$$

$$P_{11}^{(n)} = \frac{1}{5} + \left(\frac{1}{2}\right)^n \left\{ \frac{4}{5} \cos n\pi/2 - \frac{2}{5} \sin n\pi/2 \right\}$$

Repeated eigenvalues Repeated eigenvalues can occur. Canonical form is either

$$\Lambda_1 = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \quad \text{or} \quad \Lambda_2 = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

distinction is geometric.

For Λ_1, eigenvectors are $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ & $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

eigenvectors span $\mathbb{R}^2$.

$$\Lambda_1^n = \begin{pmatrix} \lambda^n & 0 \\ 0 & \lambda^n \end{pmatrix};$$

general form is $\alpha \lambda^n$

For Λ_2, only eigenvector is $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

$e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is generalized eigenvector;

$Pe_2 = \lambda e_2 + e_1.$

$$A_1^2 = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{pmatrix}$$

$$A_1^3 = \begin{pmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{pmatrix}\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} \lambda^3 & 3\lambda^2 \\ 0 & \lambda^3 \end{pmatrix}$$

$$A_1^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix}$$

general form is

$$a\lambda^n + bn\lambda^{n-1} = \lambda^n\left(a + \frac{b}{\lambda}n\right)$$

$$= \lambda^n(\hat{a} + \hat{b}n)$$

<u>Thus</u>: If λ is a double eigenvalue, include terms of form $\lambda^n(a+bn)$

Example $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

$$0 = \det(\lambda I - P) = \det \begin{pmatrix} \lambda - 1/2 & -1/2 \\ -1/2 & \lambda - 1/2 \end{pmatrix}$$

$$= (\lambda - 1/2)^2 - \frac{1}{4} = \lambda^2 - \lambda = \lambda(\lambda - 1)$$

$\lambda = \{0, 1\}$ ~ not surprising; 1 must be an eigenvalue, and since P is not full rank, 0 must also be an eigenvalue

$$P^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$P^2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = P$$

$$P^n = P$$

not surprising; independent coin flips

$$P_{11}^{(n)} = a + b\,0^n = \begin{cases} a + b & \text{if } n = 0 \\ a & \text{if } n \geq 1 \end{cases}$$

$$\left.\begin{aligned} 1 &= P_{11}^{(0)} = a + b \\ \frac{1}{2} &= P_{11}^{(1)} = a \end{aligned}\right\} \Longrightarrow \begin{aligned} a &= \frac{1}{2} \\ b &= \frac{1}{2} \end{aligned}$$

$$P_n^{(n)} = \begin{cases} 1 & \text{if } n = 0 \\ \frac{1}{2} & \text{if } n \geq 1 \end{cases}$$