The Lecture Contains :

- Markov-Bernoulli Chain
- Note
- Assignments
- Random Walks which are correlated
- Actual examples of Markov Chains
- Examples
- Use of Eigen values and eigen vectors to calculate higher transition probabilities
- Theorem
- Simulated Annealing (based on Metropolis Algorithm)
- Concept of Markov Chain Monte Carlo (MCMC) method

Previous Next

Economics and finance : In finance and economics Markov chains are used to model a variety of different phenomena, including asset prices and market crashes. Hamilton was the first person to use this methodology successfully in finance where he found the conditional probability of regime switching models or change points such that depending on the states of the process and the corresponding transition probability values one can find the probability that the asset/option prices can be forecasted/predicted with a high degree of accuracy.

Mathematical biology : Applications in biological modeling utilize Markov modeling where a particularly population and its off springs can be models as Markov chain states and the probability that the off springs survives is given by the corresponding transition probability matrix.

Gambling : Markov chains can be used to model many games of chance such as the well know game of snakes and ladders

Concept of Eigen vector and Eigen value

Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $A^{-1} = \frac{1}{|A|}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. In that case the Eigen vectors (characteristics equation) is given by $det(A - \lambda I) = 0$

If $\tilde{q}$ is an eigen vector then we have the following set of equations given by:

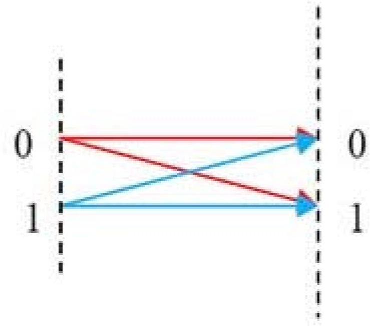
$\left.\begin{matrix} A\tilde{q}_1 = \lambda_1\tilde{q}_1 \\ A\tilde{q}_2 = \lambda_2\tilde{q}_2 \\ . \\ . \end{matrix}\right\}$, which means that $AQ = Q\Lambda$, where $\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & \cdots & \cdots & \vdots \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$.

Now as $AQ = Q\Lambda$, this implies that $A = Q\Lambda Q^{-1}$, i.e., $A \times A = (Q\Lambda Q^{-1}) \times (Q\Lambda Q^{-1}) = Q\Lambda^2 Q^{-1}$. Moreover $\Lambda^2 = \begin{bmatrix} \lambda_1^2 & 0 & 0 & 0 \\ 0 & \lambda_2^2 & 0 & 0 \\ \vdots & \cdots & \cdots & \vdots \\ 0 & 0 & 0 & \lambda_n^2 \end{bmatrix}$, hence using simple calculation we can immediately find out that $A^n = Q\Lambda^n Q^{-1}$. Thus if A is the transition matrix then A^n is the n-step transition matrix.

Previous Next

Example 1.28

Assume $P=\begin{bmatrix}\frac{8}{10} & \frac{2}{10}\\ \frac{5}{10} & \frac{5}{10}\end{bmatrix}$, for which we have the following transition diagram



What is of interest to us is the eigen value of $P=\begin{bmatrix}\frac{8}{10} & \frac{2}{10}\\ \frac{5}{10} & \frac{5}{10}\end{bmatrix}$. Using basic concepts we have det(P-I)=0,

i.e., $det\begin{bmatrix}\frac{8}{10}-\lambda & \frac{2}{10}\\ \frac{5}{10} & \frac{5}{10}-\lambda\end{bmatrix}=0$, hence $\lambda=1\ and\ \frac{3}{10}$, which implies $\lambda_1=1, q_1=\begin{bmatrix}1\\1\end{bmatrix}$ and

$\lambda_2=\frac{3}{10}, q_1=\begin{bmatrix}\frac{2}{10}\\ -\frac{5}{10}\end{bmatrix}$, i.e., for $\forall i,\ \lambda_i \le 1$, all eigen values are =1.

In case we are interested to find $P^n=Q\Lambda^n Q^{-1}$, we have $P^n=\begin{bmatrix}\frac{5}{7} & \frac{2}{7}\\ \frac{5}{7} & \frac{2}{7}\end{bmatrix}+\left(\frac{3}{10}\right)^n\begin{bmatrix}\frac{2}{7} & -\frac{2}{7}\\ -\frac{5}{7} & \frac{5}{7}\end{bmatrix}$ and as

$$n\rightarrow\infty\quad \lim_{n\rightarrow\infty}P^n=\lim_{n\rightarrow\infty}\left\{\begin{bmatrix}\frac{5}{7} & \frac{2}{7}\\ \frac{5}{7} & \frac{2}{7}\end{bmatrix}+\left(\frac{3}{10}\right)^n\begin{bmatrix}\frac{2}{7} & -\frac{2}{7}\\ -\frac{5}{7} & \frac{5}{7}\end{bmatrix}\right\}=\begin{bmatrix}\frac{5}{7} & \frac{2}{7}\\ \frac{5}{7} & \frac{2}{7}\end{bmatrix}.$$

Suppose we are at step 0, and we may be interested to find the expected time until we return to step 0. Hence if T is the time until first return, then

$P[T=0]=0$, $P[T=1]=\frac{8}{10}$, $P[T=2]=\frac{2}{10}\times\frac{5}{10}$, $P[T=3]=\frac{2}{10}\times\frac{5}{10}\times\frac{5}{10}$

$$P[T>0]=1-P[T=0]=1-0=1$$

$$P[T>1]=1-P[T=1]-P[T=0]=1-\frac{8}{10}-0=\frac{2}{10}$$

$$P[T>2]=1-P[T=2]-P[T=1]-P[T=1]=1-\frac{2}{10}\times\frac{5}{10}-\frac{8}{10}-0=\frac{1}{10}$$

$$P[T>3]=1-P[T=3]-P[T=2]-P[T=1]-P[T=0]=1-\frac{2}{10}\times\frac{5}{10}\times\frac{5}{10}-\frac{2}{10}\times\frac{5}{10}-\frac{8}{10}-0=\frac{1}{20}$$

Thus $E(T)=\sum_{n=0}^{\infty}P(T>n)=1+\frac{2}{10}+\frac{2}{10}\times\frac{1}{2}+\frac{2}{10}\times\left(\frac{1}{2}\right)^2+\cdots=\frac{7}{5}$

Note in case you start at state 1, then the expected time until return is 3.5.

Use of Eigen values and eigen vectors to calculate higher transition probabilities

Suppose you have a Markov chain with m states, where m is finite, and it is also given that the transition probability matrix is $P=\left(p_{ij}\right)=\left\|p_{ij}\right\|$, then one can very easily calculate the n step transition probability matrix, $P^n=\left(p_{ij}^{(n)}\right)=\left\|p_{ij}^{(n)}\right\|$, such that we utilize the equation, $p_{ij}^{(n)}=\sum_{\forall k} p_{ik}^{(1)}\times p_{kj}^{(n-1)}$, $n=1,2,\ldots$, where $p_{jj}^{(0)}=1$ and $p_{jk}^{(0)}=0$ for $k\neq j$. Now our main intention of this discussion is to utilize the concept of eigen values and eigen vectors to calculate $P^n=\left(p_{ij}^{(n)}\right)=\left\|p_{ij}^{(n)}\right\|$.

One must remember that for a square matrix, $A_{(m\times m)}=\left\|a_{ij}\right\|_{(m\times m)}$, the characteristics roots of the equation: $|A-\lambda_i I|=\begin{vmatrix} a_{11}-\lambda_i & a_{12} & \cdots & a_{1,m-1} & a_{1m} \\ a_{21} & a_{22}-\lambda_i & \cdots & a_{2,m-1} & a_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m-1,1} & a_{m-1,2} & \cdots & a_{m-1,m-1}-\lambda_i & a_{m-1,m} \\ a_{m,1} & a_{m,2} & \cdots & a_{m,m-1} & a_{m,m}-\lambda_i \end{vmatrix}=0$, where $i=1,2,\ldots,m$ are called the eigen values, and they are given by $\lambda_1,\lambda_2,\ldots,\lambda_m$. While $X_i=\{x_{i,1},\ldots,x_{i,m}\}^T$ / $Y_i=\{y_{i,1},\ldots,y_{i,m}\}$, $i=1,2,\ldots,m$, is the right/left eigen vector, corresponding to λ_i, $i=1,2,\ldots,m$, such that the following equation (for X_i), s, i.e., $A_{m\times m}X_{i(m\times 1)}-\lambda_i I_{(m\times m)}X_{i(m\times 1)}=0_{(m\times 1)}$, $\left(A-\lambda_i I\right)_{(m\times m)}X_{i(m\times 1)}=0_{(m\times 1)}$, i.e.,

$$\begin{pmatrix} a_{11}-\lambda_i & a_{12} & \cdots & a_{1,m-1} & a_{1m} \\ a_{21} & a_{22}-\lambda_i & \cdots & a_{2,m-1} & a_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m-1,1} & a_{m-1,2} & \cdots & a_{m-1,m-1}-\lambda_i & a_{m-1,m} \\ a_{m,1} & a_{m,2} & \cdots & a_{m,m-1} & a_{m,m}-\lambda_i \end{pmatrix}\begin{pmatrix} x_{i,1} \\ x_{i,2} \\ \vdots \\ x_{i,m-1} \\ x_{i,m} \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$ holds for $i=1,2,\ldots,m$, such that we will finally obtain

$$X_{(m\times m)}=\begin{pmatrix} x_{1,1} & x_{2,1} & \cdots & x_{m-1,1} & x_{m,1} \\ x_{1,2} & x_{2,2} & \cdots & x_{m-1,2} & x_{m,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{1,m-1} & x_{2,m-1} & \cdots & x_{m-1,m-1} & x_{m,m-1} \\ x_{1,m} & x_{2,m} & \cdots & x_{m-1,m} & x_{m,m} \end{pmatrix}$$, where the i^{th} column corresponds to the eigen vector corresponding to the i^{th} characteristics root, λ_i.

Previous Next

Module 1:Concepts of Random walks, Markov Chains, Markov Processes

Lecture 4:Markov Process

Now suppose $\lambda_1, \lambda_2, \ldots, \lambda_m$ are the m eigen values or the characteristics roots of P and also assume that $X_{(m\times m)} = \begin{pmatrix} x_{1,1} & x_{2,1} & \ldots & x_{m-1,1} & x_{m,1} \\ x_{1,2} & x_{2,2} & \ldots & x_{m-1,2} & x_{m,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{1,m-1} & x_{2,m-1} & \ldots & x_{m-1,m-1} & x_{m,m-1} \\ x_{1,m} & x_{2,m} & \ldots & x_{m-1,m} & x_{m,m} \end{pmatrix}$ and

$Y_{(m\times m)} = \begin{pmatrix} y_{1,1} & y_{2,1} & \ldots & y_{m-1,1} & y_{m,1} \\ y_{1,2} & y_{2,2} & \ldots & y_{m-1,2} & y_{m,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{1,m-1} & y_{2,m-1} & \ldots & y_{m-1,m-1} & y_{m,m-1} \\ y_{1,m} & y_{2,m} & \ldots & y_{m-1,m} & y_{m,m} \end{pmatrix}$ are the *right* and *left* eigen matrix of P

respectively, then we can easily write: $P = XDX^{-1}$, where $D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ldots & \vdots \\ 0 & 0 & \cdots & \lambda_{m-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_m \end{pmatrix}$

and $P^n = XD^nX^{-1}$, such that

$P^n = \sum_{k=1}^{m} \lambda_k^n \left\{ \frac{1}{(y_{k,1}, \ldots, y_{k,m}) \times \begin{pmatrix} x_{k,1} \\ \vdots \\ x_{k,m} \end{pmatrix}} \right\} \times \begin{pmatrix} x_{k,1} \\ \vdots \\ x_{k,m} \end{pmatrix} \times (y_{k,1}, \ldots, y_{k,m})$, where $\lambda_k^n \times \left\{ \frac{1}{(y_{k,1}, \ldots, y_{k,m}) \times \begin{pmatrix} x_{k,1} \\ \vdots \\ x_{k,m} \end{pmatrix}} \right\}$ is a scalar term, while $\begin{pmatrix} x_{k,1} \\ \vdots \\ x_{k,m} \end{pmatrix} \times (y_{k,1}, \ldots, y_{k,m})$ is a $m \times m$ matrix, such that the elements of

$P^n = \left(p_{ij}^{(n)}\right) = \left\| p_{ij}^{(n)} \right\|$ are given by $p_{ij}^{(n)} = \sum_{k=1}^{m} \lambda_k^n \times \left\{ \frac{1}{(y_{k,1}, \ldots, y_{k,m}) \times \begin{pmatrix} x_{k,1} \\ \vdots \\ x_{k,m} \end{pmatrix}} \right\} \times x_{k,i} \times y_{k,j}$, where we have

the matrix $\begin{pmatrix} x_{k,1} \\ \vdots \\ x_{k,m} \end{pmatrix} \times (y_{k,1}, \ldots, y_{k,m}) = \begin{pmatrix} x_{k,1}y_{k,1} & \cdots & x_{k,1}y_{k,m} \\ \vdots & \ddots & \vdots \\ x_{k,m}y_{k,1} & \cdots & x_{k,m}y_{k,m} \end{pmatrix}_{(m\times m)}$

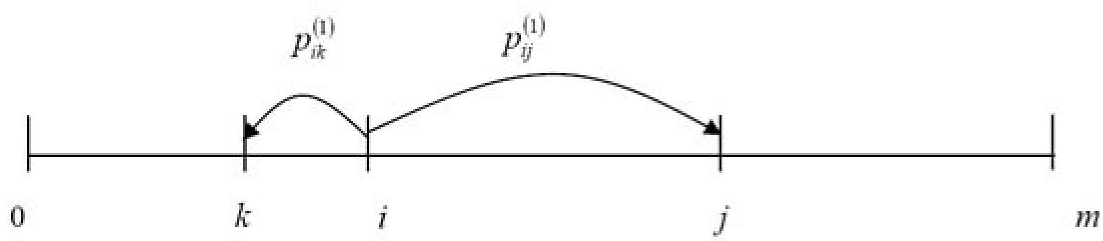
Previous Next

Assignment 1.11

If P is the transition probability matrix, while X is the eigen matrix, then show that we can express $P = XDX^{-1}$, and also prove that $P^n = XD^nX^{-1}$.

Finite irreducible Markov chain

Stationary distribution: If for a Markov chain with a transition probability matrix, $P = \begin{pmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & \ddots & \vdots \\ p_{m1} & \cdots & p_{mm} \end{pmatrix}_{(m\times m)}$ $v_j \geq 0$ and $\sum_{\forall j} v_j = 1$ holds, where v_j is a probability distribution such that in general we have $v_k = \sum_{\forall i} v_i p_{ik}^{(n)}$, $n \geq 1$, then what is of interest to us is to study the initial conditions for which as $n \to \infty$, $p_{ik}^{(n)}$ has a limiting value such that the initial state from which the stochastic process starts does not affect the value, i.e., $p_{ik}^{(n)}$ is independent of the initial state from where we start. This would very simple mean that from where ever we start the limiting values of P would have identical rows. This property is know as the property of ***ergodicity*** and the Markov chain is called ***ergodic***.



From $p_{ik}^{(1)}$and $p_{ij}^{(1)}$ and for every state we find $p_{ij}^{(1)}$, $j = 1,2,\ldots,m$, but $j \neq i$ as when $j = i$ we have the ***initial condition***.

Theorem 1.8

If state j is ***persistent***, then from every state k, we can reach j.

Previous Next

Theorem 1.9 [Ergodic theorem]

For every ***finite irreducible***, ***aperiodic*** Markov chain, with transition probability matrix, $P=\left(p_{ij}^{(1)}\right)$ we would have $\lim_{n\to\infty} p_{ik}^{(n)}=v_k$, where $v_k=\sum_{\forall i} v_i p_{ik}^{(n)}$ and also remember this ***limiting distribution*** is ***equal*** to the ***stationary distribution*** of the Markov chain.

Proof of Theorem 1.9 [Ergodic theorem]

Now for a Markov chain in which the states are ***aperiodic***, ***persistent non-null***, then for every pair, i,k we have $\lim_{n\to\infty} p_{ik}^{(n)}=\left(\frac{F_{ik}}{\mu_{kk}}\right)$, and as k is persistent, hence $F_{ik}=1$, which means that $\lim_{n\to\infty} p_{ik}^{(n)}=\left(\frac{1}{\mu_{kk}}\right)>0$ and is independent of i. Also the sums of rows should add up to 1 (in the limiting sense), i.e., $\sum_{k=1}^{N} p_{ik}^{(n)}\le 1$, where set $k=1,\ldots,N$, such that $\sum_{k=1}^{N}\lim_{n\to\infty} p_{ik}^{(n)}=\sum_{k=1}^{N}\left(\frac{1}{\mu_{kk}}\right)\le 1$, such that $\sum_{k=1}^{N} v_k\le 1$.

Moreover it is known that

(i) $p_{ik}^{(m+n)}=\sum_{\forall j} p_{ij}^{(m)}\times p_{jk}^{(n)}$

(ii) $\lim_{m+n\to\infty} p_{ik}^{(m+n)}=v_k$ are both true, hence we can write (ii) using (i) as

$$v_k=\lim_{m+n\to\infty}\sum_{\forall j} p_{ij}^{(m)}\times p_{jk}^{(n)}$$

$v_k\ge\sum_{\forall j}\left(\lim_{m+n\to\infty} p_{ij}^{(m)}\right)\times p_{jk}^{(n)}$, this we get from Fatou's Lemma, which says that for a sequence of some type of measurable function the integral (sum) of the limit of the infimum of a function is less than or equal to the limit of the infimum of the integral (sum) of the function, i.e., $\int \lim_{n\to\infty}\inf f_n(s)ds\le\lim_{n\to\infty}\inf\int f_n(s)ds$.

Example for which we can check this is (i) , (ii) $f(x)=\begin{cases}0 & x\in[-n,n]\\ 1 & otherwise\end{cases}$, where n is a natural number, (ii) $f(x)=\begin{cases}n & x\in\left(0,\frac{1}{n}\right)\\ 0 & otherwise\end{cases}$ where n is a natural number, (iii) $f(x)=\begin{cases}n & x\in\left(0,\frac{1}{n}\right)\\ 0 & otherwise\end{cases}$ where n is a real number. Just for your convenience we state the greater than inequality when we have $\int \lim_{n\to\infty}\sup f_n(s)ds\ge\lim_{n\to\infty}\sup\int f_n(s)ds$, which is the reverse of Fatou's Lemma.

Thus we have

$v_k \geq \sum_{\forall j} v_j \times p_{jk}^{(n)}$ for all n

Now if the greater than sign ***only*** holds, which would mean that we have $v_k \geq \sum_{\forall j} v_j \times p_{jk}^{(n)}$ for all n, i.e., $\sum_{\forall k} v_k > \sum_{\forall k}\sum_{\forall j} v_j \times p_{jk}^{(n)}$ for all n would hold true. But ask yourself is that possible.

Remember that sum of the pdf/pmf values at ***each point*** which is the left hand side is stated to be greater that the cumulative probability values (pdf/pmf) summed up after summing them up at teach point. But that is not possible!! Why? Ask yourself.

Hence equality has to hold. Remember the distribution which we get in the limiting case is the ***stationary distribution***.

Remember that if a Markov chain is ***irreducible*** and ***aperiodic*** and if there exits a unique ***stationary*** distribution for the Markov chain, the chain is ergodic and $v_k = \left(\frac{1}{\mu_{kk}}\right)$

Markov-Bernoulli Chain

Consider the transition probability matrix of two states as given by $\begin{bmatrix} 1-(1-c)p & (1-c)p \\ (1-c)(1-p) & (1-c)p+c \end{bmatrix}$,

with the initial distribution $P[X_0 = 0] = p_1$ and $P[X_0 = 1] = (1-p_1)$

(i) For $c = 1$, the transition probability matrix is: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, what does it mean?

(ii) For $c = 0$, the transition probability matrix is: $\begin{bmatrix} 1-p & p \\ 1-p & p \end{bmatrix}$, what does it mean?

(iii) For $0 < c < 1$, again use the same fundamental principle where we can write $p_n = p_{n-1}[(1-c)p + c] + q_{n-1}[(1-c)p]$, where $p_n = P[X_n = 1]$ and $q_n = P[X_n = 0] = 1 - p_n$. Here we can easily prove that $p_n = p$ and so for $q_n = 1 - p_n = (1-p)$.

From this above result we easily get the following

(i) $E(X_n) = \sum_{i=0}^{1}\{i \times P(X_n = i)\} = [0 \times (1-p)] + [1 \times p] = p$, in general the formulae would be $E(X_n) = \sum_{\forall i}\{i \times P(X_n = i)\}$, depending on the number of states, i.e., we have a multinomial distribution.

(ii) $V(X_n) = \sum_{i=0}^{1}\{i - E(X_n)\}^2 \times P(X_n = i) = (0-p)^2(1-p) + (1-p)^2 p = (1-p)\{p^2 + p - p^2\}$,

in general the formulae would be $V(X_n) = \sum_{\forall i}\{i - E(X_n)\}^2 \times P(X_n = i)$, depending on the number of states, i.e.,we have a multinomial distribution.

(iii) $E(X_{n-1}X_n) = \sum_{i,j=0}^{1}(ij) \times P(X_n = j, X_{n-1} = i)$

$$= \sum_{i,j=0}^{1}(ij) \times P(X_n = j | X_{n-1} = i) \times P(X_{n-1} = i) = \{(1-c)p + c\}p$$

Now we already know that

$$\operatorname{cov}(X_{n-1}, X_n) = E[\{X_{n-1} - E(X_{n-1})\}\{X_n - E(X_n)\}]$$

$$= E(X_{n-1}X_n) - E\{X_{n-1}E(X_n)\}$$

$$- E\{X_n E(X_{n-1})\} + E\{E(X_{n-1})E(X_n)\}$$

$\operatorname{cov}(X_{n-1}, X_n) = p^2 - cp^2 + cp - p^2 - p^2 + p^2 = cp(1-p)$, hence $corr(X_{n-1}, X_n) = c$

If we extend this calculation for $\operatorname{cov}(X_{n-2}, X_n), \operatorname{cov}(X_{n-3}, X_n), \ldots$ we can easily see that

(iv) $\operatorname{cov}(X_{n-k}, X_n) = c^k p(1-p)$ and $corr(X_{n-k}, X_n) = c^k$ for $k \geq 1$

Previous Next

Note

In case if one is interested to find the average, variance and covariance of S_n where $S_n = \sum_{i=1}^{n} X_i$, then we need to be careful as these trials are now dependent, unlike the simple case we already know.
So now we have the following

(i) $E(S_n) = E\left(\sum_{i=1}^{n} X_i\right) = np$

(ii) $V(S_n) = V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{i,j,i<j,i=1}^{n} \mathrm{cov}(X_i, X_j)$

$$V(S_n) = \underbrace{p(1-p) + \ldots + p(1-p)}_{n\ times} + 2 \sum_{i,j,i<j,i=1}^{n} \mathrm{cov}(X_i, X_j)$$

Now for the second term let us consider it separately, such that we have

$$2 \sum_{i,j,i<j,i=1}^{n} \mathrm{cov}(X_i, X_j) = 2 \sum_{i,j,i<j,i=1}^{n} \left\{corr(X_i, X_j) \times \sqrt{V(X_i)} \times \sqrt{V(X_j)}\right\}$$

$$2 \sum_{i,j,i<j,i=1}^{n} \left\{\frac{\mathrm{cov}(X_i, X_j)}{\sqrt{V(X_i)} \times \sqrt{V(X_j)}}\right\} = 2 \sum_{i,j,i<j,i=1}^{n} corr(X_i, X_j) = 2\{c + (c + c^2) + \mathrm{K} + (c + \mathrm{K} + c^{n-2}) + (c + \mathrm{K} + c^{n-1})\}$$

$$2 \sum_{i,j,i<j,i=1}^{n} \left\{\frac{\mathrm{cov}(X_i, X_j)}{\sqrt{V(X_i)} \times \sqrt{V(X_j)}}\right\} = 2 \times \frac{c}{(1-c)}\left\{(n-1) - c\frac{(1-c^{n-1})}{(1-c)}\right\}$$ (use GP to find this summation series)

$$2 \sum_{i,j,i<j,i=1}^{n} \{\mathrm{cov}(X_i, X_j)\} = 2 \times \sqrt{V(X_i)} \times \sqrt{V(X_j)} \times \frac{c}{(1-c)}\left\{(n-1) - c\frac{(1-c^{n-1})}{(1-c)}\right\}$$

$$2 \sum_{i,j,i<j,i=1}^{n} \{\mathrm{cov}(X_i, X_j)\} = 2 \times \sqrt{p(1-p)} \times \sqrt{p(1-p)} \times \frac{c}{(1-c)}\left\{(n-1) - c\frac{(1-c^{n-1})}{(1-c)}\right\}$$

$$2 \sum_{i,j,i<j,i=1}^{n} \{\mathrm{cov}(X_i, X_j)\} = 2 \times p(1-p) \times \frac{c}{(1-c)}\left\{(n-1) - c\frac{(1-c^{n-1})}{(1-c)}\right\}$$

Hence utilizing these two results we have

$$V(S_n) = \underbrace{p(1-p) + \ldots + p(1-p)}_{n\ times} + 2 \times p(1-p) \times \frac{c}{(1-c)}\left\{(n-1) - c\frac{(1-c^{n-1})}{(1-c)}\right\}$$

$$= np(1-p) + 2 \times p(1-p) \times \frac{c}{(1-c)}\left\{(n-1) - c\frac{(1-c^{n-1})}{(1-c)}\right\}$$

Now remember the relationship between Binomial distribution when $n \to \infty$ and $p \to 0$, such that np is a constant say, λ.

Note: Hence with, $c = 0$ we have the sequence of independent Bernoullis trails and in the limiting case we get the Poisson process

Previous Next

Assignment 1.6

Consider the same problem as given above and consider it having three (3) outcomes, such that the outcomes of any trails are dependent on the outcomes of the previous trial. The transition probability matrix is considered as $\begin{bmatrix} 1-(1-c) & (1-c)p & (1-c)(1-p) \\ 1-(1-c)(1-p) & (1-c)(1-p)p & (1-c)(1-p)^2 \\ 1-(1-c)(1-p)^2 & (1-c)(1-p)^2 p & (1-c)(1-p)^3 \end{bmatrix}$. We also know that there are three states given by 0, 1, 2, such that $P(X_0=0)=\alpha$, $P(X_0=1)=\beta$ and $P(X_0=2)=\gamma$, and $\alpha+\beta+\gamma=1$. With this information find (i) $P(X_n=0)$, $P(X_n=1)$ and $P(X_n=2)$, (ii) $E(X_n)$ and $V(X_n)$.

Assignment 1.7

Consider a different problem where we also three (3) outcomes, such that the outcomes of any trails are dependent on the outcomes of the previous trials. The transition probability matrix is given and it is $\begin{bmatrix} 1-(1-p) & (1-p)c & (1-p)(1-c) \\ 1-(1-p)(1-c) & (1-p)(1-c)c & (1-p)(1-c)^2 \\ 1-(1-p)(1-c)^2 & (1-p)(1-c)^2 c & (1-p)(1-c)^3 \end{bmatrix}$. Assume that we know that there are three states given by -1, 0, +1, such that $P(X_0=-1)=\alpha$, $P(X_0=0)=\alpha^2$ and $P(X_0=+1)=\alpha^3$, such that $\alpha+\alpha^2+\alpha^3=1$. Given this set of information, find (i) $P(X_n=-1)$, $P(X_n=0)$ and $P(X_n=+1)$, (ii) $E(X_n)$ and $V(X_n)$.

Previous Next

Random Walks which are correlated

Consider we have a sequence of random walks such that we have the transition probability matrix as give, which is $\begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$, but with the difference that we now have two states denoted as -1 and +1 only, such that $P(X_0=+1)=p_1$ and $P(X_0=+1)=(1-p_1)$. Given this we are as usual interested to find $p_n=P(X_n=+1)$ and $q_n=P(X_n=-1)$.

Hence: $p_n=P(X_n=+1,X_{n-1}=+1)+P(X_n=+1,X_{n-1}=-1)$

$$=P(X_n=+1|X_{n-1}=+1)\times P(X_{n-1}=+1)+P(X_n=+1|X_{n-1}=-1)\times P(X_{n-1}=-1)$$

$$=(1-b)\times p_{n-1}+a\times q_{n-1}=(1-b)\times p_{n-1}+a\times(1-p_{n-1})=(1-a-b)p_{n-1}+a$$

If we continue doing it we get

$p_n=\dfrac{a}{(a+b)}+\left\{p_1-\dfrac{a}{(a+b)}\right\}(1-a-b)^n$ and $p_n=1-q_n$

Given this we find

(i) $E(X_n)=\sum_{i=-1}^{1}\{i\times P(X_n=i)\}=[-1\times(1-p_n)]+[+1\times p_n]=2p_n-1$,

in general the formulae would be $E(X_n)=\sum_{\forall i}\{i\times P(X_n=i)\}$,

depending on the number of states, such that there are even number of positives and equal number of negatives, i.e., $-r,-(r-1),\ldots,(r-1),r$ and we will have the $p_{-r,n},p_{-(r-1),n},\ldots,p_{(r-1),n},p_{r,n}$, such that $p_{-r,n}+p_{-(r-1),n}+\ldots+p_{(r-1),n}+p_{r,n}=1$

(ii) $V(X_n)=\sum_{i=-1}^{1}\{i-E(X_n)\}^2\times P(X_n=i)=\{-1-(2p_n-1)\}^2(1-p_n)+\{+1-(2p_n-1)\}^2p_n$

$V(X_n)=1-(2p_n-1)^2$ and in general the formulae would be $V(X_n)=\sum_{\forall i}\{i-E(X_n)\}^2\times P(X_n=i)$,

depending on the number of states such that there are even number of positives and equal number of negatives, i.e., $-r,-(r-1),\ldots,(r-1),r$ and we will have the $p_{-r,n},p_{-(r-1),n},\ldots,p_{(r-1),n},p_{r,n}$,such that $p_{-r,n}+p_{-(r-1),n}+\ldots+p_{(r-1),n}+p_{r,n}=1$

(iii) $E(X_{n-1}X_n)=\sum_{i,j=-1}^{1}\{i-E(X_{n-1})\}\times\{j-E(X_n)\}\times P(X_n=j,X_{n-1}=i)$

$$= \sum_{i,j=-1}^{1} \{i - E(X_{n-1})\} \times \{j - E(X_n)\} \times P(X_{n-1} = i | X_n = j) \times P(X_{n-1} = i | X_n = j)$$

$$= -1(b \times p_{n-1} + a \times q_{n-1}) + 1\{(1-b) \times p_{n-1} + (1-a) \times q_{n-1}\}$$

$$= (1-2a) + 2(a-b)p_{n-1}$$

Now we already know that $\text{cov}(X_{n-1}, X_n) = E[\{X_{n-1} - E(X_{n-1})\}\{X_n - E(X_n)\}]$

$$= E(X_{n-1}X_n) - E\{X_{n-1}E(X_n)\}$$
$$- E\{X_nE(X_{n-1})\} + E\{E(X_{n-1})E(X_n)\}$$

$\text{cov}(X_{n-1}, X_n) = (1-2a) + 2(a-b)p_{n-1} - (2p_{n-1} - 1) \times (2p_n - 1)$ and hence

$$corr(X_{n-1}, X_n) = \left\{ \frac{(1-2a) + 2(a-b)p_{n-1} - (2p_{n-1} - 1) \times (2p_n - 1)}{\sqrt{1-(2p_n - 1)^2} \times \sqrt{1-(2p_{n-1} - 1)^2}} \right\}$$

Previous Next

Note

Solve the problem for $a = b$, in which case you will have

(i) $E\left(X_{n-1}X_n\right) = c$

(ii) $\mathrm{cov}\left(X_{n-1}, X_n\right) = \begin{cases} c - (2p_1 - 1)^2 c^{2n-1} & when\ p_1 \neq \frac{1}{2} \\ c & when\ p_1 = \frac{1}{2} \end{cases}$, and

$$corr\left(X_{n-1}, X_n\right) = \begin{cases} \dfrac{c - (2p_1 - 1)^2 c^{2n-1}}{\sqrt{1 - (2p_n - 1)^2} \times \sqrt{1 - (2p_{n-1} - 1)^2}} & when\ p_1 \neq \frac{1}{2} \\ c & when\ p_1 = \frac{1}{2} \end{cases}$$

Previous Next

Assignments 1.8

Consider the problem (*you have already solved a part of it*) as given, in which there are three (3) outcomes, such that the outcomes of any trails are dependent on the outcomes of the previous trial, and the transition probability matrix is given as $\begin{bmatrix} 1-(1-c) & (1-c)p & (1-c)(1-p) \\ 1-(1-c)(1-p) & (1-c)(1-p)p & (1-c)(1-p)^2 \\ 1-(1-c)(1-p)^2 & (1-c)(1-p)^2 p & (1-c)(1-p)^3 \end{bmatrix}$.

There are three (3) states given by 0, 1, 2, such that $P(X_0=0)=\alpha$, $P(X_0=1)=\beta$ and $P(X_0=2)=\gamma$, and $\alpha+\beta+\gamma=1$. With this information find (i) $E(S_n)$ and $V(S_n)$, where $S_n=\sum_{i=1}^{n} X_i$.

Assignment 1.9

Consider the problem (*you have already solved a part of it*) as given, in which there are three (3) outcomes, such that the outcomes of any trails are dependent on the outcomes of the previous trial, and the transition probability matrix is given as $\begin{bmatrix} 1-(1-p) & (1-p)c & (1-p)(1-c) \\ 1-(1-p)(1-c) & (1-p)(1-c)c & (1-p)(1-c)^2 \\ 1-(1-p)(1-c)^2 & (1-p)(1-c)^2 c & (1-p)(1-c)^3 \end{bmatrix}$.

There are three (3) states given by -1, 0, +1, such that $P(X_0=-1)=\alpha$, $P(X_0=0)=\alpha^2$ and $P(X_0=+1)=\alpha^3$, and $\alpha+\alpha^2+\alpha^3=1$. With this information find (i) $E(S_n)$ and $V(S_n)$, where $S_n=\sum_{i=1}^{n} X_i$.

Previous Next

Assignment 1.10

Consider the problem (*you have already solved a part of it*) as given, in which there are three (3) outcomes, such that the outcomes of any trails are dependent on the outcomes of the previous trial, and the transition probability matrix is given as $\begin{bmatrix} 1-(1-c) & (1-c)p & (1-c)(1-p) \\ 1-(1-c)(1-p) & (1-c)(1-p)p & (1-c)(1-p)^2 \\ 1-(1-c)(1-p)^2 & (1-c)(1-p)^2 p & (1-c)(1-p)^3 \end{bmatrix}$.

There are three (3) states given by 0, 1, 2, such that $P(X_0=-1)=\alpha$, $P(X_0=1)=\beta$ and $P(X_0=2)=\gamma$, and $\alpha+\beta+\gamma=1$. With this information find (i) $E\{X_n-E(X_n)\}^3$, (ii) $E\{X_n-E(X_n)\}^4$, (iii) $E\{S_n-E(S_n)\}^3$ and (iv) $E\{S_n-E(S_n)\}^4$, where $S_n=\sum_{i=1}^{n}X_i$.

Assignment 1.11

Consider the problem (*you have already solved a part of it*) as given, in which there are three (3) outcomes, such that the outcomes of any trails are dependent on the outcomes of the previous trial, and the transition probability matrix is given as $\begin{bmatrix} 1-(1-p) & (1-p)c & (1-p)(1-c) \\ 1-(1-p)(1-c) & (1-p)(1-c)c & (1-p)(1-c)^2 \\ 1-(1-p)(1-c)^2 & (1-p)(1-c)^2 c & (1-p)(1-c)^3 \end{bmatrix}$.

Assume that we know that there are three (3) states given by -1, 0, +1, such that $P(X_0=0)=\alpha$, $P(X_0=0)=\alpha^2$ and $P(X_0=+1)=\alpha^3$, such that $\alpha+\alpha^2+\alpha^3=1$. With this information find (i) $E\{X_n-E(X_n)\}^3$, (ii) $E\{X_n-E(X_n)\}^4$, (iii) $E\{S_n-E(S_n)\}^3$ and (iv) $E\{S_n-E(S_n)\}^4$, where $S_n=\sum_{i=1}^{n}X_i$.

Assignment 1.12

Assume we have a sequence of random walks such that the transition probability matrix is give as $\begin{bmatrix} 1-(1-c) & (1-c)q & (1-c)(1-q) \\ 1-(1-c)(1-q) & (1-c)(1-q)q & (1-c)(1-q)^2 \\ 1-(1-c)(1-q)^2 & (1-c)(1-q)^2 q & (1-c)(1-q)^3 \end{bmatrix}$, such that the outcomes of any trails are dependent on the outcomes of the previous trial and $c\sim U(0,1)$, $0\le q\le 1$. Also consider that $P(X_0=i)=p_i(1-c)$, such that sum of the probability (at the initial stage, t=0) is exactly equal to 1, i.e.,, $\sum_{\forall i}p_i(1-c)=1$ (sum of row elements is exactly equal to 1), and it is also always true that the sum of the realized values of the states at any t=n is exactly equal to six (6), i.e., $i+j+k=6$, where $i,j,k\in S$. Given this set of information, find the general formulae for (i) $p_n=P(X_n=i)$, such that $\sum_{\forall i}P_n(X_n=i)=1$, (ii) $E(X_n)$, (iii) $V(X_n)$, (iv) $E(S_n)$ & $V(S_n)$, where $S_n=\sum_{i=1}^{n}X_i$.

Previous Next

Actual examples of Markov Chains

Let us give a brief set of applications for Markov chains and the areas are:

Internet applications : Markov models can be used to generally understand the browsing characteristics of surfers, such that the web page which appears, depending on the browsing characteristics can be modeled as the state space of the browsing characteristics. Hence if we have N number of web pages which can be visited by a surfer, and each internet page has k_i number of links, then the transition probability can be given by the formulae $\left(\frac{\alpha}{k_i}+\frac{1-\alpha}{N}\right)$ for all the pages that are linked to and $\left(\frac{1-\alpha}{N}\right)$ for pages which are not linked to, where α is the transition probability parameter, e.g., the parameter c for our earlier example.

Physics : A sample set of application of Markov process in physics are *thermodynamics* and *statistical mechanics*. In general for these physical processes we try to represent the probability for the unknown and hence for specific detail about the equational form of the details of the physical system under study. For example in thermodynamics if we assume the variable, X, to be dependent on time then we may model it as a simple stochastic process where the outcome is dictated by both its state and space, such that we can denote it as $X(x,t)$. Similarly in statistical mechanics we have the rate of change of the process given as $\frac{\partial}{\partial X(x,t)} f\{X(x,t)\}$, where one may attempt to find the overall or average property of movement of the particles or in whole of the whole body using stochastic differential equations.

Queueing theory : Markov chains is also used for modeling various processes in queueing theory and statistics. In mathematical theory of communication (consider a single step process) in which the stage to which the information/communication moves can be considered as the states and the corresponding probability of the information/communication being transmitted without any los of information is given by the transition probability matrix. What we can do is to find the average rate of information flow and the correlation that message/information/communication once passed/started continues without any loss to a certain probability or certainty.

Chemistry : In crystal or carbon molecular growth (considering we are interested in finding some new combination or a new drug) such that the addition or subtraction of one known molecule of carbon or any other molecule can be consider the state and the probability that the molecule is added or subtracted is considered using it transition probability matrix values.

Statistics : A very interesting application is the use of Markov chain Monte Carlo (MCMC) process. In recent years this has revolutionized the practicability of Bayesian inference methods, allowing a wide range of posterior distributions to be simulated and their parameters found numerically.

Previous Next