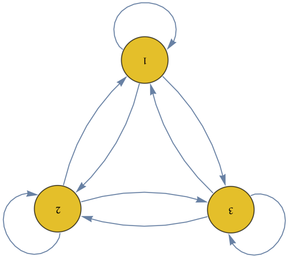
Markov chain

Example 5: Random walk over three states



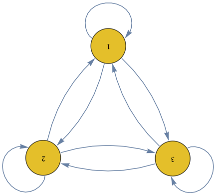
With probability 1/2, stay. With probability 1/4, go to either states.

$$K_{ij} = P(X_{t+1} = j | X_t = i).$$

Markov property: past history before X_t does not matter.

Population migration

Example 5: Random walk over three states

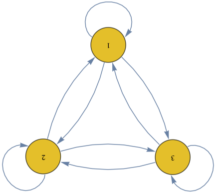


With probability $1/2$, stay. With probability $1/4$, go to either of the other two states.

$$K_{ij} = P(X_{t+1} = j | X_t = i).$$

Forward conditional probability, from cause to effect.
Imagine 1 million people migrating. At each step, for each state, half of the people stay, $1/4$ go to each of the other two states. 1 million trajectories.

Example 5: Random walk over three states



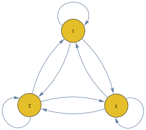
With probability 1/2, stay. With probability 1/4, go to either of the other two states.

$$K_{ij} = P(X_{t+1} = j | X_t = i).$$

$$\mathbf{K} = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}$$

Marginal probability

Example 5: Random walk over three states



With probability 1/2, stay. With probability 1/4, go to either of the other two states.

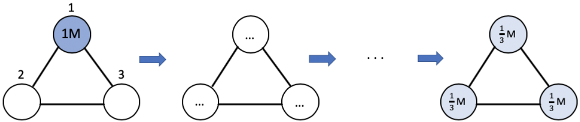
$$K_{ij} = P(X_{t+1} = j | X_t = i).$$

$$p_i^{(t)} = P(X_t = i).$$

Imagine 1 million people migrating. $p_i^{(t)}$ is the number of people (in million) in state i at time t.

$$\mathbf{p}^{(t)} = (p_1^{(t)}, p_2^{(t)}, p_3^{(t)}).$$

Example 5: Random walk over three states



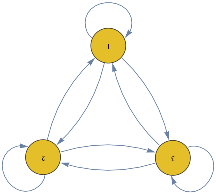
$$p_i^{(t)} = P(X_t = i).$$

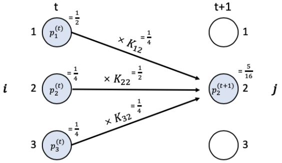
Imagine 1 million people migrating. $p_i^{(t)}$ is the number of people (in million) in state i at time t.

$$\mathbf{p}^{(t)} = (p_1^{(t)}, p_2^{(t)}, p_3^{(t)}).$$

Population migration

Example 5: Random walk over three states





$$K_{ij} = P(X_{t+1} = j | X_t = i).$$

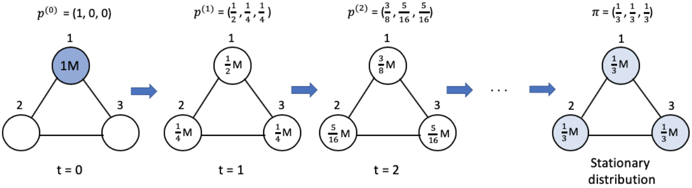
$$p_i^{(t)} = P(X_t = i).$$

$$p_j^{(t+1)} = \sum_i p_i^{(t)} K_{ij}.$$

Number of people in state j at time $t+1$ = sum number of people in state i at time t $\times$ fraction of those in i who go to j.

Stationary distribution

Example 5: Random walk over three states



$$p_j^{(t+1)} = \sum_i p_i^{(t)} K_{ij}.$$

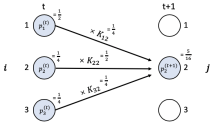
$$p_i^{(t)} \to \pi_i.$$

$$\pi_j = \sum_i \pi_i K_{ij}.$$

Stationary distribution, arrow of time.

Matrix multiplication

Example 5: Random walk over three states



$$\underbrace{\begin{array}{ccc} & p^{(t+1)} & \\ \hline & \frac{5}{16} & \\ \hline 1 & 2 & 3 \end{array}}_{} = \begin{array}{ccc} & p^{(t)} & \\ \hline \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \hline 1 & 2 & 3 \end{array}$$

$$K: \quad \begin{array}{c|ccc} i \backslash j & 1 & 2 & 3 \\ \hline 1 & & \frac{1}{4} & \\ 2 & & \frac{1}{2} & \\ 3 & & \frac{1}{4} & \end{array}$$

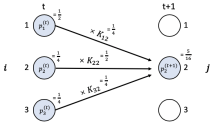
$$p_j^{(t+1)} = \sum_i p_i^{(t)} K_{ij}.$$

$$p^{(t+1)} = p^{(t)} K.$$

$$p^{(t)} = p^{(0)} K^t \to \pi.$$

Diagonalization and eigen-analysis

Example 5: Random walk over three states



$$
\begin{array}{c} p^{(t+1)} \\ \begin{array}{|c|c|c|} \hline & \frac{5}{16} & \\ \hline \end{array} \\ 1 \quad 2 \quad 3 \end{array}
=
\begin{array}{c} p^{(t)} \\ \begin{array}{|c|c|c|} \hline \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \hline \end{array} \\ 1 \quad 2 \quad 3 \end{array}
\quad
\begin{array}{c} K \\ \begin{array}{c|c|c|c} i \backslash j & 1 & 2 & 3 \\ \hline 1 & & \frac{1}{4} & \\ \hline 2 & & \frac{1}{2} & \\ \hline 3 & & \frac{1}{4} & \end{array} \end{array}
$$

Diagonalization and eigen-analysis: $K = PDP^{-1}$, D diagonal, eigenvalues.

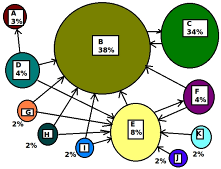
$$K^t = PDP^{-1}PDP^{-1}...PDP^{-1} = PD^tP^{-1}.$$

$$p^{(t)} = p^{(0)}K^t \to \pi.$$

Largest eigenvalue $= 1$, $1^t = 1$.
Second largest eigenvalue < 1, e.g., $.99^t \to 0$.

Google pagerank

Example 5: Random walk



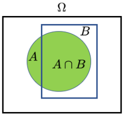
$$p_j^{(t+1)} = \sum_i p_i^{(t)} K_{ij}.$$

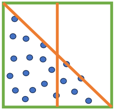
$$p_i^{(t)} \to \pi_i.$$

$$\pi_j = \sum_i \pi_i K_{ij}.$$

π_i: proportion of people who are in page i.
Popularity of i depends on the popularities of pages linked to i.

Conditional





$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

(1) Counting population: Randomly sample from subpopulation B (e.g., males).

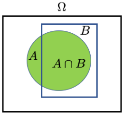
(2) Counting repetitions: When B happens, how often A happens.

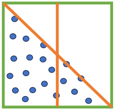
Regular prob is conditional prob: $P(A) = P(A|\Omega)$.

Fixed condition (within the same subpopulation B), conditional prob behaves like regular prob.

e.g., $P(A^c) = 1 - P(A)$; $P(A^c|B) = 1 - P(A|B)$.

Chain rule



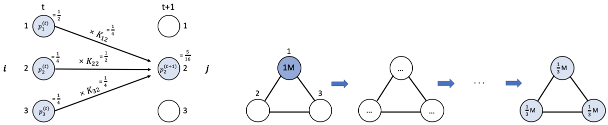


$$P(A \cap B) = P(B)P(A|B).$$

(1) Counting population: Population proportion of tall males = proportion of males × proportion of tall among males.
(2) Counting repetitions: B happens $1/2$ times. When B happens, A happens $3/4$ times. How often A and B happen together?
Generalize to chain of multiple events:

$$P(A \cap B \cap C) = P(A \cap B)P(C|A \cap B) = P(A)P(B|A)P(C|A, B).$$

Chain rule and rule of total probability



Chain rule:

$$P(X_{t+1} = j \cap X_t = i) = P(X_t = i)P(X_{t+1} = j | X_t = i)$$
$$= p_i^{(t)} K_{ij}.$$

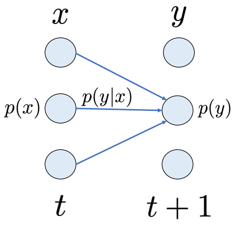
Rule of total probability:

$$P(X_{t+1} = j) = \sum_i P(X_{t+1} = j \cap X_t = i).$$

$$p_j^{(t+1)} = \sum_i p_i^{(t)} K_{ij}.$$

Add up probabilities of alternative chains of events.

Marginal, conditional and joint distributions



Marginal: $p_t(x) = P(X_t = x)$, $p_{t+1}(y) = P(X_{t+1} = y)$.
Conditional: Forward $p(y|x) = P(X_{t+1} = y|X_t = x)$.
x: cause, y: effect. $p(y|x)$: cause $\to$ effect, given or learned.
Joint: $p(x, y) = P(X_t = x, X_{t+1} = y)$.
Chain rule: $p(x, y) = p_t(x)p(y|x)$.
Rule of total probability:
$p_{t+1}(y) = \sum_x p(x, y) = \sum_x p_t(x)p(y|x)$.
Add up probabilities of alternative chains of events.