

IID and Markov Models

Gaurav Sharma

University of Rochester

Dependence in Probabilistic Models

- ▶ Fundamentally, probabilistic models model dependence
 - ▶ Single variable across "time" and multiple variables
 - ▶ Will begin with single variable across "time"
 - ▶ Discrete time random process

$$X_1, X_2, \dots X_n$$

indexed collection of RVs

- ▶ Full characterization: All Joint PMF/PDFs

$$p(x_1, x_2, \dots x_n) = \Pr\{X_1 = x_1, X_2 = x_2, \dots X_n = x_n\}$$

- ▶ Absolutely general: Can model any dependence
- ▶ How useful is it?
 - ▶ Almost useless, except in understanding assumptions underlying practical models

Independent Identically Distributed (IID)

- ▶ Simplest model for dependence
- ▶ X_1, X_2, \dots independent and identical distribution
- ▶ Factorization of Joint PMF

$$p(x_1, x_2, \dots x_n) = p(x_1)p(x_2)p(x_3) \dots p(x_n)$$

- ▶ Examples:
 - ▶ Good?
 - ▶ Bad?
- ▶ Usage Examples

Markov Models

- ▶ Allow more realistic models by incorporating dependency
- ▶ Conditional Independence
 - ▶ Past independent of future given present

$$\begin{aligned} & \Pr\{X_{n+1} = x_{n+1} \mid X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_1 = x_1\} \\ &= \Pr\{X_{n+1} = x_{n+1} \mid X_n = x_n\} \end{aligned}$$

- ▶ Factorization of Joint PMF

$$p(x_1, x_2, \dots, x_n) = p(x_1)p(x_2 \mid x_1)p(x_3 \mid x_2) \dots p(x_n \mid x_{n-1})$$

- ▶ Contrast with more general Bayes Rule based Factorization

$$\begin{aligned} & p(x_1, x_2, \dots, x_n) = \\ & p(x_1)p(x_2 \mid x_1)p(x_3 \mid x_2, x_1) \dots p(x_n \mid x_{n-1}, x_{n-2} \dots x_1) \end{aligned}$$

Time-Invariant (Homogeneous) Markov Models

- ▶ Additional simplification: transition probabilities independent of time

$$\Pr\{X_{n+1} = x_{n+1} \mid X_n = x_n\} = \Pr\{X_2 = x_{n+1} \mid X_1 = x_n\}$$

$$\Pr\{X_{n+1} = j \mid X_n = i\} = \Pr\{X_2 = j \mid X_1 = i\} = P_{ij}$$

- ▶ Finite set of discrete outcomes: $X_n \in \{1, 2, \dots, L\}$
- ▶ $L \times L$ Transition probability matrix

$$P = [P_{ij}]$$

- ▶ Markov chain terminology sometimes restricted to this setting: discrete time and finite discrete outcome space

Markov Chain Characterization

- Full characterization of time-invariant Markov chain

Recall,

$$p(x_1, x_2, \dots, x_n) = p(x_1)p(x_2 | x_1)p(x_3 | x_2) \dots p(x_n | x_{n-1}) \quad (1)$$

$$p_i^{(n)} = p_{X_n}(i)$$

$$p^{(n)} = [p_i^{(n)}]$$

- $p_{1 \times L}^{(1)}$ and $P_{L \times L}$ fully characterize the Markov chain

$$p^{(n)} = p^{(n-1)}P = p^{(1)}P^{(n-1)}$$

Markov Chain Example I

- ▶ Is an IID process a Markov process?
- ▶ For a finite-state discrete time IID process:
 - ▶ What is P ?
 - ▶ What is $\mathbf{p}^{(1)}$?

Markov Chain Example II

- ▶ Simple two state Markov Chain

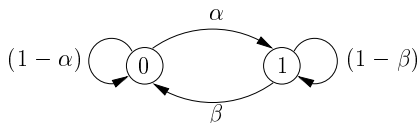


Figure: Transition Diagram representation of a Binary Markov Chain.

- ▶ What is P ?
- ▶ What is $\mathbf{p}^{(1)}$?
- ▶ Under what conditions does this degenerate into an IID process?

Time-invariance and stationarity

- ▶ Is a time-invariant Markov Chain (MC) stationary
- ▶ Depends on initial distribution
 - ▶ Example

Characterization of Finite-State Discrete Time Markov Chains

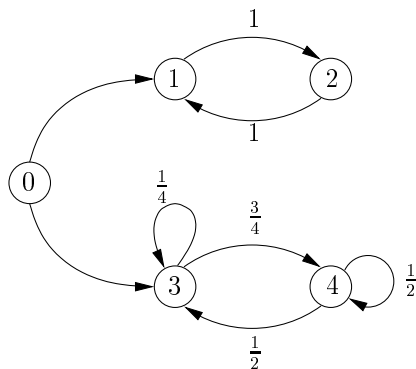


Figure: Characterization of states of a time-invariant Markov Chain.

Characterization of Finite-State Discrete Time Markov Chains

- ▶ Partitioning of States into Transient and Irreducible Sets
 - ▶ Transient states (impossible to return to from some reachable state)
 - ▶ An Irreducible set of states
 - ▶ Periodic or aperiodic (ergodic)
 - ▶ Markov Chain irreducible if all states form a single irreducible set
 - ▶ Focus particularly on irreducible ergodic MCs

Markov Model as a Representation for the Physical World

- ▶ Future state = function of current state
 - ▶ Possibly random function
 - ▶ independent of past
 - ▶ suitably chosen set of state variables should suffice
- ▶ Practical challenges
 - ▶ State space is too large
 - ▶ Not all states are observable
 - ▶ hidden/latent variables
- ▶ Observations are noisy

Summary

- ▶ Probabilistic models must tread a fine line between generality and utility
- ▶ Simplest model: IID
- ▶ Markov Models: More powerful class of models
- ▶ Important: Note distinction between distribution and temporal dependence
 - ▶ Specifically time and space of outcomes may independently be continuous or discrete