

STAT 545 Midterm Notesheet

Markov Chains — Definitions, Techniques, and Worked Examples

1. Two-State Markov Chains

Definition: A (time-homogeneous) Markov chain satisfies

$$P(X_{n+1} = j \mid X_n = i, X_{n-1}, \dots) = P_{ij}.$$

For a two-state chain $\{0, 1\}$ with

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}, \quad \alpha = (\alpha_0, \alpha_1),$$

we compute several useful quantities.

Two-Step Transitions

Key Idea: Compute $P^2 = P \cdot P$ entrywise.

$$P^2 = \begin{pmatrix} (1-p)^2 + pq & p(2-p-q) \\ q(2-p-q) & (1-q)^2 + pq \end{pmatrix}.$$

Expectation of the State

Key Idea: Since $X_n \in \{0, 1\}$,

$$E[X_n] = P(X_n = 1).$$

Thus,

$$E[X_2] = (\alpha P^2)_1 = \alpha_0 P_{01}^2 + \alpha_1 P_{11}^2.$$

General n -Step Behavior

Theorem: For $p + q > 0$, the n -step transition matrix is

$$P^n = \frac{1}{p+q} \begin{pmatrix} q + pc_n & p - pc_n \\ q - qc_n & p + qc_n \end{pmatrix}, \quad c_n = (1 - p - q)^n.$$

Therefore,

$$E[X_n] = \frac{p}{p+q} + \left(\alpha_1 - \frac{p}{p+q} \right) c_n.$$

Long-Run Limit

Key Idea: If $|1 - p - q| < 1$, then $c_n \rightarrow 0$.

$$\lim_{n \rightarrow \infty} E[X_n] = \frac{p}{p+q}.$$

Key Idea: This equals the stationary probability of state 1.

Return Times

Theorem: For an irreducible Markov chain,

$$E[T_i \mid X_0 = i] = \frac{1}{\pi_i},$$

where π is the stationary distribution.

Here,

$$\pi = \left(\frac{q}{p+q}, \frac{p}{p+q} \right), \quad E[T_1 \mid X_0 = 1] = \frac{p+q}{p}.$$

Closed-Form Expression for P^n

For the two-state Markov chain with transition matrix

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}, \quad p, q \geq 0, \quad p+q > 0,$$

the n -step transition matrix admits the closed-form solution

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

Key Idea: Key structure:

- $(1-p-q)^n$ is the *decay term* (second eigenvalue raised to power n).
- The constant terms correspond to the stationary distribution

$$\pi = \left(\frac{q}{p+q}, \frac{p}{p+q} \right).$$

- As $n \rightarrow \infty$, the exponential term vanishes whenever $|1-p-q| < 1$.

Key Idea: Fast consistency check (exam trick):

- Rows sum to 1.
- Setting $n = 0$ recovers $P^0 = I$.
- Taking $n \rightarrow \infty$ gives rows equal to π .

2. Random Walks on Graphs

Definition: A simple random walk moves uniformly to a neighboring vertex.

Transition Matrices

Key Idea: For vertex i ,

$$P_{ij} = \frac{1}{\deg(i)} \quad \text{if } i \sim j.$$

Always compute degrees first, then normalize rows.

Communication Classes

Definition: States i, j communicate if $i \rightarrow j$ and $j \rightarrow i$.

Key Idea: Disconnected components $\Rightarrow$ multiple classes.

Recurrence

Theorem: All states in a finite irreducible Markov chain are recurrent.

Period

Definition: The period of state i is

$$d(i) = \gcd\{n \geq 1 : P_{ii}^n > 0\}.$$

Key Idea: Bipartite graphs typically have period 2.

Ergodicity

Theorem: A Markov chain is ergodic iff it is irreducible and aperiodic.

Key Idea: Lazy random walks (self-loop probability > 0) are always aperiodic.

3. Absorbing Markov Chains

Definition: A state is absorbing if $P_{ii} = 1$.

Key Idea: Reorder states so transient states come first.

$$P = \begin{pmatrix} Q & R \\ 0 & S \end{pmatrix}$$

Fundamental Matrix

Theorem: The fundamental matrix is

$$F = (I - Q)^{-1}.$$

Key Idea: F_{ij} = expected number of visits to j starting from i before absorption.

Expected Time to Absorption

$$E[T \mid X_0 = i] = \sum_j F_{ij}.$$

Absorption Probabilities

Absorption probabilities = FR .

Limiting Distribution

Theorem: If all closed classes are aperiodic,

$$\lim_{n \rightarrow \infty} P^n = \begin{pmatrix} 0 & (I - Q)^{-1}R\Pi \\ 0 & \Pi \end{pmatrix},$$

where Π contains stationary distributions of closed classes.

4. Estimating Transition Matrices

Maximum Likelihood Estimation

Definition: The MLE of P is obtained by row-normalizing transition counts.

Key Idea: Small sample sizes $\Rightarrow$ unreliable estimates.

Laplace Smoothing

Definition: Add 1 to each count:

$$\hat{p}_i = \frac{N_i + 1}{\sum_j (N_j + 1)}.$$

Key Idea: Prevents zero probabilities.

Dirichlet (Empirical Bayes) Smoothing

Definition: Use a Dirichlet prior with mean μ and strength α .

Posterior mean:

$$\tilde{p}_i = \frac{N_i + \alpha\mu_i}{N + \alpha}.$$

Key Idea: Borrow strength from other rows (pooled information).