Higher Order Markov Chains and Spacey Walks

Nicolas Nytko

November 30, 2020

Markov Chains

Markov Chains are stochastic models describing a sequence of states with a probability of transitioning between each.

- Defined as matrix: $\mathbf{P} \in \mathbb{R}^{n \times n}$, $\mathbf{P}_{ij}$ gives probability of moving to state i from j, for chain with n states.
- Given distribution/state vector $\mathbf{x}^{(n)}$, one step of a random walk is represented by matrix-vector product $\mathbf{x}^{(n+1)} = \mathbf{P}\mathbf{x}^{(n)} \quad \Leftrightarrow \quad x_i^{(n+1)} = \sum_j P_{ij} x_j^{(n)}$.

Definition

Time homogeneous: each transition probability is independent of the time step t.

Definition

Irreducible: every state can be reached from every other state.

Canonical Examples

Google PageRank

- Markov chain represents the "idealized" web surfer. States are webpages and transition probability is given by number of outgoing links between pages.
- Ranking is determined by probability of being on certain page.

Board Games

- Each spot on game board is a state.
- Probability of moving between spots is given by dice, spinner, etc.

Weather

- States are possible weather conditions.
- Tomorrow's weather is randomly chosen based on today's conditions.

Steady-State Distribution

Steady state vector $\boldsymbol{\pi} \in \mathbb{R}^n$ does not change when multiplied by Markov chain:

$$\boldsymbol{\pi} = \mathbf{P}\boldsymbol{\pi}$$

- Average distribution as random walk tends to length infinity, $\lim_{k\to\infty} \mathbf{P}^k \mathbf{x}$
- Is equivalent to **Perron vector** of matrix, dominant eigenvector with nonnegative components.
- Compute iteratively with power method: $\mathbf{x}^{(k+1)} = \mathbf{P}\mathbf{x}^{(k)}$.

Remark

Unique steady states are guaranteed for **irreducible** Markov chains.

Markov Chains with Memory

What if we want our chain to **remember** where it has been? Introduce ***m*-order Markov chain**, where transition probability depends on previous m states.

- Probability is now stored as $\mathbf{P} \in \mathbb{R}^{n^m \times n}$, where rows are a m-tuple of each state permutation.
- Random step is still mat-vec, state vector $\mathbf{x}$ now stores distribution of each state permutation m-tuple.

Higher Order Markov Chains

- But, since we know about tensors we can do better.
- Transition probability is represented as $m+1$ order Tensor.
- Probability to transition given by $\mathcal{P}_{i_{(n+1)},i_{(n)},i_{(n-1)},\dots,i_{(n-m+1)}}$
  - Example for second order: $\mathcal{P}_{ijk}$ describes transition probability to state i from current state j and previous state k.
- Tensor form gives nice analysis properties — will see later.
- State/distribution is stored as order m tensor,
- Random step is given by contraction over first mode $X^{(n+1)}_{i_{(m+1)},\dots,i_{(2)}} = \sum_{i_{(1)}} P_{i_{(m+1)},\dots,i_{(1)}} X^{(n)}_{i_{(m)},\dots,i_{(1)}}$

Higher Order Steady-State Distribution

Spacey Walks

Applications, Examples

References

Austin R. Benson, David F. Gleich, and Lek-Heng Lim.
The spacey random walk: A stochastic process for higher-order data.
SIAM Review, 59(2):321–345, January 2017.

Hassan Bozorgmanesh and Masoud Hajarian.
Convergence of a transition probability tensor of a higher-order markov chain to the stationary probability vector.
Numerical Linear Algebra with Applications, 23(6):972–988, August 2016.