Optimal value of lifts in finite Markov chains

Ahmed Ali Hedayatullah, Dr. Aaron Smith

University of Ottawa Department of Mathematics

Introduction

Markov chains and its properties are a fascinating and integral part of current research done in the fields of probability and statistics. Markov chains are defined as a stochastic process that satisfies several equalities related to conditional probabilities and the usual parameters of Markov chain (Omega and P). Decreasing the time needed for a finite Markov chain to become approximately stationary, or the mixing time, was the property that was focused on in this project. The goal of this project was to determine the optimal value of the number of lifts required such that after m lifts,further lifts would give little or no improvement to overall mixing times. What makes this research so exciting is the potential prospects it has to change how computations are done, producing faster, more efficient and less costly Markov chains.

What is truly remarkable about Markov chains in general are the characteristics that define it. In particular, transition matrix P, which gives the reader an idea about certain probabilities that will occur in each state given certain conditions. It's almost as if you can glimpse into the future.

Methodology

First I deepened my understanding of Markov chains, lifts, mixing times and stationary distributions through reading research articles and textbooks written on the subject. Afterwards I attempted to find the optimal value of lifts on small and large finite chains through the statistical language R.

Important/Useful Equations

Markov property;

$$P[X_{t+1} = Y : X_t = x] = P(x, y)$$

→Probability of proceeding from state x to y is the same no matter what sequence of states precede the current state x.

Stationary distribution;

$$\pi = \pi p \ or \ \pi(y) = \sum_{x \in \sigma} \pi(x) P\,(x, y)\ \forall y \in \sigma$$

→"Finite Markov chains converge to their stationary distributions"

Mixing time;

$$t_{mix}(\epsilon) = \min(t: d(t) \leq \epsilon) \ and \ t_{mix} = t_{mix}\left(\frac{1}{4}\right)$$

→ The time needed for the chain to become approx. stationary.

Optimal Lifting

Lifting → definition: splitting each state into several states

Assume a non-negative valued function with nodes I and j as edges.

Thereom 3.2 (Dr.Chen et al.)

→ For every chain M ; ½ <inf H^ < 144C

Where the infimum is taken over all lifting of M.

Mixing time of optimal lift is at most 144C.

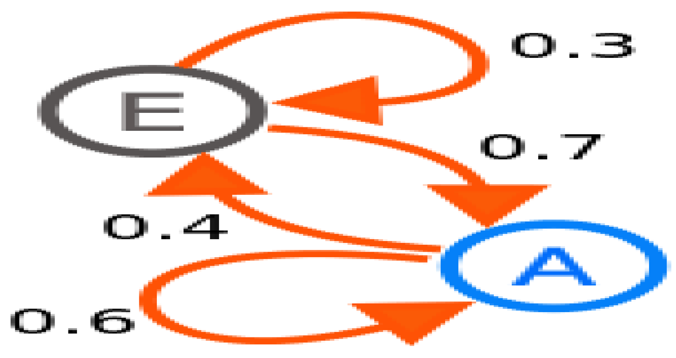
$$C = \theta\left(\frac{1}{\pi} * \frac{1}{\alpha}\right)$$

Several classes/states of Markov chains :

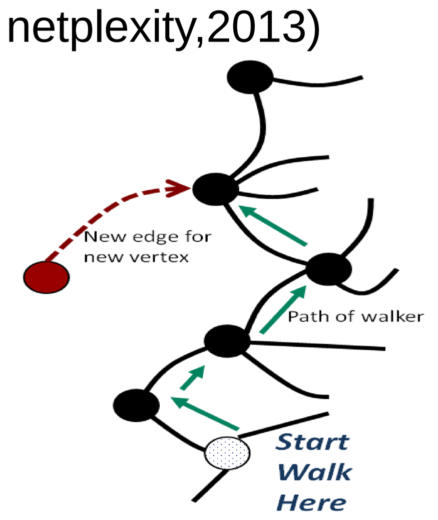
1)Essential

2) Non–essential

(MikeDehann, 2013)



(SRW, netplexity,2013)



Real Life Applications/Examples

A student has 2 ways of getting to university A & B. Each day he picks one and his choice is ONLY (holding all other factors constant) influenced by his previous days choice:

If A, then P(A) = ½ next day if B, then P(A)= ¾ next day. At day 0 (initial state) or current state x, he has no preference.

P= $\begin{matrix} A & 1/2 & 1/2 \\ B & 3/4 & 1/4 \end{matrix}$ Day 1 → (1/2, 1/2) * P = (0.625, 0.375)

Day 2 → (5/8, 3/8) *P = (19/32, 13/32)

After n steps, the chain will converge eventually to

$\begin{matrix} 0.600 & 0.400 \\ 0.600 & 0.600 \end{matrix}$

so at the end of the sequence or let's say at the end of the year the individual will likely tend towards the matrix above. This highlights stationary distributions.

Obviously this is a simple example to highlight the capability of Markov chains to predict or provide a glimpse into what someone might do.

Discussion

In the future, further research into Markov chains would prove to be useful in my opinion considering the wide range of topics that it could cover. I would like to personally learn R so that I can have a better grasp of Markov chains and statistics in general and to use that knowledge for real life scenarios such as the stock market. In regards to the current project the end goal was not reached, overall 75% of the project was completed (deepening understanding of Markov chains, linear algebra tools, and knowledge of mixing times/lifts) with the last 25% being using R to find optimal values.

Conclusion

Unfortunately I was unable to find any optimal value of lifts on finite Markov chains. Though with the knowledge I have gained throughout this opportunity I am confident that in the future given that I learn R I will be able to find some optimal values.

References

Chen, F. (n.d.). Lifting Markov Chains to Speed up Mixing. Retrieved December 20, 2015, from http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=77919F8323CC28EB9A1E91254C758368?doi=10.1.1.31.5851&rep=rep1&type=pdf

Levin, D. A. (n.d.). Markov Chains and Mixing Times. Retrieved December 20, 2015, from http://pages.uoregon.edu/dlevin/MARKOV/markovmixing.pdf

Diaconis , Holmes , Neal : Analysis of a nonreversible Markov chain sampler https://projecteuclid.org/download/pdf_1/euclid.aoap/1019487508

Haan, M. D. (2013). Cancer Research Applies a Markov Chain Monte Carlo Approach. Retrieved March 09, 2016, from http://www.decodedscience.org/cancer-research-applies-a-markov-chain-monte-carlo-approach/27675

Acknowledgements

I would like to extend my greatest appreciation and gratitude to the Undergraduate Opportunity Program (UROP) for providing me with this opportunity.

I would also like to thank Dr. Smith for his guidance and support throughout the duration of this project.

Ahmed Ali Hedayatullah

Aheda064@uottawa.ca