The Statistics of Random Walks

Habib Rehman
Imperial College London

ABSTRACT

The purpose of this research has been to identify the precise correspondence between the return distribution of a one-dimensional random walk and a two-dimensional random walk. This research also considers the correspondence between the arc length distribution of a random walker returning and the arc length distribution of the loops erased.

AIMS

1. Identify the correspondence between the return distribution of a 1D random walk and a 2D random walk.
2. Consider the correspondence between the arc length distribution of a random walker returning.

INTRODUCTION

A random walk (RW) is a stochastic process formed by the repeated summation of identically distributed (independent) random variables distributed random variables. Random walks are used to model stochastic process. This research considers random walks on an integer lattice. A Loop erased random walk (LERW) on the integer lattice is a process obtained from Random Walk by erasing the loops chronologically.

METHOD

Python code was written to generate the RWs and LERWs in systems of certain configurations. The data generated by running this code multiple times with different configurations (i.e. different Sys sizes and Sys dimensions) was used to conduct the statistical analysis.

The RW was theoretically modelled to transverse on the 2D surface of a 3D cylinder (the system) with two vertical open boundaries and to two horizontal periodic. At the open boundaries, the RW is initiated from starting boundary and ceased at the ending boundary. The RW was generated from the following function:

$$\omega_n = \left\{ \sum_{k=0}^{n} X_k \,\middle|\, X_k \neg\, 0 \vee L \vee C \right\}$$

RESULTS

The first result that my research found was that was that return times distribution for $\mathbb{Z}^2$ random walk and return times distribution for $\mathbb{Z}$ random walk were very similar. I also found that there is a difference in the distribution of the return times of a random walker (1D and 2D) and the distribution of the loops erased in LERW.

Secondly, I found that the return times distribution for $\mathbb{Z}^2$ random walk and return times distribution for $\mathbb{Z}$ random walk were very similar. A possible explanation for this is that, for a random walk in one dimension starting at n, the probability that the random walk (S_n) eventually returns to n equals $\Pr(S_n = n) = 1, n \in \mathbb{Z}$. However, the time required to return to n, averaged over all possible vertices in the system, is infinite. For a random walk in two dimensions, the survival probability $S(t)$ ultimately decays to zero

RESULTS (CONT.)

This means that a random walk is recurrent implying that it is certain to eventually return to its starting point, and indeed visit any vertex of an infinite lattice. This is also because a random walk has no memory and so it transits to a new state (i.e. resets it\' s state) every time a specific lattice vertex is transversed. Hence, recurrence also implies that every site is visited infinitely often.

Thirdly, I found that there is a correlation between the distribution of the loops erased in a random walk and the probability of return to (an arbitrary) vertex in the lattice for both systems of size 100 and 200. In fact, the combined distribution is positively skewed which implies that as the loops erased increase, the probability of returning to a vertex in the lattice decreases. The fact that the data is skewed, indicates that there is a difference arc length distribution of a random walker returning and arc length distribution of the loops erased.

DISCUSSION

A significant amount data had to collected for different system in order to draw a reliable and a more decisive conclusion. The main implication of the presence of periodic boundaries was that the size of the result was indefinite. Since the code could not be optimised for multithreading, the task for the execution of the code could not utilise more than one core. Consequently, the data collection process was very slow. Furthermore, due to continued runtime errors on the cluster, the amount of data collected was limited to the following:

- 2D RWs of system sizes of 50, 100, 200 and 400
- 1D RWs of system sizes of 50, 100, 200 and 400
- 2D LERWs of system sizes of 50, 100, 200 and 400

CONLUSIONS (&FURTHER WORK)

The research has correspondence between the return distribution of a one-dimensional random walk and a two- dimensional random walk. I can conclude that similarity in the return distribution of the one-dimensional random walk and the two-dimensional random walk is due to the survival probability of the two-dimensional random walk eventually being zero and we know that the probability of a one-dimensional random walk is already one. We can also conclude that there is a difference in the distribution of the return times of a random walker and the distribution of the loops erased in LERW.

This research can be extended by using Self Avoiding Random Walks (SARW) which are random walks that do not self-intersect during their transversal and are in a different universality class from LERW.

REFERENCES

Rick Durrett, Probability: Theory and Examples Rick Durrett January 29, 2010, vol. 49, 2010.

J. Klafter and I.M. Sokolov, First steps in random walks: From tools to applications, OUP Oxford, 2011.

G.F. Lawler and V. Limic, Random walk: A modern introduction, Cambridge University Press, 2010.

Sidney Redner and J. R. Dorfman, A Guide to First-Passage Processes, vol. 70, 2002.

Jason Schweinsberg, The loop-erased random walk and the uniform spanning tree on the four-dimensional discrete torus, Probability Theory and Related Fields 144 (2009), no. 3-4, 319–370.

ACKNOWLEDGEMENTS

I would like to express my gratitude for Professor Gunnar Pruessner for supervising research project. I would like to thank the Nuffield Foundation for funding this research.

HR