walkr

by David Kane, Andy Yao

Abstract The `walkr` package samples points using random walks from the intersection of the N simplex with M hyperplanes. Mathematically, the sampling space is all vectors x that satisfy $Ax = b$, $\sum x = 1$, and $x_i \geq 0$. The sampling algorithms implemented are hit-and-run and Dikin walk, both of which are MCMC (Monte-Carlo Markov Chain) random walks. `walkr` also provide tools to examine and visualize the convergence properties of the random walks.

Introduction

A and b represent the system of equations that we have above. Specifically, A is a $M \times N$ matrix (M variables and N constraints), and b is a $M \times 1$ vector.

Mathematical Background of Sampling Space

In this section, we go through the mathematical background needed to understand the space from which we are sampling. Specifically, we go through a few examples as well as some linear algebra tricks. The reader does not need to read this section in order to use our package or understand our sampling algorithms. However, this section should help the reader understand better what the sample space is both geometrically and mathematically.

Definition: The N-dimensional unit simplex is described by:

$$x_1 + x_2 + x_3 + ... + x_n = 1$$
$$x_i \geq 0$$

Sampling space: simple 3D case

Let's begin with the simplest case – one linear constraint in 3 dimensional space.

$$x_1 + x_3 = 0.5$$

We can express this in terms of matrix equation $Ax = b$, where:

$$A = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}, \quad b = 0.5, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

In addition, we require the solution space to be intersected with the 3D simplex:

$$\sum x_i = 1$$
$$x_i \geq 0$$

In the following graph, we draw the intersection of the two. The orange equilateral triangle represents the 3D simplex, and the blue rectangle represents the plane $w_1 + w_3 = 0.5$. The intersection of the hyperplane (blue) with the simplex (orange) is the red line segment, which is our sampling space.

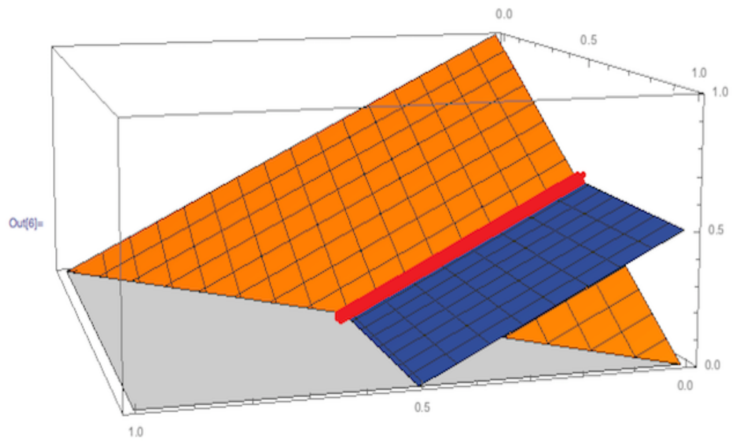

Figure 1: Intersection (red line) of Simplex and 2D hyperplane living in 3D space

Matrix Representation of Hyperplanes

Every hyperplane is described by one linear equation. Thus, a system of linear equations is the intersection of hyperplanes. In general, if we have M linear equations and N variables, then $Ax = b$ would look like:

$$A = \left.\underbrace{\begin{bmatrix} \quad & \ldots & \quad \end{bmatrix}}_{\text{N columns (variables)}}\right\}\text{M rows (constraints)}$$

Going from $Ax = b$ and the unit-simplex to $Ax \leq b$

Random Walks

Starting Points

Hit-and-run

Dikin Walk

Using `walkr`

Examining/Visualizing Results

Conclusion

Authors

David Kane
Managing Director
Hutchin Hill Capital
101 Federal Street, Boston, USA
dave.kane@gmail.com

Andy Yao
Mathematics and Physics
Williams College
Williamstown, MA, USA
ay3@williams.edu