Statistical Learning of Basketball Strategy: The Potential Field Approach

Yang Liu, Daniel Dinsdale, Seong-Hwan Jun, Creagh Briercliffe, and Jeffrey Bone
Department of Statistics, The University of British Columbia

Objectives

- SportVU technology has enabled us to track, nearly-continuously, players and ball movement in the NBA.
- We propose a statistical framework harnessing this data based on stochastic differential equations to model the movement of both the players and the ball.

Methods Overview

- We model the movement of the players and/or the basketball via a **stochastic differential equation (SDE)**.
- The expected movement direction and velocity for any location on the basketball court is described by a 3-dimensional surface called the **potential field (PF)**.
- The potential field approach was originally developed to analyze the movement of animals by learning tendencies in their velocity.

Potential Field Method

- The movement of the players and/or the basketball is modelled via the SDE:

$$\mathrm{d}r(t) = -\nabla H(r(t), t; \beta)\mathrm{d}t + \sigma(r(t), t; \theta)\mathrm{d}B(t).$$

- The function $H : \mathbb{R}^2 \to \mathbb{R}$ is the potential field.
- The negative gradient of the potential field, $-\nabla H(r(t), t; \beta)$, is interpreted as the expected velocity of the object at the location $r(t) \in \mathbb{R}^2$.
- The diffusion coefficient, $\sigma(r(t), t; \theta)$, captures the uncertainty in the movement of the object.

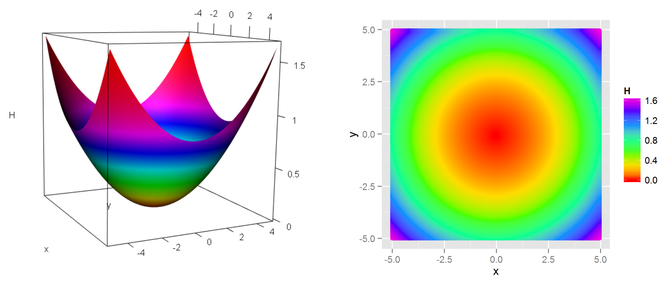


Figure: Left: a 3-D plot of a possible H function. Right: 2-D representation of this potential field using colours to indicate the values of H.

Model Fitting

- We use the Euler-Maruyama method (piecewise linear) to find approximate solution to the SDE.
- After the approximation, the parameters (β, θ) can be calculated via least squares.

Modelling the Ball or Player's Movement

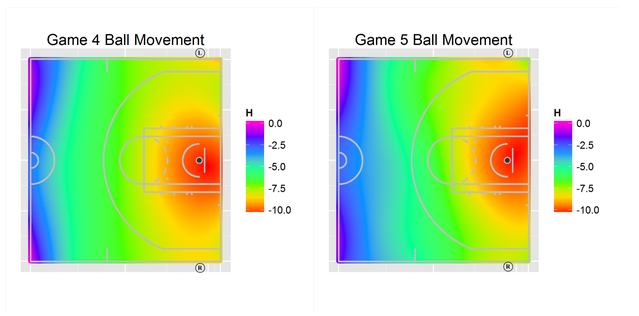
- For the ball/player movement, we consider a **tensor product spline** for the potential field,

$$H(r(t), t; \beta) = -\sum_{j=1}^{K}\sum_{k=1}^{K} \beta_{jk} S_j(x(t)) T_k(y(t)),$$

with univariate spline functions $S_1, \ldots, S_K$ and $T_1, \ldots, T_K$.
- It produces a flexible map that can capture the areas of the court that are favoured by the ball and players.

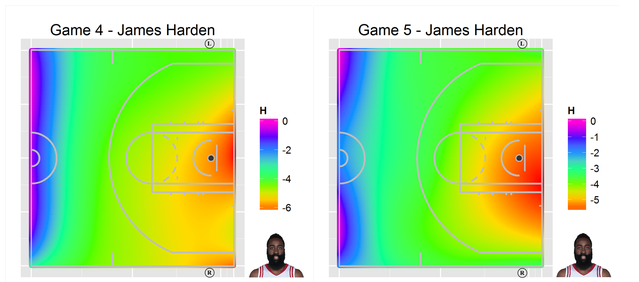
Analysis Results for Ball Movement

Our illustrations use data from the 2015 NBA Western conference semi-finals: Houston Rockets (HOU) vs Los Angeles Clippers (LAC). We use HOU's offensive plays for Games 4 (G4) and 5 (G5) as an example.



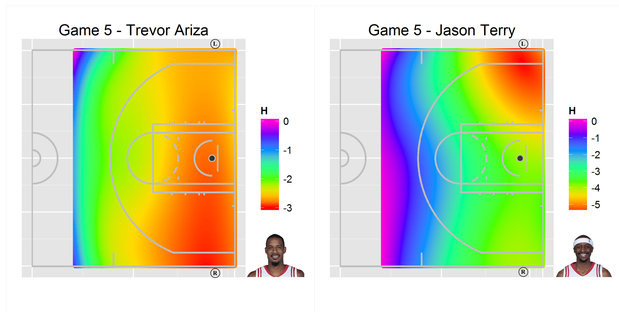
- High values around the half court line (blue) and low values towards the basket $\longrightarrow$ The ball is moving towards the basket (as it should);
- Relatively low values in the corners $\longrightarrow$ HOU likes to shoot corner threes after Harden's drives.
- Yellow areas in the corners change between games $\longrightarrow$ More 3's from the right corner in G4, compared to more left corner 3's in G5.

James Harden Movement



- Obvious discrepancy between G4 and G5: Harden moved less frequently to the right corner in G5. Presumably, Harden leaves the corner spaces to other shooters (see Ariza and Terry section).

Ariza and Terry's Movements in G5



- Both Terry and Ariza frequently go to the corners.
- Terry rarely attacks the basket; Ariza drives more regularly to the basket.
- The fields above were produced from the movement plots below:

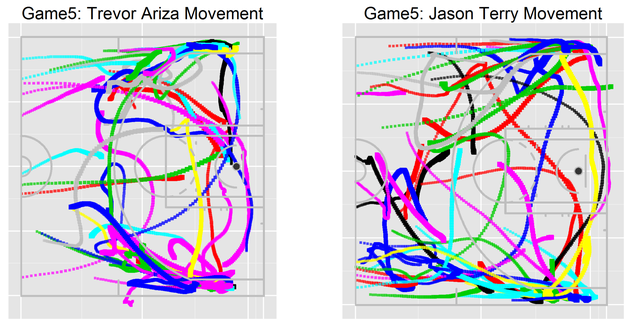


Figure: Raw movement plots for G5.

Experimental Work: PF with Covariates

We have experimented with including other players' locations as covariates in the potential field. These coefficients reflect the likeliness of one player passing the ball to another.

$$H(r(t), t; \beta) = \beta_0 \|r(t)_{\text{ball}} - r_{\text{basket}}\|_2^2 + \sum_{j=1}^{4} \beta_j \|r(t)_{\text{ball}} - r(t)_{\text{player}_j}\|_2^2,$$

G5	Ariza	Jones	Howard	Harden	Terry	Basket
Ariza	NA	-3.35	-1.51	1.90	1.03	5.38
Smith	1.14	NA	-1.23	-1.09	0.44	1.59
Howard	1.19	3.98	NA	-5.28	2.97	-6.27
Harden	0.72	-2.65	-3.70	NA	1.73	4.90
Terry	-1.21	-1.36	-3.04	2.29	NA	4.21

Conclusions

- The potential field approach is promising in summarizing and analyzing basketball strategies from large tracking datasets.
- SDEs are flexible and can be parameterized to analyze other sports.