A Probabilistic Characterization of Shark Movement Using Location Tracking Data

Samuel Ackerman, PhD (Advisers: Dr. Marc Sobel, Dr. Richard Heiberger)

Temple University Department of Statistical Science

ackerman@temple.edu

July 22, 2018
Data and motivation

- Our data come from a 2011 paper testing a VPS (acoustic telemetry) tracking transmitters implanted in 22 gray smooth-hound sharks in the tidal basin of a wetlands in Orange County, CA.
- Data are the 2-D coordinate locations of the sharks and the times they were recorded; the time gaps $\Delta t$ between observations were unequal.
- From domain knowledge, biologists believe shark behavior largely consists of either foraging (slow, meandering movement while feeding) or transiting (faster movement between feeding areas).
- Since the animal’s full trajectory—of which we have discrete realizations—depends on the movement, which depends on the behavior type, we can build a model to both estimate the movement path at times it is not observed, and infer behavior type from the movement.
- Inferring spatial distributions of behavior types is interesting as a way to identify ecologically-sensitive feeding areas.
Gray smooth-hound shark (*Mustelus Californicus*)

- Gray smooth-hound sharks are benthic (bottom-feeding) predators that typically grow to a length of about 4–5 feet.
- The Bolsa Chica wetlands have a southern outlet to the Pacific Ocean. Sharks occasionally leave the wetlands; their locations are then out of range of the receivers and thus unobserved.
State-space models (SSMs)

- **State-space models** (SSMs) are a general probabilistic method of using observations $y_t$ to sequentially model time ($t$)-evolving unobserved variables (‘states’) $x_t$.

- In robotics or tracking, $y_t$ are often sensor or location measurements.

- Observations $y_t$ are assumed to have some measurement error, and $x_t$ are noise-free values. Here, $x_t$ will include the animal’s (unobserved) true location $\zeta_t$ and movement (e.g., speed $v_t$ and direction angle) that result in the observed locations.

- Optional vector $\theta_t$ contains additional hypothesized parameters, here including the behavior type, denoted $\lambda_t$.

Sequentially update estimates of $x_t$ using new observed $y_t$:

- Prior on initial state: $x_0 \sim p(x_0)$
- State/dynamic equation: $x_t \sim \ell(x_t \mid x_{t-1}, \theta_t)$
- Measurement equation: $y_t \sim m(y_t \mid x_t, \theta_t)$
Illustrate principles of 2-D shark movement, which includes direction, using a simpler example of a robot moving along the 1-D real line.

At time $t$, its sensor measures position $z_t \in \mathbb{R}$ with error; the unknown true location at $t$ is $\zeta_t$.

Time gaps $\Delta_t$ are a constant $\Delta \Upsilon$, so $x_t$ are modeled on the same clock times as observations $y_t$. This is relaxed later for sharks.

Velocity $v_t \in \mathbb{R}$ between true $\zeta_t$ and $\zeta_{t+1}$ is random $\sim \mathcal{N}(\alpha, \sigma)$.

We will later allow the robot to have either a slow and fast mode at each $t$, analogous to the shark foraging/transiting. Thus $v_t$ will have a bimodal mixture distribution with two values of $\alpha$ to estimate. For simplicity, assume now $v_t$ is unimodal.

States $x_t = [\zeta_t \ v_t]$ are modeled sequentially by learning which values best predict the observed $y_t = z_{t+1}$. 
1-D robot SSM

State equation:
Velocities $v_t$ have single distribution with mean $\alpha$

$$x_t = \begin{bmatrix} \zeta_t \\ v_t \end{bmatrix} \sim \mathcal{N}_2 \left( \begin{bmatrix} \zeta_{t-1} + (\Delta \gamma)(v_{t-1}) \\ \alpha \end{bmatrix}, Q_t \right)$$

Measurement equation:

$$y_t = [z_{t+1}] \sim \mathcal{N} \left( \begin{bmatrix} \zeta_t + (\Delta \gamma)(v_t) \end{bmatrix}, R_t \right)$$
Particle filters (PFs) generate multiple ($N$) simulations (‘particles’) of the process (e.g. movement) variables $x_t = [\zeta_t \ \nu_t]$ to match the observations (locations) $y_t = z_{t+1}$.

The particle set $\{x_t^{(n)}\}_{n=1}^N$ empirically estimate a posterior distribution of guesses of the value of the true unobserved $x_t$.

At each step $t$, each particle $n$ calculates predictive density $d(\hat{y}_t | x_{t-1}^{(n)})$. Calculate a weight $w_t^{(n)}$ as $\propto d(y_t | x_{t-1}^{(n)})$, this density evaluated at the observed value $y_t$.

The higher $w_t^{(n)}$ is, the more likely particle $n$’s modeled movement $x_t^{(n)}$ is to resulted in the observed $y_t$, and so is a better guess.

Particle sets are resampled by the weights $\{w_t^{(n)}\}_{n=1}^N$, so particles with more likely estimates are favored for future prediction.
We use the **conditional dynamic linear model** (CDLM; Carvalho 2010, et al.) PF formulation for the 1-D robot. Here each particle \( n \) simulates its own values of \( \zeta^{(n)} \) and \( v_t^{(n)} \).

We use an animation to illustrate the 1-D PF. These shiny programs will be available to-be-released package animalEKF.

Click here for video of demonstration of package’s `cdlm_robot()` function.
The 1-D robot can have one of two speed modes (denoted by $\lambda_t$) at each time point $t$: slow (0) and fast (1), analogous to shark foraging and transiting.

Let $E(v_t \mid \lambda_t = \text{slow}) = \alpha_0$ and $E(v_t \mid \lambda_t = \text{fast}) = \alpha_1$, where $|\alpha_1| > |\alpha_0|$. By separately modeling movement for each behavior $\lambda_t$, we can infer which behavior (unobserved) occurred by seeing which movement type best predicts the observed location $y_t$.

For each particle $n$, $w_{t|k}^{(n)} \propto d(y_t \mid x_{t-1}^{(n)}, \lambda_t = k)$ is the predictive likelihood of observation $y_t$ if behavior $\lambda_t = k$ occurred. The $k$ for which $w_{t|k}^{(n)}$ is highest is the most probable behavior.

Particles are resampled by unconditional weights $w_{t}^{(n)} = \sum_k w_{t|k}^{(n)}$. Behavior $\lambda_t^{(n)} = k$ is drawn by the relative values of $w_{t|k}^{(n)}$ after resampling.
Two-behavior ($\lambda$) CDLM visualization

- The transition probabilities $p_{i \rightarrow j}$ between slow/fast modes $\lambda_t$ affect the conditional weights $w_{t|k}^{(n)}$.
- Click here for video of demonstration of package’s `cdlm_robot_twostate()` function of modeling a dual-mode trajectory.
In reality, unlike the robot example, telemetry observations $y_t$ typically do not occur at constant length $\Delta \Upsilon$ time gaps.

However, for proper parameter updates we want to model the movement $x_t$ at these equal time intervals. $\Delta \Upsilon$ should be large enough to represent a ‘distinct’ movement but short enough to allow finer resolution of trajectories (e.g., 2 minutes).

Denote state movement at these intervals by $x_c$ at times $\Upsilon_c$. For simplicity let $\Upsilon_0 = 0$ and $\Upsilon_c = \Upsilon_{c-1} + \Delta \Upsilon$.

For simplicity, sharks are modeled as moving in a straight line with constant speed $v_c$, depending on the behavior type $\lambda_c$, in each short time interval $(\Upsilon_c, \Upsilon_{c+1}]$. 
CDLM with interpolation

In any constant-length interval \((\gamma_c, \gamma_{c+1}]\), either there are or aren’t observations \(y_t\) recorded.

- If there are no observations, simulate movement from \(x_c\) to \(x_{c+1}\) by a single random draw of behavior \(\lambda_c \mid \lambda_{c-1}\), then drawing \(x_c \sim \ell(\cdot \mid x_{c-1}, \lambda_c = k)\) from the SSM equation.
- If at least one observation is in the interval, predict straight-line movement \(x_c \mid (x_{c-1}, \lambda_c = k)\) from \(\gamma_c\) to \(\gamma_{c+1}\) for each behavior \(k\).
- Let \(\{\Delta^*_t\}\) be the irregular time gaps from the interval \(\gamma_c\) to the first, and between each observation. For each \(y_t\), predict location \(\hat{y}_t \mid (\lambda_c = k, \Delta^*_t)\) along the straight line at the time of observation \(t\).
- Calculate behavior-conditional weights \(w_{c|k}^{(n)}\) by jointly how close observations \(\{y_t\}\) are to predictions \(\{\hat{y}_t\}\). The behavior \(k\) whose straight-line movement best jointly fits all observed locations (for which \(k w_{c|k}^{(n)}\) is highest) is the most likely.
Shark EKF with interpolation: observations $y_t$

- Linear movement from $\zeta_c$ to $\zeta_{c+1}$ for each behavior $\lambda_c \in \{0, 1\}$, to best fit observations $y_t = z_{t+1}$ and $y_{t+1} = z_{t+2}$ in the interval.
- Ellipses along lines between $\hat{\zeta}_c$ and $\hat{\zeta}_{c+1}$ centered at $(\hat{z}_{t+1}, \hat{z}_{t+2}) \mid \lambda_c$.
- Error ellipse sizes depend on time gaps $\Delta^*_t$ between observations.
Conclusion: are able to learn bimodal speed $v_t$ distribution indicative of two behavior mixture for both observed and synthetic trajectories. Also incorporate behavioral interactions between sharks to jointly model ‘schooling’ tendencies.

Use cross-validation to demonstrate that interpolation PF works better than simple non-statistical Euclidean connect-the-dots method for modeling omitted observations.

Adapt a model-fit measure like Bayes Factor to PFs to demonstrate that parameters match a given shark’s trajectory better than another’s.


