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 Δ_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 $\mathbf{y}_t$ to sequentially model time (t)-evolving unobserved variables ('states') $\mathbf{x}_t$.
- In robotics or tracking, $\mathbf{y}_t$ are often sensor or location measurements.
- Observations $\mathbf{y}_t$ are assumed to have some measurement error, and $\mathbf{x}_t$ are noise-free values. Here, $\mathbf{x}_t$ will include the animal's (unobserved) true location ζ_t and movement (e.g., speed v_t and direction angle) that result in the observed locations.
- Optional vector $\boldsymbol{\theta}_t$ contains additional hypothesized parameters, here including the behavior type, denoted λ_t.

Sequentially update estimates of $\mathbf{x}_t$ using new observed $\mathbf{y}_t$:

$$\begin{aligned} \text{Prior on initial state:} \quad \mathbf{x}_0 &\sim p(\mathbf{x}_0) \\ \text{State/dynamic equation:} \quad \mathbf{x}_t &\sim \ell(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \boldsymbol{\theta}_t) \\ \text{Measurement equation:} \quad \mathbf{y}_t &\sim m(\mathbf{y}_t \mid \mathbf{x}_t, \boldsymbol{\theta}_t) \end{aligned}$$

1-D robot SSM

- 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 ζ_t.
- Time gaps Δ_t are a constant Δ_Υ, so $\mathbf{x}_t$ are modeled on the same clock times as observations $\mathbf{y}_t$. This is relaxed later for sharks.
- Velocity $v_t \in \mathbb{R}$ between true ζ_t and ζ_{t+1} is random $\sim \mathcal{N}(\boldsymbol{\alpha}, \sigma)$.
- 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 $\boldsymbol{\alpha}$ to estimate. For simplicity, assume now v_t is unimodal.
- States $\mathbf{x}_t = \begin{bmatrix} \zeta_t & v_t \end{bmatrix}$ are modeled sequentially by learning which values best predict the observed $\mathbf{y}_t = z_{t+1}$.

1-D robot SSM

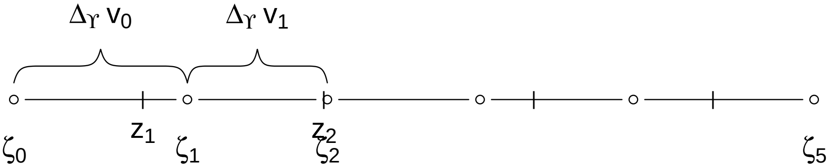
State equation:

Velocities v_t have single distribution with mean $\boldsymbol{\alpha}$

$$\mathbf{x}_t = \begin{bmatrix} \zeta_t \\ v_t \end{bmatrix} \sim \ell(\quad \mathbf{x}_t \,|\mathbf{x}_{t-1}, \boldsymbol{\theta} = \boldsymbol{\alpha} \quad) = \mathcal{N}_2 \left(\begin{bmatrix} \zeta_{t-1} + (\Delta_\Upsilon)(v_{t-1}) \\ \boldsymbol{\alpha} \end{bmatrix}, \mathbf{Q}_t \right)$$

Measurement equation:

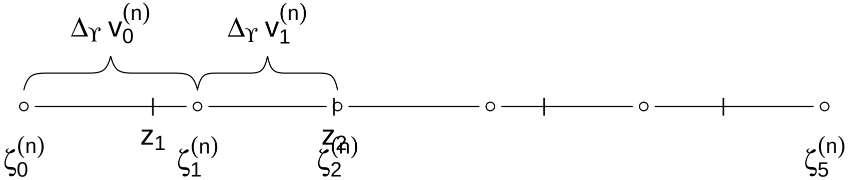
$$\mathbf{y}_t = \begin{bmatrix} z_{t+1} \end{bmatrix} \sim m(\quad \mathbf{y}_t \,|\mathbf{x}_t \quad) = \mathcal{N} \left(\begin{bmatrix} \zeta_t + (\Delta_\Upsilon)(v_t) \end{bmatrix}, \mathbf{R}_t \right)$$



Particle filters (PFs)

- Particle filters (PFs) generate multiple (N) simulations ('particles') of the process (e.g. movement) variables $\mathbf{x}_t = \begin{bmatrix} \zeta_t & v_t \end{bmatrix}$ to match the observations (locations) $\mathbf{y}_t = z_{t+1}$.
- The particle set $\{\mathbf{x}_t^{(n)}\}_{n=1}^N$ empirically estimate a posterior distribution of guesses of the value of the true unobserved $\mathbf{x}_t$.
- At each step t, each particle n calculates predictive density $d(\hat{\mathbf{y}}_t \mid \mathbf{x}_{t-1}^{(n)})$. Calculate a weight $w_t^{(n)}$ as $\propto d(\mathbf{y}_t \mid \mathbf{x}_{t-1}^{(n)})$, this density evaluated at the observed value $\mathbf{y}_t$.
- The higher $w_t^{(n)}$ is, the more likely particle n's modeled movement $\mathbf{x}_t^{(n)}$ is to resulted in the observed $\mathbf{y}_t$, 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 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.

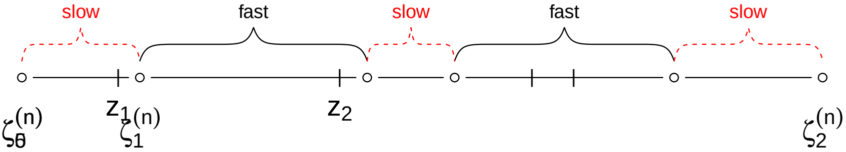


1-D robot with two 'behaviors'

- The 1-D robot can have one of two speed modes (denoted by λ_t) at each time point t: slow (0) and fast (1), analogous to shark foraging and transiting.
- Let $\mathrm{E}(v_t \mid \lambda_t = \text{slow}) = \boldsymbol{\alpha_0}$ and $\mathrm{E}(v_t \mid \lambda_t = \text{fast}) = \boldsymbol{\alpha_1}$, where $|\alpha_1| > |\alpha_0|$. By separately modeling movement for each behavior λ_t, we can infer which behavior (unobserved) occurred by seeing which movement type best predicts the observed location $\mathbf{y}_t$.
- For each particle n, $w_{t|\boldsymbol{k}}^{(n)} \propto d(\mathbf{y}_t \mid \mathbf{x}_{t-1}^{(n)}, \boldsymbol{\lambda_t = k})$ is the predictive likelihood of observation $\mathbf{y}_t$ if behavior $\boldsymbol{\lambda_t = k}$ occurred. The $\boldsymbol{k}$ for which $w_{t|\boldsymbol{k}}^{(n)}$ is highest is the most probable behavior.
- Particles are resampled by unconditional weights $w_t^{(n)} = \sum_k w_{t|\boldsymbol{k}}^{(n)}$. Behavior $\lambda_t^{(n)} = \boldsymbol{k}$ is drawn by the relative values of $w_{t|\boldsymbol{k}}^{(n)}$ after resampling.

Two-behavior (λ) CDLM visualization

- The transition probabilities $p_{i \to j}$ between slow/fast modes λ_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.



CDLM with interpolation

- In reality, unlike the robot example, telemetry observations $\mathbf{y}_t$ typically do not occur at constant length Δ_Υ time gaps.
- However, for proper parameter updates we want to model the movement $\mathbf{x}_t$ at these equal time intervals. Δ_Υ 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 $\mathbf{x}_c$ at times Υ_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 λ_c, in each short time interval $(\Upsilon_c, \Upsilon_{c+1}]$.

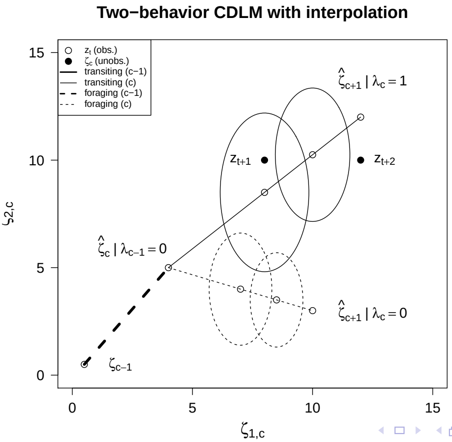
CDLM with interpolation

In any constant-length interval $(\Upsilon_c, \Upsilon_{c+1}]$, either there are or aren't observations $\mathbf{y}_t$ recorded.

- If there are no observations, simulate movement from $\mathbf{x}_c$ to $\mathbf{x}_{c+1}$ by a single random draw of behavior $\lambda_c \mid \lambda_{c-1}$, then drawing $\mathbf{x}_c \sim \ell(\cdot \mid \mathbf{x}_{c-1}, \boldsymbol{\lambda_c = k})$ from the SSM equation.
- If at least one observation is in the interval, predict straight-line movement $\mathbf{x}_c \mid (\mathbf{x}_{c-1}, \boldsymbol{\lambda_c = k})$ from Υ_c to Υ_{c+1} for each behavior $\boldsymbol{k}$.
- Let $\{\Delta_t^*\}$ be the irregular time gaps from the interval Υ_c to the first, and between each observation. For each $\mathbf{y}_t$, predict location $\hat{\mathbf{y}}_t \mid (\boldsymbol{\lambda_c = k}, \Delta_t^*)$ along the straight line at the time of observation t.
- Calculate behavior-conditional weights $w_{c|\boldsymbol{k}}^{(n)}$ by jointly how close observations $\{\mathbf{y}_t\}$ are to predictions $\{\hat{\mathbf{y}}_t\}$. The behavior k whose straight-line movement best jointly fits <u>all</u> observed locations (for which k $w_{c|\boldsymbol{k}}^{(n)}$ is highest) is the most likely.

Shark EKF with interpolation: observations $\mathbf{y}_t$

- Linear movement from ζ_c to ζ_{c+1} for each behavior $\lambda_c \in \{0, 1\}$, to best fit observations $\mathbf{y}_t = \mathbf{z}_{t+1}$ and $\mathbf{y}_{t+1} = \mathbf{z}_{t+2}$ in the interval.
- Ellipses along lines between $\hat{\zeta}_c$ and $\hat{\zeta}_{c+1}$ centered at $(\hat{\mathbf{z}}_{t+1}, \hat{\mathbf{z}}_{t+2}) \mid \lambda_c$.
- Error ellipse sizes depend on time gaps Δ_t^* between observations.



Conclusions and continuations

- 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.

Selected references

1. Ackerman, Samuel. “A probabilistic characterization of shark movement using location tracking data.” Temple University dissertation, 2018.
2. Ackerman, Samuel. “`animalEKF`: Simulation and estimation of EKFs for animal movement.” R package, 2018.
3. Carvalho, Carlos M., Johannes, Michael S., Lopes, Hedibert F., and Nicholas G. Polson. “Particle learning and smoothing.” Statistical Science, 2010.
4. Doucet, Arnaud. “On sequential simulation-based methods for Bayesian filtering.” Technical report, University of Cambridge. 1998
5. Espinoza, Mario, Farrugia, Thomas J., and Christopher G. Lowe. “Habitat use, movements and site fidelity of the gray smooth-hound shark in a newly restored southern California estuary.” Journal of Experimental Marine Biology and Ecology, 2011.
6. Särkkä, Simo. “Bayesian Filtering and Smoothing.” Cambridge University Press, 2013.