A Bayesian Penalized Hidden Markov Model for Colony-Level Interactions

Meridith Bartley, Ephraim Hanks, David Hughes Departments of Statistics and Entomology

C. pennsylvanicus Ants

- Ideal system for studying social organisms in a controlled environment.
- Colonies housed by and experiments conducted by the Hughes Lab at PSU.

Ant Interaction Data

- Oral exchange of nutrients called **trophallaxis**, allows transfer of diseases.
- 4 hours of monitoring of the black carpenter ant, *Camponotus pennsylvanicus*.
- Feeding interactions recorded at 1 second intervals.
- 79 ants engaged in 246 trophallaxis events.
- Observed "pulses" of high trophallaxis rates.

The Problem

High temporal resolution of the observations results in overfit stochastic process that switches at nearly every observation.

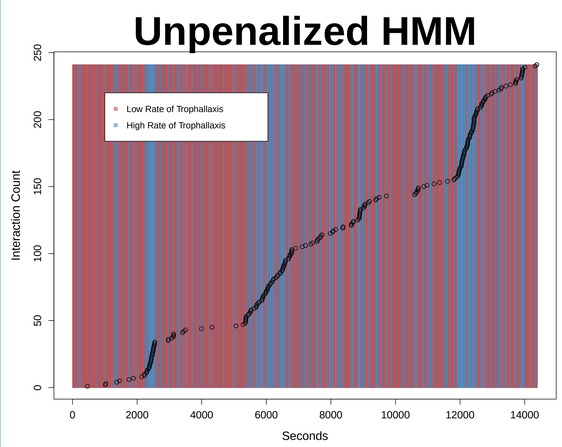
References

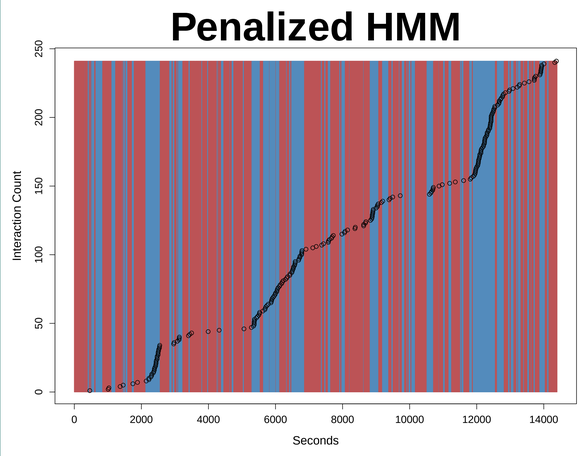
[1] Lauren E. Quevillon, Ephraim M. Hanks, Shweta Bansal, and David P. Hughes. Social, spatial, and temporal organization in a complex insect society. *Sci. Rep.*, 5:13393, 2015.

[2] Meridith L. Bartley, Ephraim M. Hanks, and David P. Hughes. Penalizing stochastic process models, with application to colony-level behavioral switching. In prep.

Penalized HMM for Trophallaxis Data

An approach for using high-resolution feeding data (N_t) to infer colony-level behavior states (X_t) within a Hidden Markov Model framework with a Bayesian penalty on the stochastic process.





DATA N_t: # of feeding events starting at time t.

$$N_t = N_{Lt} + \tilde{N}_{Ht}\mathbb{I}_{\{X_t=H\}}$$

OBSERVATIONAL MODEL

$$N_{Lt} \sim \text{Pois}(\lambda_L) \quad \tilde{N}_{Ht} \sim \text{Pois}(\tilde{\lambda}_H)$$

$$\Rightarrow N_t \sim \text{Pois}(\lambda_L + \tilde{\lambda}_H\mathbb{I}_{X_t=H})$$

PRIORS

$$\lambda_L \sim \text{Gamma}(a,b) \quad \tilde{\lambda}_H \sim \text{Gamma}(c,d)$$

LATENT STATE PROCESS X_t: state of trophallaxis rates at time t. $X_t \in \{L, H\}$

$$P(X_{t+1} = j | X_t = i) = \mathbf{P}_{ij}; \quad i,j \in \{L,H\}$$

Discrete time transition matrix ($\mathbf{P}$) from continuous time rate matrix ($\mathbf{R}$)

$$\mathbf{R} = \begin{pmatrix} 0 & \gamma_{LH} \\ \gamma_{HL} & 0 \end{pmatrix} \rightarrow \mathbf{P} = \begin{pmatrix} 1-p_{LH} & \gamma_{LH}e^{-\gamma_{LH}} \\ \gamma_{HL}e^{-\gamma_{HL}} & 1-p_{HL} \end{pmatrix}$$

PENALIZED STOCHASTIC PROCESS PRIORS

Ridge Prior

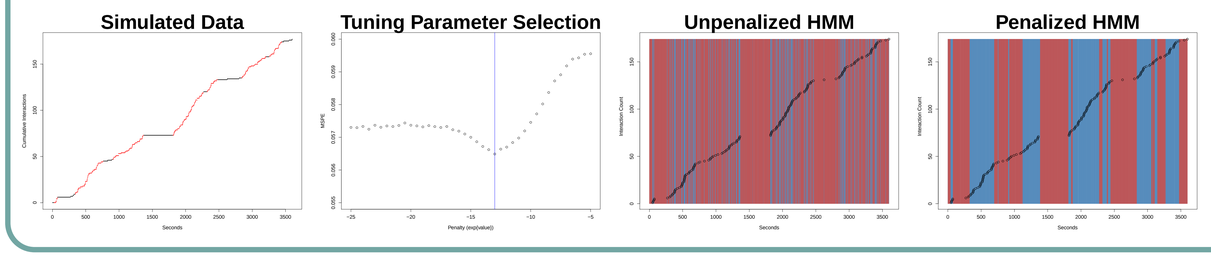
$$\gamma_{ij} \sim \text{H. Norm}(0,\tau)$$

LASSO Prior

$$\gamma_{ij} \sim \text{Exp}\left(\frac{1}{\tau}\right)$$

$$i,j \in \{L,H\}$$

Simulation Study



Penalized Markov Chain

Posterior Mode of MC:

$$\arg\max[\gamma_{LH}, \gamma_{HL} | X_1, \ldots, X_T]$$

RIDGE PRIOR

$$\arg\max\left\{\sum_{t=1}^{T} \log \mathbf{P}_{X_t, X_{t+1}} - \frac{1}{\tau}(\gamma_{LH}^2 + \gamma_{HL}^2)\right\}$$

LASSO PRIOR

$$\arg\max\left\{\sum_{t=1}^{T} \log \mathbf{P}_{X_t, X_{t+1}} - \frac{1}{\tau}(\gamma_{LH} + \gamma_{HL})\right\}$$

One Step Ahead Prediction

Models compared via one step ahead prediction.

$$\hat{Y}_{t+1} = E[Y_{t+1}|Y_{1:t}] = \sum_{k\in\{L,H\}} \lambda_k p_{ik} P(X_t = j | Y_{1:t})$$

$$\text{ERROR} = \textstyle\sum_t (\hat{Y}_t - Y_t)^2$$

Results

- $\hat{\lambda}_L = .6$ events per minute
- $\hat{\lambda}_H = 1.2$ events per minute
- Colony is in High trophallaxis state 40% of time observed.

Future Work

- Adding covariates (e.g. foragers entering the chamber) into penalized model makes testing biological hypotheses possible.
- Examining ant data under 3 (or more) state Penalized HMM

Acknowledgements

Funding from NSF EEID 1414296.