# Penalized estimation of hidden Markov models for behavioral state switching

Meridith L. Bartley
Twitter: @AlwaysScientist

In collaboration with: Ephraim M. Hanks, David A. Hughes

July 5, 2018

Pennsylvania State University

Funding Source: NSF EEID 1414296

# Hidden Markov Models

- HMMs are useful for many things (animal movement and behavior)
- Data are getting easier and easier to collect
- This can lead to problems.

# Study Species: *Camponotus pennsylvanicus*

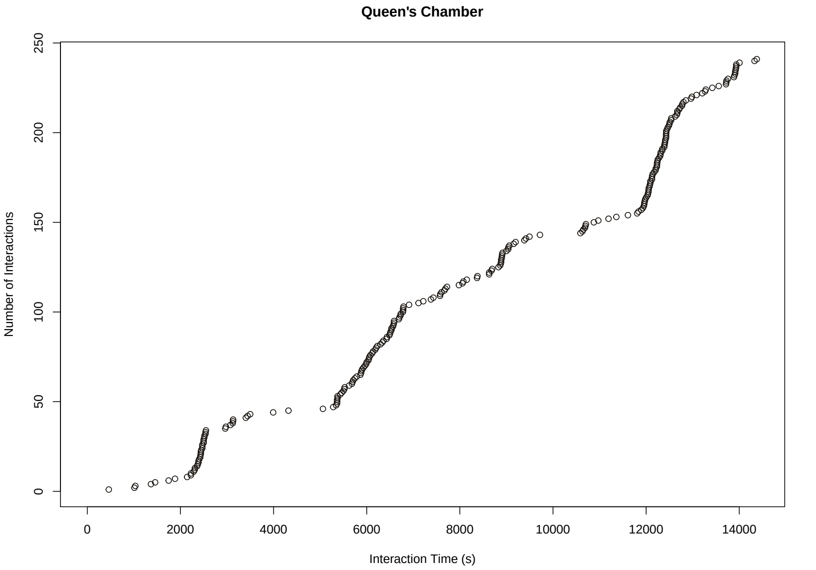
**Carpenter Ant**

- Ideal system for studying social organisms in a controlled environment.
- Colonies housed by and experiments conducted by the Hughes Lab at PSU.
- Oral exchange of nutrients, **trophallaxis**, allows transfer of diseases.

# Ant Interaction Data

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

# Cumulative Interactions over 4 Hours



# Poisson HMM

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

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

OBSERVATIONAL MODEL

$$N_t | X_t \sim \mathrm{Pois}(\lambda_{X_t}).$$

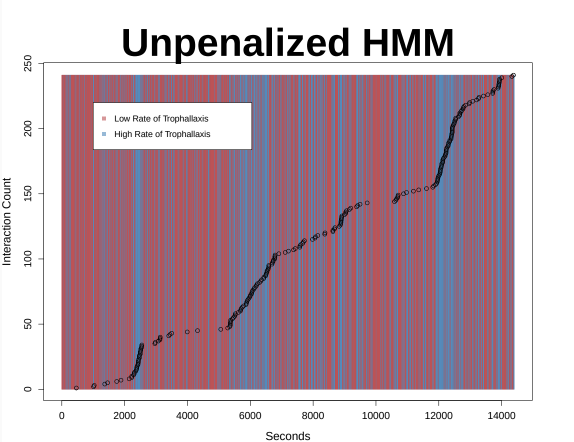
PRIORS

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

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

$$\mathbf{p}'_{\ell} \sim \mathrm{Dirichlet}(\boldsymbol{\theta}_{\ell})$$

# The Problem




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

# Data Augmentation

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

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

Data Augmentation: (Maintain Identifiability)

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

OBSERVATIONAL MODEL

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

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

# Penalized HMM for Trophallaxis Data

PRIORS

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

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

**Approach:** Model discrete time transition matrix ($\mathbf{P}$) as a function of continuous time rate parameters.

CONTINUOUS TIME RATE MATRIX

$$\mathbf{R} = \begin{pmatrix} 0 & \gamma_{LH} \\ \gamma_{HL} & 0 \end{pmatrix}$$

# Converting between DTMC and CTMC

$$p_{ij} = \Pr(\text{Colony remains in state } i \text{ for 1 second before switching to state } j)$$
$$= \Pr(i \to j \text{ after 1 second}) \Pr(\text{wait time in state } i \text{ is 1 second})$$

$$p_{ij} = \frac{\gamma_{ij}}{\gamma_{i\cdot}} \gamma_{i\cdot} e^{-\gamma_{i\cdot} 1} = \gamma_{ij} e^{-\gamma_{i\cdot}}$$

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

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

# Penalized HMM for Trophallaxis Data

PENALIZED STOCHASTIC PROCESS PRIORS

Ridge Prior

$$\gamma_{ij} \sim \text{H. Norm}(0, \tau) \quad i, j \in \{L, H\}$$

**Posterior Mode of MC:**

$$\arg\max[\gamma_{LH}, \gamma_{HL} | X_1, \ldots, X_T] = \arg\max \left\{ \sum_{t=1}^{T} \log \mathbf{P}_{X_t, X_{t+1}} - \frac{1}{\tau}(\gamma_{LH}^2 + \gamma_{HL}^2) \right\}$$

# Full Penalized HMM (2-state)

OBSERVATIONAL MODEL

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

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

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

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

PRIORS

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

PENALIZED STOCHASTIC PROCESS PRIORS

Ridge Prior

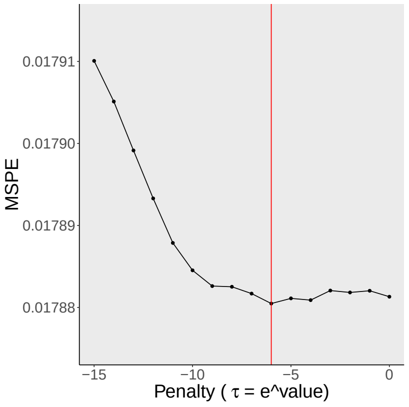
$$\gamma_{ij} \sim \text{H. Norm}(0, \tau) \quad i, j \in \{L, H\}$$

# One Step Ahead Prediction

Choosing $\tau$ via one step ahead prediction.

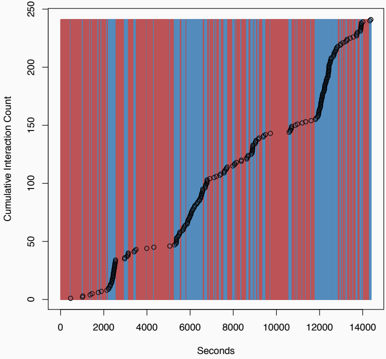
$$MSPE(\tau) = E\left[\sum_t (\hat{N}_t - N_t)^2 | \boldsymbol{N}\right]$$

$$\hat{N}_{t+1} = E\left[N_{t+1} | X_t\right] = \sum_{k=1}^{n} \lambda_k P_{X_t,k}$$

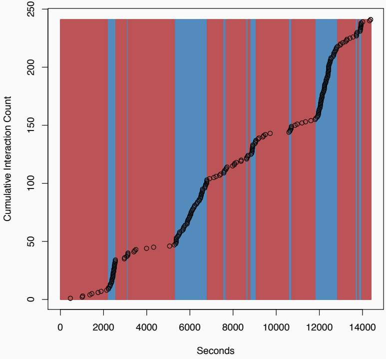


# Results - 2 State HMM

## UNPENALIZED HMM

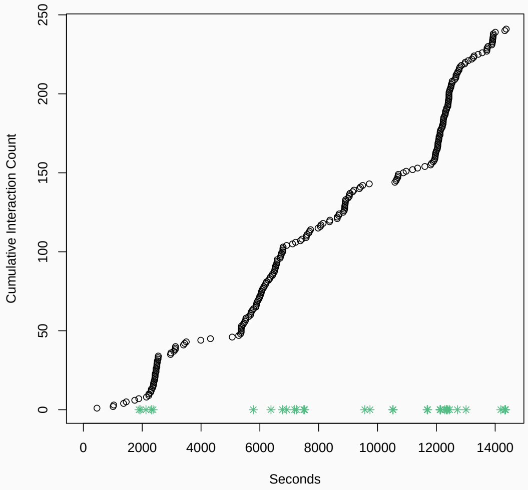



## PENALIZED HMM




- $\hat{\lambda}_L = 0.34$ events per minute.
- $\hat{\lambda}_H = 3.0$ events per minute.
- Colony is in High trophallaxis state 25% of time observed.

# Adding Biological Covariate(s)




Green symbol denotes time at which an ant enters into the nest chamber.

# Penalized HMM with Covariate

Let $\boldsymbol{w}_t$ be a vector of covariates (not including an intercept). Then let

$$\gamma_{ijt} = e^{\mu_{ij} + \boldsymbol{w}_t' \boldsymbol{\beta}_{ij}}.$$

In the 2-state setting,

$$\log(\gamma_{ijt}) = \mu_{ij} + \beta_{ij} \left( \frac{1}{w_t^{\alpha} + 1} \right)$$

$$\gamma_{LHt} = e^{\mu_{LH}} e^{\beta_{LH} \left( \frac{1}{w_t^{\alpha} + 1} \right)}$$

$$\gamma_{HLt} = e^{\mu_{HL}} e^{\beta_{HL} \left( \frac{1}{w_t^{\alpha} + 1} \right)}$$

where $w_t$ is the time since an ant has entered the nest chamber.

# Penalized HMM with Covariate

PENALIZED STOCHASTIC PROCESS PRIOR

Ridge Prior

$$e^{\mu_{ij}} c(\beta_{ij}) \sim \text{H. Norm}(0, \tau) \Rightarrow e^{\mu_{ij}} | \beta_{ij} \sim \text{H. Norm}\left(0, \frac{\tau}{c(\beta_{ij})^2}\right)$$

$$\begin{aligned} c(\beta_{ij}) &\propto \text{E}[\# i \rightarrow j | \mu_{ij}] \\ &\propto \frac{1}{T} \sum_{t=1}^{T} \gamma_{ijt} \\ &= \frac{1}{T} \sum_{t=1}^{T} e^{\beta_{ij}\left(\frac{1}{w_t^{\alpha+1}}\right)} \end{aligned}$$

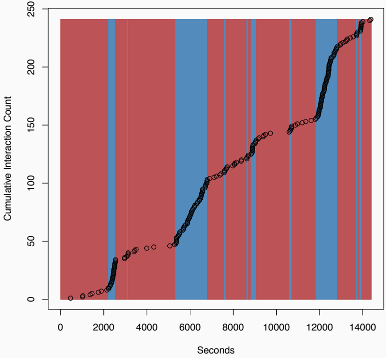
OTHER PRIORS

$$\beta_{ij} \sim \text{N}(0, 10)$$

$$\alpha \sim \text{N}(0, 10)$$

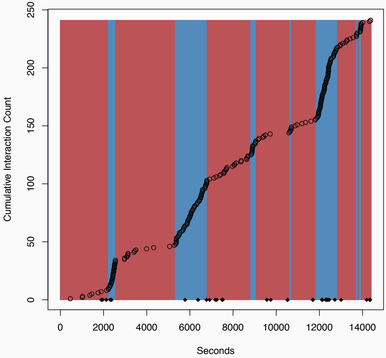
# Results - 2 State Penalized HMM with Covariate

## 2-STATE PENALIZED HMM




- $\hat{\lambda}_{LH} = 0.37$ events per minute.
- $\hat{\lambda}_{HL} = 3$ events per minute.

## 2-STATE PENALIZED HMM WITH COVARIATE




- $e^{\hat{\beta}_{LH}} = -0.33$
- $e^{\hat{\beta}_{HL}} = -0.29$

# Summary

- Basic HMM model is overfitting high-frequency data.
- Penalizing transition rates allows for 'smoothing' of the stochastic process.
- Adding covariates into penalized model makes testing biological hypotheses possible.
- Ants entering the chamber found not to be significant covariate.