Environmetrics 00, 1–7

DOI: 10.1002/env.XXXX

This is an Example of an Article Title

M. L. Bartley[a∗] and E. Hanks[a]

Summary:

Keywords: Bayesian estimation, hidden Markov model, penalization, covariate, ant trophallaxis

1. INTRODUCTION

2. POISSON HIDDEN MARKOV MODELS

2.1. Basic Model

To start, we developed a simple, baseline model for starting times of ant trophallaxis interactions. We expect, based on prior observations, that this food sharing behavior will differ according to changes in the surrounding behavioral state. We might expect, for example, that forager ants returning to the nest with nutrients might induce a flurry of high feeding activity, followed by a period of lower interaction rates. This combination of observed ant trophallaxis interactions within a colony and unobserved periods of relatively high or low rates of colony-level trophallaxis motivates the used of a hidden Markov model. For a basic n-state Markov-switching model the latent state process, here colony-level feeding activity denoted by X_t, the state at time t will only depend on what state the colony as in at time $t-1$. The state transition probabilities for this process are denoted

[a] *Department of Statistics, PA 16801, USA*

[b] *Building, Institute, Street Address, City, Code, Country*

[∗] *Correspondence to: M. L. Bartley, Department of Statistcs, The Pennsylvania State University, University Park, PA, U.S.A. E-mail: bartley@psu.edu*

by $Pr(X_{t+1} = j|X_t = i) = p_{ij}$ for $i, j = 1, \dots, n$. These probabilities are summarized in the transition probability matrix, $\mathbb{P}$, with p_{ij} being the entry in row i and column j. Each row of $\mathbb{P}$, denoted by $\underline{p}_\ell$ for each ℓ in $1, \dots, n$ is distributed as

$$\underline{p}_\ell \sim \text{Direchlet}(\underline{\theta}_\ell) \tag{1}$$

where each row of θ is a real numbered vector that acts as weights of the resulting probabilities.

In the application to the ant colony interactions, we will consider models with $n = 2$, corresponding to relatively ”high” and ”low” periods of trophallaxis interactions beginning. With this restriction to two states, we may expect that the transition probabilities of staying in each state are relatively high, while transition probabilities of changing states are relatively low. Similarly, in ant colonies, we would expect the behaviors to persist for some time before switching states. Dependent on our non-observed Markov chain, X_t, we may determine the distribution of the observable level of the model.

Here, number of ant pairs starting a trophallaxis interactions are discretely valued, and non-negative. Thus, we consider Poisson distributions for the state-dependent process, denoted by N_t. As we assume individual pairings of ants engaging in trophallaxis, we consider the number of pairs beginning this interaction at time t to not depend on the number at $t - 1$, or indeed any $t - \Delta t$. Here we consider only two underlying states so we may split this process by state such that

$$N_t = N_{Lt} + \tilde{N}_{Ht}\mathbb{I}\{X_t = H\} \tag{2}$$

where N_{Lt} becomes a baseline number of interactions starting at time t, and N_{ht} is the increase in interactions starting while the ant colony is in behavioral state, $X_t = H$. This start process is only dependent on the underlying state process and so we may parameterize the Poisson distribution in terms of a interaction starting rate, λ such that

$$N_t|X_t = L \sim \text{Poisson}(\lambda_L) \tag{3}$$

$$N_t|X_t = H \sim \text{Poisson}(\lambda_L + \tilde{\lambda}_H) \tag{4}$$

Note that λ_i depends on the behavioral state that the ant colony is in at time t. We estimate these rates within our model, both λ parameters are distributed as gamma distributions with separate hyperparameters such that

$$\lambda_L \sim \text{Gamma}(a, b) \tag{5}$$

$$\tilde{\lambda}_H \sim \text{Gamma}(c, d) \tag{6}$$

Conditioning on some starting colony-level behavioral state, $X_1 = x_1$, the likelihood of this model is

$$\begin{aligned}\mathscr{L} &= [\{X\}, \{\lambda_L, \tilde{\lambda}_H\}, \mathbb{P}|\{N_L, N_H\}] \\ &= \prod_{t=1}^{T} [N_{Lt}|X_t, \lambda_L][N_{Ht}|X_t = H, \lambda_L, \tilde{\lambda}_H][X_1|\mathbb{P}][X_t|X_{t-1}, \mathbb{P}][\lambda_L][\tilde{\lambda}_H][\underline{p}_\ell]\end{aligned} \tag{7}$$

Fitting this model...

Gibbs Sampler for Basic HMM Model

1. Initialize $\{\underline{X}^{(0)}, \lambda_L^{(0)}, \tilde{\lambda}_H^{(0)}, \mathbb{P}^{(0))}\}$
2. For i = 1 to n.mcmc
   (a) Calculate $\lambda_H = \lambda_L + \tilde{\lambda}_H$
   (b) Sample $X_1^{(i)} \sim \text{Mult}(1, \underline{\eta}_1)$
   (c) Sample $X_t^{(i)} \sim \text{Mult}(1, \underline{\eta}_t)$ where $\underline{\eta}_k = \lambda_k \exp^{-\lambda_k} \delta_k \mathbb{P}_{k,X_2^{(0)}}$
   (d) Sample $X_T^{(i)} \sim \text{Mult}(1, \underline{\eta}_T)$
   (e) Record m_{ij} = number of times states transition from state $i \rightarrow j$
   (f) Sample $\underline{p}_L^{(i)} \sim \text{Dir}(\underline{\theta}_L + \underline{m}_L)$
   (g) Sample $\underline{p}_H^{(i)} \sim \text{Dir}(\underline{\theta}_H + \underline{m}_H)$
   (h) Augment data by splitting N_t into $\{N_{Lt}, N_{Ht}\}$

$$\left\{ \begin{pmatrix} N_{Lt} \\ N_{Ht} \end{pmatrix} \right\} \sim \begin{cases} \begin{pmatrix} N_t \\ 0 \end{pmatrix} & X_t = L \\ \text{Mult}(N_t, (\lambda_L, \tilde{\lambda}_H)) \rightarrow \begin{pmatrix} N_{Lt} \\ N_{Ht} \end{pmatrix} & X_t = H \end{cases} \tag{8}$$

   (i) Sample $\lambda_L^{(i)} \sim \text{Gam}\left(\sum_{t=1}^{T} N_{Lt} + a, T + b \right)$
   (j) Sample $\tilde{\lambda}_H^{(i)} \sim \text{Gam}\left(\sum_{t:X_t=H} N_{Ht} + c, \sum \underline{m}_H + d \right)$

2.2. Penalized Model

The data show clear switches between fast and slow modes of trophallaxis; however, fitting a standard hidden Markov model (HMM) results in an estimated hidden state process that is overfit to this high resolution data, as the state process fluctuates an order of magnitude more quickly than is biologically reasonable. To counter this overfitting, we propose a novel approach for penalized estimation of HMMs through a Bayesian ridge prior on the state transition rates. This penalty induces smoothing, limiting the rate of state switching to ensure more biologically feasible results.

Now, rather than sampling the transitions probabilities directly as in the basic model, these probabilities are now a function of state switching rates denoted by $\underline{\gamma} = (\gamma_{LH}, \gamma_{HL})$.

$$\mathbb{P}^{(\Delta t)} = \begin{pmatrix} 1 - p_{LH} & \gamma_{LH} \exp\{-\gamma_{LH}\Delta t\} \\ \gamma_{HL} \exp\{-\gamma_{HL}\Delta t\} & 1 - p_{HL} \end{pmatrix} \tag{9}$$

With the distribution of $\underline{\gamma}$ we may induce smoothing of these state switching rates. Two possibilities were considered for this distribution.

$$\underline{\gamma} \sim \text{H. Norm}(0, \sigma^2) \tag{10}$$

$$\underline{\gamma} \sim \text{Exp}(\tau) \tag{11}$$

Both σ^2 and τ act as penalty parameters when estimating $\underline{\gamma}$.

The likelihood of the above HMM with penalized estimation is now

$$\begin{aligned} \mathscr{L} &= [\{X\}, \{\lambda_L, \tilde{\lambda}_H\}, \{\gamma_{LH}, \gamma_{HL}\} | \{N_L, N_H\}] \\ &= \prod_{t=1}^{T} [N_{Lt} | X_t, \lambda_L][N_{Ht} | X_t = H, \lambda_L, \tilde{\lambda}_H][X_1 | \mathbb{P}(\underline{\gamma})][X_t | X_{t-1}, \mathbb{P}(\underline{\gamma})][\lambda_L][\tilde{\lambda}_H][\underline{\gamma}] \end{aligned} \tag{12}$$

Fitting this model...

Metropolis-Hasting MCMC for Penalized HMM Model

1. Initialize $\{\underline{X}^{(0)}, \lambda_L^{(0)}, \tilde{\lambda}_H^{(0)}, \gamma_L^{(0))}, \gamma_H^{(0)}\}$
2. For i = 1 to n.mcmc
   (a) Calculate $\mathbb{P}^{(i-1)} = f(\underline{\gamma}^{(i-1)})$
   (b) Calculate $\lambda_H^{(i-1)} = \lambda_L^{(i-1)} + \tilde{\lambda}_H^{(i-1)}$
   (c) Propose $\log(\underline{\gamma})^* \sim \mathrm{N}(\log(\gamma^{(i-1)}), \Sigma)$
   (d) Calculate $\underline{\gamma}^* = \exp(\log(\underline{\gamma}^{(*)}))$
   (e) Sample $u \sim \mathrm{Unif}(0, 1)$
   (f) If $u < A(\underline{\gamma}^{(i-1)}, \underline{\gamma}^*) = \min\left\{1, \frac{p(\underline{\gamma}^*|\underline{\gamma}^{(i-1)})}{p(\underline{\gamma}^{(i-1)}|\underline{\gamma}^*)}\right\}$
   $\underline{\gamma}^{(i)} = \underline{\gamma}^*$
   Else,
   $\underline{\gamma}^{(i)} = \underline{\gamma}^{(i-1)}$
   (g) Sample $X_1^{(i)} \sim \mathrm{Mult}(1, \underline{\eta}_1)$
   (h) Sample $X_t^{(i)} \sim \mathrm{Mult}(1, \underline{\eta}_t)$ where $\underline{\eta}_k = \lambda_k \exp^{-\lambda_k} \delta_k \mathbb{P}_{k, X_2^{(0)}}$
   (i) Sample $X_T^{(i)} \sim \mathrm{Mult}(1, \underline{\eta}_T)$
   (j) Record m_{ij} = number of times states transition from state $i \to j$
   (k) Augment data by splitting N_t into $\{N_{Lt}, N_{Ht}\}$

$$\left\{\begin{pmatrix} N_{Lt} \\ N_{Ht} \end{pmatrix}\right\} \sim \begin{cases} \begin{pmatrix} N_t \\ 0 \end{pmatrix} & X_t = L \\ \mathrm{Mult}(N_t, (\lambda_L, \tilde{\lambda}_H)) \to \begin{pmatrix} N_{Lt} \\ N_{Ht} \end{pmatrix} & X_t = H \end{cases} \tag{13}$$

   (l) Sample $\lambda_L^{(i)} \sim \mathrm{Gam}\left(\sum_{t=1}^{T} N_{Lt} + a, T + b\right)$
   (m) Sample $\tilde{\lambda}_H^{(i)} \sim \mathrm{Gam}\left(\sum_{t:X_t=H} N_{Ht} + c, \sum \underline{m}_H + d\right)$

2.3. Penalized Model with Covariate(s)

3. SIMULATION STUDY

4. CASE STUDY: CARPENTER ANT LAB STUDY

5. RESULTS

6. DISCUSSION

ACKNOWLEDGEMENTS

REFERENCES

Akash FJ, Tirky T. 1988. Proper multivariate conditional autoregressive models for spatial data analysis. *Biometrics* **196**: 173.

APPENDIX

[Figure A.1 about here.]

[Table A.1 about here.]

FIGURES

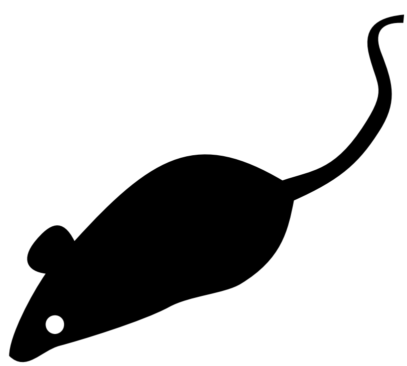

Figure A.1. Test Appendix figure

TABLES

Table A.1. Test Appendix Table

Test Table