SCALABLE BAYESIAN MULTINOMIAL LOGISTIC-NORMAL MODELS FOR THE ANALYSIS OF SEQUENCE COUNT DATA
### Sample Collection and Storage

![Diagram of sample collection](image)

### DNA Extraction

![DNA extraction](image)

### PCR Amplification

![PCR amplification](image)

### Sequencing

![Sequencing machine](image)

### Table: Framing Sequence Count Data

<table>
<thead>
<tr>
<th></th>
<th>Species 1</th>
<th>Species 2</th>
<th>Species 3</th>
<th>Species 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>23</td>
<td>53</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Sample 2</td>
<td>69</td>
<td>64</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Sample 3</td>
<td>33</td>
<td>100</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>Sample 4</td>
<td>5</td>
<td>63</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>Sample 5</td>
<td>76</td>
<td>80</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>Sample 6</td>
<td>58</td>
<td>7</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>Sample 7</td>
<td>10</td>
<td>87</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Sample 8</td>
<td>31</td>
<td>89</td>
<td>72</td>
<td></td>
</tr>
</tbody>
</table>

Adapted from Hamady. et al., *Nature Methods*, 2008
COMPOSITION

COMPOSITION: A CONTROVERSIAL TOPIC

It's all relative: analyzing microbiome data as compositions

Gregory B. Gloor PhD a, Jia Rong Wu BSc b, Vera Pawlowsky-Glahn PhD b, Juan José Egozcue PhD c

Microbiome Datasets Are Compositional: And This Is Not Optional

Gregory B. Gloor1, Jean M. Macklaim2, Vera Pawlowsky-Glahn2 and Juan J. Egozcue3

NO ITS NOT

Susan Holmes @SherlockpHolmes · 5 Apr 2018
Replying to @timtriche @samclifford and 2 others

Absolutely not, microbiome data are not compositional and those methods don’t apply, although it does apply to geostat data and other situations when one has a whole of exactly the same size. In microbiome data you have to control for different amounts of bacteria.
## Challenges of Composition

<table>
<thead>
<tr>
<th></th>
<th>Species 1</th>
<th>Species 2</th>
<th>Species 3</th>
<th>Species 4</th>
<th>Species 5</th>
<th>Species 6</th>
<th>Species 7</th>
<th>Species 8</th>
<th>Species 9</th>
<th>Species 10</th>
<th>Species 11</th>
<th>Species 12</th>
<th>Species 13</th>
<th>Species 14</th>
<th>Species 15</th>
<th>Species 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>23</td>
<td>53</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 2</td>
<td>69</td>
<td>64</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 3</td>
<td>33</td>
<td>100</td>
<td>68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 4</td>
<td>5</td>
<td>63</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 5</td>
<td>76</td>
<td>80</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 6</td>
<td>58</td>
<td>7</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 7</td>
<td>10</td>
<td>87</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 8</td>
<td>21</td>
<td>89</td>
<td>73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHALLENGES OF COMPOSITION

<table>
<thead>
<tr>
<th>Sample</th>
<th>Species 1</th>
<th>Species 2</th>
<th>Species 3</th>
<th>Species 4</th>
<th>Species 5</th>
<th>Species 6</th>
<th>Species 7</th>
<th>Species 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>23</td>
<td>53</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 2</td>
<td>69</td>
<td>64</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 3</td>
<td>33</td>
<td>100</td>
<td>68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 4</td>
<td>5</td>
<td>63</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 5</td>
<td>76</td>
<td>80</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 6</td>
<td>58</td>
<td>7</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 7</td>
<td>10</td>
<td>87</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 8</td>
<td>21</td>
<td>88</td>
<td>73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Row Sums are known to be arbitrary
**CHALLENGES OF COMPOSITION**

<table>
<thead>
<tr>
<th></th>
<th>Species 1</th>
<th>Species 2</th>
<th>Species 3</th>
<th>Species 4</th>
<th>Species 5</th>
<th>Species 6</th>
<th>Species 7</th>
<th>Species 8</th>
<th>Species 9</th>
<th>Species 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>23</td>
<td>53</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 2</td>
<td>69</td>
<td>64</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 3</td>
<td>33</td>
<td>100</td>
<td>68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 4</td>
<td>5</td>
<td>63</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 5</td>
<td>76</td>
<td>80</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 6</td>
<td>58</td>
<td>7</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 7</td>
<td>10</td>
<td>87</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 8</td>
<td>21</td>
<td>83</td>
<td>73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Row Sums are known to be arbitrary

Common practice is to "normalize" (convert to **percentages** by dividing by row totals)
CHALLENGES OF COMPOSITION

Row Sums are known to be arbitrary

Common practice is to "normalize" (convert to percentages by dividing by row totals)

Percentages = Relative Abundances = Compositions
Row Sums are known to be arbitrary

Common practice is to "normalize" (convert to **percentages** by dividing by row totals)

**Percentages** = Relative Abundances = Compositions

\[ \text{B+L+R} = k \]

And all Positive
CHALLENGES OF COMPOSITION

Row Sums are known to be arbitrary

Common practice is to "normalize" (convert to percentages by dividing by row totals)

Percentages = Relative Abundances = Compositions

B+L+R=k

And all Positive

Example of problem: If B goes up, L+R must go down
HOW DO YOU DEAL WITH COMPOSITION?
HOW DO YOU DEAL WITH COMPOSITION?

**ALR**

\[(x, y) = \left( \log \frac{L}{R}, \log \frac{B}{R} \right)\]

**CLR**

\[(x, y, z) = \left( \log \frac{L}{(LBR)^{1/3}}, \log \frac{B}{(LBR)^{1/3}}, \log \frac{R}{(LBR)^{1/3}} \right)\]
ZEROS AND COUNTING

A PROBLEM WITH THE COMPOSITIONAL PERSPECTIVE

\[ \log \frac{0}{x} = -\infty \]
A PROBLEM WITH THE COMPOSITIONAL PERSPECTIVE

\[
\log \frac{0}{x} = -\infty \\
\log \frac{x}{0} \ldots \text{Oh Shit...}
\]
The data is count data

A zero count can be because a taxa (e.g., species) had low, but non-zero, abundance.
The data is count data

A zero count can be because a taxa (e.g., species) had low, but non-zero, abundance.

Model Random Counting
(e.g., negative binomial or Poisson)
The data is count data
A zero count can be because a taxa (e.g., species) had low, but non-zero, abundance.

Model Random Counting
(e.g., negative binomial or Poisson)

Yet often models each taxa as independent.
VIEWING AS RANDOM SAMPLING

Sample Collection and Storage

DNA Extraction

PCR Amplification

Sequencing

Adapted from Hamady. et al., *Nature Methods*, 2008
BACKGROUND

PROBLEM WITH MULTIVARIATE RANDOM SUBSAMPLING

System 1

System 2
BACKGROUND

PROBLEM WITH MULTIVARIATE RANDOM SUBSAMPLING

System 1

System 2

Random sampling induces a competition to be counted (count compositional)
### Extracting More Information from Counts

<table>
<thead>
<tr>
<th>Samples</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taxa 1</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>Taxa 2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Taxa 3</td>
<td>100</td>
<td>1</td>
</tr>
</tbody>
</table>
Bayesian multinomial models reflect intuition we want.
**MULTINOMIAL-LOGISTIC NORMAL**

\[ Y \sim \text{Multinomial}(\pi) \]
\[ \pi \sim \text{Logistic Normal}(\rho, \Xi) \]

\[ Y \sim \text{Multinomial}(\pi) \]
\[ \pi = \text{ILR}^{-1}(\eta) \]
\[ \eta \sim \text{Multivariate Normal}(\mu, \Sigma) \]

- Handles Zeros and Competition-to-be-counted
- Allows positive and negative covariation between taxa
- Models Multiplicative Errors

ILR = "Isometric Log-Ratio" Transform
**BACKGROUND**

**MODELING TIME-EVOLUTION**

\[ Y_t \sim \text{Multinomial}(\pi_t) \]

\[ \pi_t = \text{ILR}^{-1}(\eta_t) \]

\[ \eta_t = F_t' \theta_t + \nu_t \quad \nu_t \sim N(0, V_t) \]

\[ \theta_t = G_t \theta_{t-1} + \omega_t \quad \omega_t \sim N(0, W_t) \]

\[ \theta_0 \sim N(m_0, C_0) \]

\[ V_1, \ldots, V_T, W_1, \ldots, W_T \sim p(\xi) \]
THE COMPUTATIONAL BOTTLENECK

10 Taxa with 650 Samples
As measured by Time to Effective Sample size of 2000

- Metropolis-within-Gibbs → >2 months

- Now on order of milliseconds to seconds.

Can even scale to 5K x 20K, ~ 1.4 days run-time
KEY IDEA

Goal
KEY IDEA

Goal
KEY IDEA

Goal

Marginal
KEY IDEA

Goal

Marginal
KEY IDEA

Goal

Marginal

Conditional
KEY IDEA

(1) I have found that a huge class of models have identical marginal forms

(2) I have found a highly accurate approximation for this marginal form

(3) These models often have conditionals that are easy to sample from.
MARGINALLY LATENT MATRIX-T PROCESSES MODELS

MARGINALLY LTP MODELS
Matrix Normal Process

\[ \eta \sim N(B, K, A) \]

- Matrix of Observed Data (Real Valued)
- Mean Matrix
- Covariance over Row space
- Covariance over Column space

Silverman JD, Roche K, et al. 2019. *arXiv*
Matrix T-Process

\[ \eta \sim T(u, B, K, A) \]

- Matrix of Observed Data (Real Valued)
- Degrees of Freedom
- Mean Matrix
- Covariance over Row space
- Covariance over Column space

LATENT MATRIX–T PROCESS (LTP)

\[ Y \sim f(\pi) \]
\[ \pi = \phi^{-1}(\eta) \]
\[ \eta \sim T(u, B, K, A) \]
AN EXAMPLE OF A MARGINALLY LTP MODEL

\[ Y_t \sim \text{Multinomial}(\pi_t) \]
\[ \pi_t = \text{ILR}^{-1}(\eta_t) \]
\[ \eta_t \sim N(M_t, \Sigma) \]
\[ M \sim N(0, \Sigma, \Gamma) \]
\[ \Gamma_{t,s} = \text{RBF}(t,s) \]
\[ \Sigma \sim \text{IW}(\Xi, u) \]

Multinomial Logistic Normal Process

Count Noise
Additional Noise
Smoothed State
Unknown Covariance Between Log-Ratios
MARGINALLY LATENT MATRIX-T PROCESSES

FOR TIME-SERIES ANALYSIS
MARGINALLY LATENT MATRIX-T PROCESSES

A FEW MORE EXAMPLES

Generalized Multivariate Dynamic Linear Models

\[ Y \sim f(\pi) \]
\[ \pi = \phi^{-1}(\eta) \]
\[ \eta_t^T = F_t^T \Theta_t + \nu_t^T, \quad \nu_t \sim N(0, \gamma_t \Sigma) \]
\[ \Theta_t = G_t \Theta_{t-1} + \Omega_t, \quad \Omega_t \sim N(0, W_t, \Sigma) \]
\[ \Theta_0 \sim N(M_0, C_0, \Sigma) \]
\[ \Sigma \sim IW(\Xi, u) \]

Generalized Multivariate Conjugate Linear Models

\[ Y \sim f(\pi) \]
\[ \pi = \phi^{-1}(\eta) \]
\[ \eta_{j.} \sim N(\Lambda X_{.j}, \Sigma) \]
\[ \Lambda \sim N(\Theta, \Sigma, \Gamma) \]
\[ \Sigma \sim IW(\Xi, u) \]

And Many More ...

MARGINALLY LATENT MATRIX-T PROCESSES

MULTINOMIAL LOGISTIC NORMAL MODELS WITH MARGINAL LAPLACE APPROXIMATION

C++, Eigen (+MKL)
R Interface using Rcpp

Extensively Unit Tested against Independent Implementations

MULTINOMIAL LOGISTIC NORMAL MODELS WITH MARGINAL LAPLACE APPROXIMATION

C++, Eigen (+MKL)
R Interface using Rcpp
Extensively Unit Tested against Independent Implementations

MULTINOMIAL LOGISTIC NORMAL MODELS - BUT FAST

Benchmarking - Kim Roche

MULTINOMIAL LOGISTIC NORMAL MODELS – BUT FAST

Efficient
~ 5 orders of magnitude faster than HMC
~ 1-2 orders of magnitude faster than Variational Bayes (VB)

Efficient

~ 5 orders of magnitude faster than HMC
~ 1-2 orders of magnitude faster than Variational Bayes (VB)

Accurate

- Point Estimation Accuracy (estimating posterior mean) is nearly perfect over all tested conditions (in contrast VB breaks down when many taxa)
- Uncertainty quantification (estimating posterior variance) only found to break down when > 93% zeros in dataset. (in contrast VB breaks down often)

Bayesian Multinomial Logistic Normal Models through Marginally Latent Matrix–T Processes

Justin D. Silverman, Kimberly Roche, Zachary C. Holmes, Lawrence A. David, Sayan Mukherjee

(Submitted on 27 Mar 2019 (v1), last revised 1 Apr 2019 (this version, v3))

Bayesian multinomial logistic–normal (MLN) models are popular for the analysis of sequence count data (e.g., microbiome or gene expression data) due to the complex covariance structure. However, existing implementations of MLN models are limited to handling small data sets due to the non-conjugacy of the model. We introduce MLN models which can be written as marginally latent matrix–t process (LTP) models. Marginally LTP models describe a flexible class of generalize series models. We develop inference schemes for Marginally LTP models and, through application to MLN models, demonstrate that our inference schemes are magnitude faster than MCMC.
Wife and Collaborator (MERCK Biostatistics)
Rachel Silverman

University of Montana
Alex Washburne

NYU
Jamie Morton

UCLA
Liat Shenhav
Eran Halperin

Duke University
Lawrence David
Sayan Mukherjee
Kim Roche
Rachael Bloom
Heather Durand
Sharon Jiang
Brianna Petrone
Zach Holmes
Jeff Letourneau
Max Villa
Kevin Zhu
Eric Dallow

U. de Girona
Vera Pawlowsky-Glahn

U. de Catalunya Polytechnic
Juan Jose Egozcue

University of Western Ontario
Greg Gloor

University of Notre Dame
Johannes R Björk
Elizabeth Archie
STRAY / MARGINALLY LATENT MATRIX-T PROCESSES

BUT WHAT ABOUT THE CONDITIONALS?

Generalized Multivariate Conjugate Linear Models

\[ Y \sim f(\pi) \]
\[ \pi = \phi^{-1}(\eta) \]
\[ \eta_j \sim N(\Lambda X_j, \Sigma) \]
\[ \Lambda \sim N(\Theta, \Sigma, \Gamma) \]
\[ \Sigma \sim IW(\Xi, \nu) \]

This is just the Solution to Bayesian Multivariate Linear Regression

\[ v_N = v + N \]
\[ \Gamma_N = (XX^T + \Gamma^{-1})^{-1} \]
\[ \Lambda_N = (\eta X^T + \Theta \Gamma^{-1}) \Gamma_N \]
\[ \Xi_N = \Xi + (\eta - \Lambda_N X)(\eta - \Lambda_N X)^T + (\Lambda_N - \Theta) \Gamma^{-1} (\Lambda_N - \Theta)^T \]
\[ p(\Sigma|\eta, X) = IW(\Xi_N, v_N) \]
\[ p(\Lambda|\Sigma, \eta, X) = N(\Lambda_N, \Sigma, \Gamma_N). \]
STRAY / MARGINALLY LATENT MATRIX-T PROCESSES

BUT WHAT ABOUT THE CONDITIONALS?

Generalized Multivariate Dynamic Linear Models

\[ Y \sim f(\pi) \]
\[ \pi = \phi^{-1}(\eta) \]
\[ \eta_t^T = F_t^T \Theta_t + \nu_t^T, \quad \nu_t \sim N(0, \gamma_t \Sigma) \]
\[ \Theta_t = G_t \Theta_{t-1} + \Omega_t, \quad \Omega_t \sim N(0, W_t, \Sigma) \]
\[ \Theta_0 \sim N(M_0, C_0, \Sigma) \]
\[ \Sigma \sim IW(\Xi, u) \]

B.2.1 Filtering Recursions for MDLM Model

1. Posterior at \( t - 1 \):
\[ p(\Theta_{t-1}, \Sigma, H_{t-1}^T) \sim N(M_{t-1}, C_{t-1}, \Sigma) \]
\[ p(\Sigma|H_{t-1}^T) \sim IW(\Xi_{t-1}, \nu_{t-1}) \]

2. Priors at \( t \):
\[ a_t = G_t m_{t-1} \]
\[ R_t = G_t C_{t-1} G_t^T + W_t \]
\[ p(\Sigma|H_t^T) \sim IW(\Xi_t, \nu_t) \]
\[ p(\Theta_{t-1}|\Sigma, H_{t-1}^T) \sim N(a_t, R_t, \Sigma) \]

3. One-step ahead forecast at \( t \):
\[ f_t^T = F_t^T a_t \]
\[ q_t = \gamma_t + F_t^T R_t F_t \]
\[ p(\Sigma|H_t^T) \sim IW(\Xi_{t-1}, \nu_{t-1}) \]
\[ p(\Theta_{t-1}|\Sigma, H_{t-1}^T) \sim N(f_t, q_t) \]

4. Posterior at \( t \):
\[ \eta_t = \eta_t^T - f_t^T \]
\[ S_t = R_t F_t \]
\[ m_t = a_t + S_t \eta_t \]
\[ C_t = R_t - q_t S_t S_t^T \]
\[ v_t = \nu_{t-1} + 1 \]
\[ \Xi_t = \frac{1}{v_t} \nu_{t-1} \Xi_{t-1} + \frac{\eta_t \eta_t^T}{v_t} \]
\[ p(\Sigma|H_t^T) \sim IW(\Xi_t, v_t) \]
\[ p(\Theta_{t-1}|\Sigma, H_{t-1}^T) \sim N(m_t, C_t, \Sigma) \]

B.2.2 Simulation Smoothing Recursion

The recursions provided here follow directly from Prado and West [39, p. 268]

1. Sample \( \Sigma \sim IW(\Xi_t, v_t) \) and then \( \Theta_T \sim N(M_t, C_t, \Sigma) \).
2. For each time \( t \) from \( T - 1 \) to 0, sample \( p(\Theta_t|\Theta_{t+1}, H_T^T) \sim N(M_t^*, C_t^*, \Sigma) \) where
\[ Z_t = C_t G_{t+1}^T R_{t+1}^{-1} \]
\[ M_t^* = M_t + Z_t(\theta_{t+1} - a_{t+1}) \]
\[ C_t^* = C_t - Z_t R_{t+1} Z_t^T \].
BUILDING A FRAMEWORK

BENCHMARKING RESULTS

Efficiency (Seconds per Effective Sample)

Error in Point Estimation

Error in Uncertainty Quantification

D # Taxa

N # Samples

Q # Covariates

HMC Uncollapsed
HMC Collapsed
VB Collapsed
LA Collapsed
(stray)
STRAY

REAL DATA
\[ Y \sim f(\pi) \]
\[ \pi = \phi^{-1}(\eta) \]
\[ \eta \sim T(u, B, K, A) \]

\[ f = \prod_{t=1}^{T} \text{Multinomial}(\pi_t) \]
\[ \phi = \text{ILR} \]
\[ B = 0_{D-1} \]
\[ K_{i,j} = \kappa^2 \exp(-\gamma^2 [d_{\text{phylo}}(i, j)]^2) \]
\[ A_{t,s} = \alpha^2 \exp(-\rho^2 (t - s)^2) \]

LATENT MATRIX-T PROCESSES

NON-LINEAR TIME-SERIES MODEL FOR MICROBIOME

Acidaminococcaceae
Bacteroidaceae
Desulfovibrionaceae
Enterobacteriaceae
Fusobacteriaceae
Lachnospiraceae
Porphyromonadaceae
Rikenellaceae
Ruminococcaceae
Synergistaceae