~HP 公開版~

細田至温

早稲田大学理工学術院 招聘研究員

September 20, 2025

細田至温 早稲田大学

#### 目次

- 1 Umibato: estimation of time-varying microbial interaction using continuous-time regression hidden Markov model
- 2 BayesianSSA: a Bayesian statistical model based on structural sensitivity analysis for predicting responses to enzyme perturbations in metabolic networks
- 3 本日のまとめ

**Umibato** 

Bioinformatics, 37, 2021, i16–i24 doi: 10.1093/bioinformatics/btab287 ISMR/FCCB 2021



# Umibato: estimation of time-varying microbial interaction using continuous-time regression hidden Markov model

Shion Hosoda<sup>1,2,\*</sup>, Tsukasa Fukunaga<sup>1,3</sup> and Michiaki Hamada<sup>1,2,4,\*</sup>

<sup>1</sup>Department of Electrical Engineering and Bioscience, Graduate School of Advanced Science and Engineering, Waseda University, Tokyo 169-8555, Japan, <sup>2</sup>Computational Bio Big-Data Open Innovation Laboratory, AIST-Waseda University, Tokyo 169-8555, Japan, <sup>3</sup>Department of Computer Science, Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan and <sup>4</sup>Graduate School of Medicine, Nippon Medical School, Tokyo 113-8602, Japan

<sup>\*</sup>To whom correspondence should be addressed.

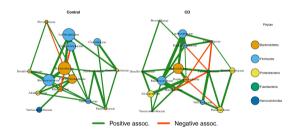
#### Outline

- Umibato: estimation of time-varying microbial interaction using continuous-time regression hidden Markov model
  - ■背景

- 手法
- 結果
- まとめ
- BayesianSSA: a Bayesian statistical model based on structural sensitivity analysis for predicting responses to enzyme perturbations in metabolic networks
- 3 本日のまとめ

#### 微生物相互作用は重要な研究対象

例えば、微生物相互作用が疾患を持つ場合と持たない場合で異なることが示唆されている (McGregor, Labbe, and Greenwood, 2020)



微生物相互作用を微生物叢データから推定する手法:

BavesianSSA

- 相関ベースの手法
- 一般化ロトカ・ヴォルテラ方程式 (gLVE) ベースの手法

6 / 75

#### 有向ネットワークは微生物間の共生関係を明らかにできる

gLVE ベースの手法では、以下のような有向ネットワークを推定できることがメ リット

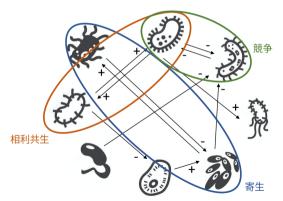
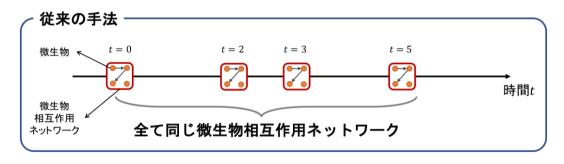


Figure: https://didier.gonze.web.ulb.be/LEUVEN2017/practicals.html

早稲田大学 研究事例の紹介

#### 従来の手法は全ての時点で同じ微生物相互作用を仮定している

BavesianSSA



疑問:この仮定は妥当か?

細田至温

早稲田大学

研究事例の紹介

References

#### 環境の条件に依存して微生物相互作用ネットワークは変化する

異なる栄養条件下で微生物相互作用が異なることが報告されている (Embree et al.. 2015)







特に栄養条件が変化する環境(例:腸内細菌叢)などではこの仮定は致命的にな りうる

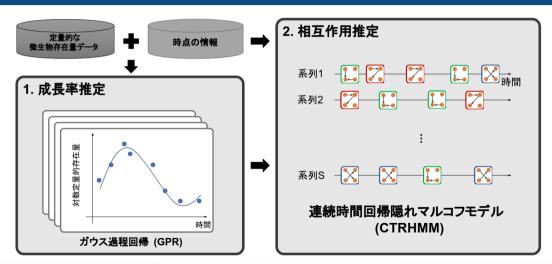
→ 時間変化する微生物相互作用ネットワークの推定を行いたい

細田至温

#### Outline

- 1 Umibato: estimation of time-varying microbial interaction using continuous-time regression hidden Markov model
  - ■背景
  - 手法
  - 結果
  - まとめ
- 2 BayesianSSA: a Bayesian statistical model based on structural sensitivity analysis for predicting responses to enzyme perturbations in metabolic networks
- 3 本日のまとめ

#### Umibato 概観



細田至温 早稲田大学

#### gLVE ベースの微生物相互作用推定

一般化ロトカ・ヴォルテラ方程式は以下の式・

$$\frac{dx_i(t)}{dt} = \left(\phi_{i,0} + \sum_{j=1}^{M} \phi_{i,j} x_j(t)\right) x_i(t). \tag{1}$$

 $x_i(t)$ :時間 t での微生物 i の絶対存在量

 $\phi_{i,j}(j>0)$ : 微生物 j から微生物 i への影響パラメタ

 $\phi_{i,0}$ :微生物 i の growth パラメタ

 $x_i(t)$  はデータとして与えられているものとする

細田至温

早稲田大学

研究事例の紹介

手法

#### 実際には、以下の常微分方程式に基づいてパラメタを推定

$$y_i(t) = \phi_{i,0} + \sum_{j=1}^{M} \phi_{i,j} x_j(t).$$
 (2)

ここで,成長率  $y_i(t)$  を以下のように定義

$$y_i(t) = \frac{d \ln x_i(t)}{dt} = \frac{dx_i(t)}{dt} \frac{1}{x_i(t)}.$$
 (3)

補完などにより  $y_i(t)$  を得ることができれば、線型回帰で  $\Phi$  を推定可能

細田至温

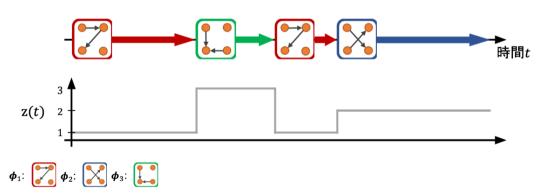
早稲田大学

研究事例の紹介

September 20, 2025

#### 本研究では gLVE に時間変化するパラメタを導入する

時間変化するパラメタ  $\phi(t)$  の代わりに  $\phi_{z(t)}$  を導入する (z(t) は相互作用の状態を 示す変数)



細田至温

早稲田大学

研究事例の紹介

#### 本研究では gLVE に時間変化するパラメタを導入する

相互作用の状態変数 z(t) を導入した gLVE は以下のようになる:

$$y_i(t) = \phi_{z(t),i,0} + \sum_{j=1}^{M} \phi_{z(t),i,j} x_j(t).$$
 (4)

- これを元に統計モデルを構築する
  - しかし z(t) は観測できないので確率過程でモデル化する

本日のまとめ

#### 状態変数を連続時間マルコフ連鎖でモデル化する

- z(t) を連続時間マルコフ連鎖でモデル化する
  - 現在の値のみに依存して未来の値が決まる
  - 短時間の経過であればあるほど状態が遷移しにくい
- 離散的に観測された x(t) から z(t) の値の系列を推定する
- 実用上は離散時間マルコフ連鎖と比較して以下の利点がある
  - 不均一なサンプリング間隔を自然に扱える
    - 離散時間マルコフ連鎖では欠損値を駆使しなければならない
  - ☑ 仟意の時間が経過したときの遷移確率を計算できる

September 20, 2025

#### 連続時間マルコフ連鎖

経過時間が  $\Delta t$  のとき連続時間マルコフ連鎖の遷移確率行列  $\mathbf{P}(\Delta t)$  は

$$\mathbf{P}(\Delta t) = \exp(\mathbf{Q}\Delta t),\tag{5}$$

によって計算できる

■ ただし、exp(·) は行列指数関数、Q は遷移速度行列 直感的には観測点間の確率過程を周辺化するような操作になる

#### 生成過程

最終的に構築した Continuous-time regression hidden Markov model (CTRHMM) は 以下の生成過程

```
for each state k = 1 \dots K do
       \boldsymbol{\phi}_{k,i} \sim \text{Normal}(\mathbf{0}, \lambda_i^{-1} \mathbf{I}_{M+1})
P(t) = \exp(Qt)
for each subject s = 1 \dots S do
      \mathbf{z}_{1}^{(s)} \sim \text{Multinomial}(\mathbf{1}_{K}/K, 1)
       for each observation point n = 2 \dots N_s do
             \mathbf{z}_{n}^{(s)} \sim \sum_{k} z_{n-1,k}^{(s)} \text{Multinomial} \left( \mathbf{p}_{k}(d_{n-1}^{(s)}), 1 \right)
             y_{n,i}^{(s)} \sim \sum_{k} z_{n,k}^{(s)} \text{Normal}\left(\phi_{k,i}^{T} \mathbf{x}_{n}^{(s)}, \eta \sigma_{n,i}^{(s)^{2}}\right)
```

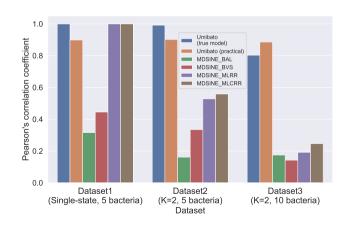
**BayesianSSA** 

References

#### Outline

- 1 Umibato: estimation of time-varying microbial interaction using continuous-time regression hidden Markov model
  - ■背景
  - 手法
  - 結果
  - まとめ
- 2 BayesianSSA: a Bayesian statistical model based on structural sensitivity analysis for predicting responses to enzyme perturbations in metabolic networks
- 3 本日のまとめ

### 人工データ実験では Umibato の性能が従来の手法より良いことが示された



細田至温

早稲田大学

研究事例の紹介

#### マウスの腸内細菌叢データを使って実データ実験を行った

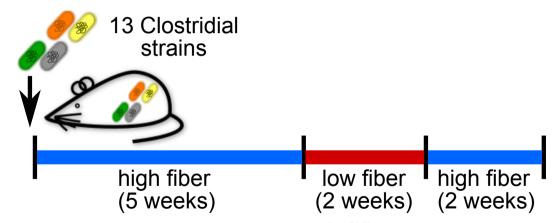
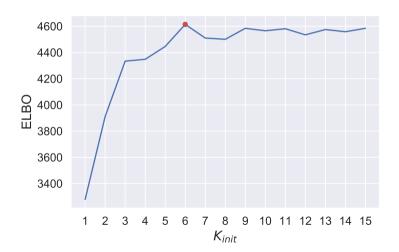


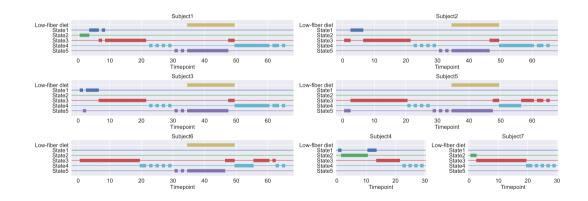
Figure: Bucci et al., 2016より引用

#### ELBO が最大を示すのは初期状態数6のときの試行であった



細田至温 早稲田大学

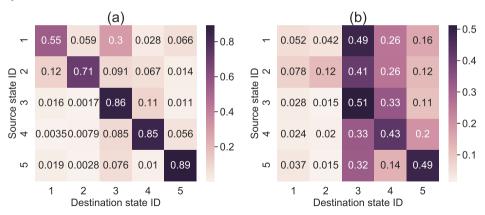
#### 食事の変化に伴って相互作用が変化していることが示唆された



#### 状態 3, 4, 5 が主な状態であることが分かる

BavesianSSA

左が一日経過した場合の遷移確率行列,右が一週間経過した場合の遷移確率行列 を示す.

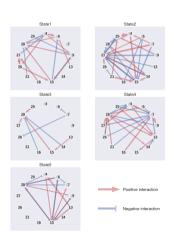


細田至温 早稲田大学

#### 微生物相互作用ネットワークは各状態で大きく異なる

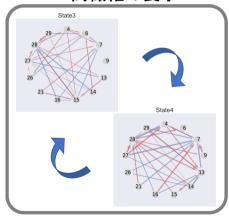
右図は各状態に対応する相互作用ネットワーク

- 各円が微生物を示す
- 例えば寄牛関係が確認できる
  - 寄生関係とは、ここでは A が B の増加に寄与し、 B が A の減少に寄与するような A と B の関係 を指す
- 特に状態5で多く見られる



#### <u>ここまでの結果から示唆される微生物相互作用の動態の概念図</u>

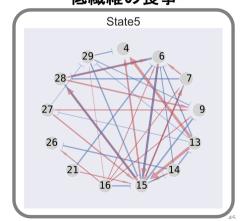
#### 高繊維の食事



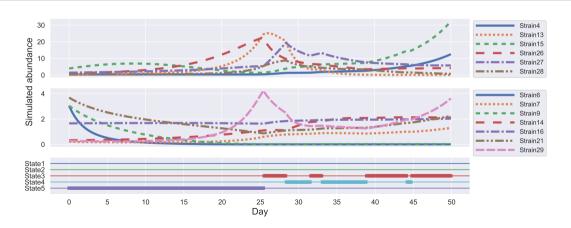
## 食事の変化



### 低繊維の食事



#### 推定されたパラメタによるシミュレーションから、長期間の低繊維 の食事は復元不可能な細菌叢多様性の減少を招くことが示唆



#### Outline

まとめ

- 1 Umibato: estimation of time-varying microbial interaction using continuous-time regression hidden Markov model
  - ■背景
  - 手法
  - 結果
  - まとめ
- 2 BayesianSSA: a Bayesian statistical model based on structural sensitivity analysis for predicting responses to enzyme perturbations in metabolic networks
- 3 本日のまとめ

#### まとめ

Umibato というアルゴリズムを提案し時間変化する微生物相互作用を推定した

- 離散的に時間変化する変数を導入して時間変化する相互作用を表現した
- 食事の変化に伴って相互作用ネットワークが変化していた
- 同じ高繊維の食事でも二つの相互作用状態で構成されていることが示唆 された

BayesianSSA

研究事例の紹介

Hosoda et al. RMC Riginformatics (2024) 25:297 https://doi.org/10.1186/s12859-024-05921-4

**BMC Bioinformatics** 

September 20, 2025

30 / 75

#### RESEARCH Open Access



BayesianSSA: a Bayesian statistical model based on structural sensitivity analysis for predicting responses to enzyme perturbations in metabolic networks

Shion Hosoda<sup>1\*</sup>, Hisashi Iwata<sup>1</sup>, Takuya Miura<sup>1</sup>, Maiko Tanabe<sup>1</sup>, Takashi Okada<sup>2</sup>, Atsushi Mochizuki<sup>2</sup> and Miwa Sato<sup>1</sup>

#### Outline

- Umibato: estimation of time-varying microbial interaction using continuous-time regression hidden Markov model
- BavesianSSA: a Bayesian statistical model based on structural sensitivity analysis for predicting responses to enzyme perturbations in metabolic networks
  - ■背景
  - 理論的背景
  - 手法
  - 結果
  - まとめ
- 3 本日のまとめ

30 / 75

#### 生体内では複数の化学反応が代謝ネットワークを構成して働いてい る

BayesianSSA

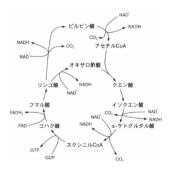


Figure: https://photosyn.jp/pwiki/index.php?TCA%E5%9B%9E%E8%B7%AF

#### 代謝ネットワークにおける応答予測は重要な問題

反応を改変したときの代謝の流れ(フラックス)の変化を知りたいことがある. 例えばバイオものづくりでは、どの遺伝子を過剰発現させれば生産量が増えるの かが分かれば効率を上げることができる.



Figure: Voigt, 2020より引用

実験すれば結果が分かるがコストがかかる → 予測できると嬉しい

細田至温 早稲田大学

本日のまとめ

#### 反応速度すべてをモデル化すれば予測が可能だが、難しい

予測する素直な方法として, 反応速度論モデルの構築がある

- 代謝ネットワークのシミュレーションにより、化学反応を促進させたときの 物質の生産量の変化を予測できる
- しかし、それには代謝ネットワークに含まれる全ての化学反応の反応速度を 同定する必要がある
- 反応速度を同定することは、関数形及び関数のパラメタを同定することであり、それを化学反応の数だけ考慮すると膨大な空間を探索することになる

細田至温早稲田大学

#### 簡単に適用できる方法として SSA がある

BavesianSSA

一方で,近年開発された構造感度解析(SSA: Structural sensitivity analysis) (Mochizuki and Fiedler, 2015) では、代謝ネットワークの構造情報のみから定性的な応答予測を行える

シミュレーションベースの方法と異なるのは,

- 量の情報は分からない
- 反応速度の関数系やパラメタが不要

という点で,適用範囲が広い

## 簡単に適用できる一方で予測結果がわからないことがある

SSA には、その適用範囲の広さから「不定」という予測結果を出すことがある

- 反応速度の情報を使わない分、それらの情報に依存して応答が変わる場合、 予測が行えない
- 反応速度の情報は、化学反応特有のものに加え、個体差や、環境情報(pH, 栄養の有無)に依存すると考えられる
- → 環境情報の統合により予測を特定の環境に特化させることで、不定を解消でき るのでは?

# 本研究では実験結果を用いて SSA に環境情報を統合することを目指す

SSA の予測に実験結果から抽出した環境情報を統合するベイズ統計モデル, BayesianSSA を開発した

このモデルは以下のような特徴を持つ

■ 摂動実験(遺伝子の過剰発現などにより反応を促進して物質の生産量を測定する)による実験結果をデータとして用いる

研究事例の紹介

- モデル内に SSA のアルゴリズムを関数として含む
- 事後分布により計算される正値性確信度により確率的な予測が可能
- 予測分布により実験結果の予測も可能

Umibato

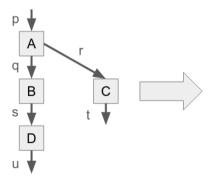
#### Outline

- Umibato: estimation of time-varying microbial interaction using continuous-time regression hidden Markov model
- BavesianSSA: a Bayesian statistical model based on structural sensitivity analysis for predicting responses to enzyme perturbations in metabolic networks
  - ■背景
  - 理論的背景
  - 手法
  - 結果
  - まとめ
- 3 本日のまとめ

Umibato

## 化学量論行列

代謝ネットワークは化学量論行列  $\nu$  と呼ばれる行列で表現される  $\nu_{m,j}$  は反応 j により起こる代謝物 m の増減を示す(図では 1 または-1 のみだが化学反応式によっては他の整数もありうる)



	p	q	r	s	t	u
Α	1	-1	-1	0	0	0
В	0	1	0	-1	0	0
С	0	0	1	0	-1	0
D	0	0	0	1	0	-1

細田至温

早稲田大学

## 代謝ネットワークの ODE

代謝ネットワークにおける代謝物濃度のシステムは以下で表される:

$$\frac{dx_m}{dt} = \sum_{j=1}^{J} \nu_{m,j} F_j(k_j, \mathbf{x}), \tag{6}$$

 $x_m$ :代謝物 m の濃度

 $u_{m,j}$ : 化学量論行列の (m,j) 要素

 $F_j(\cdot)$ : 反応速度関数(反応速度定数  $k_j$  とシステム内全ての代謝物の濃度  $\mathbf{x}$  に依存)

## 代謝ネットワークにおける用語

#### 反応フラックス

 $F_i(\cdot)$  の値を反応フラックスと呼ぶ

#### 摂動

遺伝子の過剰発現などにより  $k_i$  を  $k_i + \delta k_i$  とする操作を<mark>摂動</mark>と呼ぶ

#### 応答

システムが摂動により定常状態からまた別の定常状態に移ることで起こる、そのシステムの代謝物濃度またはフラックスの変化を<mark>応答</mark>と呼ぶ

## 代謝ネットワークにおける用語

#### 化学量論的サイクル

 $\nu z = 0, z \neq 0$  を満たすベクトル  $z \in \mathbb{R}^J$  を化学量論的サイクルと呼ぶ.

#### 定常状態のときのフラックス、パスウェイフラックス

BavesianSSA

化学量論的サイクルとなるベクトルの集合の基底を $\{c_k\}_{k}$ としたとき、定常状態 では  $F_i(\cdot) = \sum_{k} \mu_k c_{ki}$  を満たし、その時の  $\mu_k$  の値をパスウェイフラックスと呼ぶ

#### 保存量

 $\{\mathbf{z} \in \mathbb{R}^M | \mathbf{\nu}^{\mathrm{T}} \mathbf{z} = \mathbf{0}\}$  の基底を  $\{\mathbf{d}_l\}_l$  としたとき, $\lambda_l \coloneqq \sum_m x_m d_{l,m}$ ,を保存量と呼ぶ

References

#### A 行列

Umihato

理論的背景

以下の行列をA 行列と呼ぶ.

$$\mathbf{A}(\mathbf{r}) \coloneqq \left( egin{array}{c|c} \mathbf{R}(\mathbf{r}) & -\mathbf{C} \ \hline -\mathbf{D}^{\mathrm{T}} & \mathbf{0}_{L imes K} \end{array} 
ight),$$

ただし、 $\mathbf{R}(\mathbf{r})$  は (j,m) 成分が  $r_{j,m} = \partial F_j/\partial x_m$  で表され、その非ゼロ成分を  $\mathbf{r} \in \mathbb{R}^P$  として関数で表し、 $\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_K) \in \mathbb{R}^{J \times K}$ 、 $\mathbf{c}_k$  は  $\{\mathbf{z} | \nu \mathbf{z} = \mathbf{0}\}$  (化学量論的サイクルを満たす反応の組み合わせの集合)の k 番目の基底、 $\mathbf{D} = (\mathbf{d}_1, \dots, \mathbf{d}_L) \in \mathbb{R}^{M \times L}$ 、 $\mathbf{d}_l$  は  $\{\mathbf{z} | \nu^T \mathbf{z} = \mathbf{0}\}$  (保存量となる代謝物の組み合わせの集合)の l 番目の基底、とする.

細田至温 早稲田大学

## 構造感度解析(SSA: Structural sensitivity analysis)2/8

#### 構造感度解析 (Okada and Mochizuki, 2017)

全ての  $F_i(\cdot)$  が  $k_i$  について単調増加関数であり、 $\mathbf{A}(\mathbf{r})$  は正則とする.  $\mathbf{S}(\mathbf{r}):=$  $-\mathbf{A}(\mathbf{r})^{-1}$ の(m,j)要素  $S_{m,j}(\mathbf{r})$ の符号が $\mathbf{r}$ に依存しないとき,反応または保存量 iを摂動したときの代謝物濃度及びパスウェイフラックス mの応答と同じ符号 になる.

Proof. 反応  $j^*$  の摂動は, $\delta k_{j^*} > 0$  に対して  $k_{j^*}$  を  $k_{j^*} + \delta k_{j^*}$  とする操作なので,そ の操作の前後での代謝物濃度の差分  $\delta x$ , パスウェイフラックスの差分  $\delta \mu$ , 反応速 度の差分  $\{\delta F_j\}_i$  を考える.また, $\delta k_{j^*}$  は微小な値とする.反応速度の差分  $\delta F_i$  は, 摂動前の定常状態の代謝物濃度を x として,

$$\delta F_j = F_j(k_j + \delta_{j^*,j}\delta k_{j^*}, \bar{\mathbf{x}} + \delta \mathbf{x}) - F_j(k_j, \bar{\mathbf{x}}), \tag{7}$$

と表せる. ただし,  $\delta_{t',i}$  はクロネッカーのデルタである.

早稲田大学

Umihato

## <u>構造感度解析(SSA</u>: Structural sensitivity analysis)3/8

右辺第一項をテイラー展開して.

$$\delta F_{j} = \frac{\partial F_{j}}{\partial k_{j}}(k_{j}, \bar{\mathbf{x}}) \delta k_{j^{*}} \delta_{j^{*}, j} + \sum_{m=1}^{M} \frac{\partial F_{j}}{\partial x_{m}}(k_{j}, \bar{\mathbf{x}}) \delta x_{m}. \tag{8}$$

ここで、定常状態において  $\sum_i \nu_{m,j} F_j = 0$  がどの m についても成立するので、

$$\delta F_j = \sum_{k=1}^K \delta \mu_k c_{k,j},\tag{9}$$

が成立する. これを(8)に代入して,

$$\sum_{m=1}^{M} \frac{\partial F_{j}}{\partial x_{m}} (k_{j}, \bar{\mathbf{x}}) \delta x_{m} - \sum_{k=1}^{K} \delta \mu_{k} c_{k,j} = -\frac{\partial F_{j}}{\partial k_{j}} (k_{j}, \bar{\mathbf{x}}) \delta k_{j^{*}} \delta_{j^{*},j}.$$
(10)

細田至温

早稲田大学

理論的背景

簡単のため、 $r_{i,m} := \frac{\partial F_i}{\partial x_i}(k_i, \bar{\mathbf{x}})$  とおくと、

$$\sum_{m=1}^{M} r_{j,m} \delta x_m - \sum_{k=1}^{K} \delta c_{k,j} \mu_k = -\frac{\partial F_j}{\partial k_j} (k_j, \bar{\mathbf{x}}) \delta k_{j^*} \delta_{j^*,j}. \tag{11}$$

また、定義から、保存量 $\lambda_i$ は

$$\lambda_l := \sum_{m=1}^{M} d_{l,m} x_m, \tag{12}$$

であり、 $\lambda_l$ の初期値にのみ依存し、どの時間においても成り立つことから、

$$\sum_{m=1}^{M} d_{l,m} \delta x_m = 0. \tag{13}$$

細田至温 早稲田大学 Umibato

## 構造感度解析(SSA:Structural sensitivity analysis)

**A** 行列に  $\delta x$  と  $\delta \mu$  を縦に並べたベクトルを右からかけると,

$$\mathbf{A}(\mathbf{r}) \left( \frac{\delta \mathbf{x}}{\delta \boldsymbol{\mu}} \right) = \left( \begin{array}{c|c} \mathbf{R}(\mathbf{r}) & -\mathbf{C} \\ \hline -\mathbf{D}^{\mathrm{T}} & \mathbf{0}_{L \times K} \end{array} \right) \left( \frac{\delta \mathbf{x}}{\delta \boldsymbol{\mu}} \right)$$
(14)

$$= \mathbf{R}(\mathbf{r})\delta\mathbf{x} - \mathbf{C}\delta\boldsymbol{\mu} - \mathbf{D}^{\mathrm{T}}\delta\mathbf{x} + \mathbf{0}, \tag{15}$$

(11)および (13)から,

$$\mathbf{A}(\mathbf{r}) \left( \frac{\delta \mathbf{x}}{\delta \boldsymbol{\mu}} \right) = -\mathbf{e}_{j^*}. \tag{16}$$

ここで、 $\mathbf{e}_{j^\star} \in \mathbb{R}^{J+L}$  は、 $j^\star$  番目の成分が  $\delta F_{j^\star}^{(k)} = \frac{\partial F_{j^\star}}{\partial k_{i^\star}} (k_{j^\star}, \bar{\mathbf{x}}) \delta k_{j^\star}$  で、それ以外が 0 で あるベクトルである。

References

## 構造感度解析(SSA: Structural sensitivity analysis)6/8

BavesianSSA

両辺を  $\delta F_{i\star}^{(k)}$  で割ると,

$$\mathbf{A}(\mathbf{r}) \left( \frac{\delta \mathbf{x} / \delta F_{j^*}^{(k)}}{\delta \boldsymbol{\mu} / \delta F_{j^*}^{(k)}} \right) = -\mathbf{h}_{j^*}, \tag{17}$$

と表現できる.ここで、 $\mathbf{h}_{j^*}$  は  $j^*$  番目の要素が 1 でそれ以外の要素が 0 のベクトルである.

細田至温 早稲田大学

理論的背景

次に、保存量  $\lambda_{j^*}$  の摂動について考える.保存量  $\lambda_{j^*}$  の摂動は、 $\delta\lambda_{j^*}>0$  について、 $\lambda_{i^*}$  を  $\lambda_{i^*}+\delta\lambda_{i^*}$  とする操作である.(11)および (13)と同様に、

$$\sum_{m=1}^{M} r_{j,m} \delta x_m - \sum_{k=1}^{K} c_{k,j} \delta \mu_k = 0,$$
 (18)

$$\sum_{m=1}^{M} d_{l,m} \delta x_m = \delta \lambda_l \delta_{l,j^*}, \tag{19}$$

$$\mathbf{A}(\mathbf{r}) \left( \frac{\delta \mathbf{x} / \delta \lambda_{j^{\star}}}{\delta \boldsymbol{\mu} / \delta \lambda_{j^{\star}}} \right) = -\mathbf{h}_{j^{\star}}, \tag{20}$$

細田至温 早稲田大学

以上より、摂動する反応または保存量 f すべてについて考えて、 $\mathbf{A}(\mathbf{r})$  が正則とすると、

$$\frac{\delta \tilde{\mathbf{X}}}{\delta \tilde{\mathbf{M}}} = -\mathbf{A}(\mathbf{r})^{-1}\mathbf{I} = \mathbf{S}(\mathbf{r}). \tag{21}$$

 $F_j(\cdot)$  が  $k_j$  について単調増加であることから  $\delta F_{j^*}^{(k)} > 0$  なので,

$$\operatorname{sign}\left(\frac{\delta \mathbf{X}}{\delta \mathbf{M}}\right) = \operatorname{sign}(\mathbf{S}(\mathbf{r})). \tag{22}$$

 $\delta k_{j*}$  及び  $\delta \lambda_{j*}$  は微小であるが、 $\mathbf{r}$  に符号が依存しない要素については、微小でない場合においても微小な変化の積分と考えると成立することが分かる.

Umihato

理論的背景

## SSA で得られるフラックスはパスウェイフラックスだが,反応フラックスに変換することができる.反応フラックス $\mathbf{T}(\mathbf{r}) \in \mathbb{R}^{J \times (J+L)}$ は,

$$\mathbf{T}(\mathbf{r}) := \mathbf{CS}(\mathbf{r})_{M+1:M+K,1:J+L},\tag{23}$$

で表され, $T_{j,j'}$  は反応または保存量 f を摂動したときの反応フラックスの応答と同じ符号になる値を示す.

本研究では、主に反応または保存量を摂動したときの代謝物濃度または反応フラックスの応答に興味があるため、SSA からの出力として、定性的応答行列  $\mathbf{Q}(\mathbf{r})$  を以下で定義する:

$$\mathbf{Q}(\mathbf{r}) := \operatorname{sign}\left(\left(\frac{\mathbf{S}(\mathbf{r})_{1:M,1:J+L}}{\mathbf{T}(\mathbf{r})}\right)\right). \tag{24}$$

細田至温 早稲田大学

## SSA による定性的予測

計算機上は SSA を用いてどのように定性的予測を得るのか?

A(r) 及び S(r) は r に依存するが、その符号が r に依存せず一定であれば、それが 定性的予測になる、この条件を満たすか判定する方法としては以下の2つ、

- 代数的計算:  $\mathbf{r}$  を変数としたまま  $q_{m,i}(\mathbf{r})$  を計算し,  $\mathbf{r}$  の符号から  $q_{m,i}(\mathbf{r})$  $(\mathbf{Q}(\mathbf{r}))$  の (m, i) 成分)の符号を決定する
- 数値的計算:  $\mathbf{r}$  の乱数を取ることにより、 $q_{mi}(\mathbf{r})$  を計算し、 $\mathbf{r}$  の符号から  $q_{m,i}(\mathbf{r})$  の符号を決定する

いずれの場合でも, $r_{j,m}\coloneqq rac{\partial F_j}{\partial \mathbf{x}_m}(k_j,ar{\mathbf{x}})$  の符号は典型的には正(代謝物濃度の反応へ の依存関係は、多くが化学反応と基質の関係によるものであるため)なので、こ の情報から  $q_{m,i}(\mathbf{r})$  の符号を求める.

一方で、アロステリック制御のようなケースでは rim の符号は負として扱う.

September 20, 2025

Umihato

## SSA による定性的予測

具体例を示す、n > 0, n > 0 として.

- 1  $S_{m,i} = r_1 + r_2 \Rightarrow q_{m,i}(\mathbf{r}) = 1$
- $S_{m,i} = -r_1 r_2 \Rightarrow q_{m,i}(\mathbf{r}) = -1$
- $S_{m,i} = r_1 r_2 \Rightarrow q_{m,i}(\mathbf{r})$  は不定

BayesianSSA は、3 のケースに対して $\mathbf{r}$  を確率変数とみなすことで対処する手法で ある.

lacksquare これは,(主に $\left\{k_{j}\right\}_{i}$ を介して依存する)環境条件の確率を考えて特定の条件 にフォーカスすることに該当する

## SSA による定性的予測

摂動する対象である反応・保存量と、観測する対象である代謝物濃度・反応フラックスをそれぞれ総称するため、以下の語句を定義する.

#### 摂動対象

代謝ネットワークにおける化学反応及び保存量を摂動対象と呼ぶ.

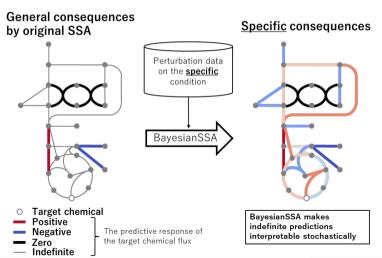
#### 観測対象

代謝ネットワークにおける代謝物濃度及び反応フラックスを観測対象と呼ぶ.

#### Outline

- Umibato: estimation of time-varying microbial interaction using continuous-time regression hidden Markov model
- BavesianSSA: a Bayesian statistical model based on structural sensitivity analysis for predicting responses to enzyme perturbations in metabolic networks
  - ■背景
  - 理論的背景
  - 手法
  - 結果
  - まとめ
- 3 本日のまとめ

## BayesianSSA 概観



## BayesianSSA

BayesianSSA が対象とする摂動実験のデータは、i番目のレコードについて、

$$y_i \coloneqq \begin{cases} 1 & \text{if increase} \\ -1 & \text{if decrease} \end{cases} \tag{25}$$

で表される.どの観測対象とどの摂動対象を選んだかは,それぞれのインデックスとして  $m_i$ ,  $j_i$  で区別する.

細田至温

早稲田大学

## BavesianSSA

提案手法 BavesianSSA ののモデルについて説明していく. まずは尤度.

#### BayesianSSA の尤度

$$p(y_i|\mathbf{r},\boldsymbol{\rho}) = \begin{cases} \rho_{m_i,j_i} & \text{if } y_i = q_{m_i,j_i}(\mathbf{r}) \\ 1 - \rho_{m_i,j_i} & \text{if } y_i \neq q_{m_i,j_i}(\mathbf{r}) \end{cases}.$$

直感的には、「SSA の予測が当たっていても、 $1-\rho_{m,j}$  の確率で外す」というよう な意図が込められている.

## BavesianSSA

### $|\mathbf{r}, ho_{m,i}|$ の事前分布

$$p(\mathbf{r} = \mathbf{r}^{(v)}) = \frac{1}{V},$$
 $p(\rho_{m,j}) = \mathcal{B}(\rho_{m,j}|\mathbf{a}, \mathbf{b}),$ 

 $\mathbf{r}^{(v)}$  は  $\mathbf{r}$  の制約にあう分布からドローした  $\mathbf{v}$  番目のサンプル, $\mathcal{B}(\cdot | \mathbf{a}, \mathbf{b})$  はベータ分 布, a,bはそれぞれベータ分布のパラメタを指す。

a,bは、a>bになるように設定する.こちらは「 $\rho_{m,i}$ は1に近くなるように」,「 $\mathbf{r}$ が制約を満たすように&ェに連続的な変化の考慮は不要」という意図が込められ ている。

BayesianSSA の事後分布・予測分布については、 ${f r}$  の事前分布に経験分布を、 ${
ho}_{m,j}$  の事前分布に共役事前分布を用いているため、解析的に計算できる. 事後分布は以下で計算できる:

$$p(\mathbf{r} = \mathbf{r}^{(v)}|\mathbf{y}) = \frac{g(\mathbf{y}, \mathbf{r}^{(v)})}{\sum_{v'=1}^{V} g(\mathbf{y}, \mathbf{r}^{(v')})},$$
(26)

$$g(\mathbf{y}, \mathbf{r}^{(v)}) := w_v \prod_{(m,j) \in \Lambda} \text{Beta}(\hat{a}_{m,j,v}, \hat{b}_{m,j,v}),$$
 (27)

$$\Lambda := \{ (m,j) \mid t_{m,j,\nu} + f_{m,j,\nu} \neq 0 \}. \tag{28}$$

 $\mathrm{Beta}(\cdot,\cdot):$ ベータ関数  $t_{m,j,v},f_{m,j,v}:\mathbf{r}^{(v)}$  による (m,j) 予測の結果が当たっていた回数と外れた回数  $\hat{a}_{m,j,v}=t_{m,j,v}+a$ ,  $\hat{b}_{m,j,v}=f_{m,j,v}+b$  予測分布については省略する.

## 正値性確信度の定義

#### 正值性確信度

 $\mathbf{r}$  の確率分布  $p^*$  に対して定義される

$$c_p(m,j) := \mathbb{E}_{p^*(\mathbf{r})}[\mathbb{I}_{q_{m,j}(\mathbf{r})=1}],$$

を正値性確信度 (positivity confidence values)と呼ぶ.

直感的には、 $\lceil g_{m,i}(\mathbf{r})$ が1であることを BayesianSSA がどれだけ確信しているか?」 を表す、実験データの予測には、実験誤差が乗った予測分布を使うべきだが、結 果の解釈にはこちらを用いる.

手法

## 使用するデータ

- Toya, Shiraki, and Shimizu, 2015による大腸菌中央代謝系のネットワークを対 象とした.
- 生物学的実験により、摂動対象に対応する、大腸菌の特定の遺伝子を過剰発 現させたデータを使用.
- 観測対象はコハク酸排出フラックス

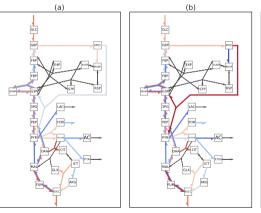
61 / 75

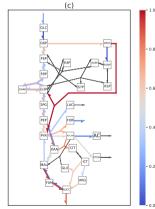
#### Outline

- 1 Umibato: estimation of time-varying microbial interaction using continuous-time regression hidden Markov model
- 2 BayesianSSA: a Bayesian statistical model based on structural sensitivity analysis for predicting responses to enzyme perturbations in metabolic networks
  - ■背景
  - 理論的背景
  - 手法
  - 結果
  - まとめ
- 3 本日のまとめ

## 疑似データによる推定結果の変化

文献に基づいてトイデータを与えてみると、正値性確信度が変化するのが見て取 れる.

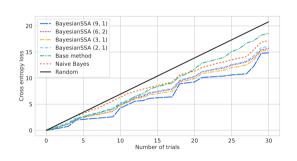




## 予測精度では BayesianSSA が最も良かった

データへのフィットをしない BavesianSSA として base method, データ へのフィットのみをする統計モデルとし て naive Bayes を比較対象とした. 以下の 結論を得た:

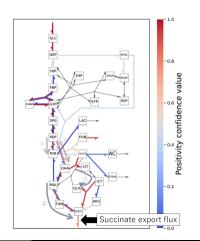
- データ使う必要は?→base method より良いので必要ある
- SSA 使う必要は?→naive Baves より 良いので必要ある
- →BayesianSSA によりデータと SSA を統 合した意味があった



## 最終的に得られた事後分布は一部解釈可能だった

最終的に得られた事後分布で,正値性確信度を計算, 解釈性を検証した

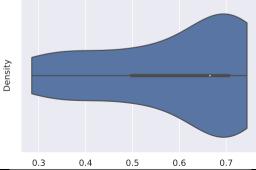
文献	反応	一致
Vemuri, Eiteman, and Altman, 2002	①Glyoxylate pathway	
van Heerden and Nicol, 2013	1) Glyoxylate patriway	•
Vemuri, Eiteman, and Altman, 2002	②Reductive pathway	./
van Heerden and Nicol, 2013	2) Reductive patriway	•
Millard et al., 1996	③PPC	×
Chatterjee et al., 2001; Liang et al., 2015	<b>4PTS</b>	×



## Out-of-sample に対する予測はランダムより良かった

Naive Bayes との比較から推測すると、BayesianSSA は新規の (m,j) に対しての予測にその他の結果を利用できている

そこで、フィットさせたデータにない (m,j) の実験結果に対する予測精度を可視化した



#### Outline

- 1 Umibato: estimation of time-varying microbial interaction using continuous-time regression hidden Markov model
- 2 BayesianSSA: a Bayesian statistical model based on structural sensitivity analysis for predicting responses to enzyme perturbations in metabolic networks
  - ■背景
  - ■理論的背景
  - 手法
  - 結果
  - まとめ
- 3 本日のまとめ

Umibato

まとめ

- 代謝ネットワークにおける理論である SSA にデータを統合する統計モデル BayesianSSA を開発した
- 実データによる精度評価ではデータと理論をそれぞれ片方のみ用いる場合より性能が良く、手法の有効性が示された
- 最終的に得られた事後分布の解釈は過去の文献と部分的に一致, 部分的に不 一致であった.

68 / 75

本日のまとめ ●00

本日のまとめ

## ベイズ統計モデルによる柔軟なモデリング

本日ご紹介した2つの研究はどちらもベイズ統計モデルの開発を主軸にしていま した.

- Umibato:時間変化する確率変数を導入
- BayesianSSA:特定の環境での関数の値を示す確率変数を導入 このような手法の開発に興味があれば、ぜひお声がけください!

細田至温 早稲田大学

70 / 75

本日のまとめ oo●

ご清聴ありがとうございました.

#### Reference I

- Bucci, Vanni et al. (June 2016). "MDSINE: Microbial Dynamical Systems INference Engine for Microbiome Time-Series Analyses". In: Genome Biology 17.1, p. 121. ISSN: 1474-760X. DOI: 10.1186/s13059-016-0980-6.
- Chatterjee, Ranjini et al. (Jan. 2001). "Mutation of the ptsG Gene Results in Increased Production of Succinate in Fermentation of Glucose byEscherichia Coli". In: Applied and Environmental Microbiology 67.1, pp. 148–154. DOI: 10.1128/AEM.67.1.148–154.2001. (Visited on 11/17/2023).
- Embree, Mallory et al. (Dec. 2015). "Networks of Energetic and Metabolic Interactions Define Dynamics in Microbial Communities". In: Proceedings of the National Academy of Sciences of the United States of America 112.50, pp. 15450–15455. ISSN: 1091-6490. DOI: 10.1073/pnas.1506034112.

#### Reference II

- Liang, Quanfeng et al. (Aug. 2015). "Comparison of Individual Component Deletions in a Glucose-Specific Phosphotransferase System Revealed Their Different Applications". In: Scientific Reports 5.1, p. 13200. ISSN: 2045-2322. DOI: 10.1038/srep13200. (Visited on 06/23/2023).
- McGregor, Kevin, Aurélie Labbe, and Celia M T Greenwood (Mar. 2020). "MDiNE: A Model to Estimate Differential Co-Occurrence Networks in Microbiome Studies". In: Bioinformatics 36.6, pp. 1840–1847. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btz824.
- Millard, C S et al. (May 1996). "Enhanced Production of Succinic Acid by Overex-pression of Phosphoenolpyruvate Carboxylase in Escherichia Coli". In: Applied and Environmental Microbiology 62.5, pp. 1808–1810. DOI: 10.1128/aem.62.5. 1808–1810.1996. (Visited on 06/23/2023).

#### Reference III

- Mochizuki, Atsushi and Bernold Fiedler (Feb. 2015). "Sensitivity of Chemical Reaction Networks: A Structural Approach. 1. Examples and the Carbon Metabolic Network". In: Journal of Theoretical Biology 367, pp. 189–202. ISSN: 0022-5193. DOI: 10.1016/j.jtbi.2014.10.025. (Visited on 04/21/2023).
- Okada, Takashi and Atsushi Mochizuki (Aug. 2017). "Sensitivity and Network Topology in Chemical Reaction Systems". In: Physical Review E 96.2, p. 022322. DOI: 10.1103/PhysRevE.96.022322. (Visited on 04/21/2023).
- Toya, Yoshihiro, Takanori Shiraki, and Hiroshi Shimizu (2015). "SSDesign: Computational Metabolic Pathway Design Based on Flux Variability Using Elementary Flux Modes". In: Biotechnology and Bioengineering 112.4, pp. 759–768. ISSN: 1097-0290. DOI: 10.1002/bit.25498. (Visited on 04/21/2023).

#### Reference IV

- van Heerden, Carel D. and Willie Nicol (Sept. 2013). "Continuous and Batch Cultures of Escherichia Coli KJ134 for Succinic Acid Fermentation: Metabolic Flux Distributions and Production Characteristics". In: Microbial Cell Factories 12.1, p. 80. ISSN: 1475-2859. DOI: 10.1186/1475-2859-12-80. (Visited on 11/16/2023).
- Vemuri, G. N., M. A. Eiteman, and E. Altman (Apr. 2002). "Effects of Growth Mode and Pyruvate Carboxylase on Succinic Acid Production by Metabolically Engineered Strains of Escherichia Coli". In: Applied and Environmental Microbiology 68.4, p. 1715. DOI: 10.1128/AEM.68.4.1715–1727.2002. (Visited on 11/21/2023).
- Voigt, Christopher A. (Dec. 2020). "Synthetic Biology 2020–2030: Six Commercially-Available Products That Are Changing Our World". In: Nature Communications 11.1, p. 6379. ISSN: 2041-1723. DOI: 10.1038/s41467-020-20122-2. (Visited on 06/14/2023).

75 / 75