

# 発現時間を考慮した遺伝子ネットワークの相互作用推定

中山 英生<sup>†</sup> 田中 寛人<sup>†</sup> 潮 俊光<sup>†</sup>

<sup>†</sup> 大阪大学 大学院基礎工学研究科  
〒 560-8531 大阪府豊中市待兼山町 1-3  
E-mail: †nakayama@hopf.sys.es.osaka-u.ac.jp

あらまし 本報告では、遺伝子ネットワークモデルの1つである連続時間スイッチングネットワークを対象とする。本モデルにおいて、遺伝子の発現パターン列から遺伝子間の相互作用パラメータを推定する方法は市瀬と合原によって提案されている。しかし、各発現パターンが持続する時間も考慮することができれば、より正確なパラメータ推定が可能になると考えられる。そこで本報告では、遺伝子の発現パターン列と各発現パターンの発現時間から、遺伝子間の相互作用パラメータを推定する方法を提案し、適用例を示す。

キーワード 遺伝子ネットワーク, 連続時間スイッチングネットワーク, 相互作用推定, 発現時間, 線形計画問題

## Estimation of Interaction in a Gene Network using Expression Times

Hideki NAKAYAMA<sup>†</sup>, Hiroto TANAKA<sup>†</sup>, and Toshimitsu USHIO<sup>†</sup>

<sup>†</sup> Graduate School of Engineering Science, Osaka University  
1-3 Machikaneyama, Toyonaka, OSAKA, 560-8531  
E-mail: †nakayama@hopf.sys.es.osaka-u.ac.jp

**Abstract** In this report, we consider a continuous-time switching network, which is one of gene network models. In this model, Ichinose and Aihara proposed a method for estimating interaction parameters among genes by using a specified gene expression pattern sequence. However, if duration time of each gene expression pattern in the specified sequence can be taken into consideration, more exact estimation will be possible. Therefore, in this report, we will propose a new method for estimating the interaction parameters by using a specified gene expression sequence and the duration time, and show an example.

**Key words** gene networks, continuous-time switching networks, estimation of interaction, expression times, linear programming

### 1. はじめに

生体内において遺伝子は mRNA に転写され、さらに翻訳を経てタンパク質が生成されている。しかし、この過程は他のタンパク質の存在・非存在によって促進されたり抑制されたりする相互作用がある。さらに、タンパク質間においても豊富な相互作用があり、その結果、遺伝子からタンパク質への発現過程は複雑なネットワーク構造をしている。例えば、図1ではタンパク質1が遺伝子3の転写を活性化させており、タンパク質2とタンパク質3が相互作用を示し遺伝子1の転写を抑制化している。このような、遺伝子ネットワーク間の構造を解析することは、遺伝子の機能理解さらには生命を理解する上で重要である。また、この相互作用にはバイナリ特性がありタンパク質濃度がある閾値を超えたときに生じる。このとき、そのタンパク質を生成する基になった遺伝子が「発現している」と呼ぶ。

遺伝子ネットワークのモデルの1つである連続時間スイッチングネットワークモデル[1]~[3]では、この遺伝子の発現データより遺伝子ネットワークの相互作用の推定方法が提案されている[4]。しかし、これは発現パターンのみから推定しているものであり、さらに各発現パターンにおける発現時間を考慮すれば、より正確なネットワーク内の相互作用推定が可能になると考えられる。

そこで本報告では、連続時間スイッチングネットワークモデルを対象とし、発現時間が与えられた周期的な発現パターン列をもつように相互作用のパラメータを求める問題を扱う。そして、これが線形計画問題に帰着できることを示し、例を示す。

本報告の構成は以下のとおりである。まず2節では遺伝子ネットワークモデルである連続時間スイッチングネットワークについて述べ、そのダイナミクスについて概説する。3節では、[4]で提案された相互作用推定問題について簡単に述べた後、

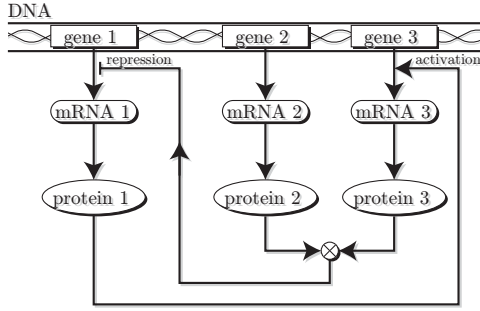


図 1 遺伝子ネットワーク

発現時間が与えられた周期的な発現パターン列を示す遺伝子ネットワークの相互作用の推定方法を提案し、適用例を示す。最後に、4 節でまとめと今後の課題を述べる。

## 2. 遺伝子ネットワーク

### 2.1 連続時間スイッチングネットワークモデル

遺伝子ネットワークモデルとして、本報告では以下で表される連続時間スイッチングネットワーク [1]~[3] を対象とする。

$$\dot{x}_i = -k_i x_i + F_i(y_1, y_2, \dots, y_n), \quad (i = 1, \dots, n).$$

ここで、 $n$  はネットワーク内の遺伝子数、 $x_i$  は遺伝子  $i$  の発現量に関する正規化された実数変数、 $k_i$  はタンパク質  $i$  の分解速度を表す正実数パラメータ、 $y_i$  は以下で定義される遺伝子  $i$  の発現・非発現を表現するバイナリ変数である。

$$y_i \triangleq \begin{cases} 0 & \text{if } x_i < 0, \\ 1 & \text{if } x_i \geq 0. \end{cases}$$

また、 $F_i: \{0, 1\}^n \rightarrow \mathbf{R}$  は全遺伝子から遺伝子  $i$  への相互作用関数であり、以下で表される。

$$F_i(y_1, y_2, \dots, y_n) = a^{(i)} + \sum_{j=1}^n w_j^{(i)} y_j + \sum_{j=1}^n \sum_{h=j+1}^n w_{jh}^{(i)} y_j y_h + \dots + w_{12\dots n}^{(i)} y_1 y_2 \dots y_n.$$

ここで、 $a^{(i)}$  は遺伝子  $i$  の基本発現量に対応するパラメータ、 $w_j^{(i)}$  はタンパク質  $j$  が遺伝子  $i$  の転写因子としての作用強度を表すパラメータ、 $y_i$  の積の項の係数はタンパク質間相互作用による遺伝子  $i$  への作用強度を表すパラメータである。

### 2.2 ダイナミクス

連続時間スイッチングネットワークの区分領域は、遺伝子発現パターンに対応するバイナリベクトル  $(y_1, y_2, \dots, y_n)$  に対応する。濃度変数  $x_i$  の状態空間では、それぞれの区分領域は状態空間のそれぞれの象限になっている。ここで発現パターン  $(y_1, \dots, y_n)$  に対応する濃度変数による状態空間中の区分線形領域を、 $R_{y_1 y_2 \dots y_n}$  と表す。

ここで、各区分領域における平衡点の候補  $e_{y_1 y_2 \dots y_n} = (x_1^*, \dots, x_n^*)$  を考える。ただし、 $x_i^* = F_i/k_i$  である。 $k_i > 0$  であることより、 $e_{y_1 \dots y_n} \in R_{y_1 \dots y_n}$  の場合には、状態は  $R_{y_1 \dots y_n}$  において  $e_{y_1 \dots y_n}$  に収束する。そのため、このような平衡点を

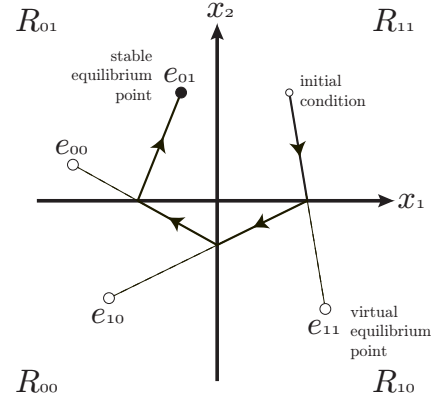


図 2 安定平衡点と仮想平衡点

安定平衡点と呼ぶ。その一方で、 $e_{y_1 \dots y_n} \notin R_{y_1 \dots y_n}$  となる場合もあり、このとき  $e_{y_1 \dots y_n}$  は実際には存在しないが、これを仮想平衡点と呼ぶ。しかし、仮想平衡点であっても状態は仮想平衡点に向かって進むため、その途中で発現パターンが切り替わり新たな安定または仮想平衡点に向かうことになる。つまり、連続時間スイッチングネットワークのダイナミクスは安定平衡点と仮想平衡点の配置に大きく依存する。

図 2 は  $n = 2$  の場合の概念図である。図中 で表された  $e_{01}$  は安定平衡点を、 で表された  $e_{00}$ ,  $e_{10}$ ,  $e_{11}$  は仮想平衡点を表す。例えば  $R_{11}$  の領域に初期値を置いた場合、その軌道は仮想平衡点  $e_{11} \in R_{10}$  に向かって進む。この場合、途中で  $x_1$  軸に衝突し発現パターンが  $(1, 1)$  から  $(1, 0)$  に変化する。発現パターンの変化に伴い、この軌道は衝突以降、仮想平衡点  $e_{10} \in R_{00}$  に向かう。次は  $x_2$  軸と衝突し、仮想平衡点  $e_{00} \in R_{01}$  に向かい、再び  $x_1$  軸と衝突し、 $R_{01}$  に入る。 $e_{01}$  は安定平衡点であるので、軌道は  $e_{01}$  に収束する。

2 つの区分領域の境界が  $(n - 1)$  次元超平面であるとき、これらの 2 領域は隣接していると呼ぶことにする。このとき、図 2 では仮想平衡点が対象となる区分線形領域に隣接した領域にあるが、一般的にはそれ以外の領域に存在しうる。このときの位相面図の例を図 3 に示す。この図では、仮想平衡点  $e_{11}$  が  $R_{00}$  に配置されており、 $R_{11}$  の隣接領域には置かれていない。このとき、 $R_{11}$  内に初期値を設定すると、その位置によって、 $(1, 1) \rightarrow (0, 1)$  または  $(1, 1) \rightarrow (1, 0)$  の 2 種類の遷移が生じうる。この 2 つの遷移が生じる初期値の境界を本報告ではセパトリクスと呼ぶ [4]。

## 3. 発現時間を考慮した相互作用推定

### 3.1 既存手法

連続時間スイッチングネットワークモデルにおいて、1 遺伝子の発現が反転する発現パターン遷移、

$$(y_1, \dots, y_i, \dots, y_n) \rightarrow (y_1, \dots, \bar{y}_i, \dots, y_n),$$

の周期的な列が与えられたとき、相互作用関数中のパラメータを推定する方法はすでに提案されている [4]。ただし、 $\bar{y}_i$  は  $y_i$  の反転パターンである。この方法は、ある配置禁止領域、

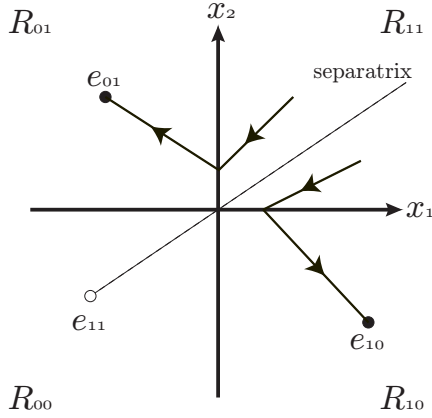


図 3 セバラトリクスの例

$$B = \{(x_1, x_2, \dots, x_n) \mid -b < x_i < b\},$$

を設け、仮想平衡点を遷移先の領域へ次のように配置する。

$$e_{y_1 y_2 \dots y_j \dots y_n} \in R_{y_1 y_2 \dots \bar{y}_j \dots y_n} \setminus B.$$

この仮想平衡点の配置を与えられた発現パターン列中の各遷移に適用することで、与えられた発現パターン列を実現させている。しかし、この方法では 1 遺伝子の発現が反転する発現パターン遷移であるので仮想平衡点は必ず隣の領域に配置されることになる。よって、ある領域からは 1 個の遷移が実現されるのみで、セバラトリクスを発生させるような遷移は実現できない。

### 3.2 提案手法

本報告では、各発現パターン遷移に対してさらに発現時間  $T$  を考慮する。つまり、発現パターン  $(y_1, \dots, y_i, \dots, y_n)$  が時間  $T$  だけ継続し、その後、 $(y_1, \dots, \bar{y}_i, \dots, y_n)$  に遷移しているものとし、これを以下のように表す。

$$(y_1, \dots, y_i, \dots, y_n) \xrightarrow{T} (y_1, \dots, \bar{y}_i, \dots, y_n).$$

発現パターン列は周期的なものとし、ある一つの遺伝子に注目すると、発現と非発現を繰り返す。例えば図 4 では、注目遺伝子  $i$  は時刻  $t^*$  にて発現する。ここから時刻  $T_1$  が経過すると、注目遺伝子以外の遺伝子において、どれかの発現が反転するためダイナミクスが変わる。さらに時刻  $T_2$  が経過したときにも同様にダイナミクスが変わり、最終的にはこれ以後、時刻  $T_3$  が経過したときに注目遺伝子  $i$  の濃度変数  $x_i$  が再び 0 となり、発現から非発現に遷移する。これを一般化して、注目遺伝子  $i$  の遷移の間にはそれ以外の遺伝子に関する  $p-1$  回の遷移が生じ、 $p$  回目に注目遺伝子の発現が反転すると考え、以下のように表す。ただし、 $y^{(j)}$  は発現パターンベクトルである。

$$y^{(1)} \xrightarrow{T_1} y^{(2)} \xrightarrow{T_2} \dots \xrightarrow{T_{p-1}} y^{(p)} \xrightarrow{T_p} y^{(p+1)}.$$

また、対象モデルにおいて発現パターン  $y^{(j)}$  が変化しない間の時刻において、注目遺伝子  $i$  に対する  $x_i$  は、

$$x_i(t) = \frac{F_i(y^{(j)})}{k_i} \left(1 - e^{-k_i t}\right) + x_i(0)e^{-k_i t},$$

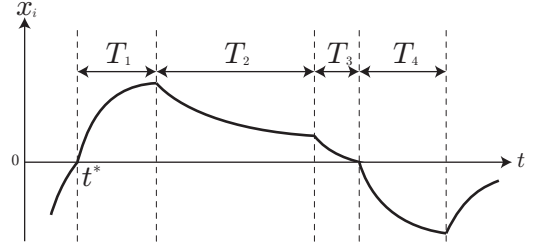


図 4 定式化の概念図

で与えられる。これは単調であり、 $y_i \rightarrow \bar{y}_i$  の遷移が生じる際には  $x_i = 0$  となるので、与えられた発現パターン列を実現するためには以下のように定式化すればよい。

$$y^{(1)} \text{ のとき: } (2y_i - 1)x_i(T_1) \geq d, \quad x_i(0) \triangleq 0,$$

$$y^{(2)} \text{ のとき: } (2y_i - 1)x_i(T_2) \geq d, \quad x_i(0) \triangleq x_i(T_1),$$

$$\vdots$$

$$\vdots$$

$$\vdots$$

$$y^{(p)} \text{ のとき: } x_i(T_p) = 0, \quad x_i(0) \triangleq x_i(T_{p-1}).$$

ただし、 $d$  は正定数であり、また、表現を簡潔にするために発現パターンが変化するたびに時刻を 0 に初期化している。また、仮定より  $y_i$  はこの間不変である。ここで、 $T_j$  と  $y^{(j)}$  は与えられているので上の不等式関係は相互作用パラメータに対する線形不等式となるため、各遺伝子の遷移関係は線形不等式の組で表されることになる。そこで、この連立線形不等式制約から適切な相互作用パラメータ値を決定するために、以下に示す評価関数を用いる。

$$E = \sum_{i=1}^n \left( |a^{(i)}| + \sum_{j=1}^n |w_j^{(i)}| + \sum_{j=1}^n \sum_{h=j+1}^n |w_{jh}^{(i)}| + \dots + |w_{1\dots n}^{(i)}| \right)$$

すなわち、相互作用パラメータの  $L_1$  ノルムの総和である。この評価関数を最小化する最適化問題を解くことにより、与えられた周期的な発現パターン列を再現するパラメータを求めることができる。また、この式に現れるすべてのパラメータ  $w$  に対して、正実数の擬似パラメータ  $w^+$ 、 $w^-$  を導入し、 $w = w^+ - w^-$  とすることで、この最適化問題を線形計画問題にすることができる。

### 3.3 適用例

例として、以下の周期的な発現パターン列を示す遺伝子ネットワークの相互作用推定を行う。

$$\begin{aligned} (0, 0, 0, 0) &\xrightarrow{1} (0, 1, 0, 0) \xrightarrow{2} (0, 1, 1, 0) \xrightarrow{3} (0, 0, 1, 0) \xrightarrow{4} \\ (1, 0, 1, 0) &\xrightarrow{5} (1, 0, 0, 0) \xrightarrow{4} (1, 0, 0, 1) \xrightarrow{3} (1, 1, 0, 1) \xrightarrow{2} \\ (0, 1, 0, 1) &\xrightarrow{1} (0, 1, 1, 1) \xrightarrow{2} (1, 1, 1, 1) \xrightarrow{3} (1, 0, 1, 1) \xrightarrow{4} \\ (1, 0, 0, 1) &\xrightarrow{3} (0, 0, 0, 1) \xrightarrow{2} (0, 0, 0, 0) \xrightarrow{1} \dots \end{aligned}$$

ただし、各タンパク質の分解速度パラメータは以下の通りである。

$$k_1 = 0.1, k_2 = 0.2, k_3 = 0.3, k_4 = 0.4.$$

今、発現パターン遷移の遺伝子 1 について注目すると、遺伝

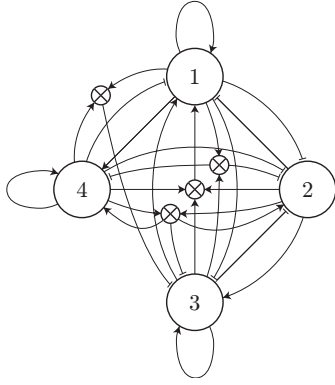


図 5 得られた相互作用関係

子 1 は,

$$0 \xrightarrow{1} 0 \xrightarrow{2} 0 \xrightarrow{3} 0 \xrightarrow{4} 1 \xrightarrow{5} 1 \xrightarrow{4} 1 \xrightarrow{3} 1 \xrightarrow{2} 0 \\ 0 \xrightarrow{1} 0 \xrightarrow{2} 1 \xrightarrow{3} 1 \xrightarrow{4} 1 \xrightarrow{3} 0 \xrightarrow{2} 0 \xrightarrow{1} \dots$$

のように遷移する．このとき， $x_1$  は，

$$x_1(t) = \frac{F_1(\mathbf{y}^{(j)})}{0.1} (1 - e^{-0.1t}) + x_1(0)e^{-0.1t},$$

で与えられる．本提案手法の特徴より  $x_1 = 0$  となる箇所，つまり，注目遺伝子の発現が反転する箇所から定式化し，発現パターン  $(1, 0, 1, 0)$  を最初とする．よって，遺伝子 1 について与えられた発現パターン列を実現するためには，以下のように定式化すればよい．

$$\begin{aligned} (1, 0, 1, 0) \text{ のとき: } & x_1(5) \geq d, & x_1(0) \triangleq 0 \\ (1, 0, 0, 0) \text{ のとき: } & x_1(4) \geq d, & x_1(0) \triangleq x_1(5) \\ (1, 0, 0, 1) \text{ のとき: } & x_1(3) \geq d, & x_1(0) \triangleq x_1(4) \\ (1, 1, 0, 1) \text{ のとき: } & x_1(2) = 0, & x_1(0) \triangleq x_1(3) \\ (0, 1, 0, 1) \text{ のとき: } & x_1(1) \leq -d, & x_1(0) \triangleq 0 \\ & \vdots & \vdots \\ (0, 0, 1, 0) \text{ のとき: } & x_1(4) = 0, & x_1(0) \triangleq x_1(3) \end{aligned}$$

他の遺伝子についても遺伝子 1 と同様に定式化し，3.2 節の評価関数を最小化する線形計画問題を  $d = 0.1$  の条件のもとで解いた結果，以下の相互作用関数を得た．

$$\begin{aligned} F_1 &= 0.0411y_1 - 0.0295y_2 + 0.0255y_3 - 0.0756y_4 \\ &\quad + 0.1248y_2y_3y_4, \\ F_2 &= 0.0903 - 0.0274y_1 - 0.1267y_3 - 0.0386y_4 \\ &\quad - 0.1170y_1y_3 + 0.1335y_2y_4, \\ F_3 &= -0.0163 - 0.0509y_1 + 0.0527y_2 + 0.0548y_3 \\ &\quad - 0.0172y_1y_4 - 0.0493y_2y_4, \\ F_4 &= -0.1213 + 0.1314y_1 + 0.0661y_4 - 0.0377y_1y_3 \\ &\quad + 0.0680y_2y_4. \end{aligned}$$

ここで得られた相互作用関係を図 5 に示す．図中， $\rightarrow$  と  $\vdash$  の矢印は矢印の向いた遺伝子の発現をそれぞれ活性化，抑制化することを意味する．さらに  $\otimes$  はタンパク質間の相互作用を意味

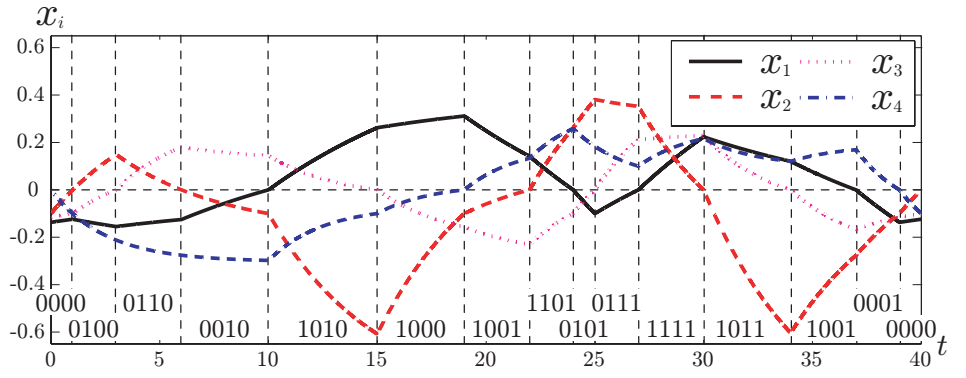


図 6 シミュレーション結果

する．例えば，図 5 において中心の  $\otimes$  に掛かる矢印は，遺伝子 2, 3, 4 が相互作用を示し遺伝子 1 の発現を活性化させていることを意味する．

また，ここで得られた相互作用パラメータを用いてシミュレーションを行った．このときの濃度変数  $x_i$  の時間変化を図 6 に示す．図中に書かれている数字は各時間区間における遺伝子の発現パターンを表している．これより，確かに与えられた発現パターン列が生じていることが確認できる．

このとき， $(1, 0, 0, 1)$  からは  $(1, 1, 0, 1)$  と  $(0, 0, 0, 1)$  の 2 つの発現パターンに遷移している．このことから， $(1, 0, 0, 1)$  ではセパトリクスが生じているのがわかる．実際， $(1, 0, 0, 1)$  における仮想平衡点は，

$$e_{1001} = (-0.345, 0.122, -0.281, 0.190) \in R_{0101},$$

となっており，隣の領域には配置されていないことがわかる．このことから，すでに提案されている手法 [4] よりも豊富なダイナミクスを生成可能な手法であるといえる．

#### 4. おわりに

遺伝子ネットワークにおいて，周期的な発現パターン列と各発現パターンにおける発現時間が与えられた場合の相互作用推定問題を扱い，これが線形計画問題に帰着できることを示した．また，本提案手法の適用例を示した．また，この例においてセパトリクスが生じていることを確認した．

今後の課題としては，本提案手法において得られた遺伝子ネットワークの安定性の解析や，実現できる周期的な発現パターン列の存在条件を調べるなどが挙げられる．

#### 文 献

- [1] L. Glass: “Classification of Biological Networks by their Qualitative Dynamics”, Journal of Theoretical Biology, Vol. 54, pp. 85-107, (1975).
- [2] T. Mestl, C. Lemay, and L. Glass: “Chaos in high-dimensional neural and gene networks”, Physica D, Vol. 98, pp. 33-52, (1996).
- [3] R. Edwards and L. Glass: “Combinatorial explosion in model gene networks”, Chaos, Vol. 10, No. 3, pp. 691-704, (2000).
- [4] 市瀬 夏洋, 合原 一幸: “遺伝子ネットワークモデルとその設計について”, 第 15 回 回路とシステム (軽井沢) ワークショップ, pp. 589-594, (2002).