周期信号に対する時変システム同定法に基づく 酵母の細胞分裂ネットワークの推定

高橋知子 (慶應義塾大学) 東 剛人 (宇都宮大学) 足立修一 (慶應義塾大学)

Estimation of cell-cycle networks in yeast based on an identification method of linear time-varying system using periodic signals

*N. Takahashi (Keio University), T. Azuma (Utsunomiya University) S. Adachi (Keio University)

Abstract— In this paper, we propose a new approach for system identification of linear time-varying systems by using periodic signals. The proposed approach is based on system identification using least-squares method for state-space models. Moreover we apply the proposed approach to an estimation problem of protein networks for cell cycle in yeast and show the efficacy of the proposed approach. Because protein wave patterns are periodic, it is able to identify not only time-invariant systems, but also time-varying systems. By using the proposed identification method of linear time-varying systems, we can contribute to deeper understanding of checkpoints in cell cycle and finding unknown phenomena.

Key Words: System identification, Linear time-varying system, Least-squares method, Systems biology, Cell cycle, Protein network.

1 はじめに

近年の分子生物学の発展により,遺伝子やタンパク質 など生命を構成する要素の理解が飛躍的に進展し,い くつかの疾患の原因遺伝子などの発見が報告されてい る.これらの研究は,原因遺伝子が発見された疾患に 対する有効な治療法の確立や薬品開発に結びつく可能 性があり,非常に重要かつ有意義なものであるが,動 的な振る舞いである生命現象を理解するには不十分で あると考えられる.そこで,生命をシステムとして理 解することを目的とした研究領域として,システムバ イオロジーが提唱され¹⁾²⁾,さまざまな研究が活発に 行われている³⁾⁴⁾.

システムバイオロジー研究における重要な対象の一 つにセルサイクル (Cell Cycle:細胞分裂周期)がある. これは細胞の活動の基幹的部分であるとともに,老化 や癌などの機構にも密接に関係していることが知られ ている.このセルサイクルの研究においては,酵母を 対象とした実験が数多く行われている.これは,酵母 が単細胞の真核生物であり、基本的な生命現象を司る機 構が人間を含む高等真核生物とよく似ているためであ る.また,酵母にはセルサイクル機構の違いから出芽 酵母と分裂酵母があり,特に出芽酵母を対象とした実 験結果が多数報告され,出芽酵母のセルサイクルに関 係したタンパク質間の化学反応を表現するタンパク質 ネットワークがいくつか示されている⁵⁾⁶⁾⁷⁾.現時点に おいても,出芽酵母のセルサイクルに関係するネット ワーク構造すべてが解明されているわけではなく,い まだ発見されていない未知のタンパク質やネットワー ク構造が存在する可能性がある.このような未知のタ ンパク質やネットワーク構造を予測する手法としては 分子生物学の実験ベースのものがあるが,実験設備の 制約などを考慮しなくてはならないため複雑なネット ワーク構造を予測することは困難であると思われる.

そこで,出芽酵母のセルサイクルにおける複雑なネットワーク構造の予測を実現するため,われわれは最小

二乗法に基づくシステム同定法を利用したタンパク質 ネットワークの推定法を提案した⁸⁾.この手法では,実 験などでタンパク質の濃度波形が得られることを考慮 して,これらの濃度波形から状態空間モデルとして記 述された線形システムを同定し,同定された係数行列 の各要素をネットワーク結合を表すと考え,ネットワー ク構造を推定する.数値実験の結果,この手法が既存 のタンパク質ネットワークを含んだものを推定可能で あり,出芽酵母のセルサイクルにおけるタンパク質ネッ トワークの推定に有効であることがわかった.しかし, この手法では,時不変の係数行列を同定するため,セ ルサイクルの一連のプロセスの中で重要な役割を果た すチェックポイントでの各タンパク質の特性までを推 定することが困難であった.これを克服するため時変 の係数行列を推定するシステム同定手法が必要である.

そこで、本稿では、出芽酵母のセルサイクルにおけ るタンパク質の濃度波形が周期性を有することに着目 し、周期信号に対する時変システムの同定法を提案す る.そして、提案法を出芽酵母のセルサイクルにおけ るタンパク質ネットワーク推定問題に適用し、その有 効性を示す、提案するシステム同定法では最小二乗法 に基づき状態空間モデルで記述された線形システムを 同定するが、入手可能な信号が周期的であることを考 慮すると線形時不変システムだけではなく、線形時変 システムの同定も可能となる.さらに、この手法を出 芽酵母のセルサイクルにおけるタンパク質ネットワー ク推定問題に適用することにより、チェックポイント で活性化するタンパク質を推定することが可能となり、 チェックポイントのより深い理解や未知現象の発見に 寄与することが期待できる.

2 周期信号に対する線形システムの同定

ー般的なシステム同定法では,同定対象の入出力デー タから対象の数学モデルを導出することを目的として いる.しかし,出芽酵母のセルサイクルが自律システ ムであり,そのタンパク質濃度波形が入手可能である ことを考慮すると,入力データが存在しなく,出力デー タのみから数学モデルを導出する必要があることがわ かる.そこで,本稿では,状態が既知であると仮定し, 状態空間モデルに対して最小二乗法に基づくシステム 同定法について考察する.

2.1 線形時不変システムの同定 以下のような周期 N の離散時間周期信号

$$\boldsymbol{x}(k) = \boldsymbol{x}(k+N), \ k = 1, 2, \cdots$$

を考える.ただし, $x(k) \in R^n$ である.そして,線形時不変システム

$$\boldsymbol{x}(k+1) = \boldsymbol{A}\boldsymbol{x}(k), \ \ k = 1, 2, \cdots,$$
 (1)

における行列 A を推定する問題を考える.いま,

$$\begin{aligned} \mathbf{Y}(k) &= \mathbf{x}(k+1) \\ \mathbf{\Phi}(k) &= \mathbf{x}(k) \\ \mathbf{\Theta} &= \mathbf{A} \end{aligned}$$

とおくと,(1)式は以下の線形回帰式で記述される.

$$\boldsymbol{Y}(k) = \boldsymbol{\Theta} \boldsymbol{\Phi}(k) \tag{2}$$

ここで,x(k)は周期Nの離散時間信号であるので,Y(k), $\Phi(k)$ と誤差信号

$$\boldsymbol{e}(k) = \boldsymbol{Y}(k) - \boldsymbol{\Theta} \boldsymbol{\Phi}(k)$$

も周期 N の信号になり,以下を満足する.

$$Y(k) = Y(k+N), \quad k = 1, 2, \cdots$$

$$\Phi(k) = \Phi(k+N), \quad k = 1, 2, \cdots$$

$$e(k) = e(k+N), \quad k = 1, 2, \cdots$$

線形回帰式に対しては最小二乗法を適用することが 可能であり,信号 $Y(k), \Phi(k)$ の周期性を考慮するとつ ぎの定理が成立する.

定理 1. (1) 式の時不変システムに対して最初の N 個 のデータが収集されたとき,(2) 式における Θ の最小 二乗推定値は

$$\hat{\boldsymbol{\Theta}}_{1} = \arg\min_{\boldsymbol{\Theta}} J_{1}$$

$$= \left(\sum_{k=1}^{N} \boldsymbol{Y}(k) \boldsymbol{\Phi}^{T}(k)\right) \left(\sum_{k=1}^{N} \boldsymbol{\Phi}(k) \boldsymbol{\Phi}^{T}(k)\right)^{-1} (3)$$

で与えられる.ただし, J1 は評価規範であり,

$$J_1 = \frac{1}{N} \sum_{k=1}^{N} \|\boldsymbol{e}(k)\|^2 \tag{4}$$

で定義される.

いま , $nN(n = 2, 3, \cdots)$ までのデータに対する評価 規範を J_n とおくと ,

$$\hat{\boldsymbol{\Theta}}_n = \arg\min_{\boldsymbol{\Theta}} J_n = \arg\min_{\boldsymbol{\Theta}} J_1 = \hat{\boldsymbol{\Theta}}_1 \tag{5}$$

が成り立つ.

証明.信号 $Y(k), \Phi(k)$ の周期性を考慮すると,評価規範は次式を満足する.

$$J_{n} = \frac{1}{nN} \sum_{k=1}^{nN} ||e(k)||^{2}$$

$$= \frac{1}{nN} \left\{ \sum_{k=1}^{N} ||e(k)||^{2} + \sum_{k=N+1}^{2N} ||e(k)||^{2} + \dots + \sum_{k=(n-1)N+1}^{nN} ||e(k)||^{2} \right\}$$

$$= \frac{1}{nN} n \sum_{k=1}^{N} ||e(k)||^{2}$$

$$= \frac{1}{N} \sum_{k=1}^{N} ||e(k)||^{2}$$

$$= J_{1} \qquad (6)$$

また, J_n を最小にする Θ の推定値 $\hat{\Theta}_n$ は

$$\hat{\boldsymbol{\Theta}}_n = \left(\sum_{k=1}^{nN} \boldsymbol{Y}(k) \boldsymbol{\Phi}^T(k)\right) \left(\sum_{k=1}^{nN} \boldsymbol{\Phi}(k) \boldsymbol{\Phi}^T(k)\right)^{-1} \quad (7)$$

で与えられ,信号 Y(k), $\Phi(k)$ の周期性を考慮すると以下の式が成立し,(5) 式が得られる.

$$\hat{\boldsymbol{\Theta}} = \left(\sum_{k=1}^{nN} \boldsymbol{Y}(k) \boldsymbol{\Phi}^{T}(k)\right) \left(\sum_{k=1}^{nN} \boldsymbol{\Phi}(k) \boldsymbol{\Phi}^{T}(k)\right)^{-1}$$
$$= \left(n \sum_{k=1}^{N} \boldsymbol{Y}(k) \boldsymbol{\Phi}^{T}(k)\right) \left(n \sum_{k=1}^{N} \boldsymbol{\Phi}(k) \boldsymbol{\Phi}^{T}(k)\right)^{-1}$$
$$= \left(\sum_{k=1}^{N} \boldsymbol{Y}(k) \boldsymbol{\Phi}^{T}(k)\right) \left(\sum_{k=1}^{N} \boldsymbol{\Phi}(k) \boldsymbol{\Phi}^{T}(k)\right)^{-1}$$
$$= \hat{\boldsymbol{\Theta}}_{1}$$
(8)

Remark 1. 定理 1 では $n(n = 1, 2, \dots)$ 周期分のデータから最小二乗法による (1) 式の同定問題について考察し,データ数は $nN(n = 1, 2, \dots)$ に限定され,データ数として $nN+1, nN+2, \dots, nN-1(n = 1, 2, \dots)$ の場合の同定問題は考慮していないが,もちろん,このような場合でも最小二乗法によって (1) 式の同定問題を解くことは可能である.

2.2 線形時変システムの同定

定理 1 における $\hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_n$ がそれぞれ $N, 2N, \dots, nN$ までのデータを用いた最小二乗法による推定値であることを考慮すると,推定値を以下のように記述することができる.

$$\hat{\boldsymbol{\Theta}}(nN) = \hat{\boldsymbol{\Theta}}_n, \quad n = 1, 2, \cdots$$

ただし, $\hat{\Theta}(nN)$ は $k = 1, 2, \cdots, nN$ までのデータに対して最小二乗法を用いて得られる推定値を表す.こ

こで, $\hat{\mathbf{\Theta}}(N)$, $\hat{\mathbf{\Theta}}(N+1)$, \cdots , $\hat{\mathbf{\Theta}}(2N)$ までを考えると, 以下のように書くことができる.

$$\hat{\boldsymbol{\Theta}}(N) = \left(\sum_{k=1}^{N} \boldsymbol{Y}(k) \boldsymbol{\Phi}^{T}(k)\right) \left(\sum_{k=1}^{N} \boldsymbol{\Phi}(k) \boldsymbol{\Phi}^{T}(k)\right)^{-1}$$
$$\hat{\boldsymbol{\Theta}}(N+1) = \left(\sum_{k=1}^{N+1} \boldsymbol{Y}(k) \boldsymbol{\Phi}^{T}(k)\right) \left(\sum_{k=1}^{N+1} \boldsymbol{\Phi}(k) \boldsymbol{\Phi}^{T}(k)\right)^{-1}$$
$$\vdots$$
$$\hat{\boldsymbol{\Theta}}(2N) = \left(\sum_{k=1}^{2N} \boldsymbol{Y}(k) \boldsymbol{\Phi}^{T}(k)\right) \left(\sum_{k=1}^{2N} \boldsymbol{\Phi}(k) \boldsymbol{\Phi}^{T}(k)\right)^{-1}$$
$$= \left(2\sum_{k=1}^{N} \boldsymbol{Y}(k) \boldsymbol{\Phi}^{T}(k)\right) \left(2\sum_{k=1}^{N} \boldsymbol{\Phi}(k) \boldsymbol{\Phi}^{T}(k)\right)^{-1}$$

$$= \left(\sum_{k=1}^{N} \boldsymbol{Y}(k) \boldsymbol{\Phi}^{T}(k)\right) \left(\sum_{k=1}^{N} \boldsymbol{\Phi}(k) \boldsymbol{\Phi}^{T}(k)\right)^{-1}$$
$$= \hat{\boldsymbol{\Theta}}(N)$$

これは最小二乗法で用いるデータを N から 2N の間 で変化させると,推定値が時間に依存して変化するこ とを表している.これよりつぎの定理を導くことがで きる.

定理 2. 線形回帰式

$$\boldsymbol{Y}(k) = \boldsymbol{\Theta}(k)\boldsymbol{\Phi}(k) \tag{9}$$

において Y(k), $\Phi(k)$ が周期 N の信号であり, $\Theta(k)$ を未知パラメータとすると,最小二乗推定値 $\hat{\Theta}(k)$ に ついて以下が成立する.

(a) $\hat{\boldsymbol{\Theta}}(nN) = \hat{\boldsymbol{\Theta}}(N), \quad n = 2, 3, \cdots$

(b)
$$\lim_{N_n \to \infty} \hat{\Theta}(N_n) = \hat{\Theta}(N)$$

証明. (a) は定理1より明らかであるので, (b) につい て証明する.N_nまでのデータに対する最小二乗法によ る推定値 $\Theta(N_n)$ は次式で与えられる.

$$\hat{\boldsymbol{\Theta}}(N_n) = \left(\sum_{k=1}^{N_n} \boldsymbol{Y}(k) \boldsymbol{\Phi}^T(k)\right) \left(\sum_{k=1}^{N_n} \boldsymbol{\Phi}(k) \boldsymbol{\Phi}^T(k)\right)^{-1}$$
(10)

ここで, $N_n = nN + N_1, (N_1 = 1, 2, \cdots, N - 1)$ とす ると

$$\sum_{k=1}^{nN+N_1} \mathbf{Y}(k) \mathbf{\Phi}^T(k) = n \sum_{k=1}^{N} \mathbf{Y}(k) \mathbf{\Phi}^T(k) + \sum_{k=1}^{N_1} \mathbf{Y}(k) \mathbf{\Phi}^T(k)$$
(11)

$$\sum_{k=1}^{nN+N_1} \mathbf{\Phi}(k) \mathbf{\Phi}^T(k) = n \sum_{k=1}^{N} \mathbf{\Phi}(k) \mathbf{\Phi}^T(k) + \sum_{k=1}^{N_1} \mathbf{\Phi}(k) \mathbf{\Phi}^T(k)$$
(12)

となり, (11) 式および (12) 式において n が十分に大き い場合には第1項目に比較して第2項目は非常に小さ くなることがわかる.したがって,

 n_{-}

$$\lim_{n \to \infty} \hat{\boldsymbol{\Theta}}(nN + N_1)$$

$$= \left(n \sum_{k=1}^{N} \boldsymbol{Y}(k) \boldsymbol{\Phi}^T(k)\right) \left(n \sum_{k=1}^{N} \boldsymbol{\Phi}(k) \boldsymbol{\Phi}^T(k)\right)^{-1}$$

$$= \hat{\boldsymbol{\Theta}}(N)$$
(13)

が成立する . $n \to \infty$ は $N_n = nN + N_1$ より $N_n \to \infty$ と等価であるので,(b)が成立する.

Remark 2. 定理 2は,周期信号 Y(k), $\Phi(k)$ に対し て最小二乗法を用いて時不変の未知パラメータを推定 する場合,その推定値 Θは1 周期分のデータに対する 推定値に収束することを示している.したがって,周期 信号に対する時不変システムの同定においては最低限 1 周期分のデータがあればよく,同定で使用するデー タ数を大きくする必要がないことがわかる.

Remark 3. 推定値 $\hat{\Theta}(k)$ は $k = N, 2N, \dots, nN$ にお いて等しくなるが,その他においては等しくない.し たがって,推定値 $\hat{\Theta}(k)$ は時変になるが,周期的ではな い.また,定理2の証明における(11)式および(12)式 の第2項目が推定値 $\Theta(nN+N_1)$ に与える影響はnが 大きくなるにつれて小さくなるので,推定値 $\hat{\Theta}(k)$ は kが大きくなるにつれて1周期分のデータに対する推 定値 $\hat{\Theta}(N)$ に漸近していく.

以上の議論を線形時変システム

$$\boldsymbol{x}(k+1) = \boldsymbol{A}(k)\boldsymbol{x}(k), \quad k = 1, 2, \cdots$$
 (14)

に対するシステム同定問題に適用することを考える. (14) 式は (9) 式の線形回帰式で記述可能であるので, 周期 N の離散時間信号 $x(k) \in \mathbb{R}^n, k = 1, 2, \cdots$ が入手 可能であれば最小二乗法を用いて A(k) の推定値が得 られ,推定されたA(k)は定理2の性質を有する.

出芽酵母のセルサイクルモデルと 3 タンパク質ネットワーク推定問題

3.1 セルサイクルの数学モデル

Chen による出芽酵母のセルサイクルモデル⁶⁾は,48 個の連立非線形微分方程式と数個のリセットルールで 表わされる数学モデルである.次節でネットワークを 推定する 48 個のなかの 9 個のタンパク質については, 以下の式で与えられる.

$$\frac{d[\text{mass}]}{dt} = k_g[\text{mass}]$$
(15)
$$\frac{d[\text{Clb2}]}{dt} = (k'_{s,b2} + k''_{s,b2}[\text{Mcm1}])[\text{mass}] + (k_{d3,c1}[\text{C2P}] + k_{di,b2}[\text{C2}]) + (k_{d3,f6}[\text{F2P}] + k_{di,f2}[\text{F2}]) - (V_{d,b2} + k_{as,b2}[\text{Sic1}] + k_{as,f2}[\text{Cdc6}])[\text{Clb2}]$$
(16)
$$\frac{d[\text{Cdh1}]}{dt} = k_{s,cdh} - k_{d,cdh}[\text{Cdh1}] + \frac{V_{a,cdh}([\text{Cdh1}]_{\text{T}} - [\text{Cdh1}])}{J_{a,cdh}([\text{Cdh1}]_{\text{T}} - [\text{Cdh1}])} - \frac{V_{i,cdh}[\text{Cdh1}]}{J_{i,cdh} + [\text{Cdh1}]}$$
(17)
$$\frac{d[\text{Cdc20}]_{\text{T}}}{dt} = k'_{s,20} + k''_{s,20}[\text{Mcm1}] - k_{d,20}[\text{Cdc20}]_{\text{T}}$$
(18)

$$\frac{d[\operatorname{Cdc20}]_{A}}{dt} = (k'_{a,20} + k''_{a,20}[\operatorname{APC-P}])([\operatorname{Cdc20}]_{T} - [\operatorname{Cdc20}]_{A} - (k_{mad2} + k_{d,20})[\operatorname{Cdc20}]_{A}$$
(19

$$\frac{d[\operatorname{APC-P}]}{dt} = \frac{k_{a,apc}[\operatorname{Clb2}](1 - [\operatorname{APC-P}])}{J_{a,apc} + 1 - [\operatorname{APC-P}]}$$
(20

$$\frac{d[\operatorname{Sic1}]}{dt} = (k'_{s,c1} + k''_{s,c1}[\operatorname{Swi5}]) + (V_{d,b2} + k_{di,b2})[\operatorname{C2}] + (V_{d,b5} + k_{di,b5})[\operatorname{C5}] + k_{pp,c1}[\operatorname{Cdc14}][\operatorname{Sic1P}] - (k_{as,b2}[\operatorname{Clb2}] + k_{as,b5}[\operatorname{Clb5}] + V_{kp,c1})[\operatorname{Sic1}] + V_{kp,c1})[\operatorname{Sic1}] + V_{kp,c1})[\operatorname{Sic1}] + (k'_{s,n2} + k''_{s,n2} \cdot [\operatorname{SBF}]) \cdot [\operatorname{mass}] - k_{d,n2} \cdot [\operatorname{Cln2}]$$
(22

$$\frac{d[\operatorname{Sic1P}]}{u} = V_{kp,c1}[\operatorname{Sic1}] - (k_{pp,c1}[\operatorname{Cdc14}] + k_{d3,c1})[\operatorname{Sic1P}]$$

$$dt = V_{kp,c1}[\text{5C1}] - (v_{pp,c1}[\text{5C1}] + v_{d3,c1})[\text{5C1}] + V_{d,b2}[\text{C2P}] + V_{d,b5}[\text{C5P}]$$
(23)

この数学モデルを Simulink で構成し,シミュレー ションを行うと,48 個の周期的なタンパク質濃度波形 が得られる.例として,(15)~(23) 式で与えられたタ ンパク質の濃度波形を Fig.1 に示す.図より,いずれも 周期が約 100 分の周期信号であることがわかる.こう して得られたタンパク質濃度から,2節で提案した周 期信号に対する線形システムの同定法を用いてタンパ ク質ネットワークを推定する.本稿で考察するタンパ ク質ネットワーク推定問題は文献⁸⁾に基づいている.

3.2 タンパク質ネットワークの推定

ここでは,2.1 節に示した時不変システム同定法と同様な手法に基づいた従来法⁸⁾と,2.2 節に示した時変システム同定法を用いる提案法の二つの方法によってネットワークを推定する.

状態変数 x(k)を, 3.1 節で述べたシミュレーション により得られるタンパク質濃度とすると, x(k) は既知 である.よって, (2) 式において

$$\Theta = A_{cell}$$

とおけば,時不変システムの場合,セルサイクルモデ ルを表わすタンパク質ネットワークは以下のような線 形回帰式で表わされる.

$$\boldsymbol{Y}(k) = \boldsymbol{A}_{cell} \boldsymbol{\Phi}(k) \tag{24}$$

また,時変システムの場合は(9)式において

$$\boldsymbol{\Theta}(k) = \boldsymbol{A}_{cell}(k)$$

とおけば

$$\mathbf{Y}(k) = \mathbf{A}_{cell}(k)\mathbf{\Phi}(k) \tag{25}$$

と表わされる.ただし, (24), (25) 式ともに

$$Y(k) = x(k+1)$$
$$\Phi(k) = x(k)$$



(a) mass

(c) Cdh1

0.6

ution at i

190 O. 2











(i) Sic1P



である.また,

	a_{11}	a_{12}	•••	a_{19}
$oldsymbol{A}_{cell} =$	a_{21}	a_{22}	• • •	a_{29}
	÷	÷	·	÷
	a_{91}	a_{92}	• • •	a_{99}

であり, *A_{cell}* の各要素はタンパク質間の影響力を表している.つまり,タンパク質ネットワーク推定問題

では,未知パラメータ O はタンパク質の結合強度を表 わす行列となる.各パラメータをタンパク質強度と呼 ぶと,その値が大きいほどそのタンパク質が及ぼす影 響力が大きいことになる.よって,パラメータの値が 十分大きければその2つのタンパク質間に結合がある と判断することができる⁸⁾.

まず, Chenのセルサイクルモデルに含まれている48 個のタンパク質のうち,セルサイクル全体に関わって いる9個のタンパク質を選び,9次元のタンパク質ネッ トワークを推定する.x(k)はその選んだ9個のタンパ ク質濃度とするので,

 $\boldsymbol{x}(k) = \begin{bmatrix} [mass] & [Clb2] & [Cdh1] \\ & [Cdc20]_T & [Cdc20]_A & [APC - P] \\ & [Sic1] & [Cln2] & [Sic1P] \end{bmatrix}^T$

となる.Fig.1を見ると,各タンパク質の濃度変化には ばらつきがあるので,濃度変化を比較するために,各タ ンパク質濃度の波形を正規化して濃度変化を1とした.

- 4 出芽酵母のセルサイクルにおけるタンパ ク質ネットワークの推定結果
- 4.1 時不変システム同定法を用いた推定結果

まず,従来法を用いた推定結果について述べる.各 出力波形のデータ数を 20000 個 (シミュレーション時間:200 min,0.01 min),すなわち $N_n = 20000$ として, (10) 式より $A_{cell} (= \hat{\Theta}(Nn))$ を計算した.また,得ら れた A_{cell} の各行に注目し,その行の絶対値の最大値 と,その最大値の約 1/1000 以上の値だけを「結合あ り」と判断し,それ以外の値を0とする.1/1000 とい う基準は理論的に導かれた数値ではないが,既に存在 が知られている既知のネットワークを推定できるよう に定めた基準である.すると, A_{cell} は以下のような行 列 A'_{cell} となる.

	9.95×10^{-1}	1.63×10^{-3}	-1.36×10^{-3}
	-1.10×10^{-1}	³ 1.00	0
	2.37×10^{-3}	-3.85×10^{-3}	3 1.00
	1.13×10^{-3}	0	-1.32×10^{-3}
$A_{cell}' =$	2.55×10^{-3}	4.43×10^{-3}	0
	0	4.45×10^{-3}	0
	0	0	0
	0	0	0
	0	0	0
	1.07×10^{-3}	0	2.04×10^{-3}
	0	0	-2.91×10^{-3}
	0	0	1.68×10^{-3}
	1.00	0	0
	-2.95×10^{-3}	1.00	-3.30×10^{-3}
	-1.21×10^{-3}	0	9.97×10^{-1}
	0	9.99×10^{-4}	-1.33×10^{-3}
	-1.23×10^{-3}	0	1.14×10^{-3}
	0	0	0
	1.26×10^{-3}	2.45×10^{-3}	0]
	0	0	0
	0	-1.19×10^{-3}	0
	0	0	9.67×10^{-4}
	0	-1.20×10^{-3}	0
	0	0	0
	1.00	0	-1.67×10^{-3}
	0	1.00	0
	2.58×10^{-3}	0	9.99×10^{-1}]



Fig. 2: An estimated 9-dimensional protein network. Known ones (black) and new ones (red). Bolder lines indicate stronger protein combinations.

A'_{cell}の対角成分がそれぞれほぼ 1 であるので , 自己 フィードバックが強いことがわかる .

得られた A'_{cell} からタンパク質ネットワークを推定し た結果を Fig.2 に示す.Fig.2 において,矢印が太いほ ど結合が強いことを表している.また,黒色の線は既 に存在が知られているネットワーク,赤色の線が今回 新たに推定された未知のネットワークである.このよ うに,既知のネットワークに加えて未知のネットワー クも推定できたことから,システム同定によってタン パク質ネットワークを推定する方法は有効であること がわかった.

4.2 時変システム同定法を用いた推定結果

2.2 節で述べた提案法を用いたときの推定結果について述べる.サンプリング周期を 0.8 秒とすると,タンパク質濃度の周期は N = 7549 なので,まず 1 周期分のデータ 0~N サンプルを用いて $\hat{\Theta}(N)$ を計算し, $A_{cell}(1)$ とした.つぎに 0 から N + m ($m = 1, 2, \cdots, 30451$) と1 サンプル毎に $A_{cell}(m) = \hat{\Theta}(N + m)$ を計算した. $A_{cell}(m)$ は 9×9 行列なので要素は 81 個あるが,例として $A_{cell}(m)$ の(1,1) 要素と (2,7) 要素の図をそれぞれ Figs.3,4 に示す.両図ともに,nN (n = 1, 2, 3, 4) においては $A_{cell}(1)$ と等しくなり,それ以外の点ではデータ数が多くなるにつれて推定値 が $A_{cell}(1)$ に漸近していくことがわかる.この結果は 定理 2 と一致しており,2.2 節で提案した時変システム 同定法が有効であることがわかった.

また, Fig.3 は *A*_{cell}(*m*) の (1,1) 要素であり, mass の自己フィードバックを示しているので, セルサイク ルを通して値はほぼ1である.そして,時間に対する 変化を見てみると,周期毎にインパルス的なふるまい が現われていることがわかる.massは細胞の大きさを 表しており,時変システムとして同定することで M 期 の開始と終了(細胞分裂の開始と終了)を判別するこ とができる.

つぎに, Fig.4 は $A_{cell}(m)$ の (2,7)要素であり, これ は Sic1 が Clb2 に与える影響力を表している. Sic1 は Clb2 を分解してその発現量を抑えているタンパク質で ある. Clb2 は特に M 期に発現量が増すタンパク質で あるので, Sic1 が Clb2 に与える影響も M 期に進むに つれて大きくなると考えられるが, Fig.4 では M 期に おいて特に Sic1 が Clb2 に与える影響が大きいことが わかる. このように,セルサイクルの各時期において,ある タンパク質が自他のタンパク質に与える影響を理解す ることができる.したがって,提案した時変システム 同定法を用いることで,セルサイクルの4つのチェック ポイントに影響を与えているタンパク質を予測できる 可能性がある.



5 おわりに

本稿では,周期信号に対する最小二乗法に基づいた 線形システム同定法を提案し,時不変システム同定法 と時変システム同定法の2つを与えた.時不変システ ムの同定では周期信号が得られる場合での最小二乗推 定値の性質を示し,その結果を利用することで時変シ ステムの同定法が新たに得られた.そして,出芽酵母 のセルサイクルに対して提案法を適用し,9次元のタン パク質ネットワーク推定した.その推定結果から,時不 変および時変システム同定法の有効性が明らかになっ た.また,時変システム同定法から,セルサイクルで 重要な役割を持つチェックポイントでの各タンパク質 の活性化の様子が理解できることがわかった.

参考文献

- H. Kitano: Systems Biology : a brief overview, Science, Vol. 295, No. 5560, pp. 1662-1664 (2002)
- H. Kitano: Computational Systems Biology, Nature, Vol. 420, No. 6912, pp. 206-210 (2002)
- 3) 北野宏明:システムバイオロジー,秀潤社(2001)
- 4) 北野宏明,東剛人:システムバイオロジ-と制御, システム/制御/情報, Vol. 48, No. 3, pp. 104-111 (2004)
- 5) J. Tyson, Modeling the Cell Division Cycle : Cdc2 and Cyclin Interactions, PNAS Vol. 88, pp. 7328-7332 (1991)
- 6) K. C. Chen, L. Calzone, A. Csikasz-Nagy, F. R.Cross, B. Novak, J. J. Tyson, Integrative analysis of cell cycle control in budding yeast, Molecular Biology of the Cell, Vol.15, pp. 3841-3862 (2004)
- 7) 東剛人,守屋央朗,北野宏明:真核生物のセルサイ クルにおける基本システムの理解(システムと生 命特集号),システム/制御/情報,Vol. 50,No. 8,pp.309-314 (2006)
- 8) 倉田, 榎戸,東,足立,酵母のセルサイクルに関す るタンパク質ネットワークの推定とロバストネス 解析,第37回制御理論シンポジウム,pp.139-142 (2008)