

線形計画マシンを用いた遺伝子ネットワーク同定: ブートストラップ法による精度推定

木村 周平^{†1} 松村 幸輝^{†1}

近年、我々は線形計画マシンを用いて遺伝子ネットワークを同定する方法を提案した。この手法は線形計画問題を解くことで遺伝子ネットワークを同定するため、計算時間が非常に短いという特長がある。しかしながら推定された遺伝子間相互作用には偽陽性が多く、推定結果を生化学実験に活用するにはまだ問題があった。そこで本研究ではブートストラップ法を利用して、遺伝子間相互作用の信頼度を評価する方法を提案する。

Inference of Genetic Networks using LPMs: Assessments of Confidence via Bootstrap Methods

SHUHEI KIMURA^{†1} and KOKI MATSUMURA^{†1}

Recently, we have proposed a genetic network inference method using linear programming machines (LPMs). As the proposed method infers genetic networks by solving linear programming problems, its computational time is very short. The networks inferred by the method however contain a number of false-positive regulations. In this study, we therefore propose a method to compute a confidence of each regulation using a bootstrap method.

1. はじめに

DNA マイクロアレイ技術などの進歩によって細胞全体レベルでの遺伝子発現パターンの計測が可能になってきている。これらの技術を活用するためには膨大な量のデータから必要な情報を抽出する方法が必要である。これらのデータからの情報抽出方法の一つとして、

遺伝子ネットワークの同定に注目が集まっている。遺伝子ネットワークの同定とは、マイクロアレイなどによって得られた遺伝子発現量の時系列データから遺伝子間相互作用を推定する問題のことである。同定した遺伝子ネットワークモデルは生化学者による仮説の生成や実験デザインに利用可能であり、また未知の遺伝子の機能を推定するためにも利用できると考えられる。その有用性から遺伝子ネットワーク同定法として、これまでに様々な手法が提案されている^{1)-3),6),7),10)-12),14)-16),19),20)}。

遺伝子ネットワーク同定に必要な計算時間の非常に短い方法として、我々は近年、線形計画マシンを用いた遺伝子ネットワーク同定法を提案した¹³⁾。この手法は判別問題を解くことによって遺伝子ネットワークを同定するという点で従来手法とは異なっている。提案手法は定義された判別問題を線形計画問題として定式化し、これを解くことによって遺伝子ネットワークを同定する。そのため本手法の計算時間は非常に短いという特長があった。しかしながらこの手法によって同定された遺伝子ネットワークには非常に多くの偽陽性相互作用が含まれるという欠点があった。

そこで本研究では線形計画マシンを用いた遺伝子ネットワーク同定法によって推定された遺伝子間相互作用の信頼度を求める方法を提案する。そのために本研究ではブートストラップ法を使用する。最後に、人工的に作成した遺伝子ネットワーク同定問題への適用を通して、信頼度の高い相互作用には確かに偽陽性が少ないことを確認する。

2. 線形計画マシンを用いた遺伝子ネットワーク同定法

本研究では遺伝子ネットワークを同定するために、線形計画マシンを用いた遺伝子ネットワーク同定法¹³⁾を利用する。そこで本章ではまず、この方法について説明する。

2.1 問題定義

線形計画マシンを用いた遺伝子ネットワーク同定法では N 個の遺伝子からなる遺伝子ネットワークの同定を、 N 個の判別問題として定式化する。判別問題とは K 個の訓練事例

$$(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_K, y_K),$$

をもとにして、その識別ルール (識別器) を構成する問題である。ただし $\mathbf{x}_k$ は k 番目の特徴ベクトルであり、 y_k はそのクラスラベルである。遺伝子 n を制御する遺伝子群を見つけるようにする場合、その訓練事例は

$$(\mathbf{X}_{|t_1}, Y_{t_1}), (\mathbf{X}_{|t_2}, Y_{t_2}), \dots, (\mathbf{X}_{|t_K}, Y_{t_K}),$$

となる。ただし $\mathbf{X}_{|t_k} = (X_1|_{t_k}, X_2|_{t_k}, \dots, X_N|_{t_k})$ は時刻 t_k における全遺伝子の発現量、 Y_{t_k} は $\mathbf{X}_{|t_k}$ の属するクラスのラベルを表す。遺伝子 n に対応する判別問題では、ある時刻における

^{†1} 鳥取大学大学院工学研究科
Graduate School of Engineering, Tottori University

全遺伝子の発現量が与えられたときに、その時刻における遺伝子 n の発現量の変化量の符号を予測することのできる予測器を得ることが目的となる。

遺伝子 n に対応する判別問題の訓練事例を生成する場合、線形計画マシンを用いた遺伝子ネットワーク同定法では遺伝子 n の発現変化量の符号に応じて Y_{t_k} に ‘+’, ‘0’, 又は ‘-’ のラベルを割り当てる。従って本研究において訓練事例を生成する際には、全ての時刻における遺伝子発現量の変化量の値を推定することが必要となる。この値は通常、観測された遺伝子発現量の時系列データから、何らかの平滑化手法を利用して推定される^{2),17),19)}。一般に推定された遺伝子発現変化量には誤差が含まれるため、文献 13) の方法では以下に従って訓練事例にクラスラベルを割り当てる: 1) もし時刻 t_k における遺伝子 n の発現量の変化量が閾値 σ 以上 (つまり $\left. \frac{dX_n}{dt} \right|_{t_k} \geq \sigma$) であれば、 Y_{t_k} にラベル ‘+’ を割り当てる; 2) 同様に $\left. \frac{dX_n}{dt} \right|_{t_k} \leq -\sigma$ の場合、 Y_{t_k} にラベル ‘-’ を割り当てる; 3) さもないければ Y_{t_k} にラベル ‘0’ を割り当てる。

2.2 識別器の学習

前述したように文献 13) の方法では N 遺伝子からなる遺伝子ネットワークを同定するために、 N 個の識別器を学習する必要がある。本節では遺伝子 n に対応する識別器を得るための方法について説明する。

2.2.1 識別関数

線形計画マシンを用いた遺伝子ネットワーク同定法では以下の識別関数を利用して入力データを分類する。

$$f(\mathbf{X}) = b + \sum_{m=1}^N w_m \log X_m. \quad (1)$$

ただし入力ベクトル $\mathbf{X} = (X_1, X_2, \dots, X_N)$ は全遺伝子の発現量、重みベクトル $\mathbf{w} = (w_1, w_2, \dots, w_N)$ および b はパラメータである。本手法では $f(\mathbf{X}) > 0$, $f(\mathbf{X}) = 0$, $f(\mathbf{X}) < 0$ のときに、それぞれ入力ベクトル $\mathbf{X}$ がクラス ‘+’, クラス ‘0’, クラス ‘-’ に属すると判断する。従ってこの判別問題の目的は、与えられた学習事例を正しく分類できるような識別関数のパラメータ $\mathbf{w}$ と b を推定することとなる (図 1)。識別関数 (1) は遺伝子発現量の対数に対して線形であり、非常に単純な形をしている。しかしながらこの判別関数は生化学反応系のモデルとしてしばしば使用される S-system¹⁸⁾ と関係があることが示されている¹³⁾。

2.2.2 パラメータ推定

妥当な識別器を得るために、識別関数 (1) のパラメータ $\mathbf{w}$ と b を推定することが必要である。文献 13) の方法では線形計画マシン (Linear Programming Machine)⁸⁾ を利用してパラメータ $\mathbf{w}$ と b の推定を行っている。これによりパラメータの推定問題は、以下の制約付き

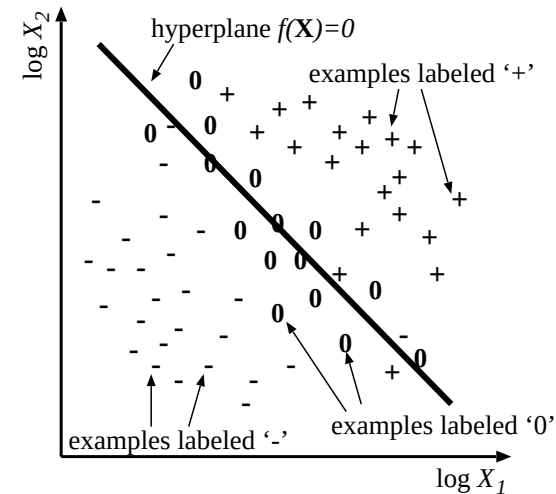


図 1 識別関数 (1) を用いた識別。

Fig. 1 A classification using the discriminant function (1).

関数最小化問題として定式化される。

$$\begin{aligned} & \text{minimize}_{\mathbf{w}, b, \xi_i^+, \xi_i^-, \eta_i^+, \eta_i^-} \sum_{m=1}^N |w_m| + C_1 \sum_{i=1}^{K^+} \xi_i^+ + C_1 \sum_{i=1}^{K^-} \xi_i^- + C_2 \sum_{i=1}^{K^0} (\eta_i^+ + \eta_i^-), \end{aligned} \quad (2)$$

subject to

$$\begin{cases} b + \sum_{m=1}^N w_m \log X_{m,i}^+ \geq 1 - \xi_i^+ & (i = 1, 2, \dots, K^+), \\ \xi_i^+ \geq 0 & (i = 1, 2, \dots, K^+), \\ b + \sum_{m=1}^N w_m \log X_{m,i}^- \leq -1 + \xi_i^- & (i = 1, 2, \dots, K^-), \\ \xi_i^- \geq 0 & (i = 1, 2, \dots, K^-), \\ b + \sum_{m=1}^N w_m \log X_{m,i}^0 \leq \eta_i^+ & (i = 1, 2, \dots, K^0), \\ \eta_i^+ \geq 0 & (i = 1, 2, \dots, K^0), \\ b + \sum_{m=1}^N w_m \log X_{m,i}^0 \geq -\eta_i^- & (i = 1, 2, \dots, K^0), \\ \eta_i^- \geq 0 & (i = 1, 2, \dots, K^0), \end{cases}$$

ただし ξ_i^+ , ξ_i^- , η_i^+ and η_i^- は補助変数である。 C_1 と C_2 は定数パラメータであり、 $C_1 = \frac{200}{N\sqrt{K}}$ と $C_2 = 0.4C_1$ が推奨されている。なお N は対象とするネットワークに含まれる遺伝子数、 K

は訓練事例の数である。 K^+, K^0, K^- はそれぞれ、クラス ‘+’, ‘0’, ‘-’ の事例数であり、 $X_{m,i}^+, X_{m,i}^0, X_{m,i}^-$ はクラス ‘+’, ‘0’, ‘-’ に属する訓練事例に含まれる i 番目の事例の遺伝子 m の発現量を表している。この問題は線形計画問題に変換することが可能であるため、容易に解くことができる。本研究では定義された線形計画問題を内点法を利用して解いている。

2.3 遺伝子間相互作用の推定

遺伝子 n に対応する識別器は、遺伝子 n を制御する遺伝子群に関する情報を持つ。本節ではこの情報を抽出するための方法について説明する¹³⁾。

遺伝子間相互作用を推定するために、感度係数がしばしば使用される¹²⁾。正の感度係数は正の制御が存在することを、負の感度係数は負の制御が存在することをそれぞれ意味する。また感度係数の絶対値の大きさは遺伝子間相互作用の強さを表す。文献 13) では以上の考え方を利用して、得られた識別器からの遺伝子間相互作用に関する情報の抽出を行う。

ある入力ベクトル $\mathbf{X}$ に対して識別関数 $f(\mathbf{X})$ の値が負であった場合、識別器は「入力ベクトル $\mathbf{X}$ はクラス -1 に属する」と判断する。このとき識別関数の重みパラメータ w_m の値が正であれば、 X_m の値を十分に増加させることで識別結果を +1 に変えることができる。本研究におけるクラスラベルは遺伝子 n の発現量の変化量の符号を表しているので、このような識別結果の変化は感度係数が正であることを示唆する。同様に、重みパラメータ w_m が負の場合、感度係数が負であることを意味する。以上のことから、識別関数における重みベクトル $\mathbf{w}$ を利用して、遺伝子間相互作用に関する情報が抽出できることが分かる。本研究では $w_m \geq T_h$ の場合、遺伝子 n は遺伝子 m から正の制御を受けていると判断する。同様に $w_m \leq -T_h$ のとき、遺伝子 n は遺伝子 m から負の制御を受けていると判断する。それ以外の場合は、遺伝子 n は遺伝子 m から制御を受けていないとみなす。ただし T_h は閾値であり、

$$T_h = \delta \max |w_l| \quad (3)$$

である。なお δ ($0 \leq \delta \leq 1$) はパラメータである。

3. ブートストラップ法による遺伝子間相互作用の信頼度評価

前章で説明したように線形計画マシンを用いた遺伝子ネットワーク同定法¹³⁾は、DNA マイクロアレイなどによって計測された遺伝子発現データのみから遺伝子ネットワークを同定する。従って同定された遺伝子ネットワークは、観測されたデータに含まれるノイズの影響を受ける。ところがこの手法は、推定された遺伝子間相互作用が本質的なものなのか、それともノイズの影響によって偶然得られたものなのかを見分けることが出来なかった。

上記の問題を解決するために本研究ではブートストラップ法⁵⁾を利用して、推定された遺伝子間相互作用の信頼度を評価する。ブートストラップ法とは、母集団に関する何らかの推定量の性質を知るための一般的な方法である。提案手法によって信頼度が低いと評価された相互作用は、ノイズの影響によってたまたま得られたものである可能性が高いと考えることが出来る。以下では遺伝子 m から遺伝子 n への制御の信頼度を評価するためのアルゴリズムを説明する。

(1) 訓練事例の複製の生成

2.1 節で生成された遺伝子 n に対応する判別問題のための K 個の訓練事例

$$(\mathbf{X}|_{t_1}, Y_{t_1}), (\mathbf{X}|_{t_2}, Y_{t_2}), \dots, (\mathbf{X}|_{t_K}, Y_{t_K}),$$

から重複を許して K 個の事例をランダムに選ぶことで、訓練事例の複製を生成する。

(2) 訓練事例の複製からの遺伝子ネットワーク同定

ステップ (1) で生成された訓練事例の複製から、前章で説明した手法を用いて遺伝子ネットワークを同定する。

(3) 遺伝子ネットワークの複数同定

上記のステップ (1) と (2) を B 回繰り返すことで、 B 個の異なる遺伝子ネットワークを得る。

(4) 遺伝子間相互作用の信頼度評価

得られた B 個のネットワークを用いて遺伝子間相互作用の信頼度を評価する。ただし遺伝子間相互作用には正の制御と負の制御の二種類があることに注意する。ステップ (3) で得られた B 個の遺伝子ネットワークの中で遺伝子 m から遺伝子 n への正の制御と負の制御の含まれているネットワークの個数をそれぞれ C_{mn}^+, C_{mn}^- としたとき、遺伝子 m から遺伝子 n への正の制御の信頼度 p_{mn}^+ と負の制御の信頼度 p_{mn}^- をそれぞれ

$$p_{mn}^+ = \frac{C_{mn}^+}{B}, \quad p_{mn}^- = \frac{C_{mn}^-}{B} \quad (4)$$

とする。

なお上記のアルゴリズムは特に遺伝子 m から遺伝子 n への制御の信頼度のみを評価しているが、例えば遺伝子 l から遺伝子 n への信頼度評価のために別の B 回の繰り返しを行う必要は必ずしもないことに注意が必要である。 B 個の遺伝子ネットワークからは、全ての遺伝子間相互作用の信頼度が同時に得られる。

4. 数値実験

提案手法によって計算された遺伝子間相互作用の信頼度が妥当なものであることを確認するために実験を行った。

4.1 実験設定

本研究では、解析対象として S-system¹⁸⁾ で記述された 30 遺伝子からなるネットワークを使用した。S-system とは以下の連立微分方程式で表現されるモデルであり、生化学反応ネットワークを表現するためにしばしば用いられている。

$$\frac{dX_n}{dt} = \alpha_n \prod_{m=1}^N X_m^{g_{n,m}} - \beta_n \prod_{m=1}^N X_m^{h_{n,m}}, \quad (n = 1, 2, \dots, N). \quad (5)$$

ただし X_n は遺伝子 n の発現量であり、 α_n , β_n , $g_{n,m}$, $h_{n,m}$ は定数パラメータである。 N は系に含まれる遺伝子の総数であり、本研究では $N = 30$ である。

全ての遺伝子の発現量の初期値をランダムに決定し、連立微分方程式 (5) を解くことにより、全遺伝子の発現量の時系列データを得ることができる。本研究では、このようにして得られた 20 セットの時系列データを測定データとして利用した。なお各データセットは 11 時点の観測からなるものとした。従って訓練データに含まれる事例数 K は $20 \times 11 = 220$ となる。生成したデータには 10% のノイズを加味した。2.1 節で説明したように訓練データを生成する際、与えられた遺伝発現の時系列データから遺伝子の発現量の変化量を推定する必要がある。そのために局所線形回帰⁴⁾ を利用した。

本実験では上記によって生成された訓練データのみから、遺伝子ネットワークの構造が正しく同定できるかどうかの確認を行った。実験は 10 試行を行った。提案手法の性能は対象とするネットワークの構造に依存する可能性があるため、試行毎に異なる構造のネットワークを対象として実験を行った。S-system はそのモデルパラメータを変えることで、異なる構造のネットワークを表現することができる。モデルのパラメータは文献 13) の方法に従って、ランダムに決定した。

遺伝子ネットワーク同定器のパラメータには文献 13) を参考にして、 $C_1 = \frac{200}{N\sqrt{K}}$, $C_2 = 0.4C_1$, $\sigma = 0.15$, $\delta = 0.05$ とした。またブートストラップ法の繰り返し回数は $B = 100$ とした。

4.2 結果

従来手法が遺伝子間相互作用の有無のみを推定する一方で、提案手法は全ての遺伝子間相互作用に対して信頼度を与える。そこで、ある閾値以上の信頼度を持つ相互作用が全て存在すると考えることで得られるネットワークをもとに、提案手法の性能を評価することにした。

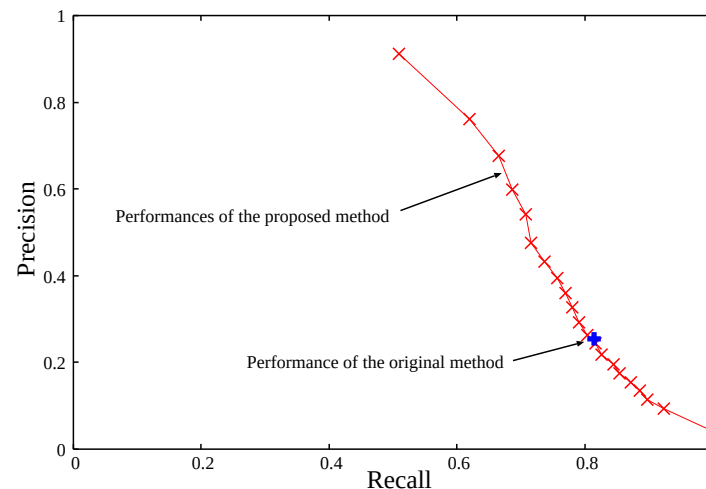


図 2 30 遺伝子のランダムネットワーク同定問題における適合度と再現率に関する性能曲線。実線: 提案手法の性能。+: オリジナルの手法の性能。

Fig. 2 Precision vs. Recall for the random network inference problems with 30 genes. Solid line: the performances of the proposed method. Plus symbol: the performance of the original method.

る。性能評価の指標には再現率 (recall) と適合率 (precision) を利用した。再現率と適合率はそれぞれ

$$\text{再現率} = \frac{TP}{TP + FN}, \quad \text{適合率} = \frac{TP}{TP + FP},$$

で定義される指標である。ただし TP , FN , FP はそれぞれ、推定されたネットワークに含まれる真陽性、偽陰性、偽陽性の相互作用の本数である。閾値を 0.95 とした場合の提案手法の再現率と適合率はそれぞれ 0.619 と 0.761 であり、閾値を 1.00 とした場合はそれぞれ 0.510 と 0.912 となった。閾値を高くするにつれて適合率が 1 に近くなることから、確かに信頼度が高いと評価された相互作用の多くが実際にターゲットネットワークに含まれていることが分かる。

提案手法の探索性能は信頼度の閾値に依存して変化するため、従来手法¹³⁾ と提案手法の性能を単純に比較することはできない。両者を比較するために PRC 曲線 (Precision-Recall Curves)¹⁵⁾ を利用した (図 2)。図 2 の実線は信頼度の閾値を変化させた場合の提案手法の探索性能を表しており、+ の点はオリジナルの手法の探索性能を表している。図より提案手法

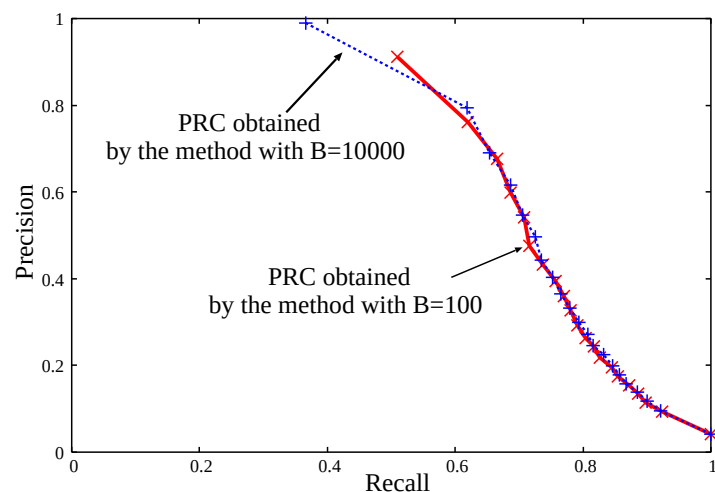


図3 $B = 100$ と $B = 10000$ の場合の提案手法の性能比較. 実線: $B = 100$. 点線: $B = 10000$.
Fig. 3 The comparison between the proposed methods with $B = 100$ and $B = 10000$. Solid line: the performances of the proposed method with $B = 100$. Dotted line: the performances of the proposed method with $B = 10000$.

はオリジナルの手法の探索性能を本質的に殆ど悪化させることなく、各遺伝子間相互作用に信頼度を割り当てることに成功していることが分かる。ただし提案手法の計算時間はオリジナルの手法の B 倍程度必要となることに注意が必要である。しかしながら $B = 100$ とした本実験においても提案手法の計算時間は平均で 21.8 分 (Pentium IV 2.8GHz) であり、十分に短いといえるだろう。さらに提案手法は並列化も容易である。

提案手法はモンテカルロ法を利用して、遺伝子間相互作用の信頼度を推定している。そのためアルゴリズムの繰り返し数 B を大きくするにつれて、信頼度の推定精度は上がる。しかしながら繰り返し数 B を 100 以上にしても、結果に大きな変化は見られなかった (図 3)。従って B の大きさは 100 程度で充分と考えられる。

5. おわりに

本研究ではブートストラップ法を利用して、遺伝子間相互作用の信頼度を評価する方法を提案した。実験により、提案手法が従来手法の探索性能を本質的に殆ど悪化させることなく、妥当な信頼度を推定できていることが確認できた。提案手法はより多くの計算を必要と

するものの、同じ量の遺伝子発現データからより多くの情報を抽出できるという点で意義がある。

本研究では、文献 13) の方法が遺伝子ネットワーク同定問題を判別問題として定式化することに着目し、訓練データをリサンプリングすることによって新たな訓練データを生成する。生成された訓練データによる学習を繰り返し行うことで、提案手法は遺伝子間相互作用の信頼度を評価する。しかしこの方法は訓練データが未知の分布からの独立なサンプルであるという仮定を必要とする。実際の訓練データは時系列データから生成したものであるから、これらが独立なサンプルであるとは言えないかも知れない。提案手法の精度向上のためには、入力データが時系列であることを考慮したアルゴリズムの設計が必要であろう。

参考文献

- 1) Akutsu, T, Miyano, S. and Kuhara, S.: Inferring Qualitative Relations in Genetic Networks and Metabolic Pathways, *Bioinformatics*, Vol.16, pp.727–734 (2000)
- 2) Cho, D.Y., Cho, K.H. and Zhang, B.T.: Identification of Biochemical Networks by S-tree based Genetic Programming, *Bioinformatics*, Vol.22, pp.1631–1640 (2006)
- 3) Chou, I.C., Martens, H. and Voit, E.O.: Parameter Estimation in Biochemical Systems Models with Alternating Regression, *Theoretical Biology and Medical Modelling*, Vol.3:25 (2006)
- 4) Cleveland, W.S.: Robust Locally Weight Regression and Smoothing Scatterplots, *J. of American Statistical Association*, Vol.79, pp.829–836 (1979)
- 5) Efron, B and Tibshirani, R.: *An Introduction to the Bootstrap*, New York: Chapman & Hall (1993)
- 6) Friedman, N., Linial, M., Nachman, I. and Pe'er, D.: Using Bayesian Networks to Analyze Expression Data, *J. of Computational Biology*, Vol.7, pp.601–620 (2000)
- 7) Gonzalez, O.R., Küper, C., Jung, K., Naval, Jr. P.C., and Mendoza, E.: Parameter Estimation using Simulated Annealing for S-system Models of Biochemical Networks, *Bioinformatics*, Vol.23, pp.480–486 (2007)
- 8) Graepel, T., Herbrich, R., Schölkopf, B., Smola, A., Bartlett, P., Müller, K.-R., Obermayer, K. and Williamson, R.: Classification on Proximity Data with LP-Machines, *Proc. of International Conference on Artificial Neural Networks'99*, pp.304–309 (1999)
- 9) Hlavacek, E.S. and Savageau, M.A.: Rules for Coupled Expression of Regulator and Effector Genes in Inducible Circuits, *J. of Molecular Biology*, Vol.255, pp.121–139 (1996)
- 10) Kikuchi, S., Tominaga, D., Arita, M., Takahashi, K. and Tomita, M.: Dynamic Modeling of Genetic Networks using Genetic Algorithm and S-system, *Bioinformatics*, Vol.19, pp.643–650 (2003)
- 11) Kimura, S., Ide, K., Kashihara, A., Kano, M., Hatakeyama, M., Masui, R., Nakagawa, N., Yokoyama, S., Kuramitsu, S. and Konagaya, A.: Inference of S-system Models of Genetic

- Networks using a Cooperative Coevolutionary Algorithm, *Bioinformatics*, Vol.21, pp.1154–1163 (2005)
- 12) Kimura, S., Sonoda, K., Yamane, S., Maeda, H., Matsumura, K. and Hatakeyama, M.: Function Approximation Approach to the Inference of Reduced NGnet Models of Genetic Networks, *BMC Bioinformatics*, Vol.9:23 (2008)
 - 13) Kimura, S., Nakayama, S. and Hatakeyama, M.: Genetic Network Inference as a Series of Discrimination Tasks, *Bioinformatics*, Vol.25, pp.918–925 (2009)
 - 14) Liu, P.K. and Wang, F.S.: Inference of Biochemical Network Models in S-system using Multiobjective Optimization Approach, *Bioinformatics*, Vol.24, pp.1085–1092 (2008)
 - 15) Margolin, A.A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Dalla-Favera, R. and Califano, A.: ARACNE: An Algorithm for the Reconstruction of Gene Regulatory Networks in a Mammalian Cellular Context, *BMC Bioinformatics*, Vol.7:S7 (2006)
 - 16) Tsai, K.Y. and Wang, F.S.: Evolutionary Optimization with Data Collocation for Reverse Engineering of Biological Networks, *Bioinformatics*, Vol.21, pp.1180–1188 (2005)
 - 17) Vilela, M., Borges, C.C.H., Vinga, S., Vanconcelos, A.T.R., Santos, H., Voit, E.O. and Almeida, J.S.: Automated Smoother for the Numerical Decoupling of Dynamics Models, *BMC Bioinformatics*, Vol.8:305 (2007)
 - 18) Voit, E.O.: *Computational Analysis of Biochemical Systems*, Cambridge: Cambridge University Press (2000)
 - 19) Voit, E.O. and Almeida, J.: Decoupling Dynamical Systems for Pathway Identification from Metabolic Profiles, *Bioinformatics*, Vol.20, pp.1670–1681 (2004)
 - 20) Yeung, M.K.S., Tegnér, J. and Collins, J.J.: Reverse Engineering Gene Networks using Singular Value Decomposition and Robust Regression, *Proc. of National Academy of Sciences of USA*, Vol.99, pp.6163–6168 (2002)