

生物システムの最適化アルゴリズム

平藤 雅之

1. はじめに

生物対象および生産手段が生物であるということは、農業の大きな特徴である。生物現象は最も複雑な自然現象であり、生物系を含む農業生産システムは極端に複雑なものである。作付計画、栽培管理、作業計画、出荷調整、施設の操業管理、作業機器の設計などでは、いずれも対象とする動植物の挙動が最も重要なファクターである。生産コストや利潤に限定した純粋に経済的な問題であっても、生物の要素は無視できない。

一方、扱う生物の種や品種の違い、土壌や気象などの環境条件、経営規模、立地条件などの違いに応じて問題の内容は大きく異なるため、個別に対処しなければならない要素が多い。わが国では、経営単位の経済的規模が小規模であることから、研究開発にかかるコスト、導入に必要なハードウェアや基礎データの収集に要するコスト、利用法を会得するまでにかかる時間的なコストなどが収支に占める割合が大きい。すなわち自己を最適化するためのコストが大きいと、目的関数には自己を最適化するのに要するコストも含めて考えなければならない。

コンピュータを農業に利用する場合では、次の目的関数 Φ を最大化する問題として考える必要がある。

$$\Phi = G - H - S - D - U \quad (1)$$

ここで、 G は最適化によって得られた利潤、 H は計算機のハードウェアのコスト、 S はソフトウェアのコスト、 D は計算に必要なデータの入手にかかるコスト、 U はユーザーがソフトの使い方をマスターするためにかかる時間的なコストである。最近では H が低下したため、パソコンの導入が即、収益増になると考えられがちである。しかし、現実にはソフトの開発に要する費用 S が非常に大きいため、 $\Phi > 0$ を実現するのはなかなか難しい。

ただし、 G の対象範囲を限定すれば S が小さい場合は

ある。たとえば、多変量解析や線形計画法などの市販のパッケージソフトを利用すれば、 $G > S$ となる問題は少なくない（フリーウェアの使用では $S = 0$ である）。しかしながら、一般ユーザーが使いこなすには理論や操作法が難解すぎるため、 $U \gg G$ となる。農家レベルでコンピュータを経営に応用した事例が少ないのは、これが主因と思われる。

さて、農業におけるコンピュータの用途は以下のようである。それぞれにおいて、カッコ内に示すようなコンピュータモデルが必要である。

- (a) 生物システムの制御（生物+環境）
- (b) 機械・施設の自動化（生物+環境+機械）
- (c) 生産方法の改善（生物+環境+機械+人間）
- (d) 生産システムの最適化（生物+環境+機械+人間+経済）

これらの用途で使われる個々のモデルの同定は、非線形関数の最適化問題に帰着される。また、(c)、(d) 自体は非線形関数の最適化問題、あるいは組合せ最適化問題になる。

以下では、キーポイントとなる複雑系の最適化方法について、生物との関連が深い最適化アルゴリズムの観点から考えてみる。

2. 非線形システムの最適化

組合せ最適化問題はホップフィールドモデル[1]のような方法を使えば非線形関数の最適化問題に帰着させることができる。逆に、非線形関数の最適化問題はパラメータの離散化によって組合せ最適化問題に変換できる。したがって、いずれかの方法でよい解法が得られれば十分である。また、最小化問題と最大化問題は評価関数を適当に変換すれば同じアルゴリズムで解くことができる。したがって、ここではこれらを特に区別しないで扱うことにする。

非線形関数の最適化は、一般に非線形関数 $\Phi(\mathbf{x}) \in \mathbf{R}$ $\mathbf{x} \in \mathbf{R}^n$ を

$$f_i(\mathbf{x}) \geq 0, \quad i=1, 2, \dots \quad (2)$$

の制約条件下で最小化する問題として表わされる。線形あるいは準線形問題では、制約条件がつかほど問題は難しくなるが、非線形問題では(2)を、

$$\Phi'(\mathbf{x}) = \Phi(\mathbf{x}) + P(f_1(\mathbf{x}), f_2(\mathbf{x}), \dots) \quad (3)$$

とすれば、単に $\Phi'(\mathbf{x})$ の最小化問題として扱うことができる。ここで、 $P(\cdot)$ は(2)の制約を満たさない場合における違反の程度を表わすペナルティ関数である。

(3)の最小化方法には多くの手法が提案されているが、興味深いことに生物現象をヒントにした方法や生物モデルを流用した方法(以下、一括して生物学的最適化アルゴリズムと呼ぶ)が意外とよい結果を示している。幸いなことに、最初に示した(a)~(d)で使われる生物モデルを作る研究は、生物学的最適化アルゴリズムとオーバーラップさせることができる。この接近方法は、生物系を含む問題((1)式が評価関数である問題)において効率的な戦略である。しかも、生物学的最適化アルゴリズムでは、これまでに蓄積された知識をアドホックに活用できる可能性がある。

最適化が最も顕著に行なわれている生物現象は、いうまでもなく進化であろう。進化モデルでは、ダーウィンの進化論が最もポピュラーであり、これは淘汰によって最適化を行なうアルゴリズムとみなせる。これを最適化アルゴリズムとして利用したものに、遺伝的アルゴリズム[2](GA: Genetic Algorithms)がある。

GAは、成分が離散値のベクトル \mathbf{x} を遺伝子型と考へ、評価関数 $g(\mathbf{x})$ を最大にする \mathbf{x} を探索するアルゴリズムである。初期状態をランダムに選んだ m 個のベクトルを元にもつ集合 $K = \{\mathbf{x}_i | i=1, 2, \dots, m\}$ としたとき、評価関数の大きい遺伝子をもつ個体ほど繁殖するという生存競争の状況を

$$n_i = f(g_i) \quad (4)$$

で表わすことにする。ここで、 n_i は \mathbf{x}_i の個体数、 g_i は \mathbf{x}_i の評価値、 f は任意の単調増加関数である。このとき、異なった遺伝子型 $\mathbf{x}_i, \mathbf{x}_j$ をもつ2個体間では、

$$p_{ij} = \frac{n_i}{N} \cdot \frac{n_j}{N} \left(N = \sum_{i=1}^m n_i \right) \quad (5)$$

の確率で遺伝子の交差が起こり、新たな2個体が発生する($\mathbf{x}_i, \mathbf{x}_j$ の成分を2分割して置換)。同時に、突然変異(ベクトル \mathbf{x}_i の成分が一定の確率でランダムに変化)によっても、新しい遺伝子型をもつ個体が発生する。

遺伝子の交差と突然変異によって生み出された新しい遺伝子型の集合を K' としたとき、これを元の集合 K に追加して、

$$K := K \cup K' \quad (6)$$

とする。 K の元の数 m を一定として、評価値が小さい遺伝子型をもつ個体は絶滅する(K から削除)。

GAではこの手続きの繰り返しによって、評価値が大きい遺伝子型の探索を行なう。上記に、遺伝学における優性遺伝、性などの概念や、生態学におけるニッチなどの効果を組み込むことも試みられている[3]。GAは評価値 g さえ得られれば適用できるため、適用範囲が非常に広い。たとえば、筆者はニューラルネットによるパターン認識結果を評価値として、植物の葉表面の顕微鏡像から気孔を強調するさいに必要な画像処理演算の探索にGAを適用し、煩雑な画像処理の自動化を行なっている[4]。

進化にもつづくアルゴリズムでは、模擬進化法(SE: Simulated Evolution)[5]も提案されている。これらはDNAレベルというよりも、むしろ染色体レベルの遺伝的なメカニズムを応用している。分子進化を参考にしたアルゴリズムとしては、DNAにおける遺伝情報のばらつきをモデル化した模擬分子進化法(SME: Simulated Molecular Evolution)[6]が提案されている。

SMEでは、DNAの転写におけるエラー分布の存在を重視し、複製されるDNAに一定の分布をもたせるようにしている。すなわち、オリジナルのDNAの塩基配列を A_i 、コピーされたDNAの塩基配列を A_i' としたとき、 A_i' はハミング距離が H までの範囲で分布し、

$$|A_i - A_i'| < H, \quad i=1, 2, \dots \quad (7)$$

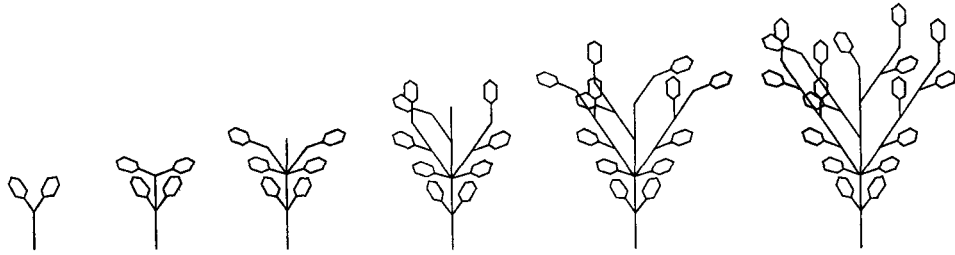
と考える。さらに、それぞれのDNA群は、種の独立性を維持するため相互に一定のハミング距離 D 以上は必ず離れていると考えて、

$$|A_i - A_j| < H, \quad i \neq j \quad (8)$$

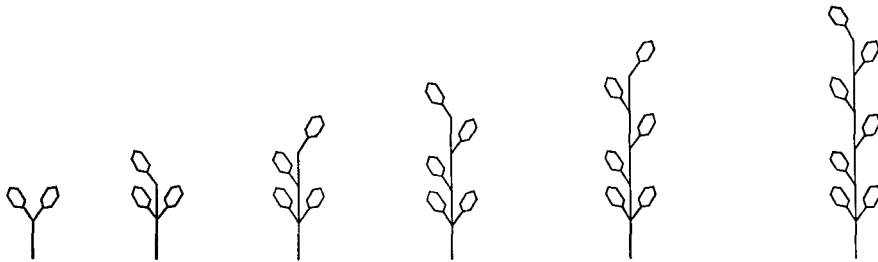
としている。GAでは不要な組合せをもった遺伝子型が多数存在して冗長であるが、SMEはこの冗長さを減らすことで探索領域を限定し、探索が効率化している。

GAやSMEは、粗結合並列プロセッサやコンピュータネットワークの環境で利用しやすいアルゴリズムであり、この面でも高速化が図れる。しかし、多少効率化しても、自由度が大きくなると手に負えなくなるという欠点がある。GAやSMEで、コンピュータのプログラムが自動作成できるかどうかを考えてみよう。

サイズが100バイトのプログラムがあるとき、たった1バイトの変異で 10^{200} 種ものプログラムが得られる。ところが、ランダムなプログラムコードの変異はバグの発生に他ならず、正常に動作するプログラムは皆無に近



光合成が活発な場合



光合成が不活発な場合

図 1 植物の生長における形態変化のコンピュータモデル (ダイズ)

い。このような問題では、GAやSMEのようなプロセスの繰り返しによって元のプログラムよりも優れたプログラムが得られる可能性は非常に少ない。これは、ダーウィンの進化論の欠点としても知られている。換言すれば、人間のDNA (約1 Gバイト) を生み出す進化メカニズムでは、もっと効率のよい最適化アルゴリズムを见いだす必要がある。

3. 対象を限定した生物モデル

前章のアルゴリズムは対象を限定していないため、対象の自由度が大きくなると破綻するが、対象の構造を限定すれば大きな自由度をもつシステムでも最適化が可能である。

ニューラルネット(Artificial neural networks)は、神経回路網を単純化したコンピュータモデルであるが、脳が多数のシナプス荷重や神経細胞をもつため、モデルに含まれる自由パラメータの数は極端に多い。このモデルでは、エネルギー関数や誤差関数の最小化によって、全パラメータが同定される。網膜におけるオプティカルフローの検出のような不良設定問題でもエネルギー関数の最小化によって解くことができる。これらに共通する最小化原理は他の生物モデルにおいても、大変、示唆に富むものである。(もっとも、パターン認識やパターン

変換などの機能を工学的に応用することへの期待の方が一般的に高いが)

ニューラルネットの学習アルゴリズムとしては、誤差逆伝搬法(BP: Back-Propagation) [7]の応用が多い。BPは最急降下法の一様であり、ローカルミニマム(2乗誤差の極小解)にトラップされるなどの問題点も残されている。それにもかかわらずBPの適用例が多いのは、多変量解析的な性質を備えているためである。

ニューラルネットに期待される機能として、学習サンプルにないパターンでも正しく推定できる機能(汎化)があげられる。特に階層型ネットワークを学習させたとき、中間層にはデータの中に含まれる情報が集約される。この場合の汎化能力は、多変量解析的機能によると考えられている[8]—[12]。ニューラルネットにおける入力パターン X_i から出力パターン Y_i への変換を f とすると、 $\{X_i, Y_i | i=1, 2, \dots\}$ をデータセットとして f を最小2乗推定する学習は、 f が線形射像のとき重回帰分析および判別分析に一致する。 f が非線形変換となるニューラルネットは、重回帰分析および判別分析、数量化理論I、II類などを非線形領域へ拡張したものに相当する。また、 f が恒等射像の場合は、主成分分析を非線形領域へ拡張したものに相当し、中間層出力値が主成分に相当する。

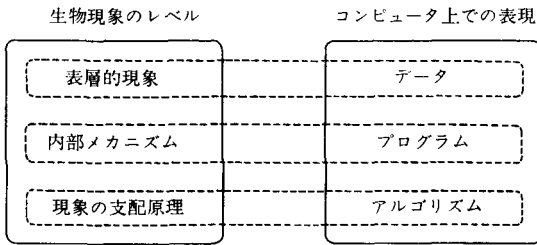


図 2 生物現象の表現形式

最近では汎化能力を向上させるテクニック[13][14]や、極小値の脱出技法[15]、連続時間リカレントネットワーク[16][17]などが提案され、機能が強化されている。また、トンネリングアルゴリズム[18]のような一般的な非線形最適化アルゴリズムとの併用や、AICなどの従来から利用されてきた統計手法との併用も可能である。

生物学的問題におけるニューラルネットの適用事例には、DNAシーケンスの解析[19]、タンパク質の2次構造の推定[20]、ランドサット画像における土地被覆分類[21]、診断型エキスパートシステム[22]、植物生長モデル[23]など(図1)をはじめ、食味計、農産物の形状判定などの産業的な応用も多い。

4. おわりに

生物現象は、図2のように3つの階層に大きく分けることができる。たとえば、植物の生長速度はCO₂濃度や光強度で変化するが、この現象の表層は、コンピュータにおいてデータの集合として表現できる。さらに、生長にかかわる光合成や呼吸などの内部のメカニズムは、プログラムコード(手続き、ルール)として扱うことができる。しかしながら、われわれが生物を見て生物らしさを強く感じる部分は、形態形成、分化、適応、進化などのような自己組織的現象である。こういった抽象的機能はアルゴリズムで表現するのが自然であり、たとえば植物の生長における適応の1つは、受光エネルギーの最大化アルゴリズムとして表現できる[24]。

生物のように極端に複雑な現象の把握にはコンピュータが不可欠であるが、一方、そのモデルの開発過程から生み出されるアルゴリズムには広範な応用の可能性が期待できる(このような研究パラダイムはArtificial Life: 人工生命[25]と呼ばれている)。しかし、生物学的最適化アルゴリズムは内容が複雑化するほど機能が向上する傾向が見られる。この傾向がつつくと、現在の単純さを失って非常に複雑なものとなるが(生物現象の複雑さの原因の1つはこの種のアルゴリズムの進化によるものだ

ろう)、生物現象とのアナロジーが保たれていれば直感的な理解しやすさという特徴は維持されると思われる。

引用文献

- [1] Hopfield, J. J. and Tank, D. W.: Neural computation of decisions in optimization problems, *Biol. Cybernetics*, **52**, 1985, 141-152.
- [2] Holland, J.H.: *Adaptation in Natural and Artificial Systems*, Ann Arbor, The Univ. of Michigan Press, 1975.
- [3] Goldberg, D. E.: *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Publishing Company, 1989.
- [4] 平藤雅之: 多モジュールMLPによる気孔開度の計測アルゴリズム, 農業気象学会・生物環境調節学会合同大会講演要旨, 1990, 302-303.
- [5] King, R.M. and Banerjee, P.: ESP: Placement by Simulated Evolution, *IEEE Trans. CAD-8-3*, 245-256(1986).
- [6] Wang, Q.: Optimization by Simulated Molecular Evolution, *Biol. Cybernetics*, **51**, 1987, 96-101.
- [7] Rumelhart, D. E., Hinton, G. E. and Williams, R.J.: Learning Internal Representations by Error Propagation. In *Parallel Distributed Processing*, Vol.1, D. E. Rumelhart, J.L. McClelland and the PDP Research Group. MIT Press, 1986, 354-361.
- [8] Asoh, H. and Otsu, N.: Nonlinear Data Analysis and Multilayer Perceptrons, *Proceedings of IJCNN 1989* in Washington D.C., II, 1989, 411-415.
- [9] Bourland, Y and Kamp, Y.: Auto-Association by Multilayer Perceptrons and Singular Value Decomposition, *Biol. Cybernetics*, **59**, 1988, 291-294.
- [10] Galliari, P., Thiria, S. and Soulie, F. F.: Multilayer Perceptron and Data Analysis, *Proceedings of IJCNN 1988* in Sandiego, Vol.1, 1988, 394-399.
- [11] 平藤雅之, 小野良孝, 小林 恭: ニューラルネットによる多変量解析とエキスパートシステム構築方法, 日本ソフトウェア科学会第5回大会論文集,

- 1989, 113-116.
- [12] Saund, E.: Dimensionality-Reduction Using Connectionist Networks, *IEEE Transactions on pattern analysis and machine intelligence*, Vol. II, No. 3, 1989, 304-314.
- [13] 石川真澄: 忘却を用いたコネクショニストモデルの構造学習アルゴリズム, *人工知能学会誌*, 5, 1990, 595-603.
- [14] Rumelhart, D. E.: Parallel Distributed Processing, *IEEE International Conference on Neural Networks*, Plenary Session, 1988.
- [15] 松葉育雄: ニューラルネットワークによる最適化計算, 計測と制御, Vol. 28, No. 12, 1990, 44-47.
- [16] Pearlmutter, B. A.: Learning State Space Trajectories in Recurrent Neural Networks, *Proceedings of IJCNN 1989 in Washington D.C.*, II, 1989, 365-372.
- [17] Pineda, F. J.: Generalization of Back-Propagation to Recurrent Neural Networks, *Physical Review Letters*, 59, 1987, 2229-2232.
- [18] Levy, A. V. and Montalvo, A.: The tunneling algorithm for global minimization of functions, *SIAM J. Sci. Stat. Comput.*, Vol. 6, No. 1, January, 1985, 15-29.
- [19] Qian, N. and Senowski, T. J.: Predicting the Secondary Structure of Globular Proteins Using Neural Network models, *J. Mol. Biol.*, 202, 1988, 865-884.
- [20] Lapedes, A., Barnes, C., Burks, C. Farber, R. and Sirotkin, K.: Application of Neural Networks and Other Machine Learning Algorithms to DNA Sequence Analysis, *Computer and DNA, SFI Studies in the Sciences of Complexity*, Vol. VII, Eds. G. Bell and T. Marr, Addison-Wesley, 1990.
- [21] 本條毅: リモートセンシングデータの画像解析, バイオエキスパートシステムズ, コロナ社, 1990, 134-138.
- [22] 平藤雅之: 大豆の病気診断ニューラルエキスパートシステム, バイオエキスパートシステムズ, コロナ社, 1990, 138-141.
- [23] 平藤雅之: ファジィニューラルネットワークによる植物生物のモデル, 情報処理学会研究報告, 32, 5, 1991.
- [24] Sakai, S.: A Model Analysis for the Adaptive Architecture of Herbaceous Plants, *J. Theor. Biol.*, 148, 1991, 535-544.
- [25] Langton, C. G. (Ed.): *ARTIFICIAL LIFE, The proceedings of an interdisciplinary workshop on the synthesis and simulation of living systems*, held september in Los Alamos, New Mexico, Addison-Wesley, 1987.

「論文・事例研究」の原稿募集!

ORの特徴は実践にあるといわれています。実践的な応用をぬきにした理論ということはORでは考えられません。本誌でも以前から会員の皆様からの事例研究の報告をお願いしてきましたが、まだ十分な成果をあげているとはいえません。

「論文・事例研究」は企業、研究所、大学等で実際に行なった事例を論文としてまとめたものを広く会員の皆様に紹介することを目的として作られた欄です。この論文は2人のレフリーによって正式に審査されますが、マネジメント、行政、工学等の広い分野において適用対象の新しさ、適用方法の新しさ、適用範囲の広さ等が論理的、科学的に論じられたものでありますならば、積極的に採用する方針です。皆様のご投稿をお願い申し上げます。

投稿要領: 学会原稿用紙36枚 (25字×12行) 以内 (図・表を含む)
(ワープロ可)投稿先はOR学会事務局OR誌編集委員会宛。

なお、原稿の他コピーを2部添付してください。

レフリー審査の結果、改訂をお願いしたり、採択されない場合があることをご了解ください。また、原稿は、採択・不採択にかかわらず、原本、コピーともお返しできません。(OR誌編集委員会)