

UTMOST IV オプティマイザ入門

UTMOST IVでは、現在、パラメータ抽出のために6種類のオプティマイザが用意されています。オプティマイザの選択は、時として分かりにくいものです。本稿は、6種類のオプティマイザを概観し、適切なオプティマイザを選択するための指針を提供します。

通常、最適化のタスクは、ローカル最適化とグローバル最適化の2つに分類されます。ローカル最適化では、コスト関数に、求めるべき極小値が1つだけ存在するものと仮定します。極小値の探索方法は比較的単純です。ローカル・オプティマイザは、ある初期値パラメータを始点とし、探索方向を決定し、その向きにコスト関数が最小となるポイントを探します。適切な探索方向が選択できるかどうかは、初期値パラメータの値と、そのオプティマイザの探索方向計算手法に依存します。一方、グローバル最適化では、コスト関数が複数の極小値を持つ可能性を想定します。そして、グローバル・ミニマム、すなわちコスト関数を最小化する極小値を求めます。解の候補が多いため、ローカル最適化よりもはるかに困難になります。そこで、実行可能にするため、グローバル・オプティマイザは、通常、ランダム化(確率化)手法を使用して解の候補の生成します。

最適化タスクはさらに、コスト関数に対する第1次(場合によっては第2次)導関数の有無によって分類できます。特にローカル・オプティマイザの場合、極小値の計算に導関数を使用すると、最適化時間が大幅に短縮できます。

導関数は、多くの場合、数値的に計算されます。しかし、実用上は、モデル・パラメータの導関数は存在しないか、計算コストがかかりすぎます。特に、コスト関数の計算についてSmartSpiceなどのシミュレータに依存する最適化タスクでは、シミュレータに起因する数値的ノイズにより、数値的導関数を計算するのは困難あるいは不可能になります。数値的ノイズは、シミュレータが数値的な解を求める過程で、アルゴリズム・パラメータ(ステップ・サイズや反復回数など)を内部的に使用する結果発生します。モデル・パラメータを僅かに変更しただけでも、それに合わせてアルゴリズム・パラメータが調整されます。このアルゴリズム・パラメータの変化が「ノイズ」を生み、最適化に必要な数値的導関数を計算で正確に求めることが困難あるいは不可能になります。

ここで、UTMOST IVで使用されるオプティマイザの説明に戻ります。表1は、各オプティマイザの主な特徴をまとめたものです。

ローカル・オプティマイザは2種類あります。1つはLevenberg-Marquardtアルゴリズム(以下LM法)[1][2]を、もう1つはHooke-Jeevesアルゴリズム(以下HJ法)[3]を使用します。LM法は第1次導関数を使用する、ローカル最適化のアルゴリズムです。非線形モデルで一般的ですが、ヤコビアン行列である第1次導関数がほぼ特異である場合に適します。導関数を使用するため、高速な処理が可能です。単純なモデルや、最適化するパラメータ数が少ない場合、LM法を効果的に使用できます。

オプティマイザ	ローカル/グローバル	導関数の使用	確率的手法
Levenberg-Marquardt	ローカル	○	×
Hooke-Jeeves	ローカル	×	×
Simulated Annealing	グローバル	×	○
Parallel Tempering	グローバル	×	○
Genetic Algorithm	グローバル	×	○
Differential Evolutions	グローバル	×	○

表1: UTMOST IVのオプティマイザの比較

LM法の短所は、導関数、すなわち数値的ノイズや計算コストの問題により計算が実用上困難あるいは不可能な関数を必要とする点です。その代替として、導関数を必要としないHJ法が用意されています。HJ法はパターン検索アルゴリズムの分類に属します。HJ法は導関数を用いる代わりに、コスト関数を最小化する可能性の高いパターンを探索方向に沿って探索します。

残りの4種類のUTMOST IVオプティマイザはグローバル・オプティマイザです。4種類すべてが、調査するパラメータ値の生成にランダム化(確率化)手法を採用しています。確率的であるため、同一のオプティマイザを使用して同一の最適化を実行すると、そのたびに異なる極小値に到達する可能性があります。これは確率化のメリットです。1回目の実行が失敗しても、2回目でも成功する可能性があるためです。

Simulated Annealingアルゴリズム(以下SA法)[4]は、物理システムのある温度における平衡状態を模倣したものです。ここで、コスト関数はエネルギーに対応し、「温度」は単に最適化のコントロール・パラメータです。「温度」が高いとき、高コストをもたらすパラメータ値が見つかりやすくなります。これは、最適化の初期段階において探索範囲を広げる結果となり、有利に働きます。最適化の段階が進むにつれてシステムは冷却化し、探索範囲はグローバル・ミニマム近傍に絞り込まれます。

SA法は堅牢なアルゴリズムであり、パラメータ数が多く、非線形性の強いモデルに威力を発揮します。短所は、速度が遅くなる傾向があることです。Parallel Temperingアルゴリズム(以下PT法)[5]はSA法に似たアルゴリズムですが、「温度」を徐々に下げるかわりに、複数の異なる温度におけるシステムの複製を用意します。複製を時々交換することで、システムがローカル・ミニマムに陥るのを防ぎつつ、収束を早めます。PT法は、深いローカル・ミニマムが多数ある場合に、単純なSA法よりも適しており、有効に活用できます。

残りの2種類のオプティマイザは、進化的アルゴリズムの分類に属します。SA法やPT法が物理システムを模倣しているのとは異なり、進化的アルゴリズムは生物システムにおける進化のプロセスを模倣します。進化的アルゴリズムでは、最適化問題に対する複数の解(「染色体」)を含む「集団」を用意します。各「世代」において、「集団」のうち「適応度(コスト関数と密接に関係)」に劣る個体が「選択」プロセスで淘汰され、生き残った個体が「交叉」や「突然変異」のプロセスにかけられます。「交叉」プロセスは、2つの個体(「親」)を選択し、その個体の一部を組み合わせる1つまたは複数の「子」を作成します。グローバル最適化では、「交叉」プロセスは、ローカル・ミニマムに陥ることなく、グローバルな最適解を探索し続ける効果があります。「突然変異」プロセスは、各個体の一部を変更します。最適化プロセスにおいては、「突然変異」は極小値近傍の解がその極小値に向かって収束していくように働きます。「選択」、「交叉」、「突然変異」のすべてのプロセスは、そのモデルとなった生物システム同様、ランダムな選択をとまいません。すなわち、進化的アルゴリズムは確率的アルゴリズムです。

UTMOST IVには、Genetic Algorithm (GA法)とDifferential Evolution (DE法)の2種類の進化的アルゴリズムがありま

す。GA法は開発当初[6]、生物学への相似から、ビット列またはその他の離散的な数量で解を表現していました。このような表現方法は、離散的な(組み合わせ)最適化に適しています。近年の研究[7]では、「交叉」および「突然変異」における演算子の表現は実数値まで拡張されています。

UTMOST IVのGAアルゴリズムには、いくつかの近年の研究成果が盛り込まれています。

一方、DE法[8]はGA法と異なり、実数値表現のために開発されたため、モデル・パラメータを最適化するのに適しています。他の進化的アルゴリズムと同様、各世代には複数の解が存在します。DE法とGA法の主な違いは、各世代で個体が選択および変更される方法にあります。GA法が「交叉」や「突然変異」を使用するのに対し、DE法ではまず無作為に選択した2つ(またはそれ以上)の個体の差分を計算し、その差分を、無作為に選択した別の個体に確率的に加えます。

DE法およびGA法は、グローバル・オプティマイザとしては非常に高速な処理が可能です。しかし、その高速性を発揮できるかどうかは、アルゴリズム・パラメータの調整によります。DE法では、GA法に比べてアルゴリズム・パラメータへの依存度が低くなります。

参考文献

- [1] Levenberg, K A method for the solution of Certain Non-linear Problems in Least Squares, Quart. Appl. Math. 2, 164-168, 1944.
- [2] Marquard, D An Algorithm for Least-Squares Estimation of Nonlinear Parameters, SIAM J. Appl. Math. 11, 431- 441, 1963.
- [3] R. Hooke, T. A. Jeeves, Direct Search Solution of Numerical and Statistical Problems, Journal of ACM, 8, pages 212-229, 1961.4]
- [4] S. Kirkpatrick, Gelatt C. D., Vecchi M. P. Optimization by Simulated Annealing, Science, 220 (671- 680), 1983.
- [5] C. J. Geyer, in Computing Science and Statistics Proceedings of the 23rd Symposium on the Interface, American Statistical Association, New York, 1991, p. 156.
- [6] J. Holland, Adaptation In Natural and Artificial Systems, University of Michigan Press, 1975.
- [7] A.H. Wright, Genetic Algorithms for Real Parameter Optimization, in G. Rawlins (Ed.), Foundations of Genetic Algorithms, First Workshop on the Foundations of Genetic Algorithms and Classifier Systems, 1991. Z. Michalewicz, G. Nazhiyath, M. Michalewicz, A Note on the Usefulness of Geometrical Crossover for Numerical Optimization Problems, Proceedings of the 5th Annual Conference on Evolutionary Programming, San Diego, CA, 29 February-3 March, MIT Press, Cambridge, MA, 1996.
- [8] Storn, K. Price, Differential Evolution - a Simple and Efficient Heuristic for Global Optimization over Continuous Spaces, Journal of Global Optimization, Kluwer Academic Publishers, 11(341-359), 1997; K. Francken, G.E. Gielen, M. S. J. Steyaert, An Efficient, Fully Parasitic-aware Power Amplifier Design Optimization Tool, IEEE Trans. On Circuits and Systems, 52, 8, 1526-1534, 2005.