

カーブフィッティング、モデル選択、測定

吉井 達哉 (Tatsuya Yoshii)

京都大学文学研究科哲学専修・日本学術振興会

ある変数（独立変数） X に基づいて別の変数（独立変数） Y の値を予測するという問題は、科学的探究や経営・政策決定においてしばしば生じる。例えば、気温からアイスの売上を予測したり、血圧から健康上のリスクを予測したりといった問題が考えられる。こうした予測は、誤差項 ϵ の分布を仮定すれば、 $y = f(x) + \epsilon$ という条件分布で表すことができる。カーブフィッティング問題とは、手持ちのデータに基づいて、 X と Y についての全ての可能な観測結果について最も良い予測をする関数 f を選ぶという問題である。

カーブフィッティングについて実質的な問題が持ち上がるのは「データへの当てはまり」という直感的な基準だけでは関数 f をうまく選択することができないからである。このことは「過適合」の問題として知られている。例えば、線形回帰を考えると、 n 個のデータ組 $(x_1, y_1), \dots, (x_n, y_n)$ に対して、常にそれらを全て通る $n-1$ 次関数が存在する。しかし、こうしたデータに完全に一致する「複雑な」関数を用いた予測の精度はあまり高くないことが経験的に知られている。むしろ、我々は通常データに完全に一致しなくとも十分データに近い予測を行うような「単純な」関数を用いる。「このような実践に対していかなる正当化が与えられるか？」というのが、カーブフィッティングについての哲学的問題である。特に、Goodman (1946,1954) による有名な哲学的問題である「グルーのパラドクス」はしばしばカーブフィッティング問題の一種として定式化されてきた (Priest 1976, Forster 1999)。

このようなカーブフィッティング問題に関して注目されてきたのが、赤池情報量基準 (AIC) などのモデル選択基準である。モデル選択理論の文脈では、パラメータ付けられた確率分布の族である「統計モデル」という構造が明示的に導入される。まず、具体的な統計モデルが与えられれば、カーブフィッティング問題はモデル内のパラメータの推定の問題に帰着する。この問題は、最尤推定などの標準的な手法によって比較的容易に解くことができる。従って、より重要なのは「どのモデルを用いて推定を行うべきか」というモデル選択の問題である。こうしたモデル選択の問題に答えるのが AIC などのモデル選択基準である。これらのモデル選択基準は「モデルが現在のデータにどれだけ上手く適合できるか」を評価するだけでなく、モデルの複雑さ、即ちパラメータの数に対してペナルティを課す。それ故、AIC などのモデル選択基準は、ある意味で、より単純な関数をより高く評価するという我々の帰納的实践に対する客観的な根拠を与えるものと言える (Forster and Sober 1994)。

そうすると、AIC などのモデル選択基準は、グルーのパラドクスを含む哲学的なカーブフィッティング問題を解決してくれるように見える。実際、グルー仮説が多くの人にとって奇妙に見えるのは、時刻 T という無関係な要因に言及しているからであり、モデル選択基準はグルー仮説に対してそうした「複雑さ」に基づいてペナルティを与える根拠を提供するように見える。しかし、Forster (1999) や Kieseppä (2001) といった

この問題を詳しく論じてきた哲学者らの総意によれば、AIC などのモデル選択基準は、グルーのパラドクスなどのカーブフィッティング問題を解決しない。というのも、これらの基準は、モデルの特定の階層（以下、モデル族）を前提として、そのような固定されたモデル族から最良のモデルを選択するための基準を与えるものだからである。これらの基準は「どのモデル族を使うべきか」という問題には原理的に答えることができない。例えば、実際のデータに完全に適合するような「複雑な」関数であっても、その関数のみを含むような（パラメータ 0 の）「単元集合モデル」に含めてしまうことができる。そのようなモデル族を仮定してしまえば、モデル選択基準はそうした関数を用いて予測を行うことを推奨せざるを得ない。それ故、モデル選択基準だけでは、カーブフィッティング問題を解決することはできない。

本発表では、測定理論における有意味性（*meaningfulness*）という概念（Falmagne and Narens 1983, Narens 2002）を用いることで、考慮されるべき「妥当なモデル」を制限し、こうした問題を解決する提案を行う。ある数的関係が有意味であるとは、それが尺度変換によって不変に留まることを意味する。これは、そうした数的関係を用いることで、経験的对象について、特定の尺度・表現方法から独立した主張を行うことが可能であるための条件を定めると言える。

本発表は特に、カーブフィッティング（回帰）モデルを採用する際に暗黙的に引き受けられている仮定（モデリング仮定）に注目する。これは「そのモデルのパラメータ値の中には、全ての可能なデータの挙動を適切に記述するような条件付き分布を指定するものが存在する」という仮定である。本発表の提案は、モデルに有意なモデリング仮定を持つことを要請することである。これは、モデルに対して、単に数学的に構成可能であることだけでなく、そのモデルの仮定が何らかの経験的な意味を持つことを要求する。この要請はモデル選択基準だけでは解決できなかったカーブフィッティング問題に対する解決の糸口を与える。特に、この要請は、多項式関数を用いたカーブフィッティング問題における「妥当なモデル」に対する適切な制限を与える。

Falmagne, J. C. and Narens, L. (1983). 'Scales and meaningfulness of quantitative laws.' *Synthese*, 55(3):287–325.

Forster, M. and Sober, E. (1994). 'How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions.' *The British journal for the philosophy of science*, 45(1):1–35.

Forster, M. R. (1999). 'Model selection in science: The problem of language variance.' *The British journal for the philosophy of science*, 50(1):83–102.

Goodman, N. [1946]: 'A Query on Confirmation', *Journal of Philosophy*, 43, pp. 383–85.

Goodman, N. [1955]: *Fact, Fiction, and Forecast*, Cambridge, MA: Harvard University Press.

Kieseppä, I. A. (2001). 'Statistical model selection criteria and the philosophical problem of underdetermination.' *The British journal for the philosophy of science*, 52(4):761–794.

McCullagh, P. [2002]: 'What Is a Statistical Model?', *The Annals of Statistics*, 30, pp. 1225–310.

Narens, L. (2002). *Theories of meaningfulness*. Lawrence Erlbaum Associates.