戦後日本社会学史への計算社会科学的アプローチ

――『社会学評論』1954-2015 の構造トピックモデルによる分析――

東北大学　　　瀧川裕貴

1　目的

本報告の目的は，計算社会科学の手法を用いて戦後日本社会学の諸テクストを分析することにより、戦後日本社会学史に新たな知見をもたらすことにある．　戦後の日本の社会学に関する研究は現在までに複数、存在している(例えば、富永 2004, Nakao 1998, 西原・杉野 2000, 太郎丸ほか 2009 など). これらの知見の中には比較的定説とみなされるものもあれば、意見が分かれている議論もある．後者についていえば、例えば西原・杉野(2000)は日本社会学の実証主義化というトレンドを指摘しているのに対して、太郎丸ほか(2009)はこれを否定するといった状況があげられる．方法の面でみれば、a)代表的なテクストの綿密な読解に基づく、いわば質的研究と b)学会報告の分類や研究論文タイトルに関する定量的な分析、とに区別できる．しかしながら、テクスト内容そのものを、かつ量的な手法で分析した研究は現在までにほとんど存在しない．本報告では、計算社会科学、具体的にはトピックモデルという確率的生成モデルを用いて、テクストの内容そのものを定量的に分析することにより、戦後日本社会学史に新たな知見をもたらすことを試みる．

2　方法

データとして用いたのは、『社会学評論』に掲載された学術論文である．旧字体である 1-3 号を除き，4 号（1954 年）から 65 号（21015 年）までの全論文が対象である．ただし、書評等を除いたため、合計 1507 本が対象となる．方法は、トピックモデルの拡張である構造トピックモデルを用いた．詳細は省くが、推定の対象は、文書ごとのトピック比率と文書集合に登場する全単語にわたる確率分布としてのトピックである．仮想例を用いて説明しよう．例えば、ある文書 A は「機能主義」や「社会学理論」といったトピックが高頻度で出現するトピック比率をもつとする（トピックの名前は分析者が事後的につける）．そこで、この文書 A は、例えば「機能主義」トピックが高頻度で産出する「機能」、「構造」、「体系」、「均衡」といった単語が多く使われることになるだろう．もちろん、実際にはトピックの割り当ては潜在変数なので、（しかるべく前処理をされた）学術論文の文書集合という観察データから、統計的に（ベイズ）推定するのである．

推定に際して、本研究では、R パッケージの"stm"を用いた．また、トピック数は 70 に固定した．

3　結果・結論

分析の結果，"社会学理論" 、"日本社会学"、"福祉国家・公共性"、"自己論"、"都市社会学"、"機能主義"、"回帰分析"、"語り・ストーリー"、"移民"、"人間論"、"現象学的社会学"、"構築主義"、"社会哲学" "組織・政策"　"農村"等と名づけることのできるトピックが抽出された．各年代ごとの主要トピックの推移やトピック間ネットワークを分析した結果、全体的にみると、機能主義 vs.マルクス主義の対立を経て、「意味学派」が台頭し、マルチパラダイムに至るという日本社会学の標準ストーリーは概ね再現されているようである．また、暫定的ながらある限定された意味で日本社会学の「実証主義化」（計量社会学の台頭）のトレンドも確認できる．他方で、様々なトピックの概念体系の詳細やその変化、また概念体系間の関連の定量化など、定量的手法によってはじめて明らかになったいくつかの知見も指摘できる．当日にはこれらの詳細について報告する．