

〈論文〉

## 社会思想史研究とテキストマイニング

林 直樹

**要旨** 経済学史や社会思想史研究の領野におけるテキストマイニングは有用である。ただし、一定の留保付きの環境で。従来の（外在内在問わず）文脈主義的な古典読解の手法を通じて得られる知見と組み合わせることなしには、テキストマイニング技術は宝の持ち腐れになるだろう。本稿ではアダム・スミスの諸著作を対象とするテキストマイニングを実施し、キータムが出現する「言葉の袋」を対象に単語ベクトルを構成したうえで、それらのコサイン類似度を比較した。そしてその作業を通じて、言語の分布構造という客観的実体を取り扱うだけでも文脈と意味の類似を導き出せるとする仮説が、ある程度実証されることを示す。同時に、従来型の古典読解法に立ったスミスの諸著作と河上肇『貧乏物語』との比較を通じた社会（思想史）的文脈の導出も行い、上記の実証から零れ落ちる面が小さくないことを、テキストマイニングの実践を通じて白日のもとにさらす。新旧の方法は相補的なものとならねばならないのである。

**キーワード** コサイン類似度 アダム・スミス 見えざる手 河上肇

## 1. はじめに

オックスフォード大学ニュー・カレッジから母校 LSE に転任して間もないロビンズ (Lionel Robbins, 1898-1984) は、経済学専攻学生向けの講義「経済分析の原理」において、開口一番、次のように述べた。「最近の慣行 (recent habit)」と化しつつある一般均衡分析、すなわち価値を排した多変量間の「関係性」の分析を自らは意識的に選択するけれども、総価値を取り扱う旧来型の分析に一般均衡分析が「対抗」するのではなく、「相補的 (complementary)」な関係を築くことが望ましいとも考えると (Robbins 2018, 7-8)。経済学者は価値判断を行わなくてよいと述べたつもりは自分にはなかった、経済分析の前提をなす領域において「価値」が大手を振って歩くと分析の起点において視野狭窄を引き起こしかねないこと、その認識論的事実に警鐘を鳴らしただけのつもりだったが、ずいぶん誤解され、時に非難されたと、ロビンズは後年

に回顧する（『一経済学者の自伝』第七章を見よ）。

ロビンスの抱えた問題の大きさには遠く及ばないながらも、本稿は一種の「最近の慣行」をあえて取り上げ、テキストマイニングと呼ばれるその新しい技術が社会思想史研究の領野において持ち得る射程の大きさを推し量ろうとするものである。願わくは誤解を免れ、ここで開陳した見通しが思想史学の今後をさらに多様かつ豊饒なものとする契機とならんことを。そのためには、旧方法が「旧」で終わることなく、まさにその方法論的意義を主張し続け、個々の社会的な文脈に照らしてしか究極的には解釈しえないのが言葉の意味だと繰り返し訴えていかねばならない。この訴えに耳を傾けることなしには新方法もテキストデータの表面をなぞるだけの皮相な分析結果をもたらすだけに終わりがちだという点を、いずれの方法を採る者にとっても明確な共通認識としておく必要があるだろう。

テキストマイニングを研究に、また教育に取り入れようとする動きは昨今、急激に加速しており、小峯敦編『テキストマイニングから読み解く経済学史』のような入門書まで、経済学史研究者グループの手で編み出されるまでになった。思想史学を専攻する筆者自身も情報学の研究者と組み、大学において「テキストマイニングの手法で経済古典を読み解く」と題した演習を営んできた。そこには経済学専攻、経営学専攻、情報学専攻の学生たちが集い、グループワークを中心にしながらつねづね新鮮な着想のもと古典を分析し、その過程において、自ら繙く機会は決して多くはないかもしれない種類の書籍に親しみを覚え、認識の広がりを手に行っている。本稿は足掛け四年に及ぶ当該演習を通じて得られた知見の一部を土台にしており、その意味では、筆者の実践経験に基づく成果の一つである。

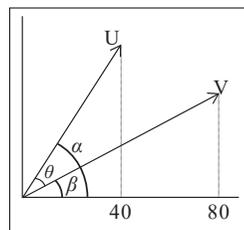
## 2. 文字列の相関を計量する簡便な方法

### 2.1 コサイン類似度の定義

テキストマイニングとは何か。その点を理解するためには、コサイン類似度を求める数学的な操作にまず触れておくのがよいだろう。一定の文字列ないしテキストをベクトルに変換し、単語ないし語群ベクトル同士の類似度（距離）を測る簡便な方法を知りさえすれば、この技術の本質的部分はつかみ取れる。

コサイン類似度とは  $\cos \theta$  の値 ( $-1 \leq \cos \theta \leq 1$ ) をベクトルの類似度合の尺度として用いるものである。二つのベクトルのなす角度  $\theta$  が小さければ小さいほどベクトルの方向性は近くなる。言い換えれば、 $\cos \theta$  の値が最大値の1に近づけば近づくほど、ベクトル間の類似度は高いと言える。数学的根拠は次の通りである。

右図のような二つのベクトルを考える。 $\alpha - \beta = \theta$  である。仮に  $\alpha = 60$  度、 $\beta = 30$  度のとき、加法定理より  $\cos \theta = \cos (\alpha - \beta) = \cos \alpha \cos \beta + \sin \alpha \sin \beta$  だから、



$\cos \theta = \frac{1}{2} \times \frac{\sqrt{3}}{2} + \frac{\sqrt{3}}{2} \times \frac{1}{2} = \frac{\sqrt{3}}{2}$ である。また、 $U(40, 40\sqrt{3})$  かつ  $V(80, \frac{80}{\sqrt{3}})$  と置くとしよう。ベクトルの大きさはピタゴラスの三平方の定理を使い、 $|U| = \sqrt{40^2 + (40\sqrt{3})^2}$ 、 $|V| = \sqrt{80^2 + (80/\sqrt{3})^2}$ と求まる。

このとき

$$\cos \alpha = \frac{40}{\sqrt{40^2 + (40\sqrt{3})^2}}, \quad \cos \beta = \frac{80}{\sqrt{80^2 + (80/\sqrt{3})^2}}, \quad \sin \alpha = \frac{40\sqrt{3}}{\sqrt{40^2 + (40\sqrt{3})^2}}, \quad \sin \beta = \frac{80/\sqrt{3}}{\sqrt{80^2 + (80/\sqrt{3})^2}}$$

$$\cos \theta = \frac{40 \times 80 + 40\sqrt{3} \times 80/\sqrt{3}}{\sqrt{40^2 + (40\sqrt{3})^2} \times \sqrt{80^2 + (80/\sqrt{3})^2}} = \frac{U \cdot V}{|U||V|} = \frac{3200 + 3200}{80 \times 160/\sqrt{3}} = \frac{\sqrt{3}}{2}$$

ここで  $U \cdot V$  はベクトルの内積である。

以上より、コサイン類似度はベクトルの大きさに依存せず、二つのベクトルが作る角度のみに依存すると分かる。ただし、ベクトルの大きさを決める各要素（上記の通り二次元であれば縦軸と横軸しかないが、より多次元もあり得る）の大きさ次第では、ある要素の値が他の相対的に値の小さい要素の特徴を押し殺してしまうケースもあり得るため注意を要する。この点には、のちに実際のテキスト分析を展開する折に再度、言及することになるだろう。

## 2.2. コサイン類似度の応用

単語ベクトルのコサイン類似度が単語の意味自体と相関するという仮定を正当化する理論が分布仮説である。この仮説によれば、同様の文脈で頻繁に共起する言葉同士は意味も似ている可能性が高い（小峯編2021, 40）。むろん、実際にはその文脈を同定すること自体に困難が伴うのだが、アメリカ合衆国の言語学者ハリス（Zellig S. Harris, 1909-92）の提唱に始まるこの仮説はテキストマイニングの大前提となっている。また、共起頻度への注目という小前提を支える理論が「言葉の袋」仮説である。文脈も意味もすべて削ぎ落とし、一定の範囲に含まれる総単語数に占める特定単語の出現数を「頻度」（出現確率）と定義したとき、その多さや、逆に少なさ（稀少性）に着目すること自体に、文意を読み取る方法として価値があると見なすものだ（同34-35）。ハリスはこう述べる。「ある人物の言葉遣い（language）が全く恒常的な反面で、経験の個性的諸相を考慮する、つまりある人物における意味の蓄えが数年をかけて成長しかつ変化する場合を考慮するなら、ないしは、ある人物が〔成長前の時点で〕自らに利用可能な言葉遣いを用いては適切に表現し得ないような思想や感じ方を懐く場合を考慮するなら、主観的経験の構造や意味の主観的世界の構造に、言語構造が必ずしも従うものではないと分かる」（Harris 1954, 151）。要するに言語には、経験主義的主観の領分に回収され尽くさない客観的実体としての側面が備わるということである。加えて「言葉遣いは単なる言葉の袋（bag of words）ではなく、その使用経路の中で形成された個々の特性を帯びた道具」に他ならないのであって、そうした「特性」を文脈と見なすならば、それと意味とのあいだに一对一の対応関係はないにせ

よ、それらは一定程度の分布的規則性ないし「分布的關係性 (distributional relations)」のもとにあると見なしてよい。したがって「ある重要なケースのいくつかにおいて、意味の一定諸相は計量可能な分布的関係性の関数 (functions) であると言明することもできるということさえ、明白になるだろう」(Ibid., 156)。簡単に言えば、意味と言葉とは文脈を介して何らかの関係を有し、かつ、その関係性は数量化可能だということになるだろう。

本稿では、分布仮説と言葉の袋仮説を承認し、これらを組み合わせて用いる。つまり一定の範囲(袋)で区切ったテキストに出現する言葉の共起頻度を比較し、単語(語群)ベクトルのコサイン類似度を計量することで、テキスト同士の関係性、すなわち意味の遠近、文脈の異同をある程度推定する。これを共起分析と呼ぶ。以下では、通常はKH Coderと呼ばれる専用アプリケーション上で処理されがちな分析をあえて手計算で行うことにより(ただし形態素解析にはKH Coderが実装するMeCabを用いた)、テキストマイニングの最基底に横たわる方法を具体的に把握することにしたい。そのことがテキストマイニングの有用性と限界を同時に知らしめ、ひいては、今後の思想史研究は何をなすべきか、という問いに対しても一定の回答を与える一助になると考えるからである。

### 3. アダム・スミスの作品をテキストマイニングする(1): 文脈を知る

#### 3.1. 「見えざる手」について

道徳哲学者にして「経済学の父」アダム・スミス(Adam Smith, 1723-90)の著書『道徳感情論』*The Theory of Moral Sentiments*(初版1759年/略号TMS)および『国富論』*An Inquiry into the Nature and Causes of the Wealth of Nations*(同1776年/略号WN)と、スミス没後に遺稿を整理した遺言執行人たちが出版した『哲学論文集』*Essays on Philosophical Subjects*(1795年)所収の「天文学史」*The History of Astronomy*(略号HA)には、それぞれ一度ずつ、invisible hand(見えざる手)という言葉が登場する。ただしHAのそれには定冠詞のtheが、TMSとWNのそれには不定冠詞のanが付されており、この事実のみからでも、この言葉の意味内容に振幅があることが容易に推測される。

グラスゴー大学教授だったマクフィ(Alec L. Macfie, 1898-1980)の短い論考「ユピテルの見えざる手」によれば、ローマの最高神ユピテル(ジュピター)の見えざる手として登場しているHAの「見えざる手」のほうは自然の過程(スミスの慣用表現にordinary [natural] course of thingsがある)に対する神の気紛れの介入を意味するのに対し、TMSとWNのそれは自然の過程を後押しする作用を意味するため、語義が全く逆転している(Macfie 1971, 595)。この逆転現象には執筆時期が関わっていると推定される。TMSとWNがスミスのグラスゴー大学教授時代(1751~63[公式には64]年)の講義ノートをベースにしていることは間違いない。これに対してHAがいつ頃書かれたものか正確には分からないが、HAがハレー彗星に言及し、それが1758年に戻って来るだろうと「予測」されているとの記述があることから、この年以前に

書かれたことは確かだろう。マクフィはさらに踏み込んで、おそらくはスミスの教授着任前、相当若い時分の著述ではないかと述べる（*Ibid.*, 598）。

要するに、HAとTMS & WNとのあいだには、相当の時間と、スミスにおける認識の深化が横たわっており、その「ユピテルに始まる長い道のり」をスミス自身の成長ないし変化として評価すべきだということである。マクフィは、スミスが人間行動と見えざる手との関係を三つの観点から眺めているとし、①正義、②社会倫理、③経済を挙げる。WNの見えざる手は端的に③の文脈に属す。各個人が自己利益を最大化するうえでは自然の秩序に委ねるのがベストだと述べ、それを支える仕組みとしての分業に言及がなされると同時に政治的介入が批判されるからである。この文脈は損益計算も可能で、ここから経済学が直接的に生まれたと言える。TMSにおける文脈は倫理的かつ総合的で、経済も含むが、WNが効率＝資源配分寄りなのに対して分配寄りである。人間の目的（幸福）は貧しい人にとっても豊かな人にとっても本来は差がないにもかかわらず、自然は人間を欺き、差があるように見せかけることで幸福への手段を様々に改良させ、その過程で徐々に貧困を撲滅していくように仕向ける。このスミスの信念が最も明確に現れているのがTMSだろう。人間がいかに倫理的過誤を犯そうとも、その過ち自体からも善が引き出されるように、自然はあらゆる要素を用いて社会を巧妙に導き、最終的正義をもたらそうとするはずだ。これがスミスの思想信条であると言ってよい（*Ibid.*, 599）。

### 3.2. 河上肇とアダム・スミス

身体の安楽と心の平静こそが幸福の核心であり、道端でひなたぼっこをしている乞食はそれらをすでに所有しているのに対し、いまだ得られない王のほうこそが日夜苦心せざるを得ないと述べた、TMSにおける「見えざる手」登場段落の結びの一文にこそ、スミスの思想は凝縮される。この乞食と王の一節には、日本のマルクス経済学者の草分けだった河上肇（1879-1946）が大正のベストセラー書『貧乏物語』（1916年に新聞連載、ロイド＝ジョージ論を付した書籍版は翌年刊行）の中で、石川啄木の詩「はたらけどはたらけど猶わが生活楽にならざりちっと手を見る」に関連付けて批判的に取り上げた例がある（河上1984, 334-35）。

「貧乏人も金持もその幸福にはさしたる相違のないものであるという説」としてスミスの主張を要約紹介した河上は、続いて、五条大橋の袂で暮らす乞食親子を題材にした京都嵯峨の臨済宗僧侶、間宮英宗（1871-1945）の講話を引く。大晦日、橋の上で立派な身なりの武士が番頭風の男に出会い、「もう一か月ばかり待ってほしい」「いつも勝手なご弁解もはや五年にも相成ります」といったやり取りの後、「平生大道狭しと威張っていくさる」側がしきりに頭を下げて謝るのを、乞食の俸が聞く。そしてあのような境遇に比べれば呑気でいられるほうがはるかにましだと自惚れて悦に入っていたところ、隣で寝そべっていた親父に、誰のおかげでその気楽さを手に入れられたと思うのだ、儂に頭を下げる、と言われた段で落ちになる。金持ちや権勢持ちゆえに幸せとは限らないことも真実だと河上は認める。だからこの乞食の話に言及したのである。しかし同時に河上は、だからと言って「過分に貧乏なのが仕合せだとはいえぬ」と問

題提起も行う。身を粉にして働いても「身心の健全なる発達を維持するに必要な物資さえ得あたわぬ」ような「強制的に受けさせられている貧乏」が存在し、それと、スミスがTMSで示唆したような「自発的貧乏」とを誤解混同してはならないと河上は告げる。乞食ですら「親様に頭を下げることを強えられる、いわんや啄木の詩の主題をなす者たちをや、と（同335-38）。

河上はスミスについて、「いやしくも斯学〔経済学〕を攻究する者にして氏の学恩を蒙らざる者はほとんどまれなり」と評し、「追慕の意」を抱くとする（同318-19）。そのうえで、道徳哲学教授だったスミスが「倫理学者としての殻を打ち割り」「有史以来いまだかつてあらざりしところのまったくあらたなる種類の学者」すなわち経済学者を努力の果てに産み落とした際、WNの中で「経済上における利己心の活動を是認することにより、ある意味において、経済上におけるいっさいの人の行為を倫理問題の埒外に推し出した」ことを（同377）、結局において非難する。その理由は、現代における「利己心の束縛なき活動」が貧富格差を拡大させている事実を目の当たりにしては「アダム・スミスの誤謬」を痛感せざるを得ず（同379-80）、今や必要なのは、ビジネスと道徳を調和させるべく「ひとたび倫理学の領域外に脱出せしわが経済学」を「ふたたび倫理学の王土内」に戻す、全く新たな取り組みに他ならないと考えたからであった（同410-11）。

以上のように、河上は『貧乏物語』においてスミスの「見えざる手」を二重の意味で、つまり第一には非自発的貧困を考慮に入れていないという意味で、第二には自由放任が生み出す過度の格差を許容するという意味で、手厳しく論難した。河上は、富者の奢侈品需要の大きさが市場を歪め、一般消費者が求める必需品供給を減らし、必需品価格を不自然に高めた結果、非自発的貧困を生んでいると解説する。この因果関係を見落としたスミスは①富者と貧者のあいだの「分配」の不平等を容認し、②貨幣価値で測った富の大きさが増えさえすればよいと論じたため、付加価値額の内訳が奢侈的な財やサービスに偏向する事態を防げなかった、そう河上は言う（同379-80）。そして貧困対策としての富者の奢侈廃止を唱えていくのである。だがマクフィを参照しながらすでに見たように、スミスはTMSで分配の改善を論じている。また、WNで説かれる効率的な資源配分とは、農⇒工⇒商の順序で投資が進み、貧者を含む雇用が最大化される時に生じる、理想的配分=分配のことに他ならない（WN 第二編第五章）。さらにスミスは生産的労働と不生産的労働を区別し、奢侈的消費にしか貢献しない後者を批判して、前者が生み出した付加価値の節約（貯蓄）⇒生産的投資を推奨してさえいる（同第三章）。

こうしたスミスの主張は、実のところ、河上のそれにきわめて近い。TMSもWNも、自由放任の名のもとに貧富格差の現状を追認する立場、つまり諦念へと導こうとする立場から貧者と富者の真の境遇差を認めなかったのではない。そうではなく、まさに、偽の境遇差に踊らされる人間の弱さを逆説的な武器として、河上がロイド=ジョージを持ち上げながら唱えた「貧乏神退治の大戦争」を先駆的に遂行しようとしていたのである。TMSが目指した「最終的正義」に資するべくWNの立論も存在すると見なせば、両著作はいずれも究極的には分配を重んじている。そしてこの見地に立てば、TMSとWNを明らかに共通の文脈が貫いていると述べ

ることができる。

### 3.3. 日本におけるスミス人気と邦訳：WN を例に

理由は一筋縄ではないにせよ、アダム・スミスはかねてより日本で人気を博してきた。1923年にはアダム・スミス生誕二百年を記念して、スミスの誕生（洗礼）日である6月5日に京都帝国大学経済学会が記念展覧会および講演会を開催している。

展覧会場にはWN初版の原著が日本全国から八冊も集まり、耳目をひく。WNの仏訳と独訳に加えて、巖復（1854-1921）による中国語訳『原富』（1901～02年）（水田2015参照）も届けられた。『道徳感情論』仏訳も出展されたい（1798年刊のコンドルセ夫人訳だろう。ちなみに同訳は1999年まで現役だった）。WNの日本語訳は、「石川暎作・三上正毅・竹内謙二諸氏の訳書」以外に「石川氏がその翻訳を最初に掲載された東京経済学講習会講義録」と、石川訳のうち自由貿易に関する箇所（WN第四編）のみを抜粋した『富国論覧要』の二種類が寄せられ、五種類が集った。陳列品目録1500部はあつという間に無くなり、当初一日だけの予定が翌6月6日も展覧会を続けねばならなくなるほどの盛況ぶりだった。5日午後には開催された講演会には一千名を超える聴衆が集まり、河上肇も第二話者として「スミスの著作」と題したスピーチを行っている（『経済論叢』第17巻第1号を見よ）。

戦後の学園闘争時の東大総長として有名な大河内一男（1905-84）はスミス学者であった。大河内監訳『国富論』全三巻（中公文庫、1978年）は、図版を交え注も充実させた名訳である。その付録「『国富論』邦訳小史」で大河内が言うように、「日本では、その近代経済の揺籃の時期にも、その後の体質の変化と軍事的エキスパンションの時期にも、第一次大戦中から大正末期へかけての『大正デモクラシー』の時期においても、さらに満州事変・日華事変、そして、その後の太平洋戦争下の苛烈な思想統制の時期においても、そしてさらに、第二次大戦後における新憲法と思想や運動の『自由』の時期においても、依然としてスミスは研究され、スミスの『国富論』は翻訳され続けてきた。まことに日本人のスミスに対する熱意と愛好の心は深く強かったと言わなければならないし、今日においても、その点は少しも変わっていない」（『国富論Ⅲ』454-55）。大河内はWNに限って述べているようだが、スミスの思想それ自体に対して、日本に生きる者たちがその時々歴史的な文脈の中で多様な接近を試み続けてきたこと、そしてその多種多様な文脈からの要請に耐えきれぬ器の広さと深さをスミスという思想家が有していたこと、それは確かだろうし、大河内が言う「今日」を21世紀に読み替えても、それは変わらないのではないかと。以下、大河内「小史」に拠りながらWN翻訳の流れを確認しておきたい。

スミスWNの抄訳は数多く存在するが、初の全訳は先に触れた石川暎作訳『富国論』第一巻（1884年）と第二巻（1885年）、そして石川・嵯峨正作共訳の第三巻（1888年）であった。それに続くのが三上正毅訳『富国論』（1910年）、その後現れたのが竹内謙二訳『富国論』全三巻（1921～23年）である。23年9月の関東大震災とその後の混乱を契機に竹内は全編を訳し直し、名を『国富論』と改めた廉価版（三巻本、1931～33年）を出版した。同様の版として青野季吉

訳『国富論』二巻本（1928～29年）も現れ、専門研究者だけでなく一般読者層のあいだに読み手を広げた。太平洋戦争直前の1940年には大内兵衛訳『国富論』五巻本の出版が始まる（44年に完結）。ロビNZの師の一人だったキャナン（Edwin Cannan, 1861-1935）による機転の利いた注釈付き校訂版を底本とした大内訳は、事実上の決定版として長く読者を獲得し続けるものの、戦後に『諸国民の富』（1959～66年）と名を変えることになり、松川七郎による改訳も施された。同じ頃、水田洋による『国富論』二巻本（1965年）が登場したことで、WN 刊行以来の、いや、より早くTMSにおける「ひなたぼっこをしている乞食」の幸せという問題提起に端を発するスミス思想の長い旅路も、いよいよ現代日本に到達する。

以上をふまえて、以下ではスミスの原典をテキストマイニングにかけ、得られた知見に対して考察を加えてみたい。

#### 4. アダム・スミスの作品をテキストマイニングする（2）：定量分析

##### 4.1. 「見えざる手」前後の言葉の袋を分析する

英語原典のHAとTMSとWNに一度ずつ登場する「見えざる手」という語の前後を一定の範囲で（おおむね同じ長さになるように）区切り、「言葉の袋」すなわち分析対象テキストを作る。HAについてはinvisible handを含む段落分（616語）、TMSについてはinvisible handを含む段落とその直前段落の後半を合わせた一・五段落分（683語）、WNについては、invisible handを含む段落と、これを挟む前一段落と後二段落を合わせた四段落分（681語）である。

三つの分析対象テキストすべてにおいてinvisible handと共起している代名詞を成分とするベクトルを作り、コサイン類似度を測定したものが下記である。数字は出現頻度（回数）を表す。

##### ①関係代名詞（5W + 1H）等

	when	where	who	whose	whom	which	what	whatever	that	how
HA	0	0	3	3	1	10	0	1	3	0
TMS	2	0	6	0	2	17	2	0	2	0
WN	1	0	3	0	0	8	6	2	0	1

ベクトルは

HA (0, 0, 3, 3, 1, 10, 0, 1, 3, 0)

TMS (2, 0, 6, 0, 2, 17, 2, 0, 2, 0)

WN (1, 0, 3, 0, 0, 8, 6, 2, 0, 1)

ベクトルの大きさは

$$|HA| = \sqrt{0^2 + 0^2 + 3^2 + 3^2 + 1^2 + 10^2 + 0^2 + 1^2 + 3^2 + 0^2} = 11.36$$

$$|TMS| = \sqrt{2^2 + 0^2 + 6^2 + 0^2 + 2^2 + 17^2 + 2^2 + 0^2 + 2^2 + 0^2} = 18.47$$

$$|WN| = \sqrt{1^2 + 0^2 + 3^2 + 0^2 + 0^2 + 8^2 + 6^2 + 2^2 + 0^2 + 1^2} = 10.72$$

ベクトルの内積は

$$HA \cdot TMS = 0 \times 2 + 0 \times 0 + 3 \times 6 + 3 \times 0 + 1 \times 2 + 10 \times 17 + 0 \times 2 + 1 \times 0 + 3 \times 2 + 0 \times 0 = 196$$

$$HA \cdot WN = 0 \times 1 + 0 \times 0 + 3 \times 3 + 3 \times 0 + 1 \times 0 + 10 \times 8 + 0 \times 6 + 1 \times 2 + 3 \times 0 + 0 \times 1 = 91$$

$$TMS \cdot WN = 2 \times 1 + 0 \times 0 + 6 \times 3 + 0 \times 0 + 2 \times 0 + 17 \times 8 + 2 \times 6 + 0 \times 2 + 2 \times 0 + 0 \times 1 = 168$$

コサイン類似度は

$$\text{Sim}(HA, TMS) = \frac{HA \cdot TMS}{|HA||TMS|} = \frac{196}{11.36 \times 18.47} = 0.934$$

$$\text{Sim}(HA, WN) = \frac{HA \cdot WN}{|HA||WN|} = \frac{91}{11.36 \times 10.72} = 0.747$$

$$\text{Sim}(TMS, WN) = \frac{TMS \cdot WN}{|TMS||WN|} = \frac{168}{18.47 \times 10.72} = 0.848$$

よって、HA から見て執筆年代に近いほど類似度が高いと言える。TMS と WN の類似度が HA と TMS のそれを下回る点は、マクフィに拠りつつすでに確認したテキストの成立事情を勘案した場合、思想史上の関係性と齟齬を来しているようである。

## ②人称代名詞等

	I	we	he	himself	she	they	themselves	it	its	itself
HA	0	1	6	0	0	8	2	12	5	0
TMS	0	6	10	3	1	16	0	15	1	1
WN	1	0	21	2	0	12	0	19	4	0

ベクトルは

$$HA (0, 1, 6, 0, 0, 8, 2, 12, 5, 0)$$

$$TMS (0, 6, 10, 3, 1, 16, 0, 15, 1, 1)$$

$$WN (1, 0, 21, 2, 0, 12, 0, 19, 4, 0)$$

ベクトルの大きさは

$$|HA| = \sqrt{0^2 + 1^2 + 6^2 + 0^2 + 0^2 + 8^2 + 2^2 + 12^2 + 5^2 + 0^2} = 16.55$$

$$|TMS| = \sqrt{0^2 + 6^2 + 10^2 + 3^2 + 1^2 + 16^2 + 0^2 + 15^2 + 1^2 + 1^2} = 25.08$$

$$|WN| = \sqrt{1^2 + 0^2 + 21^2 + 2^2 + 0^2 + 12^2 + 0^2 + 19^2 + 4^2 + 0^2} = 31.10$$

ベクトルの内積は

$$HA \cdot TMS = 0 \times 0 + 1 \times 6 + 6 \times 10 + 0 \times 3 + 0 \times 1 + 8 \times 16 + 2 \times 0 + 12 \times 15 + 5 \times 1 + 0 \times 1 = 379$$

$$HA \cdot WN = 0 \times 1 + 1 \times 0 + 6 \times 21 + 0 \times 2 + 0 \times 0 + 8 \times 12 + 2 \times 0 + 12 \times 19 + 5 \times 4 + 0 \times 0 = 470$$

$$TMS \cdot WN = 0 \times 1 + 6 \times 0 + 10 \times 21 + 3 \times 2 + 1 \times 0 + 16 \times 12 + 0 \times 0 + 15 \times 19 + 1 \times 4 + 1 \times 0 = 697$$

コサイン類似度は

$$\text{Sim}(\text{HA}, \text{TMS}) = \frac{\text{HA} \cdot \text{TMS}}{|\text{HA}| |\text{TMS}|} = \frac{379}{16.55 \times 25.08} = 0.913$$

$$\text{Sim}(\text{HA}, \text{WN}) = \frac{\text{HA} \cdot \text{WN}}{|\text{HA}| |\text{WN}|} = \frac{470}{16.55 \times 31.10} = 0.913$$

$$\text{Sim}(\text{TMS}, \text{WN}) = \frac{\text{TMS} \cdot \text{WN}}{|\text{TMS}| |\text{WN}|} = \frac{697}{25.08 \times 31.10} = 0.894$$

驚くべきことに、HA と TMS および HA と WN の類似度が同値になった。TMS と WN の類似度が最低をマークしている点も、先の関係性を正しく反映できていないようである。

続いて名詞・動詞の共起具合を見ておこう。動詞については be 動詞を除外してある。括弧内の数字は出現頻度（回数）を表す。また、頻度1の語は省略した。

名詞			動詞		
HA	TMS	WN	HA	TMS	WN
nature (7)	life (5)	industry (12)	ascribe (4)	have (7)	employ (11)
event (4)	earth (4)	produce (10)	act (3)	make (4)	attempt (5)
gratitude (4)	imagination (3)	value (7)	do (3)	consume (3)	make (5)
object (4)	inhabitant (3)	capital (5)	design (2)	divide (3)	promote (5)
being (3)	desire (2)	domestick (5)	employ (2)	maintain (3)	do (4)
pleasure (3)	greatness (2)	part (4)	excite (2)	afford (2)	have (4)
power (3)	labour (2)	society (4)	flow (2)	consider (2)	intend (4)
savage (3)	mankind (2)	support (4)	stop (2)	employ (2)	buy (2)
sentiment (3)	manner (2)	individual (3)	suppose (2)	fit (2)	direct (2)
thing (3)	means (2)	interest (3)	take (2)	keep (2)	know (2)
age (2)	nature (2)	manner (3)		oblige (2)	
appearance (2)	necessaries (2)	case (2)		produce (2)	
cause (2)	oeconomy (2)	hand (2)		promote (2)	
course (2)	order (2)	man (2)		seem (2)	
displeasure (2)	produce (2)	people (2)			
favour (2)	share (2)	profit (2)			
god (2)	view (2)	publick (2)			
man (2)		regulation (2)			
notion (2)		revenue (2)			
reverence (2)		shoemaker (2)			
storm (2)		statesman (2)			
superstition (2)		taylor (2)			
		thing (2)			

上表の通り、二つ以上のテキストに共通して現れている語は少ない。名詞は man, manner, nature, produce の四語、動詞は do, employ, have, make, promote の五語に過ぎない。そこで名詞の四語と動詞の五語、計九語を成分とするベクトルを組み、同様にコサイン類似度を求め

ておこう。

③名詞と動詞

	man	manner	nature	produce	do	employ	have	make	promote
HA	2	0	7	0	3	2	0	0	0
TMS	0	2	2	2	0	2	7	4	2
WN	2	3	0	10	4	11	4	5	5

ベクトルは

HA (2, 0, 7, 0, 3, 2, 0, 0, 0)

TMS (0, 2, 2, 2, 0, 2, 7, 4, 2)

WN (2, 3, 0, 10, 4, 11, 4, 5, 5)

ベクトルの大きさは

$$|HA| = \sqrt{2^2 + 0^2 + 7^2 + 0^2 + 3^2 + 2^2 + 0^2 + 0^2 + 0^2} = 8.12$$

$$|TMS| = \sqrt{0^2 + 2^2 + 2^2 + 2^2 + 0^2 + 2^2 + 7^2 + 4^2 + 2^2} = 9.22$$

$$|WN| = \sqrt{2^2 + 3^2 + 0^2 + 10^2 + 4^2 + 11^2 + 4^2 + 5^2 + 5^2} = 17.78$$

ベクトルの内積は

$$HA \cdot TMS = 2 \times 0 + 0 \times 2 + 7 \times 2 + 0 \times 2 + 3 \times 0 + 2 \times 2 + 0 \times 7 + 0 \times 4 + 0 \times 2 = 18$$

$$HA \cdot WN = 2 \times 2 + 0 \times 3 + 7 \times 0 + 0 \times 10 + 3 \times 4 + 2 \times 11 + 0 \times 4 + 0 \times 5 + 0 \times 5 = 38$$

$$TMS \cdot WN = 0 \times 2 + 2 \times 3 + 2 \times 0 + 2 \times 10 + 0 \times 4 + 2 \times 11 + 7 \times 4 + 4 \times 5 + 2 \times 5 = 106$$

コサイン類似度は

$$\text{Sim}(HA, TMS) = \frac{HA \cdot TMS}{|HA||TMS|} = \frac{18}{8.12 \times 9.22} = 0.240$$

$$\text{Sim}(HA, WN) = \frac{HA \cdot WN}{|HA||WN|} = \frac{38}{8.12 \times 17.78} = 0.263$$

$$\text{Sim}(TMS, WN) = \frac{TMS \cdot WN}{|TMS||WN|} = \frac{106}{9.22 \times 17.78} = 0.647$$

ここでははっきりと、先の思想史的文脈からして互いに似通っているはずの TMS と WN の類似度合の高さが見て取れる。また、成立事情に照らせば異質なテキストである HA と、残り二つのテキストとの方向性の違いも、ほぼ等しく数値化されている。分析としては成功であろうか。

形容詞と副詞の共起具合についても見ておきたい。同様に、括弧内の数字は出現頻度を表している。頻度1の語は省略した。

形容詞			副詞		
HA	TMS	WN	HA	TMS	WN
intelligent (3)	different (3)	greatest (5)	not (7)	nearly (3)	not (6)
invisible (3)	agreeable (2)	own (5)	sometimes (4)	little (2)	only (5)
irregular (3)	capable (2)	annual (3)	even (2)	only (2)	always (3)
other (3)	human (2)	other (3)	most (2)	so (2)	so (3)
proper (3)	more (2)	direct (2)	naturally (2)	thus (2)	generally (2)
agreeable (2)	natural (2)	exchangeable (2)	now (2)	well (2)	indeed (2)
own (2)	own (2)	foreign (2)	presently (2)		much (2)
same (2)	poor (2)	good (2)	therefore (2)		never (2)
	real (2)	great (2)			precisely (2)
	vain (2)	hurtful (2)			therefore (2)
	whole (2)	likely (2)			very (2)
		much (2)			
		private (2)			
		same (2)			
		useless (2)			
		whole (2)			

形容詞は agreeable, other, own, same, whole の五語、副詞は not, only, so, therefore の四語が二つ以上のテキストに現れている。これらの形容詞五語と副詞四語を成分とするベクトルを組み、類似度を比較しよう。

#### ④形容詞と副詞

	agreeable	other	own	same	whole	not	only	so	therefore
HA	2	3	2	2	0	7	0	0	2
TMS	2	0	2	0	2	0	2	2	0
WN	0	3	5	2	2	6	5	3	2

ベクトルは

HA (2, 3, 2, 2, 0, 7, 0, 0, 2)

TMS (2, 0, 2, 0, 2, 0, 2, 2, 0)

WN (0, 3, 5, 2, 2, 6, 5, 3, 2)

ベクトルの大きさは

$$|HA| = \sqrt{2^2 + 3^2 + 2^2 + 2^2 + 0^2 + 7^2 + 0^2 + 0^2 + 2^2} = 8.60$$

$$|TMS| = \sqrt{2^2 + 0^2 + 2^2 + 0^2 + 2^2 + 0^2 + 2^2 + 2^2 + 0^2} = 4.47$$

$$|WN| = \sqrt{0^2 + 3^2 + 5^2 + 2^2 + 2^2 + 6^2 + 5^2 + 3^2 + 2^2} = 10.77$$

ベクトルの内積は

$$HA \cdot TMS = 2 \times 2 + 3 \times 0 + 2 \times 2 + 2 \times 0 + 0 \times 2 + 7 \times 0 + 0 \times 2 + 0 \times 2 + 2 \times 0 = 8$$

$$HA \cdot WN = 2 \times 0 + 3 \times 3 + 2 \times 5 + 2 \times 2 + 0 \times 2 + 7 \times 6 + 0 \times 5 + 0 \times 3 + 2 \times 2 = 69$$

$$TMS \cdot WN = 2 \times 0 + 0 \times 3 + 2 \times 5 + 0 \times 2 + 2 \times 2 + 0 \times 6 + 2 \times 5 + 2 \times 3 + 0 \times 2 = 30$$

コサイン類似度は

$$\text{Sim}(HA, TMS) = \frac{HA \cdot TMS}{|HA||TMS|} = \frac{8}{8.60 \times 4.47} = 0.208$$

$$\text{Sim}(HA, WN) = \frac{HA \cdot WN}{|HA||WN|} = \frac{69}{8.60 \times 10.77} = 0.745$$

$$\text{Sim}(TMS, WN) = \frac{TMS \cdot WN}{|TMS||WN|} = \frac{30}{4.47 \times 10.77} = 0.623$$

HA と WN の類似度の高さは、否定（not）構文などの言い回しが似ているためだろう。

このように分析してきたが、①関係代名詞群、②人称代名詞群、③名詞＋動詞群、④形容詞＋副詞群の中では、③によるコサイン類似度の測定結果が、マクフィの言う意味でのテキスト本来の関係性を最も正確に反映することが分かった。では①～④すべての成分を組み合わせたベクトルを作り、コサイン類似度を測定するとどうなるか。

ベクトルは

$$HA \quad (0, 0, 3, 3, 1, 10, 0, 0, 1, 3, 0, 0, 1, 6, 0, 0, 8, 2, 12, 5, 0, 2, 0, 7, 0, 3, 2, 0, 0, 0, 2, 3, 2, 2, 0, 7, 0, 0, 2)$$

$$TMS \quad (2, 0, 6, 0, 2, 17, 2, 0, 2, 0, 0, 6, 10, 3, 1, 16, 0, 15, 1, 1, 0, 2, 2, 2, 0, 2, 7, 4, 2, 2, 0, 2, 0, 2, 0, 2, 2, 0)$$

$$WN \quad (1, 0, 3, 0, 0, 8, 6, 2, 0, 1, 1, 0, 21, 2, 0, 12, 0, 19, 4, 0, 2, 3, 0, 10, 4, 11, 4, 5, 5, 0, 3, 5, 2, 2, 6, 5, 3, 2)$$

コサイン類似度は

$$\text{Sim}(HA, TMS) = \frac{HA \cdot TMS}{|HA||TMS|} = \frac{196 + 379 + 18 + 8}{\sqrt{11.36^2 + 16.55^2 + 8.12^2 + 8.60^2} \times \sqrt{18.47^2 + 25.08^2 + 9.22^2 + 4.47^2}} = 0.787$$

$$\text{Sim}(HA, WN) = \frac{HA \cdot WN}{|HA||WN|} = \frac{91 + 470 + 38 + 69}{\sqrt{11.36^2 + 16.55^2 + 8.12^2 + 8.60^2} \times \sqrt{10.72^2 + 31.10^2 + 17.78^2 + 10.77^2}} = 0.737$$

$$\text{Sim}(TMS, WN) = \frac{TMS \cdot WN}{|TMS||WN|} = \frac{168 + 697 + 106 + 30}{\sqrt{18.47^2 + 25.08^2 + 9.22^2 + 4.47^2} \times \sqrt{10.72^2 + 31.10^2 + 17.78^2 + 10.77^2}} = 0.785$$

三つの値に大きな差はないが、これは、出現頻度が高く、いわばありふれた言い回しに相当する①と②の成分がベクトルのなす角度 $\theta$ に重く影響し、差異を見えにくくしているためである。より特徴的な部分に焦点を当てなければならない。③と④の成分（下線部）に限定すると、 $\text{Sim}(TMS, WN) = 0.639$ が最高値、 $\text{Sim}(HA, TMS) = 0.215$ が最低値となり明白な差が現れる。 $\text{Sim}(HA, WN) = 0.435$ の水準が中位なのは、前述の通り否定構文の形式的影響に過ぎないと考えられるため、TMS と WN の類似度が突出して高いと述べて差し支えない。

本稿3節で明らかにした文脈、つまり思想史研究のアプローチ法の所産に従った場合に想定される類似性と、分布仮説および「言葉の袋」仮説に立つベクトル距離から求めた類似度とは、同質・同次元の現象ではないため完全に重なるとは言えない。しかしながら、両分析手法の結果が少なくとも交差する場合があると確かめられた。後者は無価値な知見ではないと言ってよい。続いて翻訳について見てみよう。

#### 4.2. 邦訳のテキストマイニング

TMSの初の全訳は米林富男(1905-68)の偉業だが、タイトルの *moral sentiments* を「道徳情操」とした点は後続の賛同を得られなかった(米林富男『道徳情操論』1969~70年)。世界に名を知られたスミス研究者、水田洋による『道徳感情論』(2003年)は学術的価値も十分と言える。近年では高哲男訳『道徳感情論』(2013年)と村井章子・北川知子訳『道徳感情論』(2014年)が相次いで出版され、特に村井+北川訳は職業翻訳家の作品らしく読みやすい。米林、水田、高はいずれも学者を本業とする人々である。

比較的近年のWN翻訳としては、先述した大河内訳の他に、水田洋と杉山忠平が訳した『国富論』(2000~01年)が存在する。これらの学者訳に職業翻訳家が挑戦したものが、「国の豊かさの本質と原因についての研究」という原著本来のタイトルを副題として添えた山岡洋一訳『国富論』(2007年)だろう。最近では、高が『道徳感情論』に続いて新訳『国富論』(2020年)を完成させ、学界を沸かせている。

さて、TMSとWNの各種邦訳をテキストマイニングにかけると、どのような特徴が浮かび上がるだろうか。英語原典を分析した際と同様に、*invisible hand* を含む文脈を一定の範囲で区切り、「言葉の袋」を取り出そう。袋の大きさを文字数で表現すれば、TMSについては(句読点を含み)米林訳1974字、水田訳1793字、高訳1675字、村井+北川訳1547字である。またWNの場合は大河内訳1703字、水田訳1749字、山岡訳1503字、高訳1757字となる。英語原典として見れば全く同一の長さを対象にしているにもかかわらず邦訳の文字数が大きく異なり、特に職業翻訳家のそれが顕著に少なくなる傾向が見られる。それは、読みやすさを重視した翻訳プロフェッショナルならではのテクニックのなせる業だと言えようか。なお、紙幅の都合上、名詞とサ変名詞についての分析結果のみ掲載する。

TMS：YBは米林訳，Mは水田訳，Tは高訳，MKは村井+北川訳を指す

名詞				サ変名詞			
YB	M	T	MK	YB	M	T	MK
人々(13)	人びと(10)	人々(8)	大地(4)	生活(4)	生活(6)	生活(6)	労働(3)
土地(8)	自身(3)	人間(6)	人々(3)	生産(4)	維持(3)	維持(3)	運営(2)
事物(4)	住民(3)	富者(5)	土地(3)	想像(4)	使用(3)	生産(3)	活用(2)
自分(4)	人間(3)	自身(4)	胃袋(2)	欺瞞(3)	消費(3)	提供(3)	策略(2)
分け前(4)	大地(3)	自分(4)	観点(2)	消費(3)	想像(3)	改良(2)	満足(2)
快樂(3)	見方(2)	価値(3)	金持ち(2)	分配(3)	分割(3)	考察(2)	
人類(3)	公道(2)	住人(3)	手段(2)	満足(3)	分配(3)	考慮(2)	
富者(3)	自分(2)	大地(3)	大勢(2)	改良(2)	運動(2)	支配(2)	
胃袋(2)	社会(2)	地球(3)	邸宅(2)	企図(2)	改良(2)	除外(2)	
観点(2)	手段(2)	地主(3)	農地(2)	苦勞(2)	考察(2)	消費(2)	
自身(2)	人類(2)	土地(3)	分け前(2)	所有(2)	生産(2)	想像(2)	
人間(2)	地位(2)	必需(3)	本質(2)	心配(2)	提供(2)	分配(2)	
人口(2)	土地(2)	手段(2)	欲望(2)	促進(2)	欲求(2)	満足(2)	
地球(2)	必需(2)	分け前(2)		配置(2)	労働(2)	利用(2)	
必需(2)	分け前(2)	欲望(2)		分割(2)		労働(2)	
貧乏人(2)				労働(2)			
欲望(2)							

名詞については「人々」「土地」「自分」「分け前」「人類」「富者」「胃袋」「観点」「自身」「人間」「地球」「必需」「欲望」「大地」「手段」の15語，サ変名詞については「生活」「生産」「想像」「消費」「分配」「満足」「改良」「分割」「労働」「維持」「考察」「提供」の12語，計27語の出現頻度を成分とするベクトルを組み，類似度を測定する。

	人々	土地	自分	分け前	人類	富者	胃袋	観点	自身	人間	地球	必需	欲望
YB	13	8	4	4	3	3	2	2	2	2	2	2	2
M	0	2	2	2	2	0	0	0	3	3	0	2	0
T	8	3	4	2	0	5	0	0	4	6	3	3	2
MK	3	3	0	2	0	0	2	2	0	0	0	0	2

	大地	手段	生活	生産	想像	消費	分配	満足	改良	分割	労働	維持	考察	提供
0	0	4	4	4	3	3	3	2	2	2	0	0	0	
3	2	6	2	3	3	3	0	2	3	2	3	2	2	
3	2	6	3	2	2	2	2	2	0	2	3	2	3	
4	2	0	0	0	0	0	2	0	0	3	0	0	0	

ベクトルは（下線部=サ変名詞，以下同）

YB (13, 8, 4, 4, 3, 3, 2, 2, 2, 2, 2, 2, 0, 0, 4, 4, 4, 3, 3, 3, 2, 2, 2, 0, 0, 0)

M (0, 2, 2, 2, 2, 0, 0, 0, 3, 3, 0, 2, 0, 3, 2, 6, 2, 3, 3, 3, 0, 2, 3, 2, 3, 2, 2)

T (8, 3, 4, 2, 0, 5, 0, 0, 4, 6, 3, 3, 2, 3, 2, 6, 3, 2, 2, 2, 2, 2, 0, 2, 3, 2, 3)

MK (3, 3, 0, 2, 0, 0, 2, 2, 0, 0, 0, 0, 2, 4, 2, 0, 0, 0, 0, 2, 0, 0, 3, 0, 0, 0)

ベクトルの大きさは

$$|YB| = \sqrt{13^2 + 8^2 + 4^2 + 4^2 + 3^2 + 3^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 4^2 + 4^2 + 4^2 + 3^2 + 3^2 + 3^2 + 2^2 + 2^2 + 2^2} \\ = 19.95 \text{ (17.63, 9.33)}$$

$$|M| = \sqrt{2^2 + 2^2 + 2^2 + 2^2 + 3^2 + 3^2 + 2^2 + 3^2 + 2^2 + 6^2 + 2^2 + 3^2 + 3^2 + 3^2 + 2^2 + 3^2 + 2^2 + 3^2 + 2^2 + 2^2} = 12.33 \text{ (7.14, 10.05)}$$

$$|T| = \sqrt{8^2 + 3^2 + 4^2 + 2^2 + 5^2 + 4^2 + 6^2 + 3^2 + 3^2 + 2^2 + 3^2 + 2^2 + 6^2 + 3^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 3^2 + 2^2 + 3^2} \\ = 17.20 \text{ (14.32, 9.54)}$$

$$|MK| = \sqrt{3^2 + 3^2 + 2^2 + 2^2 + 2^2 + 2^2 + 4^2 + 2^2 + 2^2 + 3^2} = 8.19 \text{ (7.35, 3.61)}$$

ベクトルの内積は

$$YB \cdot M = 0 + 16 + 8 + 8 + 6 + 0 + 0 + 0 + 6 + 6 + 0 + 4 + 0 + 0 + 0 + 24 + 8 + 12 + 9 + 9 + 0 + 4 + 6 + 4 + 0 + 0 + 0 = 130 \text{ (54, 76)}$$

$$YB \cdot T = 104 + 24 + 16 + 8 + 0 + 15 + 0 + 0 + 8 + 12 + 6 + 6 + 4 + 0 + 0 + 24 + 12 + 8 + 6 + 6 + 6 + 4 + 0 + 4 + 0 + 0 + 0 \\ = 273 \text{ (203, 70)}$$

$$YB \cdot MK = 39 + 24 + 0 + 8 + 0 + 0 + 4 + 4 + 0 + 0 + 0 + 0 + 4 + 0 + 0 + 0 + 0 + 0 + 0 + 6 + 0 + 0 + 6 + 0 + 0 + 0 = 95 \text{ (83, 12)}$$

$$M \cdot T = 0 + 6 + 8 + 4 + 0 + 0 + 0 + 0 + 12 + 18 + 0 + 6 + 0 + 9 + 4 + 36 + 6 + 6 + 6 + 6 + 0 + 4 + 0 + 4 + 9 + 4 + 6 = 154 \text{ (67, 87)}$$

$$M \cdot MK = 0 + 6 + 0 + 4 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 12 + 4 + 0 + 0 + 0 + 0 + 0 + 0 + 6 + 0 + 0 + 0 + 0 = 32 \text{ (26, 6)}$$

$$T \cdot MK = 24 + 9 + 0 + 4 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 4 + 12 + 4 + 0 + 0 + 0 + 0 + 4 + 0 + 0 + 6 + 0 + 0 + 0 = 67 \text{ (57, 10)}$$

コサイン類似度は

$$\text{Sim}(YB, M) = \frac{YB \cdot M}{|YB||M|} = \frac{130}{19.95 \times 12.33} = 0.528 \left( \frac{54}{17.63 \times 7.14} = 0.429 \right) \left( \frac{76}{9.33 \times 10.05} = \underline{0.811} \right)$$

$$\text{Sim}(YB, T) = \frac{YB \cdot T}{|YB||T|} = \frac{273}{19.95 \times 17.20} = 0.796 \left( \frac{203}{17.63 \times 14.32} = 0.804 \right) \left( \frac{70}{9.33 \times 9.54} = \underline{0.786} \right)$$

$$\text{Sim}(YB, MK) = \frac{YB \cdot MK}{|YB||MK|} = \frac{95}{19.95 \times 8.19} = 0.581 \left( \frac{83}{17.63 \times 7.35} = 0.641 \right) \left( \frac{12}{9.33 \times 3.61} = \underline{0.356} \right)$$

$$\text{Sim}(M, T) = \frac{M \cdot T}{|M||T|} = \frac{154}{12.33 \times 17.20} = 0.726 \left( \frac{67}{7.14 \times 14.32} = 0.655 \right) \left( \frac{87}{10.05 \times 9.54} = \underline{0.907} \right)$$

$$\text{Sim}(M, MK) = \frac{M \cdot MK}{|M||MK|} = \frac{32}{12.33 \times 8.19} = 0.317 \left( \frac{26}{7.14 \times 7.35} = 0.495 \right) \left( \frac{6}{10.05 \times 3.61} = \underline{0.165} \right)$$

$$\text{Sim}(T, MK) = \frac{T \cdot MK}{|T||MK|} = \frac{67}{17.20 \times 8.19} = 0.476 \left( \frac{57}{14.32 \times 7.35} = 0.542 \right) \left( \frac{10}{9.54 \times 3.61} = \underline{0.290} \right)$$

WN：Oは大河内訳，Mは水田（杉山）訳，Yは山岡訳，Tは高訳を指す

名詞				サ変名詞			
O	M	Y	T	O	M	Y	T
自分(17)	自分(13)	自分(13)	自分(10)	勤労(8)	勤労(14)	生産(11)	生産(11)
社会(10)	価値(7)	社会(8)	自身(8)	生産(8)	生産(12)	労働(11)	労働(9)
価値(7)	国内(7)	利益(8)	価値(7)	活動(5)	使用(9)	努力(5)	組織(7)
資本(7)	最大(6)	価値(7)	国内(7)	増進(5)	意図(5)	意図(4)	意図(5)
国内(5)	資本(6)	各人(6)	産業(7)	意図(4)	推進(5)	交換(3)	促進(5)
最大(5)	利益(6)	国内(6)	最大(6)	維持(4)	規制(3)	規制(2)	利用(5)
産業(5)	産業(5)	資本(5)	資本(5)	交換(4)	交換(3)	供給(2)	維持(4)
利益(5)	社会(5)	年間(3)	利益(5)	規制(3)	指示(2)	指示(2)	交換(3)
自身(4)	個人(3)	目的(3)	社会(4)	労働(2)	判断(2)	注文(2)	購入(3)
外国(2)	自身(3)	外国(2)	人々(4)			配慮(2)	規制(2)
権限(2)	種類(3)	権限(2)	個人(3)			判断(2)	指導(2)
国産(2)	一般(2)	雇い主(2)	方法(3)			予想(2)	製造(2)
仕立(2)	権威(2)	仕立て(2)	一般(2)				努力(2)
市場(2)	公共(2)	市場(2)	外国(2)				判断(2)
種類(2)	仕立(2)	収入(2)	個人(2)				
収入(2)	私人(2)	政治(2)	公共(2)				
政治(2)	収入(2)	民間(2)	仕立(2)				
目的(2)	政治(2)		収入(2)				
利潤(2)	利潤(2)		政治(2)				
			利潤(2)				

名詞については「自分」「社会」「価値」「資本」「国内」「最大」「産業」「利益」「自身」「外国」「権限」「仕立」「市場」「種類」「収入」「政治」「利潤」「個人」「一般」「公共」の20語，サ変名詞については「勤労」「生産」「意図」「維持」「交換」「規制」「労働」「指示」「判断」「努力」の10語，計30語の出現頻度を成分とするベクトルを組み，類似度を測定する。

	自分	社会	価値	資本	国内	最大	産業	利益	自身	外国
O	17	10	7	7	5	5	5	5	4	2
M	13	5	7	6	7	6	5	6	3	0
Y	13	8	7	5	6	0	0	8	0	2
T	10	4	7	5	7	6	7	5	8	2

権限	仕立	市場	種類	収入	政治	利潤	個人	一般	公共
2	2	2	2	2	2	2	0	0	0
0	2	0	3	2	2	2	3	2	2
2	0	2	0	2	2	0	0	0	0
0	2	0	0	2	2	2	3	2	2

勤労	生産	意図	維持	交換	規制	労働	指示	判断	努力
8	8	4	4	4	3	2	0	0	0
14	12	5	0	3	3	0	2	2	0
0	11	4	0	3	2	11	2	2	5
0	11	5	4	3	2	9	0	2	2

ベクトルは

$$O = (17, 10, 7, 7, 5, 5, 5, 5, 4, 2, 2, 2, 2, 2, 2, 2, 2, 0, 0, 0, 8, 8, 4, 4, 4, 3, 2, 0, 0, 0)$$

$$M = (13, 5, 7, 6, 7, 6, 5, 6, 3, 0, 0, 2, 0, 3, 2, 2, 2, 3, 2, 2, 14, 12, 5, 0, 3, 3, 0, 2, 2, 0)$$

$$Y = (13, 8, 7, 5, 6, 0, 0, 8, 0, 2, 2, 0, 2, 0, 2, 2, 0, 0, 0, 0, 11, 4, 0, 3, 2, 11, 2, 2, 5)$$

$$T = (10, 4, 7, 5, 7, 6, 7, 5, 8, 2, 0, 2, 0, 0, 2, 2, 2, 3, 2, 2, 0, 11, 5, 4, 3, 2, 9, 0, 2, 2)$$

ベクトルの大きさは

$$\begin{aligned} |O| &= \sqrt{17^2 + 10^2 + 7^2 + 7^2 + 5^2 + 5^2 + 5^2 + 5^2 + 4^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 8^2 + 8^2 + 4^2 + 4^2 + 4^2 + 3^2 + 2^2} \\ &= 28.71 (25.20, 13.75) \end{aligned}$$

$$\begin{aligned} |M| &= \sqrt{13^2 + 5^2 + 7^2 + 6^2 + 7^2 + 6^2 + 5^2 + 6^2 + 3^2 + 2^2 + 3^2 + 2^2 + 2^2 + 2^2 + 2^2 + 3^2 + 2^2 + 2^2 + 14^2 + 12^2 + 5^2 + 3^2 + 3^2 + 2^2 + 2^2} \\ &= 29.44 (21.82, 19.77) \end{aligned}$$

$$|Y| = \sqrt{13^2 + 8^2 + 7^2 + 5^2 + 6^2 + 8^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 11^2 + 4^2 + 3^2 + 2^2 + 11^2 + 2^2 + 2^2 + 5^2} = 27.04 (20.66, 17.44)$$

$$\begin{aligned} |T| &= \sqrt{10^2 + 4^2 + 7^2 + 5^2 + 7^2 + 6^2 + 7^2 + 5^2 + 8^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 2^2 + 3^2 + 2^2 + 2^2 + 11^2 + 5^2 + 4^2 + 3^2 + 2^2 + 9^2 + 2^2 + 2^2} \\ &= 26.72 (21.21, 16.25) \end{aligned}$$

ベクトルの内積は

$$\begin{aligned} O \cdot M &= 221 + 50 + 49 + 42 + 35 + 30 + 25 + 30 + 12 + 0 + 0 + 4 + 0 + 6 + 4 + 4 + 4 + 4 + 0 + 0 + 0 \\ &\quad + 112 + 96 + 20 + 0 + 12 + 9 + 0 + 0 + 0 + 0 = 765 (516, 249) \end{aligned}$$

$$\begin{aligned} O \cdot Y &= 221 + 80 + 49 + 35 + 30 + 0 + 0 + 40 + 0 + 4 + 4 + 0 + 4 + 4 + 0 + 0 + 0 + 0 + 0 + 0 \\ &\quad + 0 + 88 + 16 + 0 + 12 + 6 + 22 + 0 + 0 + 0 = 619 (475, 144) \end{aligned}$$

$$\begin{aligned} O \cdot T &= 170 + 40 + 49 + 35 + 35 + 30 + 35 + 25 + 32 + 4 + 0 + 4 + 0 + 0 + 4 + 4 + 4 + 0 + 0 + 0 \\ &\quad + 0 + 88 + 20 + 16 + 12 + 6 + 18 + 0 + 0 + 0 = 631 (471, 160) \end{aligned}$$

$$\begin{aligned} M \cdot Y &= 169 + 40 + 49 + 30 + 42 + 0 + 0 + 48 + 0 + 0 + 0 + 0 + 0 + 4 + 4 + 0 + 0 + 0 + 0 + 0 + 0 + 132 + 20 + 0 + 9 + 6 + 0 + 4 + 4 + 0 \\ &= 561 (386, 175) \end{aligned}$$

$$\begin{aligned} M \cdot T &= 130 + 20 + 49 + 30 + 49 + 36 + 35 + 30 + 24 + 0 + 0 + 4 + 0 + 0 + 4 + 4 + 4 + 9 + 4 + 4 \\ &\quad + 0 + 132 + 25 + 0 + 9 + 6 + 0 + 0 + 4 + 0 = 612 (436, 176) \end{aligned}$$

$$\begin{aligned} Y \cdot T &= 130 + 32 + 49 + 25 + 42 + 0 + 0 + 40 + 0 + 4 + 0 + 0 + 0 + 0 + 4 + 4 + 0 + 0 + 0 + 0 \\ &\quad + 0 + 121 + 20 + 0 + 9 + 4 + 99 + 0 + 4 + 10 = 597 (330, 267) \end{aligned}$$

コサイン類似度は

$$\text{Sim}(O, M) = \frac{O \cdot M}{|O||M|} = \frac{765}{28.71 \times 29.44} = 0.905 \left( \frac{516}{25.20 \times 21.82} = 0.938 \right) \left( \frac{249}{13.75 \times 19.77} = \underline{0.916} \right)$$

$$\text{Sim}(O, Y) = \frac{O \cdot Y}{|O||Y|} = \frac{619}{28.71 \times 27.04} = 0.797 \left( \frac{475}{25.20 \times 20.66} = 0.912 \right) \left( \frac{144}{13.75 \times 17.44} = \underline{0.601} \right)$$

$$\text{Sim}(O, T) = \frac{O \cdot T}{|O||T|} = \frac{631}{28.71 \times 26.72} = 0.823 \left( \frac{471}{25.20 \times 21.21} = 0.881 \right) \left( \frac{160}{13.75 \times 16.25} = \underline{0.716} \right)$$

$$\text{Sim}(M, Y) = \frac{M \cdot Y}{|M||Y|} = \frac{561}{29.44 \times 27.04} = 0.705 \left( \frac{386}{21.82 \times 20.66} = 0.856 \right) \left( \frac{175}{19.77 \times 17.44} = \underline{0.508} \right)$$

$$\text{Sim}(M, T) = \frac{M \cdot T}{|M||T|} = \frac{612}{29.44 \times 26.72} = 0.778 \left( \frac{436}{21.82 \times 21.21} = 0.942 \right) \left( \frac{176}{19.77 \times 16.25} = \underline{0.548} \right)$$

$$\text{Sim}(Y, T) = \frac{Y \cdot T}{|Y||T|} = \frac{597}{27.04 \times 26.72} = 0.826 \left( \frac{330}{20.66 \times 21.21} = 0.753 \right) \left( \frac{267}{17.44 \times 16.25} = \underline{0.942} \right)$$

以上のように、TMS 邦訳では米林訳と高訳、次いで水田訳と高訳の類似度が目立って高く、WN 邦訳では大河内訳と水田訳のペアを頂点として、山岡訳と高訳、大河内訳と高訳の両ペアがほぼ同値で続く（ただし、他のペアもおしなべて高水準である）。なお、ここでは省いた形容動詞・動詞・形容詞・副詞を用いた分析では、TMS の場合は水田訳と高訳が最も類似、それに米林訳と水田訳が続く、WN の場合は大河内訳と高訳が突出して高い値をマークする。いずれの場合も上位ペアのカウンターパートに登場する高訳は、翻訳語の選択にさほど特徴がないということになるのか。また、一般的に言って直近の先行訳に訳語選定が影響されやすいという仮説も成り立つだろう。むしろ、invisible hand 前後の「言葉の袋」を対象を限ったうえでの話に過ぎないことには、十分留意しておかねばならない。

## 5. 結語

本稿では、スミスが「見えざる手」という言葉をその中で用いた三つの原典、すなわち「天文学史」『道徳感情論』『国富論』のそれぞれから切り出した「言葉の袋」のコサイン類似度を比較し、特徴的な一定の品詞と関連付けた場合において、『道徳感情論』と『国富論』が著しく近い関係にある事実を導き出した。両著は何らかの共通の文脈のもとにある可能性が高い、と確かめられたことになる。かたや、こうして類似性が裏付けられた『道徳感情論』と『国富論』の同じ「言葉の袋」を対象にした諸翻訳テキストのコサイン類似度を計量してみると、値が1に近い高水準に集まる傾向を持つのは当然のことながら、中でも一部の翻訳については「訳語の選定において特徴が少ない」蓋然性や（概して）「直近の先行訳の影響を受けやすい」蓋然性が高いと知られた。ただし、切り出すテキストの範囲をより広げ、分析対象とする総単語数（データ量）を増やした際には、翻訳テキストの比較に基づいたこの後段の知見は修正を迫られる可能性がある。

テキストマイニングの方法により導出された上記前段の「客観的」文脈を、思想史的アプロー

チを通じて見出される文脈と照らし合わせても、違和感はない。しかし、その「違和感」自体が、思想史研究者個々の主観的ないし集合主観的な価値判断に合う／合わないによって感得されるものに過ぎないとしたら、問題はそこにこそ横たわっている。テキストマイニングという「最近の慣行」に棹差すことの本当の意義は、控えめな言い方にはなるが、むしろその全域に及ぶものとはなりえないにせよ、テキスト分析には確固たる客観性を帯びた領域が存在するという事実を、思想史専業の研究集団に対して改めて注意喚起するところにあるのではないだろうか。

## 参考文献

- Harris, Z. S. 1954. Distributional Structure. *WORD*, vol. 10, issue 2-3.
- Macfie, A. L. 1971. The Invisible Hand of Jupiter. *Journal of the History of Ideas*, vol. 32, no. 4.
- Robbins, L. 2018. *Lionel Robbins on the Principles of Economic Analysis*, edited by Susan Howson. Routledge.
- Smith, A. 1976. *The Theory of Moral Sentiments*, edited by D. D. Raphael and A. L. Macfie. Clarendon Press.
- Smith, A. 1976. *An Inquiry into the Nature and Causes of the Wealth of Nations*, edited by W. B. Todd. 2 vols. Clarendon Press.
- Smith, A. 1982. *Essays on Philosophical Subjects*, edited by W. P. D. Wightman and J. C. Bryce. Liberty Fund.
- スミス1969-70. 米林富男訳『道徳情操論（上・下）』未来社.
- スミス1978. 大河内一男監訳『国富論（Ⅰ・Ⅱ・Ⅲ）』中公文庫.
- スミス1993. 水田洋ほか訳『アダム・スミス 哲学論文集』名古屋大学出版会.
- スミス2000-01. 水田洋監訳・杉山忠平訳『国富論（1・2・3・4）』岩波文庫.
- スミス2003. 水田洋訳『道徳感情論（上・下）』岩波文庫.
- スミス2007. 山岡洋一訳『国富論：国の豊かさの本質と原因についての研究（上・下）』日本経済新聞出版.
- スミス2013. 高哲男訳『道徳感情論』講談社学術文庫.
- スミス2014. 村井章子・北川知子訳『道徳感情論』日経BP.
- スミス2020. 高哲男訳『国富論（上・下）』講談社学術文庫.
- ロビンズ2009. 田中秀夫監訳『一経済学者の自伝』ミネルヴァ書房.
- 岩波データサイエンス刊行委員会編2016. 『岩波データサイエンス Vol.2』岩波書店.
- 河上肇1984. 住谷一彦編『河上肇』中公パックス.
- 京大経済学会1923. 「アダム・スミス生誕二百年記念会記事」『経済論叢』17巻1号.
- 小峯敦編2021. 『テキストマイニングから読み解く経済学史』ナカニシヤ出版.
- 牧野邦昭 [2010] 2020. 『戦時下の経済学者：経済学と総力戦』中公選書.
- 水田洋2015. 「巖復（1854-1921）による『国富論』中国訳（1901-2）について」『日本学士院紀要』70巻1号.