柳井先生と多変量解析

高根芳雄

December 10, 2014

ABSTRACT

Late Professor Yanai has contributed to many fields ranging from aptitude diagnostics, epidemiology, and nursing to psychometrics and statistics. This paper reviews some of his accomplishments in linear algebra and multivariate analysis through his collaborative work with the present author, along with some untold episodes for the inception of key ideas underlying the work. The various topics covered include constrained principal component analysis, extensions of Khatri's lemma, the Wedderburn-Guttman theorem, ridge operators, decompositions of the total association between two sets of variables, and ideal instruments. A common thread running through all of them is projectors and singular value decomposition (SVD), which are the main subject matters of a recent monograph by Yanai, Takeuchi, and Takane (2011).

Key Words: projectors, singular value decomposition, Khatri's lemma, the Wedderburn-Guttman theorem, ridge operators, principal component analysis, canonical correlation analysis, confounding variables, propensity scores, instrumental variables
キーワード： 射影子、特異値分解、カトリの補助定理、ウェダーバーン・ガットマンの定理、リッジ演算子、主成分分析、正準相関分析、交絡変数、傾向得点、操作変数.

所属：Adjunct Professor in the Department of Psychology, University of Victoria, British Columbia, Canada
連絡先：5173 Del Monte Avenue, Victoria, BC, Canada, V8Y 1X3. Email: yoshio.takane@mcgill.ca

1　プロローグ

私は大学に入った頃社会心理学に興味を持っていた。ところが3年になって心理学科に進学した直後殆んど偶然といってもよい成行きで柳井先生と知り合いになった。そして先生の影響の下に計量心理学、多変量解析の研究に進むことになった。それ以来50年近い月日が流れているが、いまだお釈迦様の手の平を飛んでいる孫悟空のようにその影響下から逃れられない。それどころか最近ますます影響が深くなって来ている気がしてならない。そんな中で今回柳井先生の業績を振り返る役を仰せつかった。それは私にとってこの上もない名誉であるが、若干重荷でもあった。今回初めて柳井先生のホームページを覗かせていただいたが、そこには先生がこれまで出版された本やその他の刊行物が7つのカテゴリーに分けてリストされ、その1つ1つに数十から百を超える文献が掲載されている。7つのカテゴリーは　1.単行本、2.教育心理学・心理学・テスト研究、3.医学・疫学・看護学関連論文、4.統計学・多変量解析関連論文、5.大学入試関連論文、6.国際学会発表論文、7.学会・研究会での講演の多岐に渡っている。私にはとてもその全貌をカバーする知識も力量もないので本文ではそのうち4に関するもの、特に私が共著者として割合深く関係したものを中心にその誕生のエピソードなどを交えながら取り上げて行きたい。柳井先生との共著出版物は今回改めて数えたところ単行本も含めて15にのぼり（本文末の共著リストを参照）、最初の1篇（肥田野他、1971）を除いては全て上記の4に関するものである。このうち 1976 年初版の「多変量解析法」（柳井・高根、1976）を除いては 1990 年以降のものばかりである。これが柳井先生の影響が最近ますます強くなって来ていると感じる所以であるが、おそらく若い頃は柳井先生と同じようなことをやっても到底太刀打ち出来る筈がなく、何となくそれを避けて来たためではないかと思われる。因みに「多変量解析法」で私が担当したのは多次元尺度法と非計量多変量解析法の部分で、いづれも柳井先生の得意とする分野とは一線を画している。

このような事態に転機が訪れたのは 1980 年代の後半であるが、本題に入る前に本論文で頻繁に使われる記号を整理しておきたい。矩形行列 $\mathbf{X}$ の列ベクトルの張る空間を $\mathrm{Sp}(\mathbf{X})$, その直交補空間を $\mathrm{Ker}(\mathbf{X}')$ で表す。$\mathrm{Sp}(\mathbf{X})$ への直交射影子を

$$\mathbf{P}_X = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}', \tag{1}$$

$\mathrm{Ker}(\mathbf{X}')$ への直交射影子を

$$\mathbf{Q}_X = \mathbf{I} - \mathbf{P}_X = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}' \tag{2}$$

で表す。ここで $^{-}$ は一般化逆行列を表す記号である。$\mathbf{P}_X$ と $\mathbf{Q}_X$ は対称 ($\mathbf{P}'_X = \mathbf{P}_X$, $\mathbf{Q}'_X = \mathbf{Q}_X$) で、べき等 ($\mathbf{P}^2_X = \mathbf{P}_X$, $\mathbf{Q}^2_X = \mathbf{Q}_X$)、かつ互いに直交する ($\mathbf{P}_X\mathbf{Q}_X = \mathbf{Q}_X\mathbf{P}_X = \mathbf{O}$) ことが知られている。この2つの射影子は回帰分析で基準変数ベクトル $\mathbf{y}$ を説明変数行列 $\mathbf{X}$ によって説明出来る部分 $\mathbf{P}_X\mathbf{y}$ と出来ない部分 $\mathbf{Q}_X\mathbf{y}$ に分解するのによく用いられる。柳井先生は多変量解析を変数ベクトルの存在する全空間を射影子によって意味のある部分空間に分けて行く方法であると位置づけておられたが (竹内・柳井,1972; Takeuchi, Yanai, and Mukherjee, 1982)、その基本は上で定義された2つの射影子にあると言っても過言ではない。$\mathbf{P}_X$ と $\mathbf{Q}_X$ を足し合わせると $\mathbf{I}$ となり、これは $\mathbf{I}$ の担う全空間を $\mathbf{P}_X$ と $\mathbf{Q}_X$ が担う2つの部分空間に直交分解することに他ならない。

上の2つの射影子を若干一般化したものとして次のように定義されるK直交射影子がある。

$$\mathbf{P}_{X/K} = \mathbf{X}(\mathbf{X}'\mathbf{K}\mathbf{X})^{-}\mathbf{X}'\mathbf{K}, \tag{3}$$

$$\mathbf{Q}_{X/K} = \mathbf{I} - \mathbf{P}_{X/K} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{K}\mathbf{X})^{-}\mathbf{X}'\mathbf{K}. \tag{4}$$

ここで $\mathbf{K}$ は計量行列と呼ばれる $\mathrm{rank}(\mathbf{KX}) = \mathrm{rank}(\mathbf{X})$ を満たす任意の非負定値行列である。$\mathbf{P}_{X/K}$, $\mathbf{Q}_{X/K}$ には $\mathbf{P}_X$, $\mathbf{Q}_X$ 同様次のような性質がある。K対称 ($(\mathbf{KP}_{X/K})' = \mathbf{KP}_{X/K}$, $(\mathbf{KQ}_{X/K})' = \mathbf{KQ}_{X/K}$)、べき等 ($\mathbf{P}^2_{X/K} = \mathbf{P}_{X/K}$, $\mathbf{Q}^2_{X/K} = \mathbf{Q}_{X/K}$)、K直交 ($\mathbf{P}'_{X/K}\mathbf{KQ}_{X/K} = \mathbf{Q}'_{X/K}\mathbf{KP}_{X/K} = \mathbf{O}$)。K直交射影子は重み付最小2乗法で重要な役割を果たす。

本文で用いる射影子は上記の2種類のみであるが、回帰分析で操作変数推定法を用いた時に有効な斜交射影子については Yanai (1990)、Takane and Yanai (1999) などを参照されたい。

2　制約付き主成分分析

柳井先生が考えていらっしゃった多変量解析には後年 (柳内・竹内、1983; Yanai, Takeuchi, and Takane, 2011) 特異値分解 (Singular Value Decomposition (SVD)) というもう一つの柱が加えられた。SVDはある空間の中でその空間を最もよく表す部分空間を求めるのに便利な計算手法で、主成分分析の基礎となる。制約付き主成分分析 (Takane and Shibayama, 1991; Takane and Hunter, 2001, 2011; Takane, 2013) は射影子を用いてデータ行列をその行と列に関する外部情報によって分解し、分解された個々の行列に主成分分析を適用する方法である。分解された個々の行列には固有の意味があり、それぞれの部分から固有の意味を持つ成分が抽出される。いまデータ行列を $\mathbf{Y}$, $\mathbf{Y}$ の行に関する外部情報（制約）行列を $\mathbf{G}$、列に関する外部制約行列を $\mathbf{H}$ で表すと、これらの外部情報に則した $\mathbf{Y}$ の分解は直交射影子 $\mathbf{P}_G$、$\mathbf{Q}_G$, $\mathbf{P}_H$, $\mathbf{Q}_H$ を用いて次のように書き表せる。

$$\mathbf{Y} = \mathbf{P}_G\mathbf{Y}\mathbf{P}_H + \mathbf{Q}_G\mathbf{Y}\mathbf{P}_H + \mathbf{P}_G\mathbf{Y}\mathbf{Q}_H + \mathbf{Q}_G\mathbf{Y}\mathbf{Q}_H. \tag{5}$$

（あるいはK直交射影子を用いても同様の分解が得られる。）(5) の右辺第1項は $\mathbf{Y}$ のうち $\mathbf{G}$ によっても $\mathbf{H}$ によっても説明出来る変動、第2項は $\mathbf{H}$ によっては説明出来るが $\mathbf{G}$ によっては説明出来ない変動、第3項は逆に $\mathbf{G}$ によっては説明出来るが $\mathbf{H}$ によっては説明出来ない変動、第4項は $\mathbf{G}$ によっても $\mathbf{H}$ によっても説明出来ない変動部分に相当する。制約付き主成分分析ではこの様に分解された右辺の各項に個別に主成分分析を適用し、各項の中で一番重要な変動部分を取り出す手法である。

さらに $\mathbf{G}$ が $\mathbf{G} = [\mathbf{M}, \mathbf{N}]$ のように2つの部分行列から成るときは $\mathbf{P}_G$ をさらに分解することにより、データ行列のより細かな分解が可能である。$\mathbf{P}_G$ の主な分解には次のようなものがある。

1) $\mathbf{P}_G = \mathbf{P}_M + \mathbf{P}_N \Longleftrightarrow \mathbf{M}'\mathbf{N} = \mathbf{O}$.（$\Longleftrightarrow$ は同値関係を示す。）
2) $\mathbf{P}_G = \mathbf{P}_M + \mathbf{P}_{Q_M N} = \mathbf{P}_N + \mathbf{P}_{Q_N M}$.
3) $\mathbf{P}_G = \mathbf{P}_{M/Q_N} + \mathbf{P}_{N/Q_M} \Longleftrightarrow \mathrm{rank}(\mathbf{G}) = \mathrm{rank}(\mathbf{M}) + \mathrm{rank}(\mathbf{N})$.
4) $\mathbf{P}_G = \mathbf{P}_{GA} + \mathbf{P}_{G(G'G)^{-}B} \Longleftrightarrow \mathbf{A}'\mathbf{B} = \mathbf{O}, \mathrm{Sp}(\mathbf{A}) \oplus \mathrm{Sp}(\mathbf{B}) = \mathrm{Sp}(\mathbf{G}')$.

このうち 1) から 3) までは) は Rao and Yanai (1979)、4) は Yanai (1990) による。1) は $\mathbf{M}$ と $\mathbf{N}$ が直交する時、2) は直交しない時、$\mathbf{M}$, $\mathbf{N}$ のどちらか一方を先に当てはめ残りの一方を残差に当てはめる時、3) は直交しない $\mathbf{M}$, $\mathbf{N}$ を同時に当てはめる時に得られる。4) は $\mathbf{G}$ にかかる重み $\mathbf{C}$ に $\mathbf{C} = \mathbf{AC}^{*}$、あるいは $\mathbf{B}'\mathbf{C} = \mathbf{O}$ の制

約がかかる時に得られる分解である。$\mathbf{H}$ が $\mathbf{G}$ と同じような構造を持つとき、$\mathbf{P}_H$ にも同様の分解が成り立つことは明らかであろう。制約付き主成分分析はこうしてさらに細分化されたデータ行列にSVDを適用しその主成分を抽出する。

こうして見ると制約付き主成分分析はまさに柳井先生が多変量解析の柱と考えていらっしゃった2つの技法を機能的に組み合わせた分析法と言うことができる。制約付き主成分分析では柳井先生との直接の協力関係が一度もなかった。にもかかわらず、本文でこの手法を真っ先に取り上げたのはこういう理由による。実は柳井先生と制約付き主成分分析の結びつきは上で述べたよりももっと深く、(5) の右辺の2項と4項を一緒に主成分分析する方法は柳井先生がずっと以前提案された偏主成分分析 (Yanai, 1970) に他ならず、この点でも柳井先生の考えが先行していたことを示している。

3　カトリの補助定理とその拡張

この論文 (Yanai and Takane, 1992) の端緒は 1989 年岡山で開催された行動計量学会からの帰りの新幹線で柳井先生と御一緒したことだった。当時私は上で述べた制約付き主成分分析の特別の場合にあたる制約付き対応分析の手法間の関連に興味を持っていた。文献を調べてみると制約付き対応分析は制約のかけ方によって大きく2種類に分類出来ることがわかった。1つはCCA (Canonical Correspondence Analysis; ter Braak, 1986) に代表されるグループで $\mathbf{U}$ を分割表の行側の表現行列として $\mathbf{A}$ を $\mathbf{U}$ が存在すべき部分空間の基底を表す行列とすると

$$\mathbf{U} = \mathbf{A}\mathbf{U}^* \tag{6}$$

と表せる。ここで $\mathbf{U}^*$ は $\mathbf{A}$ にかかる重み行列を表す。(列側の制約も同様にして考えられるがここでは簡単のため行側の制約のみを考慮する。)もう1つはCALC (Canonical Analysis with Linear Constraints; Böckenholt and Böckenholt, 1990) に代表されるグループで、$\mathbf{U}$ が存在すべき部分空間の直交補空間を指定する。この補空間を表す行列を $\mathbf{B}$ で表すと制約条件は

$$\mathbf{B}'\mathbf{U} = \mathbf{O} \tag{7}$$

と表される。$\mathbf{A}$ と $\mathbf{B}$ が互いに直交補空間を張るとき2つの制約条件は一致すると考えられるが、私はそれを直接証明するのに腐心していたのである。これは

$$\mathbf{P}_A = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' = \mathbf{I} - \mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}' = \mathbf{Q}_B \tag{8}$$

でほぼ自明な関係であるが、対応分析で使う特殊な計量行列の下ではどうなるのかその時はまだよく分かっていなかった。ところが柳井先生との会話の中で上の関係に類した関係式が Khatri (1966) の成長曲線モデルの論文の中で使われているのを思い出した。東京に着いてから詳しく調べて見ると,その定理はカトリの補助定理と呼ばれ

$\mathbf{A}$ $(p \times r)$, $\mathbf{B}$ $(p \times (p-r))$ をそれぞれ階数 r, $p-r$ の $\mathbf{A}'\mathbf{B} = \mathbf{O}$ を満たす行列とすると

$$\mathbf{I} = \mathbf{A}(\mathbf{A}'\mathbf{K}\mathbf{A})^{-1}\mathbf{A}\mathbf{K} + \mathbf{K}^{-1}\mathbf{B}(\mathbf{B}'\mathbf{K}^{-1}\mathbf{B})^{-1}\mathbf{B}' \tag{9}$$

が成り立つというもので、これを用いるとCCAとCALCの関係が直ちに証明できることがわかった。その結果を報告したのが Takane, Yanai, and Mayekawa (1991) の論文である。(この補助定理は (8) に還元することにより簡単に証明できる。また (9) の両辺に前から $\mathbf{K}$ をかけて

$$\mathbf{K} = \mathbf{KA}(\mathbf{A}'\mathbf{KA})^{-1}\mathbf{AK} + \mathbf{B}(\mathbf{B}'\mathbf{K}^{-1}\mathbf{B})^{-1}\mathbf{B}' \tag{10}$$

の形で提示されることもある。なお、$\mathbf{K}$ と $\mathbf{K}^{-1}$ は交換出来ることに注意。)この補助定理は Khatri の論文では成長曲線モデルの最尤解でQ型の射影子をP型に書き換えるのに用いられたが (Seber, 1994, pp. 474-495; Takane and Zhou, 2012)、その後共分散構造分析における漸近理論 (Shapiro, 1986), 一般線形モデルの残差共分散行列の逆行列の表現 (LaMotte,2007; Verbyla,1990) などにも用いられ非常に有用な定理の一つである。

柳井先生は Khatri の補助定理をさらに一歩進めて次のように拡張なさった (Yanai and Takane, 1992)。$\mathbf{A}$ $(p \times r)$, $\mathbf{B}$ $(p \times (p-r))$ をそれぞれ階数 r, $p-r$ の行列、$\mathbf{M}$, $\mathbf{N}$ を非負定値行列とし、次の三つの条件を満たすものと仮定する。

(i) $\mathbf{A}'\mathbf{MNB} = \mathbf{O}$.
(ii) $\mathrm{rank}(\mathbf{MA}) = \mathrm{rank}(\mathbf{A})$.
(iii) $\mathrm{rank}(\mathbf{NB}) = \mathrm{rank}(\mathbf{B})$.

すると次の等式が成り立つ。

$$\mathbf{I} = \mathbf{A}(\mathbf{A}'\mathbf{MA})^{-}\mathbf{A}'\mathbf{M} + \mathbf{NB}(\mathbf{B}'\mathbf{NB})^{-}\mathbf{B}'. \tag{11}$$

証明は Yanai and Takane (1992) を参照のこと。この定理は $\mathbf{M} = \mathbf{K}$, $\mathbf{N} = \mathbf{K}^{-1}$ の時、Khatri の元の定理に一致する。Khatri の補助定理は別の方向にも拡張されていて、例えば Khatri (1990) は $\mathbf{K}$ が必ずしも対称ではなく、正則でもないが、$\mathrm{Sp}(\mathbf{B}) \subset \mathrm{Sp}(\mathbf{K})$, $\mathrm{Sp}(\mathbf{B}) \subset \mathrm{Sp}(\mathbf{K}')$ を満たす正方行列の時、

$$\mathbf{K} = \mathbf{KA}(\mathbf{A}'\mathbf{KA})^{-}\mathbf{AK} + \mathbf{B}(\mathbf{B}'\mathbf{K}^{-}\mathbf{B})^{-}\mathbf{B}' \tag{12}$$

が成り立つことを示した。また Takane and Hunter (2011; Takane, 2013, pp. 140-143; また次節を参照のこと) は $\mathbf{K}$ が矩形行列の場合に拡張している。また Khatri の補助定理の一般化としては前節で述べた 4) もその一つに数えられる。4) を計量行列が $\mathbf{K}$ の場合に拡張すると $\mathbf{P}_{G/K} = \mathbf{P}_{GA/K} + \mathbf{P}_{G(G'KG)^{-}B/K}$ が得られ、ここで $\mathbf{G} = \mathbf{I}$ と置くと Khatri の補助定理が得られる。

4 ウェダーバーン・ガットマン（WG）の定理とその拡張

この定理に関する論文は3部作 (Takane and Yanai, 2005、2007, 2009) になっているが、最初の2つはもともと1つの論文だった。あまり長いので後に分割することになった。3つ目の論文はもともと書く予定ではなかったが、我々の最初の論文が間違っているという論文 (Galantai, 2007) が出たためその反論を書かないわけには行かなくなった。これらの論文のもとになったウェダーバーン・ガットマンの定理とは次のような定理である。$\mathbf{Y}$ を階数 r を持つ $n \times p$ 行列とする。また,$\mathbf{A}$ $(n \times s)$, $\mathbf{B}$ $(p \times s)$ を

$\mathbf{A}'\mathbf{YB}$ が非特異であるような行列とする。すると

$$\begin{aligned}\mathrm{rank}(\mathbf{Y}_1) &= \mathrm{rank}(\mathbf{Y}) - \mathrm{rank}(\mathbf{YB}(\mathbf{A}'\mathbf{YB})^{-1}\mathbf{A}'\mathbf{Y}) \\ &= \mathrm{rank}(\mathbf{Y}) - \mathrm{rank}(\mathbf{A}'\mathbf{YB}) = r - s \end{aligned} \tag{13}$$

が成り立つ。但し

$$\mathbf{Y}_1 = \mathbf{Y} - \mathbf{YB}(\mathbf{A}'\mathbf{YB})^{-1}\mathbf{A}'\mathbf{Y} \tag{14}$$

である。この定理は Wedderburn (1934) がまず $s = 1$ の場合を証明し、その後 Guttman (1944) によって $s > 1$ の場合に拡張された。Guttman (1944) はこの定理を Lagrange の定理と呼び Wedderburn (1934) を引用しているが、後者には Lagrange への引用はなく出典は不明のままである。そこで本文では Hubert 等 (2000) にならい上記の定理をWGの定理と呼ぶのことにする。Guttman (1957) はさらに (13) の等式が成り立つためには $\mathbf{Y}_1$ は (14) の形でなければならないことを示した。

Guttman (1944) は上記の定理をいわゆる行列階数法を用いて証明している。この方法はちょっと興味深い方法なので下に再録することにしよう。この方法は特定の行列にその行列の階数を変えないブロック操作を施してその行列の階数式を導き出し、同じ行列に違うブロック操作を施して別の階数式を導き出す。どちらの操作も元の行列の階数を変えないから二つの階数式は等しくなければならない。Tian はこれを利用して実に多様な階数の公式を生み出している (例えば、Takane, Tian, and Yanai, 2007a, b)。私は長いことこの方法は Khatri (1961) に由来するものと思っていた。それが 1944 年に既に Guttman によって用いられていたというのは実に興味深い。

$$\mathbf{C} = \begin{bmatrix} \mathbf{I}_s & (\mathbf{A}'\mathbf{YB})^{-1}\mathbf{A}'\mathbf{Y} \\ \mathbf{YB} & \mathbf{Y} \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ -\mathbf{YB} & \mathbf{I} \end{bmatrix},$$

$$\mathbf{F} = \begin{bmatrix} \mathbf{I} & -(\mathbf{A}'\mathbf{YB})^{-1}\mathbf{A}'\mathbf{Y} \\ \mathbf{O} & \mathbf{I} \end{bmatrix}$$

と定義すると

$$\mathbf{ECF} = \begin{bmatrix} \mathbf{I}_s & \mathbf{O} \\ \mathbf{O} & \mathbf{Y}_1 \end{bmatrix} \tag{15}$$

が得られる。これより

$$\mathrm{rank}(\mathbf{C}) = s + \mathrm{rank}(\mathbf{Y}_1) \tag{16}$$

が導かれる。一方

$$\mathbf{G} = \begin{bmatrix} \mathbf{I} & -(\mathbf{A}'\mathbf{YB})^{-1} \\ \mathbf{O} & \mathbf{I} \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ -\mathbf{B} & \mathbf{I} \end{bmatrix}$$

と定義すると、

$$\mathbf{GCH} = \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{Y} \end{bmatrix}. \tag{17}$$

従って

$$\mathrm{rank}(\mathbf{C}) = \mathrm{rank}(\mathbf{Y}) \tag{18}$$

が導かれる。(16) と (18) を合わせると (13) が導かれる。なお、Yanai et al. (2011, Excercise 1.5) も参照されたい。

ところで私が最初この定理に興味を持ったのは 1998 年イリノイ大学で開催された Psychometric Society の Meeting での Hubert 等 (Hubert, Meulman, and Heiser, 2000) の発表であった。この発表は Chu, Funderlic, and Golub (1995) を初めとする線形計算の専門家達が心理計量学における Guttman の業績を無視しているのを批判するものであったが、私は発表が終わったところで1つ質問をした。それはWGの定理で $\mathbf{A}'\mathbf{Y}\mathbf{B}$ が特異行列の時、その逆行列を一般化逆行列で置き換えられないかというものであった。私は何となく出来そうな気がしていたが、Hubert は多分出来ないだろうと答えた。いずれもその時点では確証がなく私はその後しばらくこの問題に没頭することになった。結果的には二人とも半分正解で、半分間違っていたことになる。「出来るには出来るが、条件が必要」というのが結論であった。

私は最初この問題を専ら行列の階数の加算性（減算性）の問題としてとらえていた。すなわち

$$\mathrm{rank}(\mathbf{Y} - \mathbf{YB}(\mathbf{A}'\mathbf{YB})^{-}\mathbf{A}'\mathbf{Y}) = \mathrm{rank}(\mathbf{Y}) - \mathrm{rank}(\mathbf{YB}(\mathbf{A}'\mathbf{YB})^{-}\mathbf{A}'\mathbf{Y}) \qquad (19)$$

が成立する条件を見つければ事足りると思い込んでいた。この思い込みの中には

$$\mathrm{rank}(\mathbf{YB}(\mathbf{A}'\mathbf{YB})^{-}\mathbf{A}'\mathbf{Y}) = \mathrm{rank}(\mathbf{A}'\mathbf{YB}) \qquad (20)$$

が常に成り立つという先入観も含まれていた。ところがだいぶ後になって、これが間違いであることに気付かされた。Tian and Styan (2009) が上で述べた行列階数法を用いて

$$\mathrm{rank}(\mathbf{Y} - \mathbf{YB}(\mathbf{A}'\mathbf{YB})^{-}\mathbf{A}'\mathbf{Y}) = \mathrm{rank}(\mathbf{Y}) - \mathrm{rank}(\mathbf{A}'\mathbf{YB}) \qquad (21)$$

は無条件で成り立つことを証明したからである。私は幸いなことにこの論文の preprint version を 2004 年に見せてもらった。そこで (20) には条件が必要なこと、その条件は (19) が成り立つ条件と同じであることが分かった。

私は最初

$$\mathbf{C} = \mathbf{B}(\mathbf{A}'\mathbf{YB})^{-}\mathbf{A}' \qquad (22)$$

とおいた時、$\mathbf{YC}$ がべき等、ないしは $\mathbf{CY}$ がべき等というのがこの条件ではないかと推測していた。ところが十分性はすぐに証明出来たものの必要性がなかなか証明出来なかった。そこで柳井先生に相談することになったが、それでも必要性はなかなか証明出来なかった。そうこうしているうちにどうも私の最初の考えは間違っているのではないかという疑義が生じそれでは反例を作って見てはどうかということになった。すると反例が出来たのである。どうりで必要性が証明できなかった筈である。最初の考えが間違いだとわかると正しい必要十分条件に辿り着くのは比較的簡単だった。その条件は

$$\mathbf{YCYCY} = \mathbf{YCY} \qquad (23)$$

というもので、明らかに $(\mathbf{YC})^2 = \mathbf{YC}$ や $(\mathbf{CY})^2 = \mathbf{CY}$ の条件よりも弱い条件となっている。必要十分条件がわかるとそれと同等の条件の洗い出しにかかった。その中で特に面白いと思ったのは $\mathbf{YCYY}^{-}$ ないしは $\mathbf{Y}^{-}\mathbf{YCY}$ がべき等というものだった。さらに Takane and Yanai (2005) の論文ではもっと強い条件も含めて条件間のヒエラルキーを作り、それらをまとめて示すことにした。実際のデータ解析場面では多くの場合十分条件だけで十分だからである。ところが幸か不幸かこのことが第3の論文を書く伏線になった。

(19) とそれが成り立つ必要十分条件をまとめて一般化WG定理と呼ぶ。Takane and Yanai (2007) は一般化WGの定理にまつわる条件を product singular value decomposition (PSVD; これは行列積のSVDを当該の行列の積を取らず個々の行列の分解から求める方法である)を用いて表現したものである。ところがこの論文の審査中に Takane and Yanai (2005) の結果は間違いであるという論文が出た。Galantai (2007) がそれである。Galantai はおそらくこの2つ目の論文を査読して我々の最初の論文を知ったのではないかと思われるが、一般化WG定理が成立する必要十分条件は

$$\mathbf{CYC} = \mathbf{C} \tag{24}$$

($\mathbf{C}$ は $\mathbf{Y}$ の反射型一般化逆行列)であるとする Cline and Funderlic (1979) の結果を鵜呑みにし、我々が Takane and Yanai (2005) で述べた十分条件の多くが必要条件でもあると主張した。明らかに (24) が成り立てば $\mathbf{YC}$ や $\mathbf{CY}$ のべき等性も、(23) の条件も成り立つが、逆は成り立たない。そこで Takane and Yanai (2009) では反例を掲げて Galantai (2007) の間違いを指摘した。実は我々は Takane and Yanai (2005) を書くとき既に Cline and Funderlic (1979) の誤りに気が付いていたのであるが、はっきり誤りであるとは書かなかった。彼らが我々の論文の査読者になる可能性もあったからである。

WGの定理は $\mathbf{Y}$ の分解にまつわる行列の階数の関係を表した公式である。しかし、この定理の前提となる $\mathbf{Y}$ の分解そのものも結構興味深い。すなわち

$$\mathbf{Y} = \mathbf{YB}(\mathbf{A}'\mathbf{YB})^-\mathbf{A}'\mathbf{Y} + (\mathbf{Y} - \mathbf{YB}(\mathbf{A}'\mathbf{YB})^-\mathbf{A}'\mathbf{Y}) \tag{25}$$

をデータ行列 $\mathbf{Y}$ の分解方式を示すものと考える。するといろいろ面白い分解が導出できる。事実、Takane and Hunter (2011) はこの分解にのっとって新しい制約付き主成分分析の方法を提案している。この分解の第2項目はQ型の射影子を用いているが、これをP型の射影子に書き直すことが出来る。$\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$ を次の条件を満たす行列とする。$\mathrm{Sp}(\tilde{\mathbf{A}}) \subset \mathrm{Sp}(\mathbf{Y})$, $\mathrm{Sp}(\tilde{\mathbf{B}}) \subset \mathrm{Sp}(\mathbf{Y}')$,

$$\mathrm{rank}(\mathbf{A}'\mathbf{YB}) + \mathrm{rank}(\tilde{\mathbf{B}}'\mathbf{Y}^-\tilde{\mathbf{A}}) = \mathrm{rank}(\mathbf{Y}), \tag{26}$$

$$\mathbf{A}'\mathbf{YY}^-\tilde{\mathbf{A}} = \mathbf{A}'\tilde{\mathbf{A}} = \mathbf{O}, \tag{27}$$

$$\tilde{\mathbf{B}}'\mathbf{Y}^-\mathbf{YB} = \tilde{\mathbf{B}}'\mathbf{B} = \mathbf{O}. \tag{28}$$

すると、(25) は次のように書き換えられる。

$$\mathbf{Y} = \mathbf{YB}(\mathbf{A}'\mathbf{YB})^-\mathbf{A}'\mathbf{Y} + \tilde{\mathbf{A}}(\tilde{\mathbf{B}}'\mathbf{Y}^-\tilde{\mathbf{A}})^-\tilde{\mathbf{B}}'. \tag{29}$$

証明は Takane (2013, pp. 141-142) を参照されたい。(29) はカトリの定理の矩形版ということも出来る。

5　リッジ演算子

柳井先生とリッジ演算子についての論文を書くきっかけは 2000 年代の中頃、私が多変量解析における正則化 (regularization) に興味を持ったせいである。このうち多変量解析手法に実際に正則化法を組み入れ、それを評価する論文は柳井先生を

煩わせるまでもなく、私は大部分自分の学生と一緒に書いた。ところがこれらの論文を書いているうちに共通したテーマであるリッジ演算子について一つまとまった論文を書いた方が良いのではないかと思うようになった。それは数学的にもきちんとした論文が望ましいということで柳井先生にご協力をお願いした (Takane and Yanai, 2008; Takane 2008)。この節ではそのリッジ演算子について述べることにしよう。

$\mathbf{X}$ を $n \times p$ の行列とする。この時、

$$\mathbf{R}_X(\lambda) = \mathbf{X}(\mathbf{X}'\mathbf{X} + \lambda\mathbf{P}_{X'})^{-}\mathbf{X}' \tag{30}$$

をリッジ演算子と呼ぶ。ここで $\mathbf{P}_{X'} = \mathbf{X}'(\mathbf{X}\mathbf{X}')^{-}\mathbf{X}$ は $\mathrm{Sp}(\mathbf{X}')$ への直交射影行列（$\mathbf{X}$ が非特異であれば、$\mathbf{P}_{X'} = \mathbf{I}$）、$\lambda$ (≥ 0) はリッジ・パラメータと呼ばれる。この形の行列は線形回帰分析モデル $\mathbf{y} = \mathbf{X}\mathbf{c} + \mathbf{e}$ で回帰係数ベクトル $\mathbf{c}$ をリッジ最少2乗基準

$$\phi_\lambda(\mathbf{c}) = \mathrm{SS}(\mathbf{y} - \mathbf{X}\mathbf{c}) + \lambda\mathrm{SS}(\mathbf{c})_{P_{X'}} \tag{31}$$

を最小化するよう定める時に得られる。ここで $\mathrm{SS}(\mathbf{y} - \mathbf{X}\mathbf{c}) = (\mathbf{y} - \mathbf{X}\mathbf{c})'(\mathbf{y} - \mathbf{X}\mathbf{c})$, $\mathrm{SS}(\mathbf{c})_{P_{X'}} = \mathbf{c}'\mathbf{P}_{X'}\mathbf{c} = \mathbf{c}'\mathbf{c}$ を表す。（ここでは一般性を失うことなく $\mathrm{Sp}(\mathbf{c}) \subset \mathrm{Sp}(\mathbf{X}')$ であると仮定する。）また $\mathbf{R}_X(\lambda)$ は $\lambda = 0$ の時 $\mathbf{P}_X = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'$, すなわち $\mathrm{Sp}(\mathbf{X})$ への直交射影子に還元する。次にリッジ計量行列を次のように定義する。

$$\mathbf{K}_X(\lambda) = \mathbf{P}_X + \lambda(\mathbf{X}\mathbf{X}')^{+}. \tag{32}$$

ここで $^{+}$ はムーア・ペンローズ逆行列を表し、$(\mathbf{X}\mathbf{X}')^{+} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{+2}\mathbf{X}'$ が成り立つ。但し $(\mathbf{X}'\mathbf{X})^{+2} = ((\mathbf{X}'\mathbf{X})^{+})^2$ であることに注意。すると $\mathbf{R}_X(\lambda)$ はリッジ計量行列を用いて次のように書き換えることができる。

$$\mathbf{R}_X(\lambda) = \mathbf{X}(\mathbf{X}'\mathbf{K}_X(\lambda)\mathbf{X})^{-}\mathbf{X}'. \tag{33}$$

次に

$$\mathbf{S}_X(\lambda) = \mathbf{I} - \mathbf{R}_X(\lambda) \tag{34}$$

と定義すると、$\mathbf{R}_X(\lambda)$, $\mathbf{S}_X(\lambda)$ は直交射影子 $\mathbf{P}_X$, $\mathbf{Q}_X = \mathbf{I} - \mathbf{P}_X$ と似たような性質を持つことが分かる。主なものとしては

$$\mathbf{R}_X(\lambda)\mathbf{K}_X(\lambda)\mathbf{R}_X(\lambda) = \mathbf{R}_X(\lambda), \tag{35}$$

$$\mathbf{R}_X(\lambda) - \mathbf{R}_X(\lambda)^2 = \mathbf{R}_X(\lambda)\mathbf{S}_X(\lambda) \geq \mathbf{O}, \tag{36}$$

$$\mathbf{R}_X(\lambda)\mathbf{K}_X(\lambda) = \mathbf{P}_X \tag{37}$$

などが掲げられる。さらに $\mathbf{X} = [\mathbf{M}, \mathbf{N}]$ とすると第2節で示した $\mathbf{P}_G$ の分解と同様の分解が得られる。

1) $\mathbf{R}_X(\lambda) = \mathbf{R}_M(\lambda) + \mathbf{R}_N(\lambda) \Longleftrightarrow \mathbf{M}'\mathbf{N} = \mathbf{O}$.
2) $\mathbf{R}_X(\lambda) = \mathbf{R}_M(\lambda) + \mathbf{R}_{S_M(\lambda)N}(\lambda) = \mathbf{R}_N(\lambda) + \mathbf{R}_{S_N(\lambda)M}(\lambda)$. ここで $\mathbf{R}_{S_M(\lambda)N}(\lambda) = \mathbf{S}_M(\lambda)\mathbf{N}(\mathbf{N}'\mathbf{S}_M(\lambda)\mathbf{N} + \lambda\mathbf{P}_{N'})^{-}\mathbf{N}'\mathbf{S}_M(\lambda)$, $\mathbf{R}_{S_N(\lambda)M}(\lambda) = \mathbf{S}_N(\lambda)\mathbf{M}(\mathbf{M}'\mathbf{S}_N(\lambda)\mathbf{M} + \lambda\mathbf{P}_{M'})^{-}\mathbf{M}'\mathbf{S}_N(\lambda)$.
3) $\mathbf{R}_X(\lambda) = \mathbf{R}_{M/S_N(\lambda)}(\lambda) + \mathbf{R}_{N/S_M(\lambda)}(\lambda) \Longleftrightarrow \mathrm{rank}(\mathbf{X}) = \mathrm{rank}(\mathbf{M}) + \mathrm{rank}(\mathbf{N})$. ここで $\mathbf{R}_{M/S_N(\lambda)}(\lambda) = \mathbf{M}(\mathbf{M}'\mathbf{S}_N(\lambda)\mathbf{M} + \lambda\mathbf{P}_{M'})^{-}\mathbf{M}'\mathbf{S}_N(\lambda)$, $\mathbf{R}_{N/S_M(\lambda)}(\lambda) = \mathbf{N}(\mathbf{N}'\mathbf{S}_M(\lambda)\mathbf{N} + \lambda\mathbf{P}_{N'})^{-}\mathbf{N}'\mathbf{S}_M(\lambda)$.

4) $\mathbf{R}_X(\lambda) = \mathbf{R}_{XA}(\lambda) + \mathbf{R}_{X(X'K(\lambda)X)^- B}(\lambda)$. ここで $\mathbf{A}'\mathbf{B} = \mathbf{O}$, $\mathrm{Sp}(\mathbf{A}) \oplus \mathrm{Sp}(\mathbf{B}) = \mathrm{Sp}(\mathbf{X}')$.

これらの関係式の証明には柳井先生による $\mathbf{P}_G$ の分解(第2節 1) から 4))の証明が大いに役に立った。

これらの結果を用いると従来最小2乗法によって行われていた冗長性分析 (Redundancy Analysis) のパラメータ推定をほぼそのままの形でリッジ最小2乗法でも行えることが分かった。冗長性分析は多変量回帰分析で回帰係数行列 $\mathbf{C}$ に $\mathrm{rank}(\mathbf{C}) = r$ のような階数制約が課される場合に相当するが、最小2乗解ではまず制約なしの解を求め、それにSVDを適用することによって階数制約下の解が求められるが、全く同じことがリッジ最小2乗法でも成り立つ。これは通常の冗長性分析に限らず偏冗長性分析、制約付き冗長性分析の場合でも成り立つ。より詳しくは Takane and Jung (2008, 2009) を参照されたい。

なお上記のリッジ演算子、リッジ計量行列は次のように拡張出来る (Takane, 2008)。$\mathbf{L}$, $\mathbf{W}$ を $\mathrm{Sp}(\mathbf{L}) \subset \mathrm{Sp}(\mathbf{X}')$, $\mathrm{rank}(\mathbf{W}\mathbf{X}) = \mathrm{rank}(\mathbf{X})$ を満たす任意の非負定値行列として

$$\mathbf{R}_X^{(W,L)}(\lambda) = \mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X} + \lambda\mathbf{L})^- \mathbf{X}'\mathbf{W}. \tag{38}$$

これを一般化リッジ演算子と呼ぶ。

$$\mathbf{K}_X^{(W,L)}(\lambda) = \mathbf{P}_X + \lambda\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^- \mathbf{L}(\mathbf{X}'\mathbf{W}\mathbf{X})^- \mathbf{X}'\mathbf{W} \tag{39}$$

と定義すると、一般化リッジ演算子は次のように書き換えられる。

$$\mathbf{R}_X^{(W,L)}(\lambda) = \mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{K}_X^{(W,L)}(\lambda)\mathbf{X})^- \mathbf{X}'\mathbf{W}. \tag{40}$$

一般化リッジ演算子の場合も、通常のリッジ演算子と同様の性質が成り立つことが知られている。

6 制約付き正準相関分析

制約付き主成分分析は射影子を用いてデータ行列を直交分解し、得られた個々の行列を主成分分析するものであった。従ってまずはデータ行列の分解が第一の課題である。それに対し正準相関分析では2つのデータ行列間の関係を分析する。ところが私は最初単純に制約付き正準相関分析でも2つのデータ行列を別々に直交分解し、それぞれの分解から興味のある部分データを一つずつ選んでその間の正準相関分析を行えばよいと思っていた (Takane and Hwang, 2002)。ところがこれだと全ての可能な対を作って関連度を求めそれらを合計しても全体の関連度とは一致しない。2つのデータ行列 $\mathbf{X}$, $\mathbf{Y}$ 間の関連度は $\mathrm{tr}(\mathbf{P}_X\mathbf{P}_Y)$ で定義されるが、$\mathbf{X}$, $\mathbf{Y}$ の直交分解によっては $\mathrm{tr}(\mathbf{P}_X\mathbf{P}_Y)$ の加法分解が得られないのである。これは $\mathbf{X} = \mathbf{M} + \mathbf{N}$、$\mathbf{M}'\mathbf{N} = \mathbf{O}$ であっても、$\mathbf{P}_X = \mathbf{P}_M + \mathbf{P}_N$ が成り立つとは限らないことからも明らかである。(これによく似た状況として $\mathbf{X} = [\mathbf{M}, \mathbf{N}]$, $\mathbf{M}'\mathbf{N} = \mathbf{O}$ がある。この場合は確かに $\mathbf{P}_X = \mathbf{P}_M + \mathbf{P}_N$ が成り立つ。) 従って $\mathrm{tr}(\mathbf{P}_X\mathbf{P}_Y)$ の加法分解にはデータ行列の直交分解でなく、射影子 $\mathbf{P}_X$, $\mathbf{P}_Y$ の直交分解が必要になる。それを実現したのが Takane, Yanai, and Hwang (2006) の論文である。この論文では $\mathbf{G}$, $\mathbf{H}$ をそれぞれ $\mathbf{X}$ の行側、列側の制約行列として、第2節で述べた射影子の直交分解のうち 2) と

4) を組み合わせて次のような2つの直交分解を導き出した。

1. 行列 $\mathbf{A}, \mathbf{B}, \mathbf{W}$ をそれぞれ $\mathrm{Sp}(\mathbf{A}) = \mathrm{Ker}(\mathbf{H}'\mathbf{X}'\mathbf{P}_G\mathbf{X})$, $\mathrm{Sp}(\mathbf{B}) = \mathrm{Ker}(\mathbf{H}'\mathbf{X}'\mathbf{Q}_G\mathbf{X})$, $\mathrm{Sp}(\mathbf{W}) = \mathrm{Ker}(\mathbf{X}'\mathbf{G})$ を満たす行列とする。すると次の直交分解が成り立つ。

$$\mathbf{P}_{[X,G]} = \mathbf{P}_{P_G XH} + \mathbf{P}_{P_G XA} + \mathbf{P}_{Q_G XH} + \mathbf{P}_{Q_G XB} + \mathbf{P}_{GW}. \tag{41}$$

2. 行列 $\mathbf{K}$, $\mathbf{U}$, $\mathbf{V}$ をそれぞれ $\mathrm{Sp}(\mathbf{K}) = \mathrm{Ker}(\mathbf{H}'\mathbf{X}'\mathbf{X})$, $\mathrm{Sp}(\mathbf{U}) = \mathrm{Ker}(\mathbf{G}'\mathbf{X}\mathbf{H})$, $\mathrm{Sp}(\mathbf{V}) = \mathrm{Ker}(\mathbf{G}'\mathbf{X}\mathbf{K})$ を満たす行列とする。すると次の直交分解が成り立つ。

$$\mathbf{P}_{[X,G]} = \mathbf{P}_{P_{XH}G} + \mathbf{P}_{XHU} + \mathbf{P}_{P_{XK}G} + \mathbf{P}_{XKV} + \mathbf{P}_{Q_X G}. \tag{42}$$

$\mathbf{Y}$ にも同様の分解が考えられるから、両方の分解を組み合わせ多様な制約付き正準相関分析が可能となる。なお正準相関分析そのものは $\mathbf{X}$ 側の射影子と $\mathbf{Y}$ 側の射影子の積にSVDを適用することによって成される。

7 エピローグ

以上柳井先生との共同研究の足跡を辿ってきた。先生の構想力は壮大で私はいまだその枠外に一歩も出られないでいる。私は若干不甲斐なさを感じつつも、いつか breakthrough が来る日を信じてこれまで来た道を歩んでいくつもりである。道はまだ遠く、前進しなければならない。最後に先生がもっと長生きしていらっしゃったらもう一つの共著論文になったかも知れない着想を述べて本文の締めくくりとしたい。

ものごとの原因を見定めるのは統計学の大きな課題の一つである。よく知られているように無作為化が不可能な時、変量の相関関係のみから原因を探るのには数々の落とし穴がある。その第一は交絡変数 (confounding variables) の問題である。$\mathbf{x}$ から $\mathbf{y}$ を予測する時に $\mathbf{x}$ 以外の変数、いわゆる交絡変数 $\mathbf{U}$ の影響をどのように排除したらよいかという問題である。これは $\mathbf{x}$ から $\mathbf{y}$ を予測する回帰式に説明変数として $\mathbf{U}$ を含めればよい。この回帰式を

$$\mathbf{y} = \mathbf{x}a_1 + \mathbf{U}\mathbf{c} + \mathbf{e}_1 \tag{43}$$

とすると $\mathbf{x}a_1$ の最小2乗解は

$$\mathbf{x}\hat{a}_1 = \mathbf{P}_{x/Q_u}\mathbf{y} \tag{44}$$

で与えられる。一方、$\mathbf{x}$ の $\mathbf{U}$ に対する回帰分析を考えると

$$\mathbf{x} = \mathbf{U}\mathbf{d} + \mathbf{e}_2. \tag{45}$$

$\mathbf{U}\mathbf{d}$ の最小2乗解は

$$\mathbf{U}\hat{\mathbf{d}} = \mathbf{P}_U\mathbf{x} \tag{46}$$

で与えられる。$\mathbf{P}_U\mathbf{x}$ を線形傾向得点と呼ぶことにしよう。また (45) からの残差ベクトルは $\mathbf{Q}_U\mathbf{x}$ で与えられる。このベクトルは $\mathbf{x}$ のうち $\mathbf{U}$ によっては説明出来ない部分の変動を表す。いま (43) の回帰式で $\mathbf{U}$ の代わりに $\mathbf{P}_U\mathbf{x}$ を用いると

$$\mathbf{y} = \mathbf{x}a_2 + \mathbf{P}_U\mathbf{x}b + \mathbf{e}_3 \tag{47}$$

が得られる。この回帰式の $\mathbf{x}a_2$ の最小2乗解は

$$\mathbf{x}\hat{a}_2 = \mathbf{P}_{x/Q_{P_U x}}\mathbf{y} \tag{48}$$

で与えられる。ここで $\mathbf{Q}_{P_U x} = \mathbf{I} - \mathbf{P}_U\mathbf{x}(\mathbf{x}'\mathbf{P}_U\mathbf{x})^{-1}\mathbf{x}'\mathbf{P}_U$ である。従って、

$$\mathbf{Q}_{P_U x}\mathbf{x} = \mathbf{x} - \mathbf{P}_U\mathbf{x}(\mathbf{x}'\mathbf{P}_U\mathbf{x})^{-1}\mathbf{x}'\mathbf{P}_U\mathbf{x} = \mathbf{Q}_U\mathbf{x}. \tag{49}$$

これより

$$\mathbf{P}_{x/Q_{P_U x}}\mathbf{y} = \mathbf{P}_{x/Q_U}\mathbf{y} \tag{50}$$

が得られ、(48) は (44) と一致する。これが、直接 $\mathbf{U}$ を使わなくとも、傾向得点で条件づければ $\mathbf{U}$ の影響を排除できることの根拠である。しかも $\mathbf{P}_U\mathbf{x}$ は一変数であるためその得点上での等化がし易いという利点がある。

ところが最近次のような性質を持つ操作変数 $\mathbf{z}$ によって因果関係を探る方法が盛んになってきている。

1) $\mathbf{z}'\mathbf{U} = \mathbf{0}$（$\mathbf{z}$ と $\mathbf{U}$ は無関係）.
2) $\mathbf{z}'\mathbf{x} \neq 0$（$\mathbf{z}$ と予測変数 $\mathbf{x}$ には相関がある）.
3) $\mathbf{z}'\mathbf{Q}_{[U,x]}\mathbf{y} = 0$（$\mathbf{z}$ は $\mathbf{x}$ を通してのみ $\mathbf{y}$ への予測力がある）.

操作変数 $\mathbf{z}$ は交絡変数 $\mathbf{U}$、あるいは線形傾向得点 $\mathbf{P}_U\mathbf{x}$ と一体どのような関係にあるのだろうか。いま仮に

$$\mathbf{z} = c\mathbf{Q}_U\mathbf{x} \tag{51}$$

と置くと、1), 2) が成り立つ事はすぐ分かる。ここで c は規準化の定数である。(51) で定義される $\mathbf{z}$ が 3) を満たすことは次のようにして示される。

$$\begin{aligned}
(1/c)\mathbf{z}'\mathbf{Q}_{[U,x]}\mathbf{y} &= \mathbf{x}'\mathbf{Q}_U\mathbf{Q}_{[U,x]}\mathbf{y} \\
&= \mathbf{x}'\mathbf{Q}_{[U,x]}\mathbf{y} \\
&= \mathbf{x}'\mathbf{Q}_U\mathbf{Q}_{x/Q_U}\mathbf{y} \\
&= \mathbf{x}'\mathbf{Q}_U(\mathbf{I} - \mathbf{P}_{x/Q_U})\mathbf{y} \\
&= \mathbf{x}'\mathbf{Q}_U\mathbf{y} - \mathbf{x}'\mathbf{Q}_U\mathbf{x}(\mathbf{x}'\mathbf{Q}_U\mathbf{x})^{-1}\mathbf{x}'\mathbf{Q}_U\mathbf{y} = 0.
\end{aligned} \tag{52}$$

（$\mathbf{Q}_{[U,x]} = \mathbf{Q}_U\mathbf{Q}_{x/Q_U}$ であることに注意 (Takane and Zhou, 2012)。）さらに $\mathbf{y} = \mathbf{x}a_3 + \mathbf{e}_4$ の操作変数推定量は

$$\mathbf{x}\hat{a}_3 = \mathbf{P}_{x/P_z}\mathbf{y} = \mathbf{P}_{x/Q_U}\mathbf{y} \tag{53}$$

で与えられ、これは $\mathbf{P}_z = \mathbf{Q}_U\mathbf{x}(\mathbf{x}'\mathbf{Q}_U\mathbf{x})^{-1}\mathbf{x}'\mathbf{Q}_U$, $\mathbf{x}'\mathbf{P}_z = \mathbf{x}'\mathbf{Q}_U$ だから (44), (48) に一致する。逆に $\mathbf{Q}_z = \mathbf{I} - \mathbf{P}_z = \mathbf{U}\mathbf{U}'$（但し $\mathbf{U}'\mathbf{U} = \mathbf{I}$）と置くと、$\mathbf{P}_U\mathbf{x} = \mathbf{Q}_z\mathbf{x}$ が成り立つ。従って (51) がある意味で理想的な操作変数であると考えられる。

共著論文、単行本

肥田野直・柳井晴夫・塗師斌・繁桝算男・高根芳雄　(1971). 検査の尺度構成に関する方法論的研究. 教育心理学研究、19 巻 1 号、37-51.

Takane, Y., Tian, Y., and Yanai, H. (2007a). On reverse order laws for least-squares g-inverses and minimum norm g-inverses of a matrix. *Aequationes Mathematicae*, **73**, 56-70.

Takane, Y., Tian, Y., and Yanai, H. (2007b). On constrained g-inverses of matrices and their properties. *Annals of the Institute of Statistical Mathematics*, **59**, 807-820.

Takane, Y., and Yanai, H. (1999). On oblique projectors. *Linear Algebra and Its Applications*, **289**, 297-310.

Takane, Y., and Yanai, H. (2005). On the Wedderburn-Guttman theorem. *Linear Algebra and Its Applications*, **410**, 267-278.

Takane, Y., and Yanai, H. (2007). Alternative characterizations of the extended Wedderburn-Guttman theorem. *Linear Algebra and Its Applications*, **422**, 701-711.

Takane, Y., and Yanai, H. (2008). On ridge operators. *Linear Algebra and Its Applications*, **428**, 1778-1790.

Takane, Y., and Yanai, H. (2009) On the necessary and sufficient condition for the extended Wedderburn-Guttman theorem. *Linear Algebra and Its Applications*, **430**, 2890-2895.

Takane, Y., Yanai, H., and Hwang, H. (2006). An improved method for generalized constrained canonical correlation analysis. *Computational Statistics and Data Analysis*, **50**, 221-241.

Takane, Y., Yanai, H., and Mayekawa, S. (1991). Relationships among several methods of linearly constrained correspondence analysis. *Psychometrika*, **56**, 667-684.

柳井晴夫・高根芳雄 (1976). 多変量解析法.　東京:朝倉書店、新版 (1985).

Yanai, H., and Takane, Y. (1992). Canonical correlation analysis with linear constraints. *Linear Algebra and Its Applications*, **176**, 75-82.

Yanai, H., and Takane, Y. (2007). Matrix methods and their applications to factor analysis. In S. Y. Lee (Ed.), *Handbook of latent variable and related models*, (pp. 345-366). Amsterdam: Elsevier.

Yanai, H., Takane, Y., and Ishii, H. (2006). Non-negative determinant of a rectangular matrix: Its definition and applications to multivariate analysis. *Linear Algebra and Its Applications*, **417**, 259-274.

Yanai, H., Takeuchi, K., Takane, Y. (2011). *Projection matrices, generalized inverse matrices, and singular value decomposition.* New York: Springer.

その他の引用文献

Böckenholt, U., and Böckenholt, I. (1990). Canonical analysis of contingency tables with linear constraints. *Psychometrika*, **55**, 633-639.

Chu, M. T., Funderlic, R. E., and Golub, G. H. (1995). A rank-one reduction formula and its applications to matrix factorizations. *SIAM Review*, **37**, 512-530.

Cline, R. E., and Funderlic, R. E. (1979). The rank of a difference of matrices and associated generalized inverses. *Linear Algebra and its Applications*, **24**, 185-215.

Galantai, A. (2007). A note on generalized rank reduction. *Acta Mathematica Hungarica*, **116**, 239-246.

Guttman, L. (1944). General theory and methos for matric factoring. *Psychometrika*, **9**, 1-16.

Guttman, L. (1957). A necessary and sufficient formula for matric factoring. *Psychometrika*, **22**, 79-81.

Hubert, L., Meulman, J., and Heiser, W. J. (2000). Two purposes of matrix factorization: A historical appraisal. *SIAM Review*, **42**, 68-82.

Khatri, C. G. (1961). A simplified approach to the derivation of the theorems on the rank of a matrix. *Journal of the Maharaja Sayajirao University of Baroda*, **10**, 1-5.

Khatri, C. G. (1966). A note on a MANOVA model applied to problems in growth curve. *Annals of the Institute of Statistical Mathematics*, **18**, 75-86.

Khatri, C. G. (1990). Some properties of BLUE in a linear model and canonical correlations associated with linear transformations. *Journal of Multivariate Analysis*, **34**, 211-226.

LaMotte, L. R. (2007). A direct derivation of the REML likelihood function. *Statistical Papers*, **48**, 321-327.

Rao, C. R., and Yanai, H. (1979). General definition and decomposition of projectors and some applications to statistical problems. *Journal of Statistical Planning and Inference*, **3**, 1-17.

Seber, G. F. A. (1984). *Multivariate observations.* New York: Wiley.

Shapiro, A. (1986). Asymptotic theory of overparameterized structural models. *Journal of the American Statistical Association*, **81**, 142-149.

高根芳雄 (1995). 制約付き主成分分析法. 東京：朝倉書店。

Takane, Y. (2008). More on regularization and (generalized) ridge operators. In K. Shigemasu, A. Okada, T. Imaizumi, and T. Hoshino (Eds.), *New trends in psychometrics*, (pp. 443-452). Tokyo: University Academic Press.

Takane, Y. (2013). *Constrained principal component analysis and related techniques.* Boca Raton, FL: Chapman and Hall/CRC Press.

Takane, Y., and Hunter, M. A. (2001). Constrained principal component analysis: A comprehensive theory. *Applicable Algebra in Engineering, Communication, and Computing*, **12**, 391-419.

Takane, Y., and Hunter, M. A. (2011). New family of constrained principal component analysis (CPCA). *Linear Algebra and Its Applications*, **434**, 2539-2555.

Takane, Y., and Hwang, H. (2002). Generalized constrained canonical correlation analysis. *Multivariate Behavioral Research*, **37**, 163-195.

Takane, Y., and Jung, S. (2008). Regularized partial and/or constrained redundancy analysis. *Psychometrika*, **73**, 671-690.

Takane, Y., and Jung, S. (2009). Regularized nonsymmetric correspondence analysis. *Computational Statistics and Data Analysis*, **53**, 3159-3170.

Takane, Y., and Shibayama, T. (1991). Principal component analysis with external information on both subjects and variables. *Psyhometrika*, **56**, 97-120.

Takane, Y., and Zhou, L. (2012). On two expressions of the MLE for a special case of the extended growth curve models. *Linear Algebra and Its Applications*, **436**, 2567-2577.

竹内啓・柳井晴夫 (1972). 多変量解析の基礎ー線形空間への射影による方法ー. 東京：東洋経済新報.

Takeuchi, K., Yanai, H., and Mukherjee, B. N. (1982). *The foundation of multivariate analysis.* New Delhi: Wiley Eastern and New York: Halsted Press.

ter Braak, C. J. F. (1986). Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis. *Ecology*, **67**, 1167-1179.

Tian, Y., and Styan, G. P. H. (2009). On some matrix equalities for generalized inverese with applications. *Linear Algebra and Its Applications*, **430**, 2716-2733.

Verbyla, A. P. (1990). A conditional derivation of residual maximum likelihood. *Australian Journal of Statistics*, **32**, 227-230.

Wedderburn, J. H. M. (1934). *Lectures on matrices.* Colloquium Publication, Vol. 17, Providence: American Mathematical Society.

Yanai, H. (1970). Factor analysis with external criteria. *Japanese Psychological Research*, **12**, 143-153.

Yanai, H. (1990). Some generalized forms of least squares g-inverse, minimum norm g-inverse and Moore-Penrose inverse matrices. *Computational Statistics and Data Analysis*, **10**, 251-260.

柳井晴夫・竹内啓(1983). 射影行列・一般逆行列・特異値分解. 東京:東大出版会.