Multicollinearityと因果関係 : 賃銀計算における O.R. (水谷一雄博士記念號)

後尾, 哲也

(Citation)
国民経済雑誌,104(2):73-89

(Issue Date)
1961-08

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(JaLCDOI)
https://doi.org/10.24546/00167751

(URL)
https://hdl.handle.net/20.500.14094/00167751

Multicollinearityと因果関係

——賃銀計算におけるO. R.——

後　尾　哲　也

I　序

我々経済統計学者にとって Multicollinearity は1つの悪魔である。経済統計学者や社会統計学者にとって，折角苦労を重ねた計算の結果がこの小さな悪魔のために一朝にして崩れさることも少くない。しかし，Multicollinearityという難問は果して我々経済統計学者の乗りこえ難い困難だろうか。

Multicollinearityという言葉が経済統計学上の問題となったのは1930年代初期のことに属する。ラグナー・フリシュ（R. Frisch）が econometrics における古典的最小自乗法の無批判的適用に批判の矢を射て以来，既に30年近くの歳月を経た。この間に econometric method は飛躍的な進歩を示し，Tinbergen-Wold の recursive analysis や Haavelmo-Koopmans の linear shock model analysis を生み出したことは econometrician ならば，何人たりとも既に知るところである。しかし，Multicollinearityの問題は部分的にはこれらの方法によって解決されたとは云え，安全なる解決に到達する道は尚遠く嶮しいものがある。

このように Multicollinearity は我々 econometrician の前に立ちはだかる1つの大きな障害物ではあろうが，そこまで恐れねばならないものでもないように思われる。我々が計算しようとする経済変数とその函数関係の意味を充分に考察して，適当な手段をとれば，避けえられる場合も少なくないのである。経済変数に機械的に最小自乗法を適用することは極めて危険であり，警戒を要することであるにしても，適当な方法によって避けうる場合も少なくない。本

論文の目的はO・R・の1例として俸給計算をした場合に Multicollinearity を避けえた実例を示すとともに，Multicollinearity を避ける1つの方法を述べたいと考えている。勿論，これから述べるような方法が決して万能薬であるというわけではない。ある条件をみたせば Multicollinearity を避ける方法があることを示すに過ぎないものである。Multicollinearity は Frisch の当時より今日に到るまで，未だ今日の問題であり，Multicollinearity に対する万能薬は未だ発見されてはいない。我々，econometrician は Multicollinearity に対し万全の策を講じねばならないとともに，方法論的にもそれに対処する方法を将来も探し求めねばならないであろう。

I　Multicollinearity の定義と意味

経済統計学や econometrics で Multicollinearity を問題としなければならない原因は申すまでもなく，passive nature of economic observation にある。経済観測のこの性質は Multicollinearity という問題を引き起す原因となったばかりか，Haavelmo-Koopmans の linear shock model analysis への道を econometric method がたどらねばならなかった原因でもあった。それは自然科学における統計的方法にみられない社会科学の統計的方法の特殊性であり，これまた社会科学の統計的方法特有の困難をひきおこすものなのである。我々はふり返ってフリッシュの Multicollinearity の定義から説明してゆくこととしよう。

いま，経済変数 x_1, x_2, ……, x_n があるとする。この経済変数が2つ以上の線型関係をもつとき，経済変数 x_1, x_2, ……, x_n は Multicollinear であるという。これがフリッシュに従った経済変数の Multicollinearの定義である。この Multicollinearity の定義に従って Multicollinear な経済変数間の関係を単一の一次式によって推定すれば，non-sense な統計的結果をうるという理由は比較的に簡単なことである。しかし，それにも拘らず，Multicollinear な経済変数に古典的最小自乗法を無批判的に適用すれば，non-sense な結果をうるとい

うことを発見したのは R. Frisch の偉大な功績である。更にこの命題の基礎が今日の econometric method の発展の原動力となったことを思えば尚更である。

しかし，ひるがえって考えてみるに，R. Frisch の Multicollinearity に関する命題は幾分その基礎にあいまいな点がみうけられる。というのは R. Frisch は probability approach を採用しなかったからである。近代的統計的方法の基礎は確率模型を基礎としてのみその正確なる意義を明らかにすることができるということは我々 econometrician にとって既に周知の事実である。R. Frisch は Multicollinearity を論ずるに当り，確率論的接近の方法を採用せず，基礎となる確率模型を明示しなかったため，その数学的な精密さにおいて欠くるところがあったことは事実である。R. Frisch が暗々裡に想定していた確率模型は次のようなものであったといわれている。

いま，経済変数 y_1, x_1, x_2, ……, x_n が存在するとし，y_1, x_1, x_2, ……, x_n は

$$
\begin{aligned}
y_1 &= y_1' + u_1 \\
x_1 &= x_1' + u_2 \\
x_2 &= x_2' + u_3 \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
x_n &= x_n' + u_{n+1}
\end{aligned}
$$

から構成されていると仮定される。y_1', x_1', x_2', ……, x_n' は理論変数であり，u_1, u_2, ……, u_{n+1} は確率変数で，観測誤差を示すものとする。それ故，y_1, x_1, x_2, ……, x_n なる経済変数は観察変数である。確率変数 $u_1, u_2, \cdots\cdots, u_{n+1}$ は相互に独立であると仮定される。

クープマンスやハーベルモの想定によれば，Frisch がMulticollinearity の理論を展開したときにその基礎として想定していた確率模型は上記のような誤差模型（error model）であろうと考えられている。しかし，ハーベルモの指適するごとく，このような模型の上に，経済変数 y_1, x_1, x_2, ……, x_n 間に

$$y_1 = a_1x_1 + a_2x_2 + \cdots\cdots + a_nx_n$$

なる理論式を最小自乗法によって当はめることは，最小自乗法の最大の利点であるマルコフの定理を適用することができなくなるという難点を生じてくる。

このため，えられた統計的推定値も最良線型不偏推定値でないことになる。しかも，それどころか，不偏推定値ですらない。このように，近代統計学の観点から云えば， Frisch の想定していたといわれる誤差模型は極めて使用しにくいものでもあるし，推定値に関して，近代統計学から何らの保証もえられぬことになる。

R. Frisch は Multicollinearity の問題を解決するためにバンチ・マップ法と呼ばれる図解方式を考案した。バンチ・マップ法の目的は真実は独立ではないのに変数の観測誤差にかくされて，観測変数間にみせかけの独立性が保たれているとき，その間の従属性を明らかにすることにあった。バンチ・マップ法は判定方式に幾分の主観的な要素がはいることを避けえないという意味で批判にさらされることもあるとはいえ，充分に目的を達しうる便利な方法である。

しかし，確率論的接近からみたその意義はフリッシュの誤差模型では必ずしも明らかなものではない。この観点から Multicollinearity をマルコフの定理と関係付けようとする試みがハーベルモによってなされている。Multicollinearity をマルコフの定理と関係付ける方がその理論的意義をより判明にできようし，また実用上も応用範囲を広めることができよう。以下少しくハーベルモの観点に従って論を進めることにしよう。

マルコフの定理は次のような仮定の上に立っている。

いま，経済変数 y，x_1，x_2，……，x_n があるとし，これらに関するN期間観察された変数の値を y_t，x_{1t}，x_{2t}，……，x_{nt} $(t=1, 2, \cdots\cdots, N)$ とする。変数 x_{1t}，x_{2t}，……，x_{nt} は指定変数であり，y_t は

$$y_t = y_t' + u_t$$

である。但し，y_t' は指定変数で，u_t は確率変数であり，y_t のみが観測できうるものとする。その上，経済変数 y_t' と x_{1t}，x_{2t}，……，x_{nt} の間には次の

ような線型関係があるとする。

$$y_t' = a_1x_{1t} + a_2x_{2t} + \cdots\cdots + a_nx_{nt} \quad (t=1,\ 2,\ \cdots\cdots,\ N)$$

よって

$$y_t = a_1x_{1t} + a_2x_{2t} + \cdots\cdots + a_nx_{nt} + u_t \quad (t=1,\ 2,\ \cdots\cdots,\ N)$$

とすることができる。確率変数 u_1, u_2, ……, u_t は互に独立で,

$$E(u_i)=0 \quad (i=1,\ 2,\ \cdots\cdots,\ N)$$

$$E(u_i^2)=\sigma^2 \quad (i=1,\ 2,\ \cdots\cdots,\ N)$$

である。最後に, $n \times N$ 行列

$$X=\begin{pmatrix} x_{11} & x_{22}\cdots\cdots\cdots\cdots x_{1N} \\ x_{21} & x_{22}\cdots\cdots\cdots\cdots x_{2N} \\ \cdots\cdots & \cdots\cdots\cdots\cdots \\ \cdots\cdots & \cdots\cdots\cdots\cdots \\ x_{n1} & x_{n2}\cdots\cdots\cdots\cdots x_{nN} \end{pmatrix}$$

の階数が n なることが仮定される。

マルコフの定理によれば，以上の仮定の下で，最小自乗法，換言すれば,

$$\sum_{t=1}^{N}(y_{1t}-a_1x_{1t}-a_2x_{2t}\cdots\cdots a_nx_{nt})^2=\text{min.}$$

によって，最良線型不偏推定値を求めることができる。

これがマルコフの定理の意味するところのものである。Multicollinearity はマルコフの定理と関係付けることにより，その理論的意義を最もよく明らかにすることができる。いま変数 x_1, x_2, ……, x_n の間に精密な線型関係があるとすれば，行列 X の階数は n より小となり，マルコフの定理の仮定は崩れ，統計的推定自体が不可能となる。 Multicollinearity の問題はこのような場合にあるのではない。かくされた誤差によって，行列 X の階数が n で，しかも変数 x_1, x_2, ……, x_n の間に何らかの線型関係が伏在しているときである。このとき，マルコフの定理により，最良線型不偏推定値を得たとしても， Multicollinearity のため，事実上 non-sense な推定結果をうることになる。これを示すのが，パラメーター a_1, a_2, ……, a_n の推定値 $\hat{a}_1$, $\hat{a}_2$, ……, $\hat{a}_n$ の分散

である。これを計算してみることにしよう。

$$\sum_{t=1}^{N}(y_t-a_1x_{1t}-a_2x_{2t}\cdots\cdots-a_nx_{nt})^2$$

を最小ならしめれば，

$$\sum_{t=1}^{N}x_{it}(y_t-\sum_{j=1}^{n}\hat{a}_jx_{jt})=0 \quad (j=1,\ 2,\ \cdots\cdots,\ n)$$

をうる。

$$m_{ij}=\sum_{t=1}^{N}x_{it}x_{jt}$$
$$m_{oi}=\sum_{t=1}^{N}y_tx_{it}(t,\ j,\ =1,\ 2,\ \cdots\cdots,\ n)$$
$$m_{oo}=\sum_{t=1}^{N}y_t{}^2$$

とおけば，方程式は

$$m_{oj}-\sum_{j=1}^{n}\hat{a}_im_{ij}=0 \quad (j=1,\ 2,\ \cdots\cdots,\ n)$$

となる。行列 X の階数が n かなることより，行列

$$M=\begin{pmatrix} m_{11} & m_{12}\cdots\cdots\cdots\cdots m_{1n} \\ m_{21} & m_{22}\cdots\cdots\cdots\cdots m_{2n} \\ \cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ m_{n1} & m_{n2}\cdots\cdots\cdots\cdots m_{nn} \end{pmatrix}=X'X$$

の階数は n である。行列の M 逆行列を

$$N=\begin{pmatrix} n_{11} & n_{12}\cdots\cdots\cdots\cdots n_{1n} \\ n_{21} & n_{22}\cdots\cdots\cdots\cdots n_{2n} \\ \cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ n_{n1} & n_{n2}\cdots\cdots\cdots n_{nn} \end{pmatrix}$$

で示せば，方程式の解は

$$\hat{a}_i=\sum_{j=1}^{n}n_{ij}m_{oj}$$

で表わされる。いま $\hat{a}_i$ の期待値をとれば，

$$E(\hat{a}_i)=\sum_{i=1}^{n}n_{ij}E(m_{oj})$$

$$=\sum_{j=1}^{n} n_{ij}\sum_{t=1}^{N} x_{jt}E(y_t)$$
$$=\sum_{j=1}^{n} n_{ij}\sum_{t=1}^{N} x_{jt}\sum_{k=1}^{n} a_k x_{kt}$$
$$=\sum_{j=1}^{n} n_{ij}\sum_{k=1}^{n} a_k \sum_{t=1}^{N} x_{jt}x_{kt}$$
$$=\sum_{j=1}^{n} n_{ij}\sum_{k=1}^{n} a_k m_{jk}$$
$$=\sum_{k=1}^{n} a_k \sum_{j=1}^{n} n_{ij}m_{jk}=a_i$$

となり，$\hat{a}_i$ が不偏推定値なることが証明される。ついで，

$$\hat{a}_i-a_i=\sum_{j=1}^{n} n_{ij}m_{oj}-a_i$$
$$=\sum_{j=1}^{n} n_{ij}m_{oj}-\sum_{j=1}^{n} n_{ij}\sum_{k=1}^{n} a_k m_{kj}$$
$$=\sum_{j=1}^{n} n_{ij}[\sum_{t=1}^{N} x_{jt}y_t-\sum_{k=1}^{n} a_k\sum_{t=1}^{N} x_{kt}x_{jt}]$$
$$=\sum_{t=1}^{N}\sum_{j=1}^{n} n_{ij}x_{it}(y_t-\sum_{k=1}^{n} a_k x_{kt})$$

となる。$y_t(t=1, 2, \cdots\cdots, n)$ は互に独立で

$$E(y_t-\sum_{t=1}^{N} a_k x_{kt})=\sigma^2$$

であるから

$$E(\hat{a}_i-a_i)^2=\sum_{t=1}^{N}(\sum_{j=1}^{n} n_{ij}x_{jt})^2\sigma^2$$
$$=\sigma^2\sum_{t=1}^{N}\sum_{j=1}^{n}\sum_{k=1}^{n} n_{ij}n_{ik}x_{jt}x_{kt}$$
$$=\sigma^2\sum_{j=1}^{n}\sum_{k=1}^{n} n_{ij}n_{ik}m_{jk}$$
$$=n_{ii}\sigma^2$$

となる。かくして，$\hat{a}_i$ の分散は

$$E(\hat{a}_i-a_i)^2=n_{ii}\sigma^2$$

である。

経済変数 $x_1, x_2, \cdots\cdots, x_n$ の間に独立性がなければ，容易にわかるようにこの分散は無限大となり，推定値が定まらなくなる。これが Multicollinearity の原始的な形態である。しかし，実際の統計計算でこのようなことが起るのは稀である。というのは，何らかの誤差のため，隠された従属性の場合が多いからである。このような場合には a_i の推定値は一義に定まるが，その分散が大きく，推定値が統計学上無意味に終る危険性が大きいわけである。**Multicollinearity**

の統計理論的基礎はこのように簡単なものであり，決して難解といえるものではない。しかし我々 econometrician が統計計算をするとき，これに充分の注意を払わなければ，統計計算を一朝にして無意味に帰せしめる場合がままあるのである。私はこれから，某社で実際に行った賃銀計算の O.R. を例にとって，或る場合に Multicollinearity を避けうる方法を1例述べてみたいと思う。

Ⅲ　賃銀計算における Multicollinearity

某社から依頼をうけて，俸給計算をしたとき，Multicollinearity を避けえた実例をこれから述べてみたいと思う。某社は戦前より比較的長い歴史のある会社ではあったが，戦後急速な成長を成し遂げたため，多数の中途採用者を抱えていた。しかも，大学，高校新卒の採用は戦後少し経過して以後のことであったので，年配者の殆んどは中途採用者であった。このような状態であったためどのような要素にいくらの俸給が支払われているかがわからず，俸給の真の実態が非常に把握しにくく，ために人事政策の立案に支障を来すようなこともままあった。私は某社の依頼によって，俸給の実態分析を試みた。俸給の支払われている要素として，初任給，勤続年数，年令，能力の4つが考えられる。申すまでもなく初任給は学校新規採用者に支払われる俸給で，社員の俸給の出発点となる。日本の給与は勤続すれば，その年数に応じて定期昇給が行われ，初任給を出発点として給与が増加してゆくのが通常である。このようにして，初任給と勤続給が俸給の構成部分となる。学校新規卒業者にとっては問題はないのであるが，中途採用者は同じ勤続年数を勤めていても，学校卒業と同時に入社した人達より，身分が同じであれば，給与が高いのが普通である。これは中途採用者の方が年令が高いからである。かくして，年令給という要素が中途入行者の俸給のなかに存在することになる。最後に我々の考慮しなければならない要素は能力である。凡らく我が国の会社の殆んどは人事考課制度をもっており，その考課点数如何によって，会社の定期昇給時の昇給額を加減することが行なわれている。これが能力給部分である。我が国の年功序列形賃銀にはこ

の外に家族手当や役職手当，職務手当等が加えられるのが普通であるが，この分析ではこれらを除外した本俸だけを対象として統計学的分析を行なった。

まづ，我々は本俸部分を4つの要素，初任給部分，勤続給部分，年令給部分，能力給部分に分離することから始めた。このため，Regression Analysis を採用することとした。

a_0：初任給部分

a_1：勤続1年当り勤続給

a_2：年令1年当り年令給

a_3：考課点1点当り能力給

x_1：勤続年数

x_2：年令

x_3：考課点数

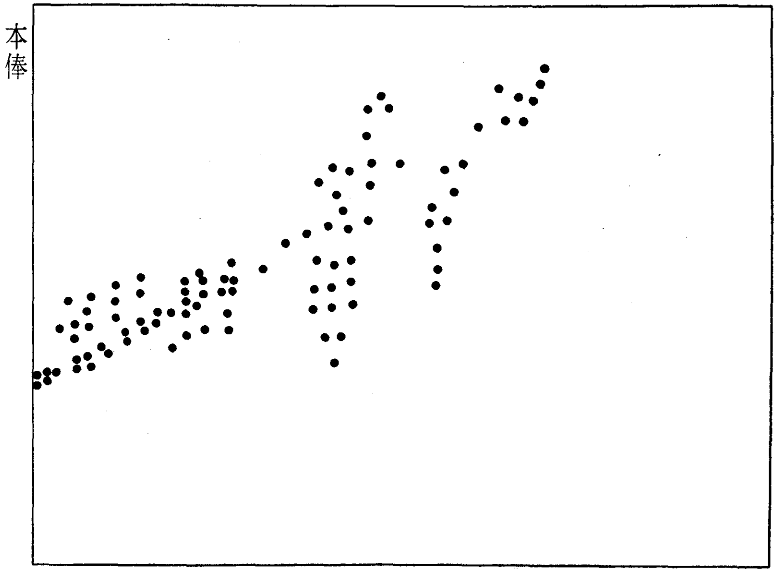
y：本俸

として，Regression

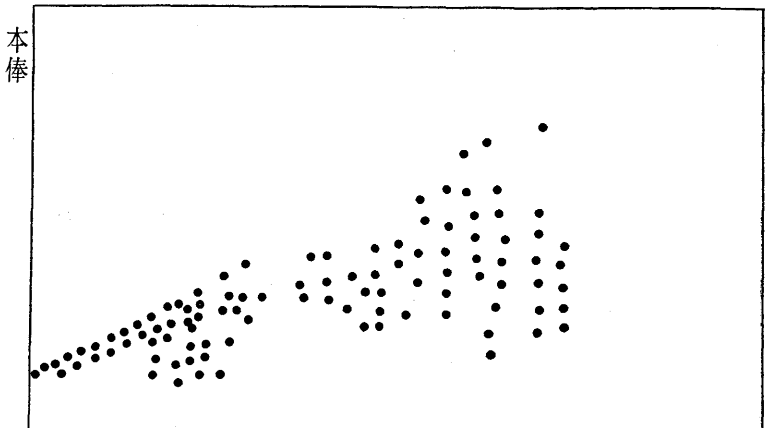
$$y=a_0+a_1x_1+a_2x_2+a_3x_3$$

なる一次式で計算することを最初考えてみた。しかし，ここで問題となったのは Multicollieairty の存在であった。勤続年数 x_1 と年令 x_2 との間に相関はないか，勤続年数 x_1 と考課点数 x_3 との間に相関はないかということであった。勤続年数がある程度存在するということは当然年令の高いことをも示すわけであるから，勤続年数と年令の間に可成りの相関があるだろうことは当然予想されることだからである。念のため，勤続年数と本俸，年令と本俸，勤続年数と年令をプロットして3個の図表を構成してみた。これが第1図，第2図，第3図である。（これらの表は真実のものではない。例として適当な変更を加えたものである。経営上の秘密を守るため真実の図表を掲載することは遠慮することにした。以後掲げる図表は全て皆この意味の修正がほどこされている。）この図表をみても年令と勤続年数の間に可成りの相関のあることがわかる。このため，勤続年数と共に年令という要素を Regression のなかに入れたのでは果

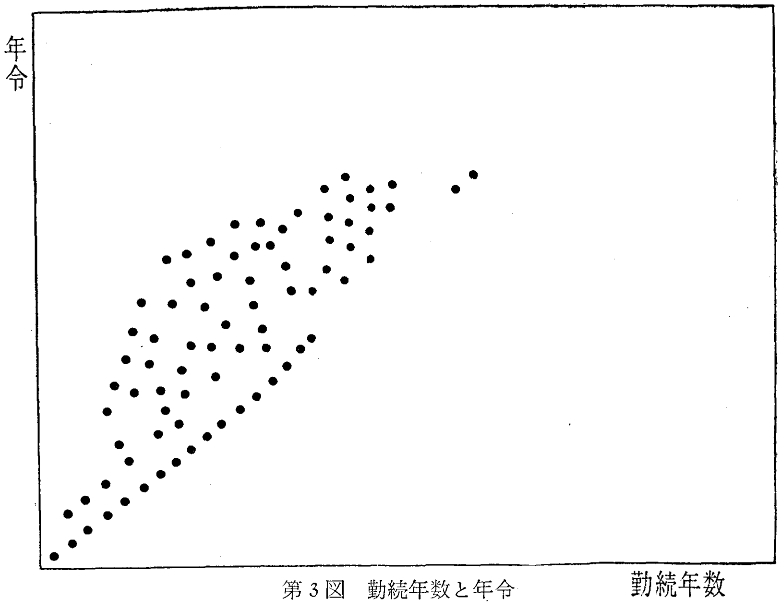
して有意義な分析ができるかどうか疑わしい。そこで，年令という変数を変更



第１図　勤続年数と本俸



第２図　年 令 と 本 俸



第3図　勤続年数と年令

して，学校卒業後入社までに経過した年数を x_2' として，これを前歴年数と名付けこれを年令の代りに分析に加えることにした。この結果，えられた一次式は

$$y=a_0+a_1x_1+a_2'x_2'+a_3x_3$$

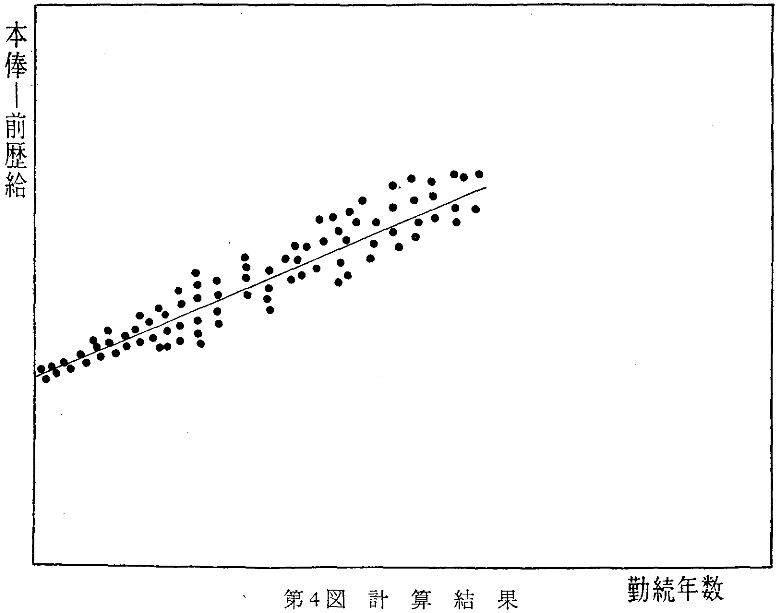
a_2'：前歴年数一年当り給与

x_1'：前歴年数

と修正された。申すまでもなく，前歴年数は学校卒業と同時に入社した社員には無関係で中途採用者だけに関係するものである。また，勤続年数と前歴年数との間にさしたる相関もないだろうということは理論上当然予期できるところである。事実，2変数で

$$y=a_0+a_1x_1+a_2'x_2'$$

を試算してみた結果，充分実用に耐える数字がえられたのみか，勤続年数と前歴年数との間に殆んど相関がないことがわかった。このように2変数で試験的に計算してみた結果が第4図である。この図には横軸に勤続年数をとり，縦軸



第4図 計 算 結 果

は本俸から前歴給を引いたもの，即ち，

$$y-a_2{}'x_2{}'$$

を示している。この試算の結果，我々の分析方法は充分実用に耐えることが証明された。

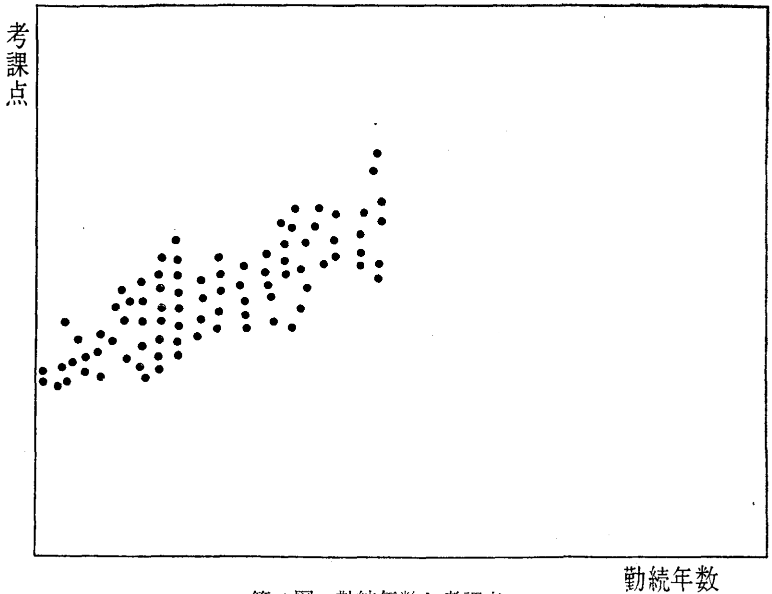
しかし，考課点数を分析に加えると，勤続年数と考課点数との間に相関がないかどうかが問題となった。勿論考課点数が本俸に関係するのは1時期の考課点数だけではない。過去の勤続期間にえた考課点数によって能力給が決まってくる。このため，考課点数 x_3 は過去3年間の平均を用いた。まづ，勤続年数と考課点数との間の相関図表をつくってみた。それが第5図である。この図からみると，勤続年数と考課点数とが或る程度相関をもっていることがわかる。そこで，考課点数 x_3 を勤続年数 x_1 によって修正し，直交化することにした。

$$x_3=b_0+b_1x_1$$

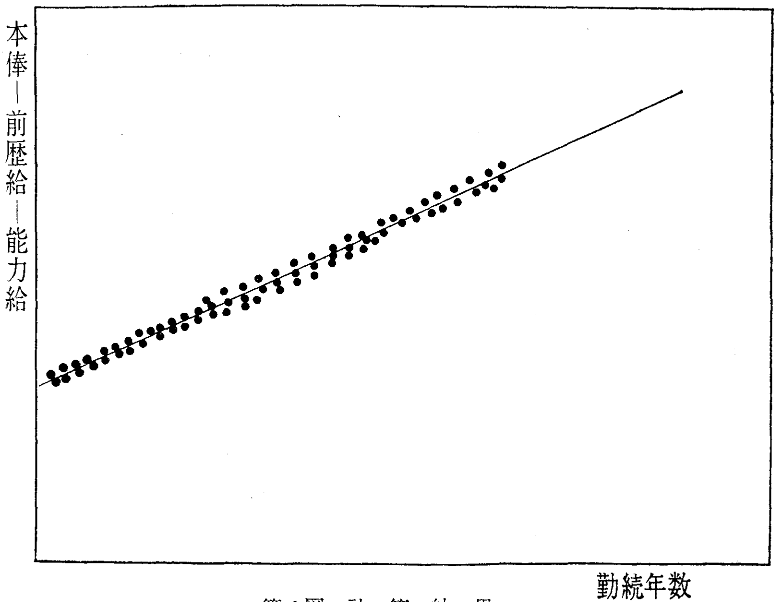
なる Regression line を求め，

$$x_3{}'=x_3-b_0-b_1x_1$$

なる変数 x_3' をもって修正考課点数とした。



第5図 勤続年数と考課点



第6図 計 算 結 果

この結果，我々の理論式は

$$y=a_0x_1+a_2'x_2'+a_3'x_3'$$

となった。この理論式によって実際に計算してみた結果第6図がえられた。この図も横軸には勤続年数 x_1 を，縦軸には本俸から前歴給と能力給を差引いたもの，即ち

$$y-a_2'x_2'-a_3'x_3'$$

がとられている。この図をみても，この分析が充分に実用に耐えることがわかる。

このような分析方法を用いて，男女別，学歴別にわけて，係長以下の給与分析を完了した。課長以上の役職者は上の分析から一応除外して，役職者は別な分析式を用いて計算を行なった。それは役職在任年数を理論式のなかに投入した。計算の理論式は

$$y=a_0+a_1x_1+a_2'x_2'+a_3''x_3''+a_4x_4$$

a_4：役職在任1年当りの役職昇給額

x_4：役職在任年数

とした。しかも役職在任年数 x_4 と考課点数 x_3 との間に相関があったため，考課点数を勤続年数と役職在任年数とで修正した。それは

$$x_3=b_0+b_1x_1+b_2x_4$$

なる Regression を構成し，

$$x_3''=x_3-b_0-b_1x_1-b_2x_2$$

を修正考課点として計算を行なった。この方式によって，これまた実用に耐える理論式を計算することができたのはいうまでもないことである。このようにして，バンチ・マップ法を用いた場合のように Multicollinear のある変数を理論式から除外してしまうことは必要でもなければ，望ましいことでもない。適当な変数の変更や修正によって Multicollinearity を避けることができる場合があるからである。この分析で用いた Multicollinearity を避ける方法の意義を次節で考察してみよう。それによって，この方法の適用範囲もまた自ら定

まるはずだからである。

Ⅳ 上述の方法の意義

まず，年令 x_2 を前歴年数 x_2' に変更した意義を考えよう。年令をよく考えてみると，

年令＝学校卒業年令＋前歴年数＋勤続年数

となる。大学卒とか高校卒等の学校卒業年令は多少の差はあれ通常大部分の人は一定である。これをとZすれば，

$$x_2=Z+x_2'+x_1$$

となり，x_2 と x_1 とが相関関係をもつのは x_2 のちに x_1 を含んでいるからである。そこで，x_1 に当る部分や全てに共通する部分を引き去った残部を変数とすることが linear regression analysis において望ましいことだということはいうまでもないことである。これは変数を定義的に変更することによって，Multicollinearity を避けえた実例である。

ついで，考課点数 x_3 を修正考課点数に修正した意義はどこにあるのだろうか。一般的について我が国の会社では考課点数が勤続年数に正の相関をもっているようである。これは日本の給与体系が所謂年功序列型になっているところからきているのであろうと思われる。それ故，考課点数のなかには個人の能力差だけではなく，勤続による年功を加味されていると考えることができる。この計算を試みた会社でもはっきりとこのことが認められる。それ故，

考課点数＝能力差＋勤続による年功

と考えることができる。今，能力差を r，勤続による年功を s とすれば，

$$x_3=r+s$$

となる。それ故，考課点数が勤続年数と相関をもつのは勤続による年功 s の部分のためである。これを除去するため，

$$x_3=b_0+b_1x_1+r$$

なる仮定をおき，

$$x_3'=r$$

と考えたのである。

このようにすれば，

$$\frac{1}{N}\sum_{t=1}^{N}x_{3t}'=\frac{1}{N}\sum_{t=1}^{N}(x_{3t}-b_0-b_1x_{1t})$$
$$=\frac{1}{N}\sum_{t=1}^{N}x_{3t}-b_0-b_1\frac{1}{N}\sum_{t=1}^{N}x_{1t}$$
$$=\bar{x}_3-\bar{x}_3+b_1\bar{x}_1-b_1\bar{x}_1=0$$
$$\bar{x}_3=\frac{1}{N}\sum_{t=1}^{N}x_{3t}$$
$$\bar{x}_1=\frac{1}{N}\sum^{N}x_{1t}$$
$$b_0=\bar{x}_3-b_1\bar{x}_1$$

となるから，

$$\frac{1}{N}\sum_{t=1}^{N}x_3{}'_t(x_{1t}-\bar{x}_1)$$
$$=\frac{1}{N}\sum_{t=1}^{N}(x_{3t}-b_0-b_1x_{1t})(x_{1t}-\bar{x}_1)$$
$$=\frac{1}{N}\sum x_{3t}x_{1t}-b_0\frac{1}{N}\sum_{t=1}^{N}x_{1t}-b\frac{1}{N}\sum_{t=1}^{N}x_{1t}{}^2$$
$$-\bar{x}_1\frac{1}{N}\sum_{t=1}^{N}x_{3t}+b_0\bar{x}_1+b_1\bar{x}_1\frac{1}{N}\sum_{t=1}^{N}x_{1t}$$
$$=\frac{1}{N}\sum_{t=1}^{N}x_{3t}x_{1t}-b_0\bar{x}_1-b\frac{1}{N}\sum_{t=1}^{N}x_{1t}{}^2$$
$$-\bar{x}_1\bar{x}_3+b_0\bar{x}_1+b_1\bar{x}_1{}^2$$
$$=\left[\frac{1}{N}\sum_{t=1}^{N}x_{3t}x_{1t}-\bar{x}_1\bar{x}_3\right]-b_1\left[\frac{1}{N}\sum_{t=1}^{N}x_{1t}{}^2-\bar{x}_1{}^2\right]$$

しかるに，

$$b_1=\frac{\frac{1}{N}\sum_{t=1}^{N}x_{3t}x_{1t}-\bar{x}_1\bar{x}_3}{\frac{1}{N}\sum_{t=1}^{N}x_{1t}{}^2-\bar{x}_1{}^2}$$

であるから，

$$\frac{1}{N}\sum_{t=1}^{N}x_{3t}'(x_{1t}-\bar{x}_1)=0$$

となり，x_3' と x_1 とは相互に直交化することになる。このような修正によって，Multicollinearityを避けることができたのである。しかしながら，よく考えてみれば，この方法を適用できたのは問題となった2変数間に判つきりとした因果関係が存在するためである。考課点数は勤続年数とともに増加するが，その

逆は成立しない。考課点数が高いから勤続年数がより長くなるという因果関係は考えることもできない。このように，Multicollinear な変数でもその間に一方的な因果関係が存在すれば，理論式のなかに変数を修正してとり入れることができることがわかる。

V 結び

以上我々は賃銀計算の O. R. を例にとりながら，Multicollinearity を避けうる方法を述べることができた。それとともにこの方法の適用範囲をも定めることができた。それは Multicollinear な変数の間に一方的な因果関係が存在すれば，変数の変更または修正によって，有意義な関係式を推定しうるということである。このように単一方程式接近において，フリッシュの考案したバンチ・マップ法のように，Multicollinear な変数の一方を有害変数として除外しなくてもよい場合も存在しうるわけである。しかし，この方法も前に述べた通り万能薬ではない。Multicollinear な変数間に相互作用があれば，適用不能となる。我々 econometrician の取り扱う経済変数間には相互作用をもつ Multicollinear な変数も少なくないであろう，いやそのような場合の方が多いかも知れないのである。このような場合は本論に述べた方法の適用範囲外である。これにて本論を終る。