回帰分析と線形代数

丸山，祐造

(Citation)
経済学・経営学学習のために,2021(後期号):1-16

(Issue Date)
2021-10-01

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(URL)
https://hdl.handle.net/20.500.14094/E0042528

回帰分析と線形代数

丸　山　祐　造

1　はじめに

本稿では，回帰分析と最小 2 乗法の記述統計的な側面について，線形代数と関連付けながら説明します。概ね 2020 年度第 3Q に担当した経営学部科目「経営数学」で扱った内容に基づいています。

回帰分析は説明変数を用いて目的変数を予測する手法であり，最も重要な多変量解析手法です。説明変数が 1 つの場合を単回帰分析といい，2 つ以上ある場合を重回帰分析といいます。単回帰分析は，経営学部科目「経営統計」など入門レベルの統計学で必ず扱われるので，本稿を読み始めた学生さんであれば理解していることでしょう。ただし，目的変数に関連する説明変数が 1 つだけという想定は単純すぎるので，実際のデータ解析においては重回帰分析の出番が圧倒的に多いのです。

重回帰分析では，一般のサイズのベクトルや行列が登場するために，苦手意識を持ったり，理解を諦めてしまう経営学部生が多いかもしれません。経営学部生の 1 年次の必修科目「線形代数入門」の教科書を見ると，3 次元までのベクトルや行列について，固有値・固有ベクトルに至る様々な概念を学んで計算に習熟するようです。またベクトルについては，大きさと向きを持った量であることが強調されています。しかし，統計学やデータサイエンスの視点からは，

- ベクトルを「複数の数値の組」，行列を「複数の『成分を縦に並べた同じサイズのベクトル』を横に束ねて表形式に整列した数値」，として見なすこと

・一般のサイズのベクトルの和・スカラー倍，内積を怖れずに扱えること

が重要です。回帰分析はこのようなベクトルや行列の見方と非常に親和性があり，一般のサイズのベクトルの和・スカラー倍と内積（及びそれらの行列による表記）に慣れることでより深い理解が得られます。

ここでは，説明変数が p 個の重回帰分析と線形代数の関係を初歩から説明するほどのスペースはないので，説明変数を２個に限定し，以下のような問いを設定してそれに答えるというスタイルにします。「重回帰分析の最小二乗解が $(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$ で与えられる，と説明されるとき，『$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ の逆行列が存在すれば』と決まり文句がつきます。しかし，$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ の逆行列が存在しない場合には最小二乗解はどうなるのでしょうか？」という問いです。$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ が逆行列を持たない場合が，実データの重回帰分析において特段重要というわけではありません。しかし，その理解には線形代数の視点が不可欠なので，回帰分析をより深く理解することに繋がるのです。特に，説明変数が２個の場合は，最小二乗解が（分散，共分散，相関係数などの）基本統計量を用いて具体的に表現出来るのでイメージが湧きやすいと思います。

２節・３節では，線形代数を前提とせずに説明変数が２個の場合の重回帰分析を解説します。４節で必要最小限のベクトル・行列の知識を整理した後，５節・６節・７節で線形代数の立場から２節と３節の結果を解釈します。それでは始めましょう。

2　重回帰分析と最小二乗法 I

最も単純な重回帰分析として説明変数が２つの場合を考えます。説明変数 x, z と目的変数 y が近似的に線形関係

$$y \approx a + bx + cz \tag{1}$$

を満たすならば，x_0, z_0 が与えられたとき $a + bx_0 + cz_0$ によって y の値を予測

可能でしょう。そこで既に得られている n 組のデータ

$$(\{x_1, z_1\}, y_1), \ldots, (\{x_n, z_n\}, y_n)$$

に基づいて，式（1）における適切な a, b, c を求めましょう。1 つの自然な考え方は，現実の値 y_i と予測値 $a+bx_i+cz_i$ の差の絶対値（$i=1, \ldots, n$ は，観測対象の個体を表す添字）

$$|a+bx_1+cz_1-y_1|, |a+bx_2+cz_2-y_2|, \ldots, |a+bx_n+cz_n-y_n|$$

が，全体的な意味で小さくなるような a, b, c を求めることです（図 1)。数学的には，その平方和

$$Q(a, b, c) = \sum_{i=1}^{n} (a+bx_i+cz_i-y_i)^2 \tag{2}$$

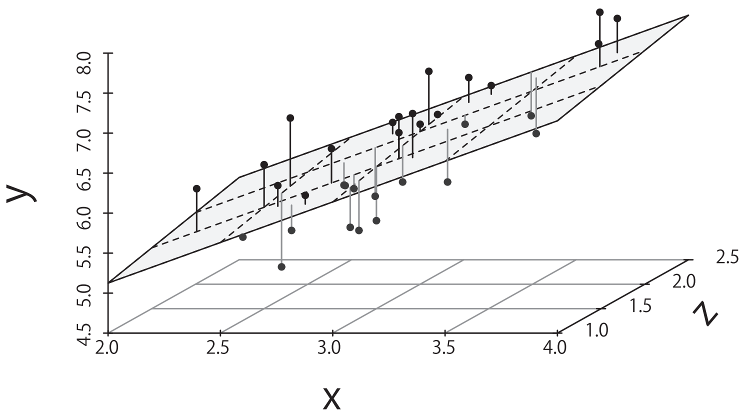

図 1　散布図と重回帰分析

を最小にする a, b, c を求めるのが容易であり，$Q(a, b, c)$ の最小化を最小二乗法といいます。

3 変数関数 $Q(a, b, c)$ の最小化問題においても，微分が重要な役割を果たします。$Q(a, b, c)$ の（b, c を固定したもとでの）a に関する微分と（a, c を固定したもとでの）b に関する微分及び（a, b を固定したもとでの）c に関する微分を，それぞれ $=0$ とおいた連立方程式

$$\begin{cases} \dfrac{\partial}{\partial a} Q(a,\, b,\, c) = 0 \\[2ex] \dfrac{\partial}{\partial b} Q(a,\, b,\, c) = 0 \\[2ex] \dfrac{\partial}{\partial c} Q(a,\, b,\, c) = 0 \end{cases} \tag{3}$$

の解が Q を最小化する解の候補になります。微分の基本的性質として，「関数の和 $f_1(x)+f_2(x)$ の微分は $f_1'(x)+f_2'(x)$ である」ことや，「定数 α に対して合成関数 $f(\alpha x)$ の微分は $\alpha f'(\alpha x)$ である」ことを用いると

$$\frac{\partial}{\partial a} Q(a,\, b,\, c) = 2\sum_{i=1}^{n}(a+bx_i+cz_i-y_i)$$

$$\frac{\partial}{\partial b} Q(a,\, b,\, c) = 2\sum_{i=1}^{n} x_i(a+bx_i+cz_i-y_i)$$

$$\frac{\partial}{\partial c} Q(a,\, b,\, c) = 2\sum_{i=1}^{n} z_i(a+bx_i+cz_i-y_i)$$

が得られます。従って３変数 $a,\, b,\, c$ に関する連立方程式（3）は

$$\begin{cases} \sum_{i=1}^{n}(a+bx_i+cz_i-y_i) = 0 \\ \sum_{i=1}^{n}x_i(a+bx_i+cz_i-y_i) = 0 \\ \sum_{i=1}^{n}z_i(a+bx_i+cz_i-y_i) = 0 \end{cases} \tag{4}$$

と書けます。連立方程式（4）の解 $\hat{a},\, \hat{b},\, \hat{c}$ を最小化する解の「候補」と言いましたが，$\hat{a},\, \hat{b},\, \hat{c}$ は実際に $Q(a,\, b,\, c)$ を最小化します。任意の実数 $a,\, b,\, c$ に対して，

$$\sum_{i=1}^{n}(a+bx_i+cz_i-y_i)^2 \geq \sum_{i=1}^{n}(\hat{a}+\hat{b}x_i+\hat{c}z_i-y_i)^2 \tag{5}$$

が成り立つためです。以下の４番目の等号で $\hat{a},\, \hat{b},\, \hat{c}$ が連立方程式（4）の解であることを使っています。

$$\begin{aligned} &\sum_{i=1}^{n}(a+bx_i+cz_i-y_i)^2 \\ &\quad = \sum_{i=1}^{n}\{a+bx_i+cz_i-(\hat{a}+\hat{b}x_i+\hat{c}z_i)+(\hat{a}+\hat{b}x_i+\hat{c}z_i)-y_i\}^2 \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \{(a-\hat{a})+(b-\hat{b})x_i+(c-\hat{c})z_i+(\hat{a}+\hat{b}x_i+\hat{c}z_i-y_i)\}^2 \\
&= \sum_{i=1}^{n} \{(a-\hat{a})+(b-\hat{b})x_i+(c-\hat{c})z_i\}^2+\sum_{i=1}^{n}(\hat{a}+\hat{b}x_i+\hat{c}z_i-y_i)^2 \\
&\quad +2(a-\hat{a})\sum_{i=1}^{n}(\hat{a}+\hat{b}x_i+\hat{c}z_i-y_i)+2(b-\hat{b})\sum_{i=1}^{n}x_i(\hat{a}+\hat{b}x_i+\hat{c}z_i-y_i) \\
&\quad +2(c-\hat{c})\sum_{i=1}^{n}z_i(\hat{a}+\hat{b}x_i+\hat{c}z_i-y_i) \\
&= \sum_{i=1}^{n} \{(a-\hat{a})+(b-\hat{b})x_i+(c-\hat{c})z_i\}^2+\sum_{i=1}^{n}(\hat{a}+\hat{b}x_i+\hat{c}z_i-y_i)^2 \\
&\geq \sum_{i=1}^{n}(\hat{a}+\hat{b}x_i+\hat{c}z_i-y_i)^2
\end{aligned}
$$

3　最小二乗解とその一意性 I

この節では，連立方程式（4）の解を明示的に表現しましょう。まず記法を準備します。x の平均，x と y の共分散を

$$\bar{x}=\frac{\sum_{i=1}^{n}x_i}{n},\ s_{xy}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n} \tag{6}$$

のように書きます。このとき x の分散は s_{xx} と表現できます。また実データにおける現実的な想定として，x, y, z の分散は正，つまり

$$s_{xx}>0,\ s_{yy}>0,\ s_{zz}>0 \tag{7}$$

とします。x と y の相関係数 r_{xy} は，分散 s_{xx}, s_{yy} と共分散 s_{xy} を用いて

$$r_{xy}=\frac{s_{xy}}{\sqrt{s_{xx}}\sqrt{s_{yy}}} \tag{8}$$

で与えられます。

連立方程式（4）の第 1 式の両辺を n で割ると，（6）で与えられる平均の定義より

$$a+b\bar{x}+c\bar{z}-\bar{y}=0 \tag{9}$$

を得ます。（9）式を連立方程式（4）の第 2 式と第 3 式に代入すると，

$$\begin{cases} \sum_{i=1}^{n} x_i(bx_i - b\bar{x} + cz_i - c\bar{z} - y_i + \bar{y}) = 0 \\ \sum_{i=1}^{n} z_i(bx_i - b\bar{x} + cz_i - c\bar{z} - y_i + \bar{y}) = 0 \end{cases} \tag{10}$$

を得ます。また（6）で与えられる平均の定義から，恒等的に

$$\sum_{i=1}^{n} (x_i - \bar{x}) = 0$$

が成立し，同様に

$$\sum_{i=1}^{n} \{b(x_i - \bar{x}) + c(z_i - \bar{z}) - (y_i - \bar{y})\} = 0 \tag{11}$$

も恒等的に成立します。「(11) 式の両辺に $\bar{x}$ をかけて，（10）の第１式から引く」及び「(11) 式の両辺に $\bar{z}$ をかけて，（10）の第２式から引く」ことにより，

$$\begin{cases} \sum_{i=1}^{n} (x_i - \bar{x})\{b(x_i - \bar{x}) + c(z_i - \bar{z}) - (y_i - \bar{y})\} = 0 \\ \sum_{i=1}^{n} (z_i - \bar{z})\{b(x_i - \bar{x}) + c(z_i - \bar{z}) - (y_i - \bar{y})\} = 0 \end{cases}$$

を得ます。さらに（6）で与えられる共分散の定義から，連立方程式（10）は

$$\begin{cases} s_{xx}b + s_{xz}c - s_{xy} = 0 \\ s_{xz}b + s_{zz}c - s_{yz} = 0 \end{cases} \tag{12}$$

と書き直せます。

連立方程式（12）を解きましょう。まず $s_{xx}s_{zz} \neq s_{xz}^2$ の場合を考えます。相関係数の定義（8）から，$|r_{xz}| < 1$ と言い換えられます。このとき，連立方程式（12）の解は

$$\hat{b} = \sqrt{\frac{s_{yy}}{s_{xx}}}\frac{r_{xy} - r_{xz}r_{yz}}{1 - r_{xz}^2} \quad \hat{c} = \sqrt{\frac{s_{yy}}{s_{zz}}}\frac{r_{yz} - r_{xz}r_{xy}}{1 - r_{xz}^2}$$

で与えられます（(8) で与えられる相関係数の定義を繰り返し使うことで計算が確認できます)。連立方程式（4）の第１式である式（9）と合わせて

$$\begin{aligned} \hat{b} &= \sqrt{\frac{s_{yy}}{s_{xx}}}\frac{r_{xy} - r_{xz}r_{yz}}{1 - r_{xz}^2} \\ \hat{c} &= \sqrt{\frac{s_{yy}}{s_{zz}}}\frac{r_{yz} - r_{xz}r_{xy}}{1 - r_{xz}^2} \\ \hat{a} &= \bar{y} - \hat{b}\bar{x} - \hat{c}\bar{z} \end{aligned} \tag{13}$$

が，連立方程式（4）の一意な解として得られます。また（5）より $Q(a, b, c)$ を最小にします。

次に $|r_{xz}| = 1$ の場合を考えましょう。この場合は，$(x_1, z_1), \ldots, (x_n, z_n)$ が直線上に並ぶので，切片 l と傾き $k \neq 0$ が存在して

$$z_i = l + kx_i \quad i = 1, \ldots, n \tag{14}$$

が成立します。さらに $\bar{z} = l + k\bar{x}$ に注意すると，（14）は

$$(z_i - \bar{z}) = k(x_i - \bar{x}) \quad i = 1, \ldots, n \tag{15}$$

と書き換えられます。（15）を連立方程式（10）に代入すると，

$$\begin{cases} s_{xx}(b+kc) - s_{xy} = 0 \\ k\{s_{xx}(b+kc) - s_{xy}\} = 0 \end{cases}$$

を得ます。従って，任意の実数 t に対して，

$$\begin{aligned} \hat{c} &= t \\ \hat{b} &= -kt + \frac{s_{xy}}{s_{xx}} \\ \hat{a} &= \bar{y} - \hat{b}\bar{x} - \hat{c}\bar{z} \end{aligned} \tag{16}$$

が連立方程式（4）の解として $Q(a, b, c)$ を最小にします。t の任意性により，$|r_{xz}| = 1$ の場合には最小二乗解が無限個あることが分かります。

4 ベクトルと行列

この節では，2 節，3 節の議論を線形代数の観点から理解するために，ベクトルと行列に関連する必要最小限の概念を導入します。n 次元ベクトルは

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

のような n 個の実数の組です。ベクトルの和とスカラー倍は

$$\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1+v_1 \\ u_2+v_2 \\ \vdots \\ u_n+v_n \end{pmatrix}, \quad c\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} cu_1 \\ cu_2 \\ \vdots \\ cu_n \end{pmatrix} \tag{17}$$

と定義されます。

2つの n 次元ベクトル $\boldsymbol{u}, \boldsymbol{v}$ の内積は，成分同士の積和

$$\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} = \sum_{i=1}^{n} u_i v_i$$

と定義されます。この定義から

$$\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} = \boldsymbol{v}^{\mathrm{T}}\boldsymbol{u},\ (c\boldsymbol{u})^{\mathrm{T}}\boldsymbol{v} = c\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v},\ \boldsymbol{u}^{\mathrm{T}}(\boldsymbol{v}+\boldsymbol{w}) = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}+\boldsymbol{u}^{\mathrm{T}}\boldsymbol{w} \tag{18}$$

が従います。$\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} = 0$ のとき，$\boldsymbol{u}$ と $\boldsymbol{v}$ は直交するといいます。同じベクトル同士の内積

$$\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u} = \sum_{i=1}^{n} u_i u_i = \sum_{i=1}^{n} u_i^2 \tag{19}$$

はユークリッドノルムの二乗であり，$\|\boldsymbol{u}\|^2$ と表現します。(18) を繰り返し使うと，実数 s, t に対する展開公式 $(s+t)^2 = s^2+2st+t^2$ に対応する関係

$$\begin{aligned} \|\boldsymbol{u}+\boldsymbol{v}\|^2 &= (\boldsymbol{u}+\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{u}+\boldsymbol{v}) = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}+2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}+\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} \\ &= \|\boldsymbol{u}\|^2+2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}+\|\boldsymbol{v}\|^2 \end{aligned} \tag{20}$$

を示すことが出来ます。特に $\boldsymbol{u}$ と $\boldsymbol{v}$ が直交する場合の関係式

$$\|\boldsymbol{u}+\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2 \tag{21}$$

はピタゴラスの定理と呼ばれます。

$n\times q$ 行列 $\boldsymbol{W}$ は q 本の n 次元ベクトルにより

$$\boldsymbol{W} = (\boldsymbol{w}_1\ \boldsymbol{w}_2\ \ldots\ \boldsymbol{w}_q) \tag{22}$$

で定義されます。従って，n 次元ベクトルは行列の特別な場合 $(q = 1)$ と見なせます。行列 $\boldsymbol{W}$ は nq 個の実数が n 行 q 列の形に整列していて，上から i 行，左から j 列の値を (i, j) 成分といいます。行列 $\boldsymbol{W}$ に対して，(i, j) 成分を (j, i) 成分に配置して出来る $q\times n$ 行列を行列 $\boldsymbol{W}$ の転置行列といい $\boldsymbol{W}^{\mathrm{T}}$ と書き

ます。行列の特別な場合である n 次元ベクトル $\boldsymbol{w}$ を転置すると，n 個の成分が横に配置された $\boldsymbol{w}^{\mathrm{T}}=(w_1\ w_2\ \ldots\ w_n)$ になります。これは $1\times n$ 行列とみなせます。もちろん，（22）で表現される行列 $\boldsymbol{W}$ の転置行列は

$$\boldsymbol{W}^{\mathrm{T}}=\begin{pmatrix}\boldsymbol{w}_1^{\mathrm{T}}\\ \boldsymbol{w}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{w}_q^{\mathrm{T}}\end{pmatrix} \tag{23}$$

で与えられます。

$n\times q$ 行列の $\boldsymbol{W}=(\boldsymbol{w}_1\ \boldsymbol{w}_2\ \ldots\ \boldsymbol{w}_q)$ と $n\times r$ 行列 $\boldsymbol{Z}=(\boldsymbol{z}_1\ \boldsymbol{z}_2\ \ldots\ \boldsymbol{z}_r)$ について，行列の積 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{Z}$ は，その (i, j) 成分が $\boldsymbol{w}_i$ と $\boldsymbol{z}_j$ の内積で与えられる $q\times r$ 行列

$$\boldsymbol{W}^{\mathrm{T}}\boldsymbol{Z}=\begin{pmatrix}\boldsymbol{w}_1^{\mathrm{T}}\\ \boldsymbol{w}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{w}_q^{\mathrm{T}}\end{pmatrix}(\boldsymbol{z}_1\ \boldsymbol{z}_2\ \ldots\ \boldsymbol{z}_r)=\begin{pmatrix}\boldsymbol{w}_1^{\mathrm{T}}\boldsymbol{z}_1 & \boldsymbol{w}_1^{\mathrm{T}}\boldsymbol{z}_2 & \cdots & \boldsymbol{w}_1^{\mathrm{T}}\boldsymbol{z}_r\\ \boldsymbol{w}_2^{\mathrm{T}}\boldsymbol{z}_1 & \boldsymbol{w}_2^{\mathrm{T}}\boldsymbol{z}_2 & \cdots & \boldsymbol{w}_2^{\mathrm{T}}\boldsymbol{z}_r\\ \vdots & \vdots & & \vdots\\ \boldsymbol{w}_q^{\mathrm{T}}\boldsymbol{z}_1 & \boldsymbol{w}_q^{\mathrm{T}}\boldsymbol{z}_2 & \cdots & \boldsymbol{w}_q^{\mathrm{T}}\boldsymbol{z}_r\end{pmatrix} \tag{24}$$

として定義されます。行列の積 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{Z}$ においては，$\boldsymbol{W}^{\mathrm{T}}$ の列数と $\boldsymbol{Z}$ の行数が一致していることが必要です。そして，一致していれば $\boldsymbol{w}_i$ と $\boldsymbol{z}_j$ のサイズが同じになり内積 $\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{z}_j$ が定義できます。このことは，（24）のように積が定義できることと整合的です。

（22）で与えられる $n\times q$ 行列 $\boldsymbol{W}$ を構成する n 次元ベクトルの線形和 $\sum_{j=1}^{q}c_j\boldsymbol{w}_j$ は，q 次元ベクトル（$q\times 1$ 行列）$\boldsymbol{c}=(c_1\ c_2\ \ldots\ c_q)^{\mathrm{T}}$ を用いた行列の積 $\boldsymbol{W}\boldsymbol{c}$ で表現できます。念の為，成分を具体的に表記して確認しましょう。$\boldsymbol{W}$ の (i, j) 成分を w_{ij} とします。このとき線形和は

$$c_1\boldsymbol{w}_1+c_2\boldsymbol{w}_2+\cdots+c_q\boldsymbol{w}_q=\begin{pmatrix}\sum_{j=1}^{q}c_jw_{1j}\\ \sum_{j=1}^{q}c_jw_{2j}\\ \vdots\\ \sum_{j=1}^{q}c_jw_{nj}\end{pmatrix} \tag{25}$$

です。ところで，$\boldsymbol{W}$ の第 i 行からなる q 次元ベクトル $\tilde{\boldsymbol{w}}_i^{\mathrm{T}}=(w_{i1}\ w_{i2}\ \ldots\ w_{iq})$

により，$\boldsymbol{W}$ は

$$\boldsymbol{W} = \begin{pmatrix} \tilde{\boldsymbol{w}}_1^{\mathrm{T}} \\ \tilde{\boldsymbol{w}}_2^{\mathrm{T}} \\ \vdots \\ \tilde{\boldsymbol{w}}_n^{\mathrm{T}} \end{pmatrix}$$

と表記されます（転置行列 $\boldsymbol{W}^{\mathrm{T}}$ を表記した（23）と混乱しないで下さい）。（24）で与えられる行列の積の定義より

$$\boldsymbol{W}\boldsymbol{c} = \begin{pmatrix} \tilde{\boldsymbol{w}}_1^{\mathrm{T}} \\ \tilde{\boldsymbol{w}}_2^{\mathrm{T}} \\ \vdots \\ \tilde{\boldsymbol{w}}_n^{\mathrm{T}} \end{pmatrix} (\boldsymbol{c}) = \begin{pmatrix} \tilde{\boldsymbol{w}}_1^{\mathrm{T}}\boldsymbol{c} \\ \tilde{\boldsymbol{w}}_2^{\mathrm{T}}\boldsymbol{c} \\ \vdots \\ \tilde{\boldsymbol{w}}_n^{\mathrm{T}}\boldsymbol{c} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{q} c_j w_{1j} \\ \sum_{j=1}^{q} c_j w_{2j} \\ \vdots \\ \sum_{j=1}^{q} c_j w_{nj} \end{pmatrix}$$

が従うので，$\boldsymbol{W}\boldsymbol{c}$ は（25）で与えられる線形和に一致する，つまり

$$c_1\boldsymbol{w}_1 + c_2\boldsymbol{w}_2 + \cdots + c_q\boldsymbol{w}_q = \boldsymbol{W}\boldsymbol{c} \tag{26}$$

が分かります。ここでは，一般のサイズのベクトル・行列に習熟するステップにおいて，（必要に応じて）成分表記に戻ることの重要性を強調するために，やや丁寧に説明しました。

本節の最後に行列の積と転置行列に関する関係を紹介します。（24）のように行列の積 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{Z}$ が定義できるとき，(j, i) 成分が $\boldsymbol{w}_i$ と $\boldsymbol{z}_j$ の内積で与えられる $r \times q$ 行列 $\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}$ も同様に定義できます。これは行列 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{Z}$ の転置行列なので，行列の積と転置行列に関する関係

$$\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W} = (\boldsymbol{W}^{\mathrm{T}}\boldsymbol{Z})^{\mathrm{T}} \tag{27}$$

が成立します。

5　重回帰分析と最小二乗法Ⅱ

この節では，2節の内容を4節で導入したベクトルと行列で書き換えます。与えられた n 組のデータ

$$(\{x_1, z_1\}, y_1), \cdots, (\{x_n, z_n\}, y_n)$$

に対して，n 次元ベクトル $\boldsymbol{y}$, $\mathbf{1}_n$, $\boldsymbol{x}$, $\boldsymbol{z}$ と $n\times 3$ 行列 $\boldsymbol{X}$，及び 3 次元ベクトル $\boldsymbol{b}$ を

$$\boldsymbol{y}=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \mathbf{1}_n=\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \boldsymbol{z}=\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}, \boldsymbol{X}=(\mathbf{1}_n\ \boldsymbol{x}\ \boldsymbol{z}),\ \boldsymbol{b}=\begin{pmatrix} a \\ b \\ c \end{pmatrix} \tag{28}$$

のように定義します。(7) における x, y, z の分散が正であるという想定により，「$\boldsymbol{x}$, $\boldsymbol{z}$, $\boldsymbol{y}$ が $\mathbf{1}_n$ のスカラー倍ではないこと」が保証されます。

(2) で定義される $Q(a, b, c)$ における n 個の成分から成るベクトルは，(17), (26) 及び (28) により

$$\begin{aligned}\begin{pmatrix} a+x_1b+z_1c-y_1 \\ a+x_2b+z_2c-y_2 \\ \vdots \\ a+x_nb+z_nc-y_n \end{pmatrix} &= a\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}+b\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}+c\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}-\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \\ &= a\mathbf{1}_n+b\boldsymbol{x}+c\boldsymbol{z}-\boldsymbol{y} \\ &= \boldsymbol{X}\boldsymbol{b}-\boldsymbol{y}\end{aligned}$$

と書けます（特に 3 番目の等号は (26) より成立します)。そして (19) により $Q(a, b, c)=\|\boldsymbol{X}\boldsymbol{b}-\boldsymbol{y}\|^2$ が従います。また，連立方程式 (4) は

$$\begin{aligned}\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} &= \begin{pmatrix} \sum_{i=1}^{n}(a+bx_i+cz_i-y_i) \\ \sum_{i=1}^{n}x_i(a+bx_i+cz_i-y_i) \\ \sum_{i=1}^{n}z_i(a+bx_i+cz_i-y_i) \end{pmatrix} = \begin{pmatrix} \mathbf{1}_n^{\mathrm{T}}(\boldsymbol{X}\boldsymbol{b}-\boldsymbol{y}) \\ \boldsymbol{x}^{\mathrm{T}}(\boldsymbol{X}\boldsymbol{b}-\boldsymbol{y}) \\ \boldsymbol{z}^{\mathrm{T}}(\boldsymbol{X}\boldsymbol{b}-\boldsymbol{y}) \end{pmatrix} \\ &= \boldsymbol{X}^{\mathrm{T}}(\boldsymbol{X}\boldsymbol{b}-\boldsymbol{y})\end{aligned} \tag{29}$$

と書けます。連立方程式 (29) の解 $\hat{\boldsymbol{b}}$ と任意の 3 次元ベクトル $\boldsymbol{b}=(a\ b\ c)^{\mathrm{T}}$ に対して，(20) と (27) を用いると

$$\begin{aligned}\|\boldsymbol{X}\boldsymbol{b}-\boldsymbol{y}\|^2 &= \|\boldsymbol{X}\boldsymbol{b}-\boldsymbol{X}\hat{\boldsymbol{b}}+\boldsymbol{X}\hat{\boldsymbol{b}}-\boldsymbol{y}\|^2\\ &= \|\boldsymbol{X}\boldsymbol{b}-\boldsymbol{X}\hat{\boldsymbol{b}}\|^2+\|\boldsymbol{X}\hat{\boldsymbol{b}}-\boldsymbol{y}\|^2+2(\boldsymbol{b}-\hat{\boldsymbol{b}})^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}(\boldsymbol{X}\hat{\boldsymbol{b}}-\boldsymbol{y})\\ &= \|\boldsymbol{X}\boldsymbol{b}-\boldsymbol{X}\hat{\boldsymbol{b}}\|^2+\|\boldsymbol{X}\hat{\boldsymbol{b}}-\boldsymbol{y}\|^2\\ &\geq \|\boldsymbol{X}\hat{\boldsymbol{b}}-\boldsymbol{y}\|^2 \end{aligned} \tag{30}$$

を得るので，連立方程式（29）の解 $\hat{\boldsymbol{b}}$ が $Q(a, b, c)$ を最小にすることが分かります。

6　最小二乗解とその一意性Ⅱ

この節では，5節で得た最小二乗解（連立方程式（29）の解 $\hat{\boldsymbol{b}}$）の表現と一意性について，3節の結果と対応させながら考察します。$\boldsymbol{X}\boldsymbol{b}$ で $\boldsymbol{b}$ の各成分が実数全体を動くことで作られる n 次元ベクトルの集合

$$\{\boldsymbol{X}\boldsymbol{b} : \boldsymbol{b} \in \mathcal{R}^3\} = \{a\mathbf{1}_n+b\boldsymbol{x}+c\boldsymbol{z} : a \in \mathcal{R},\ b \in \mathcal{R},\ c \in \mathcal{R}\}$$

を $\mathbf{1}_n, \boldsymbol{x}, \boldsymbol{z}$ が張る線形部分空間といい，$<\mathbf{1}_n, \boldsymbol{x}, \boldsymbol{z}>$ と表記しましょう。また $<\mathbf{1}_n, \boldsymbol{x}, \boldsymbol{z}>$ の中で $\boldsymbol{y}$ とのユークリッド距離が最も近い元を $\boldsymbol{y}$ の $<\mathbf{1}_n, \boldsymbol{x}, \boldsymbol{z}>$ への正射影といい，$\hat{\boldsymbol{y}}$ と表記しましょう。この観点で（30）を見ると，最小二乗法は正射影 $\hat{\boldsymbol{y}}$ を求める手続きであると言い換えることが出来ます。また，正射影

$$\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{b}} = \hat{a}\mathbf{1}_n+\hat{b}\boldsymbol{x}+\hat{c}\boldsymbol{z}$$

を表現する係数ベクトル $\hat{\boldsymbol{b}}$ が最小二乗解であると言えます。

次に，最小二乗解の一意性について検討しましょう。3節とは逆に，一意性がなく最小二乗解が無限個あることが分かっている場合，つまり（14）の場合から考えましょう。（14）は

$$\boldsymbol{z} = l\mathbf{1}_n+k\boldsymbol{x},\ l \in \mathcal{R},\ k \neq 0 \tag{31}$$

のように，3本のベクトル $\mathbf{1}_n, \boldsymbol{x}, \boldsymbol{z}$ に関する線形式として表記されます。（31）のような線形関係が成り立つとき，3本のベクトル $\mathbf{1}_n, \boldsymbol{x}, \boldsymbol{z}$ は線形従属であるといいます。このとき行列の積 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ について考察しましょう。まず，（24）で

与えられる行列の積の定義より

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}=\begin{pmatrix}\mathbf{1}_n^{\mathrm{T}}\mathbf{1}_n & \mathbf{1}_n^{\mathrm{T}}\boldsymbol{x} & \mathbf{1}_n^{\mathrm{T}}\boldsymbol{z}\\ \boldsymbol{x}^{\mathrm{T}}\mathbf{1}_n & \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} & \boldsymbol{x}^{\mathrm{T}}\boldsymbol{z}\\ \boldsymbol{z}^{\mathrm{T}}\mathbf{1}_n & \boldsymbol{z}^{\mathrm{T}}\boldsymbol{x} & \boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}\end{pmatrix}$$

のように 3×3 行列が得られます。そして，（31）に由来する関係

$l\times$(1 行目の横ベクトル)$+k\times$(2 行目の横ベクトル)$=$(3 行目の横ベクトル)
$l\times$(1 列目の縦ベクトル)$+k\times$(2 列目の縦ベクトル)$=$(3 列目の縦ベクトル)

が成立するので，経営学部で学ぶ「線形代数入門」の学修内容からも $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ は逆行列を持たないことが分かるでしょう。

以下では，（31）のような線形関係が成り立つ，すなわち $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ が逆行列を持たない場合に，正射影 $\hat{a}\mathbf{1}_n+\hat{b}\boldsymbol{x}+\hat{c}\boldsymbol{z}$ を表現する係数ベクトル $(\hat{a}\ \hat{b}\ \hat{c})^{\mathrm{T}}$ が最小二乗解であるという観点から，最小二乗解の具体的な表現を求めましょう。まず複数のベクトルが張る線形部分空間には，包含関係

$$\begin{aligned}<\mathbf{1}_n,\boldsymbol{x},\boldsymbol{z}> &= \{a\mathbf{1}_n+b\boldsymbol{x}+c\boldsymbol{z} : a\in\mathcal{R}, b\in\mathcal{R}, c\in\mathcal{R}\}\\ &\supset \{a\mathbf{1}_n+b\boldsymbol{x} : a\in\mathcal{R}, b\in\mathcal{R}\}\\ &=<\mathbf{1}_n,\boldsymbol{x}>\end{aligned}$$

が成立することに注意しましょう。ただし，線形従属である 3 本のベクトルが生成する線形部分空間 $<\mathbf{1}_n,\boldsymbol{x},\boldsymbol{z}>$ の任意の元 $a\mathbf{1}_n+b\boldsymbol{x}+c\boldsymbol{z}$ は，（31）を考慮すると

$$a\mathbf{1}_n+b\boldsymbol{x}+c(l\mathbf{1}_n+k\boldsymbol{x})=(a+cl)\mathbf{1}_n+(b+ck)\boldsymbol{x}$$

のように（$\boldsymbol{z}$ を使わずに）$<\mathbf{1}_n,\boldsymbol{x}>$ の元として表現されるのです。従って

$$<\mathbf{1}_n,\boldsymbol{x},\boldsymbol{z}>=<\mathbf{1}_n,\boldsymbol{x}>$$

が成り立ちます。つまり，（31）のもとでは，重回帰分析の正射影は x のみを説明変数とした単回帰分析で得られる正射影（$\boldsymbol{y}$ の線形部分空間 $<\mathbf{1}_n,\boldsymbol{x}>$ への正射影）

$$\hat{\boldsymbol{y}}=\left(\bar{y}-\bar{x}\frac{s_{xy}}{s_{xx}}\right)\mathbf{1}_n+\frac{s_{xy}}{s_{xx}}\boldsymbol{x} \tag{32}$$

に一致するということです（(32) の $\mathbf{1}_n$ と $\boldsymbol{x}$ の係数は単回帰分析の最小二乗解です)。この正射影は，任意の実数 t，及び全ての成分が 0 の n 次元ベクトル $\mathbf{0}_n$ に対して，

$$\begin{aligned}\hat{\boldsymbol{y}} &= \left(\bar{y}-\bar{x}\frac{s_{xy}}{s_{xx}}\right)\mathbf{1}_n+\frac{s_{xy}}{s_{xx}}\boldsymbol{x}+t\mathbf{0}_n \\ &= \left(\bar{y}-\bar{x}\frac{s_{xy}}{s_{xx}}\right)\mathbf{1}_n+\frac{s_{xy}}{s_{xx}}\boldsymbol{x}+t(\boldsymbol{z}-l\mathbf{1}_n-k\boldsymbol{x}) \\ &= \left(\bar{y}-\bar{x}\frac{s_{xy}}{s_{xx}}-tl\right)\mathbf{1}_n+\left(\frac{s_{xy}}{s_{xx}}-tk\right)\boldsymbol{x}+t\boldsymbol{z} \qquad (33)\end{aligned}$$

のように $\mathbf{1}_n$, $\boldsymbol{x}$, $\boldsymbol{z}$ の線形和として表現できます。(33) における係数ベクトル

$$\hat{\boldsymbol{b}} = \begin{pmatrix}\bar{y}-\bar{x}\frac{s_{xy}}{s_{xx}}-tl & \frac{s_{xy}}{s_{xx}}-tk & t\end{pmatrix}^{\mathrm{T}}$$

が最小二乗解であり，3 節の（16）でも指摘した通り，t の任意性から無限個の解があるわけです。正射影は一意性を持ちますが，3 本のベクトルが線形従属ならば，正射影を表現する係数ベクトルは一意性を持たないのです。

一方，3 本のベクトル $\mathbf{1}_n$, $\boldsymbol{x}$, $\boldsymbol{z}$ が（31）のような線形関係を持たない場合，線形独立であるといいます。このとき，3×3 行列 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ は

$$(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} = \begin{pmatrix}1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1\end{pmatrix}$$

を満たす逆行列 $(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}$ を持ちます。逆行列 $(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}$ を連立方程式 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{b} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$ の両辺の左からかけることで最小二乗解の一意な表現

$$\hat{\boldsymbol{b}} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$$

を得ます。この 3 次元ベクトルの各成分は 3 節の（13）で導出されています。

7　残差ベクトルと重回帰分析の可視化

最後に残差ベクトル

$$\boldsymbol{e} = \boldsymbol{y} - \hat{\boldsymbol{y}} = \boldsymbol{y} - X\hat{\boldsymbol{b}}$$

の性質を整理しましょう。前節で説明したように最小二乗解 $\hat{\boldsymbol{b}}$ が一意性を持たない場合でも正射影 $\hat{\boldsymbol{y}} = X\hat{\boldsymbol{b}}$ には一意性があります。従って，$\boldsymbol{e}$ にも一意性があります。連立方程式（29）に解 $\hat{\boldsymbol{b}}$ を代入すると

$$\begin{pmatrix} \mathbf{1}_n^{\mathrm{T}}\boldsymbol{e} \\ \boldsymbol{x}^{\mathrm{T}}\boldsymbol{e} \\ \boldsymbol{z}^{\mathrm{T}}\boldsymbol{e} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

を得ます。つまり $\boldsymbol{e}$ は $\mathbf{1}_n, \boldsymbol{x}, \boldsymbol{z}$ と直交します。さらに内積の性質（18）により，

$$(a\mathbf{1}_n + b\boldsymbol{x} + c\boldsymbol{z})^{\mathrm{T}}\boldsymbol{e} = a\mathbf{1}_n^{\mathrm{T}}\boldsymbol{e} + b\boldsymbol{x}^{\mathrm{T}}\boldsymbol{e} + c\boldsymbol{z}^{\mathrm{T}}\boldsymbol{e} = 0$$

が成り立つので，$\boldsymbol{e}$ は線形部分空間 $< \mathbf{1}_n, \boldsymbol{x}, \boldsymbol{z} >$ の任意の元と直交します。もちろん $\hat{\boldsymbol{y}} = \hat{a}\mathbf{1}_1 + \hat{b}\boldsymbol{x} + \hat{c}\boldsymbol{z}$ と表現される正射影と $\boldsymbol{e}$ も直交します。従って，（21）よりピタゴラス関係

$$\|\boldsymbol{y}\|^2 = \|\hat{\boldsymbol{y}} + \boldsymbol{e}\|^2 = \|\hat{\boldsymbol{y}}\|^2 + \|\boldsymbol{e}\|^2$$

が成り立ちます。これらの性質を可視化したのが図 2 です。

ところで，$\boldsymbol{e}^{\mathrm{T}}\mathbf{1}_n = 0$ は

$$\sum_{i=1}^{n} e_i = 0 \tag{34}$$

と成分表示されます。図 1 の縦線で表されるのが残差であり，（34）のように符号と長さの帳尻が合うように平面が置かれるのはイメージできると思います。一方 $\boldsymbol{e}^{\mathrm{T}}\boldsymbol{x} = 0$ およびその成分表示

$$\sum_{i=1}^{n} e_i x_i = 0$$

は，図 1 からはイメージできない性質ですが，図 2 からは明らかです。このように 3 次元空間の n 個の点（図 1）と，n 次元ベクトル $\boldsymbol{y}, \hat{\boldsymbol{y}}, \boldsymbol{e}$ と線形部分空間 $< \mathbf{1}_n, \boldsymbol{x}, \boldsymbol{z} >$ の可視化（図 2）は，いずれも重回帰分析のイメージを掴むのに重要なのです。

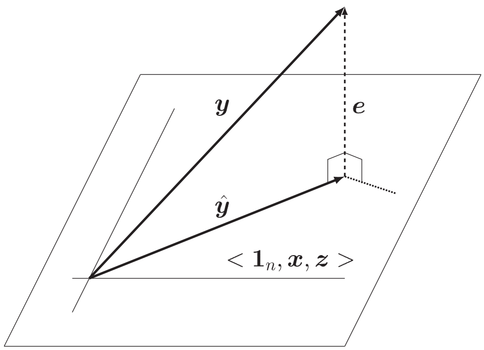

図２　重回帰分析の可視化

参考文献

佐和隆光（2020）回帰分析（新装版）　朝倉書店

竹村彰通（2007）統計（第２版）　共立出版

田中久稔（2019）計量経済学のための数学　日本評論社