経営科学（日本オペレーションズ・リサーチ学会邦文機関誌） 第14巻 第4号（1971年3月）

Iteration による最小自乗法の解法とある領域における最小自乗法[†]

野　田　竜　夫[*]

1. 序　　論

$A=(a_{ij})$ を $m\times n$ 行列，A^* を A の転置行列，$x=(x_j)$ および x^* をそれぞれ n 次元列および行ベクトル，$b=(b_i)$ を m 次元列ベクトル，R^n を n 次元 Euclid 空間とする．また A の rank が n であるとし $m>n$ とする．

このとき

$$E(x)=\sum_{i=1}^{m}(\sum_{j=1}^{n}a_{ij}x_j-b_i)^2$$

の値を最小にする $x=\bar{x}\varepsilon R^n$ を求める方法は最小自乗法として知られている．第2節でのべるように，$x=\bar{x}$ は方程式 $A^*A\bar{x}=A^*b$ で与えられる．この論文では1つの簡単な iteration を行なうことによって，$x=\bar{x}\varepsilon R^n$ が得られることを第3節で示す．さらに R^n のある部分集合 B の中で $E(x)$ の値を最小にする解が存在することを第4節で証明する．他方，$\bar{x}\varepsilon B$ の場合にこの iteration が B の中で $E(x)$ の値を最小にする解を求めるのに応用できることを第5節で示す．第6節では，第3節の iteration についての注意および数値例をのべる．

2. 最小自乗法

まず，準備としてつぎの補助定理1および2から始めよう．行列 A の rank を rank(A) とおくことにする．

補助定理 1. $\mathrm{rank}(A)=\mathrm{rank}(A^*A)=\mathrm{rank}(AA^*A)$.

証明. $\mathrm{rank}(A)=\mathrm{rank}(A^*A)$ が成り立つことはよく知られているから，$\mathrm{rank}(A^*A)=\mathrm{rank}(AA^*A)$ が成り立つことを示す．それは

$$X_1=\{x\,|\,A^*Ax=0\},$$
$$X_2=\{x\,|\,AA^*Ax=0\}$$

とおくとき $X_1=X_2$ であることを証明すればよい．任意の $x\varepsilon X_1$ に対して $A^*Ax=0$，したがっ

---

† 1969年9月8日受理.

* 富山県立大谷技術短期大学.

© 日本オペレーションズ・リサーチ学会．無断複写・複製・転載を禁ず．

て $AA^*Ax=0$. ゆえに $x\varepsilon X_2$.

逆に任意の $x\varepsilon X_2$ に対して $AA^*Ax=0$,

したがって, $x^*A^*(AA^*Ax)=0$, すなわち $\|A^*Ax\|^2=0$. A^*Ax の成分はすべて実数であることから $A^*Ax=0$. ゆえに $x\varepsilon X_1$. (証明終り)

補助定理 2. $m\times n$ 行列Aの rank が n であるとして

$$E(x)\equiv\|Ax-b\|^2=\sum_{i=1}^{m}(\sum_{j=1}^{n}a_{ij}x_j-b_i)^2$$

の値を最小にする $x=\bar{x}\varepsilon R^n$ は方程式

(2.1) $A^*A\bar{x}=A^*b$

をみたし $\bar{x}=(A^*A)^{-1}A^*b$ で与えられる.

さらに任意の $x\varepsilon R^n$ に対して

(2.2) $E(x)-E(\bar{x})=\|A(x-\bar{x})\|^2$

および

(2.3) $E(x)>E(x)$, $x\neq\bar{x}$

が成り立つ.

証明. $E(x)-E(\bar{x})$

$$\begin{aligned}&=\|Ax-b\|^2-\|A\bar{x}-b\|^2\\&=\|Ah+A\bar{x}-b\|^2-\|A\bar{x}-b\|^2\\&=\|Ah\|^2+2(h,\ A^*(A\bar{x}-b)),\end{aligned}$$

ここに $h=x-\bar{x}$ であり, (,) は内積をあらわす. また

$$\begin{aligned}\frac{1}{2}\cdot\frac{\partial E(\bar{x})}{\partial x_k}&=\sum_{i=1}^{m}a_{ik}(\sum_{j=1}^{n}a_{ij}\bar{x}_j-b_i)\\&=\sum_{j=1}^{n}(\sum_{i=1}^{m}a_{ik}a_{ij})\bar{x}_j-\sum_{i=1}^{m}a_{ik}b_i\\&=0\quad(k=1,2,\cdots\cdot,n)\end{aligned}$$

が成り立つことから

$$A^*A\bar{x}-A^*b=A^*(A\bar{x}-b)=0$$

が得られる. 補助定理1が成り立つことからこの方程式を解けば $\bar{x}=(A^*A)^{-1}A^*b$.

ゆえに

$$\begin{aligned}E(x)-E(\bar{x})&=\|Ah\|^2+2(h,0)\\&=\|Ah\|^2.\end{aligned}$$

いま $h=x-\bar{x}\neq0$ のとき

$$E(x)-E(\bar{x})=\|Ah\|^2=0$$

であるとすれば $h\neq0$ に対して $Ah=0$.

これはAの rank が n であることに矛盾する.

したがつて (2.2) より

© 日本オペレーションズ・リサーチ学会. 無断複写・複製・転載を禁ず.

$E(x)-E(\bar{x})>0$，$x \neq \bar{x}$.（証明終り）

3. R^n の中での iteration

方程式（2.1）を解かないで $E(x)$ の値を最小にする $x=\bar{x}\varepsilon R^n$ を求めるのに反復手順

(3.1) $\quad y_k+1=y_k-\alpha\ (A^*Ay_k-A^*b) \qquad (k=0,\ 1,\ 2,\ \cdots\cdot)$

を任意の適当に選ばれた初期値 y_0 に対して行なう．ここに λ は

(3.2) $$\lambda=\frac{1}{\sum_{i=1}^{m}(\sum_{j=1}^{n}a_{ij}^2)}$$

によって定められる正数である．このとき，(3.1) によってきまる点列 $\{y_k\}$ に対して定理1が成り立つ．

定理 1. $m\times n$ 行列 A の rank が n であるとして，(2.1) および (3.2) のもとに (3.1) によってきまる点列 $\{y_k\}$ に対して

(3.3) $\quad E(y_{k+1})\leqq E(y_k) \qquad (k=0,1,2,\cdots\cdot)$

および

(3.4) $\quad \|y_{k+1}-\bar{x}\|\leqq M\|y_k-\bar{x}\| \qquad (k=0,1,2,\cdots\cdot)$

が成り立つ．定数 M $(0\leqq M<1)$ は証明の中で決定される．

(3.3) の証明. $h_k=y_k-\bar{x}$ とおけば

$$h_{k+1}=y_{k+1}-\bar{x}=h_k-\lambda A^*Ah_k$$

が成り立ち，したがって (2.2) を用いれば

$$\begin{aligned}&E(y_{k+1})-E(y_k)\\&\quad=\{E(y_{k+1})-E(\bar{x})\}-\{E(y_k)-E(\bar{x})\}\\&\quad=\|Ah_{k+1}\|^2-\|Ah_k\|^2\\&\quad=\|Ah_k-\lambda AA^*Ah_k\|^2-\|Ah_k\|^2.\end{aligned}$$

ゆえに

(3.5) $$\begin{aligned}&E(y_{k+1})-E(y_k)\\&\quad=\|AA^*Ah_k\|^2\lambda^2-2\|A^*Ah_k\|^2\lambda.\end{aligned}$$

補助定理1が成り立つことから $h_k\neq 0$ に対して $\|AA^*Ah_k\|^2>0$.

また Schwarz の不等式より

$$\|AA^*Ah_k\|^2\leqq(\sum_{i=1}^{m}(\sum_{j=1}^{n}a_{ij}^2))\cdot\|A^*Ah_k\|^2$$

が成り立つので λ は

$$0<\lambda<\frac{2\|A^*Ah_k\|^2}{\|AA^*Ah_k\|^2}=\frac{2\|A^*A\rho_k\|^2}{\|AA^*A\rho_k\|^2}.$$

ここに

$$h_k=\|h_k\|\frac{h_k}{\|h_k\|}=\|h_k\|\rho_k,\quad \|\rho_k\|=1$$

© 日本オペレーションズ・リサーチ学会．無断複写・複製・転載を禁ず．

とおく．λ のこの範囲の値に対して (3.5) より

$$E(y_{k+1}) \leqq E(y_k).$$

(3.4) の証明． $h_{k+1}=h_k-\lambda A^*Ah_k$ であることから

$$\|h_{k+1}\|^2=\|h_k\|^2-2\|Ah_k\|^2\lambda+\|A^*Ah_k\|^2\lambda^2,$$

すなわち

$$\begin{aligned}\|h_{k+1}\|^2-\|h_k\|^2&=\|A^*Ah_k\|^2\lambda^2-2\|Ah_k\|^2\lambda\\&=\left(\frac{\|A^*Ah_k\|^2}{\|h_k\|^2}\lambda^2-2\frac{\|Ah_k\|^2}{\|h_k\|^2}\lambda\right)\cdot\|h_k\|^2\\&=(\|A^*A\rho_k\|^2\lambda^2-2\|A\rho_k\|^2\lambda)\cdot\|h_k\|^2.\end{aligned}$$

また Schwarz の不等式より

$$\|A^*A\rho_k\|^2\leqq\frac{1}{\lambda}\|A\rho_k\|^2$$

が成り立つことから

$$\begin{aligned}\|h_{k+1}\|^2-\|h_k\|^2&\leqq(\|A\rho_k\|^2\lambda-2\|A\rho_k\|^2\lambda)\cdot\|h_k\|^2\\&=\left(-\frac{\|A\rho_k\|^2}{\sum_{i=1}^{m}(\sum_{j=1}^{n}a_{ij}^2)}\right)\cdot\|h_k\|^2.\end{aligned}$$

したがって

$$\|h_{k+1}\|^2\leqq\left(1-\frac{\|A\rho_k\|^2}{\sum_{i=1}^{m}(\sum_{j=1}^{n}a_{ij}^2)}\right)\cdot\|h_k\|^2.$$

一方，$\|A\rho\|$ は compact set $\{\rho\,|\,\|\rho\|=1\}$ で連続であることから

$$K=\min_{\|\rho\|=1}\|A\rho\|^2$$

を満足するKが存在し，しかも補助定理 1 が成り立つことから，$K>0$．よって

$$\|h_{k+1}\|^2\leqq\left(1-\frac{K}{\sum_{i=1}^{m}(\sum_{j=1}^{n}a_{ij}^2)}\right)\cdot\|h_k\|^2.$$

いま

$$M=\sqrt{1-\frac{K}{\sum_{i=1}^{m}(\sum_{j=1}^{n}a_{ij}^2)}}$$

とおけば，Mは $0\leqq M<1$ をみたし

$$\|h_{k+1}\|\leqq M\|h_k\|.$$ (証明終り)

4.　領域 B の中での解の存在

$$B=\{x\,|\,x_1\geqq0,\ x_2\geqq0,\cdots,\ x_n\geqq0\}$$

とおく．

© 日本オペレーションズ・リサーチ学会．無断複写・複製・転載を禁ず．

定理 2. (解の存在定理). $m \times n$ 行列Aの rank がnであるとして, (2.1) のもとに領域Bにおいて, $E(x)=\|Ax-b\|^2$ の値を最小にする $x=x_0$ は

(1) $\bar{x} \varepsilon B$ ならば $x_0=\bar{x}$ である.

(2) $\bar{x} \varepsilon R^n-B$ ならば x_0 は B の境界上に存在する. ここに $\bar{x}$ は (2.1) の解である.

証明. (1)の場合は (2.3) から明らかである. $\bar{x} \varepsilon R^n-B$ ならば x_0 はBの内点にはなり得ないことを示そう. かりに x_0 が B の内点であるとすれば, 線分 $\bar{x}x_0$ が B の境界と最初に交わる点を y として

$$p=x_0-\bar{x},$$
$$q=y-\bar{x}$$

とおけば, ベクトル p, q は同一直線上にあつて $q=\alpha p$ を満足する α $(0<\alpha<1)$ が存在する. (2.2) を用いれば

$$\begin{aligned} E(y)-E(\bar{x}) &= \|Aq\|^2 \\ &= \|A(\alpha p)\|^2 \\ &= |\alpha|^2 \cdot \|Ap\|^2 \\ &< \|Ap\|^2 \\ &= E(x_0)-E(\bar{x}). \end{aligned}$$

ゆえに $E(y)<E(x_0)$. 一方, y の定義により B の境界を ∂B と書けば, $y \varepsilon \partial B \subset B$.

ゆえに $x=x_0$ は領域 B において $E(x)$ の値を最小にすることはない. したがって x_0 は B の内点にはなり得ない.

つぎに x_0 が ∂B の中に存在することを示そう. いま, $E(x)=C$ とし, 二次形式

(1) $$\begin{aligned} C &= E(\bar{x})+\|Ah\|^2 \\ &= E(\bar{x})+h^*A^*Ah, \end{aligned}$$
$$h=x-\bar{x}$$

を考える. Aの rank がnであることから実対称行列 A^*A は positive definite である. それゆえ, ある適当な正則変換 $\bar{h}=D^{-1}h$ によって

(2) $$h^*A^*Ah=\bar{h}^*D^*A^*AD\bar{h}.$$

いま, $\lambda_j (j=1,2,\cdots\cdot,n)$ を A^*A の固有値とするとき

(3) $$D^*A^*AD=\mathrm{diag}\ (\lambda_1,\ \lambda_2,\ \cdots\cdot,\lambda_n),$$
$$\lambda_j>0 (j=1,2,\cdots\cdot,n)$$

が成り立ち(1), (2)および(3)によって

(4) $$E(\bar{x})+\sum_{j=1}^{n}\lambda_j\bar{h}_j^2=C.$$

二次形式(4)は $x=\bar{x}$ を中心とする楕円面群をつくる. Cの値を次第に増加させたとき, この楕円面が ∂B と最初に共通点 x_0 をもつとき, $x=x_0$ が領域Bにおいて $E(x)$ の値を最小にする解であり, このときのCに対して, $\|Ax_0-b\|^2=C$ が成り立つ. (証明終り)

© 日本オペレーションズ・リサーチ学会. 無断複写・複製・転載を禁ず.

5. $\bar{x}$ が B に属する場合の第3節の iteration

第3節では $\bar{x}\varepsilon R^n$ および $y_k\varepsilon R^n(k=0,1,2,\cdots\cdot)$ の場合における反復手順 (3.1) を考え，その結果として (3.3) および (3.4) の不等式関係が成立した．ここでは $\bar{x}$ が B に属する場合の iteration について考える．

(1) $\bar{x}$ が B の内点である場合.

(3.4) の不等式関係から適当な正数 N_0 が存在して，$k \geqq N_0$ であるすべての k に対して

(5.1) $y_k\varepsilon B$

が成り立ち

(5.2) $\lim\limits_{k\to\infty} y_k=\bar{x}$

であることがわかる.

(2) $\bar{x}$ が B の境界点である場合.

これを2つの場合にわけて

(A1) 任意の正数Nに対し $k(\geqq N)$ が存在して $y_k\varepsilon B$ である場合には $y_k\varepsilon B$ をみたす点列 $\{y_k\}$ の部分列 $\{y_{nk}\}$ をとる．そのとき

$$y_{nk}\varepsilon B \ (k=1,2,\cdots\cdot)$$

および

$$\lim_{k\to\infty} y_{nk}=\bar{x}.$$

$E(x)$ の連続性から任意の正数 ε に対して正数 N_1 と δ が存在して $k \geqq N_1$ であるすべての k に対して

$$\|y_{nk}-\bar{x}\|<\delta \text{ であるから } E(y_{nk})-E(\bar{x})<\varepsilon$$

が成り立つ.

(A2) ある十分大きい正数 N_2 が存在して，$k \geqq N_2$ であるすべての k に対して，$y_k\varepsilon R^n-B$ である場合には，ベクトル y_k の成分のうち負である成分をすべて0でおきかえたベクトルを $\bar{y}_k$ とすれば

$$\|\bar{y}_k-\bar{x}\|<\|y_k-\bar{x}\|$$

が成立する．(3.4) の関係式から正数 N_3 が存在して $k \geqq N_3$ ならば

$$\|y_k-\bar{x}\|<\delta \text{ であるから } E(y_k)-E(\bar{x})<\varepsilon.$$

いま，$N_4=max(N_2,\ N_3)$ とおけば $k\geqq N_4$ であるすべての k に対して

(5.3) $\bar{y}_k\varepsilon\partial B\subset B$

および

(5.4) $\|\bar{y}_k-\bar{x}\|<\delta$ であるから $E(\bar{y}_k)-E(\bar{x})<\varepsilon.$

© 日本オペレーションズ・リサーチ学会. 無断複写・複製・転載を禁ず.

6. 第3節の定理1についての注意および数値例

(6.1) $y_{k+1}=y_k-\mu(A^*Ay_k-A^*b)$

$(k=0,1,2,\cdots)$

なる iteration を考える．第3節の定理1の証明をみると $0<\mu<2\lambda$ をみたす μ に対しては

(6.2) $0<\mu<\dfrac{2\|A^*A\rho_k\|^2}{\|AA^*A\rho_k\|^2}$

および

(6.3) $0<\mu<\dfrac{2\|A\rho_k\|^2}{\|A^*A\rho_k\|^2}$

が成り立つことは明らかであるが，$\mu\geqq 2\lambda$ なる μ に対しても (6.2)，(6.3) が成り立つことがある．これは行列Aに関係するものである．

A				b
.6731	−.4135	.7213	.1783	.6471
.2948	.5326	−.3471	.8272	.2538
.1238	.3267	.5197	.2690	.8933
−.6292	.9235	.3578	.4275	.2283
.7530	.1497	.2193	−.1976	.1009
.8105	−.1215	.7068	.5320	.3478

なる6×4行列A，6次元列ベクトルbに対して (6.1) の反復手順を初期値

$y_0^*=(-1.000,\ -1.000,\ -1.000,\ -1.000)$

として本学の電子計算機 OKITAC-5090C を用いて計算した．$\mu=t\lambda$ として，t を0.1から4.0まで0.1きざみで変えたとき，$t=3.5$ に対する iteration (6.1)，すなわち

$y_{k+1}=y_k-3.5\lambda(A^*Ay_k-A^*b)$

を採用するのが最良であることがわかった．

このとき

$y_k^*=(.0967,\ .1299,\ .6030,\ .3161)$

となり，$A^*A\bar{x}=A^*b$ の解

$\bar{x}^*=(.0967,\ .1299,\ .6029,\ .3160)$

とほとんど一致している．

終りに，この研究をまとめるにあたり，富山大学文理学部数学教室の田中専一郎教授には終始ご指導をいただいた．また九州大学理学部数学教室の工藤昭夫教授には二，三のご注意と激励をいただいた．また第6節の数値例は本学の応用数学科の浦山和子助手によって計算されたものである．ここに両教授および浦山和子氏に深く感謝の意を表わします．

© 日本オペレーションズ・リサーチ学会．無断複写・複製・転載を禁ず．

参 考 文 献

[1] 田中専一郎, "最小自乗法における Jacobi の Algorithm について," 京都大学数理解析研究所講究録37 (1968), 127-140.

[2] ——, "最小自乗法における Algorithm について," 京都大学数理解析研究所講究録 42 (1968), 37-51.

[3] 野田竜夫, "ある領域における最小自乗法について," 日本オペレーションズ・リサーチ学会春季研究発表会アブストラクト集 (1968), 13-14.

[4] ——, "ある領域における最小自乗法について (第2報)," 日本オペレーションズ・リサーチ学会春季研究発表会アブストラクト集 (1969), 87-88.

© 日本オペレーションズ・リサーチ学会. 無断複写・複製・転載を禁ず.