統計数理(2005)
第 53 巻 第 2 号 391–404

特集「計算推論—モデリング・数理・アルゴリズム—」
［研究ノート］

層別最小二乗法

—重み付き最小二乗法の極限—

土谷 隆[†]

（受付 2005 年 7 月 25 日；改訂 2005 年 10 月 6 日）

要　　旨

重み付き最小二乗法は通常の最小二乗法のごく単純な一般化であり，適切な座標変換を施せば通常の最小二乗法に帰着する．ところが重みの間の比を無限大にした極限−これは重みに辞書式順序による大小関係が想定されている場合として考えることもできる−は座標変換によって取り扱うことができない例外的な場合である．実は，この意味での重み付き最小二乗解の極限が存在することが知られている．これを層別最小二乗解と呼ぶ．本論文では，重み付き最小二乗解と層別最小二乗解の差のノルムを「重み付き最小二乗問題の重みの比」と「最小二乗問題を定義する係数行列に関するある種の条件数」の関数として評価する．

キーワード： 重み付き最小二乗法，条件数，内点法，層別最小二乗法．

1. 重み付き最小二乗問題とその極限

1.1 重み付き最小二乗問題

本論文では重みつき最小二乗問題

$$\min_x \sum_{i=1}^{N} d_i^2 \|A_i x - b_i\|^2 \tag{1.1}$$

を考える．ここで，各 $i=1,\ldots,N$ について，A_i は $n_i \times m$ 行列，b_i は n_i 次元ベクトル，d_i は正数とし，x は m 次元ベクトルとする．本論文ではベクトルは縦(列)ベクトルとし，ノルム記号 $\|\cdot\|$ は常にベクトルのユークリッドノルムおよびそれより誘導される作用素ノルムを表すものとする．また，$k \times k$ 単位行列を I_k と記す.（ただし，次元が明らかな場合は I とする.）そして列数が等しい行列，例えば $C_1,\ldots,C_N$ を縦に繋げた行列を $[C_1;\cdots;C_N]$ と記すことにする．ベクトルについても同様の記法を採用する．正方行列 $C_1,\ldots,C_k$ が与えられた時，第 $i \times i$ ブロックが C_i である(k ブロック)$\times$ (k ブロック)ブロック対角行列を $\mathrm{diag}(C_1,\ldots,C_k)$ と記す．

以下，

$$A = [A_1;\cdots;A_N] = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_N \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix} \tag{1.2}$$

とおく．そして A の列ベクトルは一次独立であると仮定する．

[†] 統計数理研究所：〒106-8569 東京都港区南麻布 4–6–7

重み付き最小二乗問題(1.1)の解は，$D = \mathrm{diag}(d_1 I_{n_1}, d_2 I_{n_2}, \ldots, d_N I_{n_N})$ と置くと，

$$x = (A^T D^2 A)^{-1} A^T D^2 b$$

と書ける．各 d_i を重みと呼び，$d = (d_1, \ldots, d_N)$ を重みベクトルと呼ぶ．一般性を失うことなく，重み d_i は i について単調減少であるとする．本論文では，重み間の比率が無限大となった時の最小二乗問題の解の漸近的挙動について論ずる．

2. 層別最小二乗法 ── 最小二乗法の極限 ──

最小二乗法では残差 $Ax - b$ をユークリッドノルムの意味で小さくするように x を定める．ここで，残差の特定の要素の大きさを小さくすることが特に重要であるとすれば，対応する要素に相対的に大きな重みをつけて重みつき最小二乗法を行えば良いと考えられる．ここで，重み d_i が A_i と b_i から定まる残差の部分 $A_i x - b_i$ の重要度を表すとして，$d_1, \ldots, d_N$ に，$d_1 \succ d_2 \succ \cdots \succ d_N$ という「辞書的な順序」がついているとすれば，重みつき最小二乗法の直観的な意味での極限として，『まず他の部分を無視して部分行列 A_1 に関する最小二乗法を行い，次いで A_2 に関する最小二乗法を「A_1 に関する最小二乗解の残差に影響を与えないという条件の下で」行い $\cdots$』という以下のような手続きが考えられる．これを層別最小二乗法ということにする．

層別最小二乗法

(1) $i = 1$ とする．$x^{(0)} = 0$ とする．

(2) $\min\ \|(A_i(x^{(i-1)} + \Delta x) - b_i)\|^2$ を，等号条件 $A_1 \Delta x = 0, \ldots, A_{i-1}\Delta x = 0$ の下で解く．$x^{(i-1)} + \Delta x$ を $x^{(i)}$ とする．ただし，$i = 1$ の時は，等号条件は付けずに $\|A_1 \Delta x - b_1\|^2$ の最小化問題を解く．

(3) $i = N$ ならば，$x^{(i)}$ を解として返す．さもなくば，$i := i + 1$ として，第 2 段に戻る．

ここで，b_i は A_i に対応する n_i 次元ベクトルである．

以下，

$$L_0 = R^m, \quad L_i = \{\Delta x \in R^m :\ A_1 \Delta x = 0, \ldots, A_i \Delta x = 0\} \quad (i = 1, \ldots, N)$$

と定義する．すると，第 2 段で解くべき最適化問題は

$$\min\ \|(A_i(x^{(i-1)} + \Delta x) - b_i)\|^2 \ \text{s.t.}\ \ \Delta x \in L_{i-1}$$

と記すことができる．

さて，Vavasis and Ye (1996)は，元の重み付き最小二乗問題の解において，d_i/d_{i+1} $(i = 1, \ldots, N-1)$ が無限大となる極限を取ったとき，層別最小二乗解が極限として得られることを示した．本原稿の目的は，最小二乗問題の残差空間の表現形式として基底ブロック下三角行列を導入してこの事実の別証明を与え，重み付き最小二乗解と層別最小二乗解の間の差のノルムが $\mathcal{O}((\chi_A \bar{\chi}_A^{N+2} \max(d_{i+1}/d_i))^2)$ と評価できることを示すことである．ここで，χ_A や $\bar{\chi}_A$ は係数行列 A の「ある種の条件数」である．

以下，行列 A が層 $A_1, \ldots, A_N$ からなるものであることを特に強調する時には，これを層行列と呼ぶことにし，$A = [A_1; \cdots; A_N]$ などと記すことにする．次の事実が成立する．

命題 2.1. 列 1 次独立な層行列 $A = [A_1; \cdots; A_N]$ に対する層別最小二乗法の解は一意に定まる．

証明. その作り方より，層別最小二乗法の解 x の残差は一意に定まる．残差を r とすると，

$Ax-b=r$ となり，この方程式は解を持つ．解の一意性は A の列一次独立性より従う． □

3. ブロック下三角行列と基底ブロック下三角行列

3.1 ブロック下三角行列

以下,

$$\bar{A}_i=\begin{pmatrix} A_1 \\ \vdots \\ A_i \end{pmatrix}$$

とする．$\bar{A}_i$ の階数を，$\bar{r}_i$ とする．$\bar{r}_i$ は i の単調増加関数である．また，$r_i=\bar{r}_i-\bar{r}_{i-1}$ とする．ただし，$\bar{r}_0=0$ である．正則行列 G を適当に選ぶことにより，AG^{-1} を以下のような（N ブロック）$\times$（N ブロック）のブロック下三角行列に変換することができる．これが本論文において基本的な役割を果たす.

定理 3.1. （1.2）の A に対し，$T=AG^{-1}$ が以下の性質を有するような正則行列 G が存在する．

（1） $T=AG^{-1}$ は，$N\times N$ のブロック下三角行列である．

$$T=\begin{pmatrix} T_1 \\ \vdots \\ T_N \end{pmatrix}=\begin{pmatrix} T_{11} & 0 & \cdots & 0 \\ T_{21} & T_{22} & 0 & 0 \\ \vdots & \vdots & \vdots & 0 \\ T_{N1} & T_{N2} & \cdots & T_{NN} \end{pmatrix}. \tag{3.1}$$

ここで，T_{ij} は，$n_i\times r_j$ 行列である．

（2） $\mathrm{Rank}\ T_{ii}=\bar{r}_i-\bar{r}_{i-1}=r_i$，つまり，$T_{ii}$ は列一次独立である．ここで，$\bar{r}_0=0$ とする．

（3） $k=1,\ldots,N$ について，

$$\mathrm{Im}\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_k \end{pmatrix}=\mathrm{Im}\begin{pmatrix} T_{11} & 0 & \cdots & 0 \\ T_{21} & T_{22} & 0 & 0 \\ \vdots & \vdots & \vdots & 0 \\ T_{k1} & T_{k2} & \cdots & T_{kk} \end{pmatrix}.$$

ここで，$r_i=r_{i-1}$ であっても，形式的に T_{ij} をブロック要素と見なす．

証明. $\mathrm{Im}(A_1)$ の基底となる列ベクトルを A_1 の中から r_1 本選び，これらのベクトルに対応する A の列ベクトルが一番左側にくるように A を列交換する（列交換は列基本変形で行える）．そして，A_1 の残りのベクトルを列基本変形して消去して 0 とすることにより，第 1 行ブロックを T の形に変換することができる．A を列基本変形した結果としてこの段階で得られる行列を $T^{(1)}$ と表す．

次いで $[A_1;A_2]$ について考える．$T^{(1)}$ において，第 2 列ブロックから第 N 列ブロックに関し，第 1 段と同様の列基本変形を施すことにより，$(T_{21}^{(1)}\ \cdots\ T_{2N}^{(1)})$ を $(*\ *\ 0\ \cdots\ 0)$ の形にすることができる．この段階で得られる行列を $T^{(2)}$ とする．この時，

$$\begin{pmatrix} T_{11}^{(2)} & 0 \\ T_{21}^{(2)} & T_{22}^{(2)} \end{pmatrix}=\begin{pmatrix} T_{11}^{(1)} & 0 \\ T_{21}^{(1)} & T_{22}^{(2)} \end{pmatrix}$$

は列一次独立で，そのランクは明らかに $\bar{r}_2$，そして $T_{22}^{(2)}$ には $r_2=\bar{r}_2-\bar{r}_1$ 本の 1 次独立な列ベクトルが並ぶ．$T_{22}^{(2)}$ が T_{22} となる．以下同様にして，A より列基本変形のみによって，(1)，(2)を満たすような T を作ることができる．したがって，適当な正則行列 G により，$AG^{-1}=T$ とできる．$AG^{-1}=T$ なので，(3)が成立することは自明． □

$A=[A_1;\cdots;A_N]$ に対し，定理 3.1 を満たす(3.1)の層行列 $T=[T_1;\cdots;T_N]$ を層行列 $A=[A_1;\cdots;A_N]$ に対応するブロック下三角行列と呼ぶことにする．ブロック下三角行列は Tsuchiya (1992)で内点法に現れる射影行列の漸近的挙動を解析するのに用いられ，また，Megiddo et al. (1998)においても層別最小二乗解の効率的計算のために用いられている．

以下，$k=1,\ldots,N$ に対し，

$$\bar{T}_k=\begin{pmatrix} T_{kk} & 0 & \cdots & 0 \\ T_{k+1,k} & T_{k+1,k+1} & 0 & 0 \\ \vdots & \vdots & \vdots & 0 \\ T_{N,k} & T_{N,k+1} & \cdots & T_{NN} \end{pmatrix} \tag{3.2}$$

と置く．

$T=AG^{-1}$ が定理 3.1 の条件を満たすブロック下三角行列となる場合，T は $N\times N$ ブロック行列となる．これに対応して，変数も層に対応した部分ベクトルに分けることができ，$u=[u_1;\cdots;u_N]$ と表せる $(u_i\in R^{r_i}, i=1,\ldots,N)$．ところが T が定理 3.1 の条件を満たすブロック下三角行列であるため，

$$Tu \text{ の第 } 1 \text{ 層成分から第 } k-1 \text{ 層成分が } 0 \Longleftrightarrow [u_1;\cdots;u_{k-1}]=0$$

が成立する．これより，

$$AL_{k-1}=\{Tu:[u_1;\cdots;u_{k-1}]=0\}$$

と書ける．したがって，T (の各列の一次結合)を $\mathrm{Im}(A)$ の表現として採用すると，層別最小二乗法は，以下のように，$u_1,\ldots,u_N$ を逐次定めていく，より簡単な手続きとして表せる．

ブロック下三角行列に対する層別最小二乗法

(1) $i=1$ とする．$r^{(0)}=b$.

(2) $u_i=(T_{ii}^T T_{ii})^{-1}T_{ii}^T r_i^{(i-1)}$

(3) $[r_i^{(i)};\cdots;r_N^{(i)}]=[r_i^{(i-1)};\cdots;r_N^{(i-1)}]-[T_{ii};\cdots;T_{Ni}]u_i$ とする．($r_i^{(i)}$ は層別最小二乗法の第 i 層の残差となる．)

(4) $i=N$ ならば終了；さもなくば，$i:=i+1$ として，第 2 段に戻る．

右辺ベクトル $[r_k;\cdots;r_N]$ $(r_i\in R^{n_i})$ が与えられた時に(3.2)の層行列 $\bar{T}_k=[T_k;\cdots;T_N]$ に関する層別最小二乗解を求める関数を $F(\bar{T}_k,[r_k;\cdots;r_N])$ とすると，この関数は次のように再帰的に定義できる．

層別最小二乗解を求める関数 F の再帰的定義

(1) $k=N$ ならば，$F=(T_{NN}^T T_{NN})^{-1}T_{NN}^T r_N$.

(2) そうでなければ，

$$[u_k;\cdots;u_N]=((T_{kk}^T T_{kk})^{-1}T_{kk}^T r_k; F(\bar{T}_{k+1},[r_{k+1};\cdots;r_N]-[T_{k+1,k};\cdots;T_{Nk}](T_{kk}^T T_{kk})^{-1}T_{kk}^T r_k))$$

次の命題が成立する．

命題 3.2. $T=AG^{-1}$ を $A=[A_1;\cdots;A_N]$ をブロック下三角行列とするような正則行列 G とする．$b=[b_1;\cdots;b_N]$ を右辺とし，像空間 $\mathrm{Im}(A)$ を T によって表した時の層別最小二乗解は，以下のように帰納的に表される行列 $S_L^{(k)}$ を用いて $u=S_L^{(1)}b$ として与えられる．また，元の空間での層別最小二乗解は，$x=G^{-1}S_L^{(1)}b$ で与えられる．

(i) $1\leq k\leq N-1$ の場合：

$$S_L^{(k)}=\begin{pmatrix} (T_{kk}^TT_{kk})^{-1}T_{kk}^T & 0 \\ -S_L^{(k+1)}\begin{pmatrix} T_{k+1,k} \\ \vdots \\ T_{N,k}\end{pmatrix}(T_{kk}^TT_{kk})^{-1}T_{kk}^T & S_L^{(k+1)}\end{pmatrix}.$$

(ii) $k=N$ の場合：

$$S_L^{(N)}=(T_{NN}^TT_{NN})^{-1}T_{NN}^T.$$

証明． 上述の $F(T_i,[r_i;\cdots;r_N])$ を $S_L^{(i)}[r_i;\cdots;r_N]$ と置くと，同じ漸化式に従っていることが確認できるので補題が成立する． □

$A=[A_1;\cdots;A_N]$ に対応するブロック下三角行列を $T=[T_1;\cdots;T_N]$ とする．以下，上の命題に現れる層別最小二乗解を与える行列 $S_L^{(1)}$ を S_L と記し，重みベクトル $d=(d_1,\ldots,d_N)$ に対する最小二乗解を与える行列 $(T^TD^2T)^{-1}T^TD^2$ を $S_W(d)$ と表すことにする．この時，T で像空間 $\mathrm{Im}(A)$ が表されている時のベクトル b に関する層別最小二乗解と重み付き最小二乗解の差は，

$$[S_L-S_W(d)]b$$

で与えられる．また，元の空間 x においては層別最小二乗解と重み付き最小二乗解の差は

$$G^{-1}[S_L-S_W(d)]b$$

で与えられる．そして，層別最小二乗解の残差と重み付き最小二乗解の残差の差は，

$$-T[S_L-S_W(d)]b$$

で与えられる．これらの差がどのように重みベクトル d と 4 節で導入される条件数 χ_A，$\bar{\chi}_A$ に依存するかを解析するのが本論文の目的である．

3.2 基底ブロック下三角行列

以下の性質を満たす行列 T を基底ブロック下三角行列と呼ぶ：

(1) T は縦のブロック数と横のブロック数が等しいブロック下三角行列である．
(2) T は，もし必要があれば行を適当に入れ替えた後に，(T の列数)$\times$(T の列数)の大きさの単位行列を部分行列として含む．
(3) T の各ブロック対角成分 T_{ii} は，もし必要があれば行を適当に入れ替えた後に，(T_{ii} の列数)$\times$(T_{ii} の列数)の大きさの単位行列を部分行列として含む．

T を層行列 $A=[A_1;\cdots;A_N]$ に対応するブロック下三角行列で上記(1)$\sim$(3)を満たすものとする．これを層行列 $A=[A_1;\cdots;A_N]$ に対応する基底ブロック下三角行列と呼ぶことにする．定理 3.1 に示したように，A が列一次独立であれば，適当な正則行列 G を用いて AG^{-1} を $[A_1;\cdots;A_N]$ に対応するブロック下三角行列とできるが，この行列にさらに列基本変換を施すことにより，必ず基底ブロック下三角行列を得ることができる．すなわち以下の定理が成立する．

定理 3.3. 列一次独立な層行列 $A=[A_1;\cdots;A_N]$ に対し，A の行ベクトルからなる正則行列 G で，$T=[T_1;\cdots;T_N]=AG^{-1}$ が層行列 $A=[A_1;\cdots;A_N]$ に対応する基底ブロック下三角行列となるような正則行列 G が存在する．

証明． 定理 3.1 の証明に基づき，すでに $A=[A_1;\cdots;A_N]$ が対応するブロック下三角行列 $T=[T_1;\cdots;T_N]$ に変形されているものとする．ここで T の各対角ブロック T_{ii} が列一次独立であることに注意した上で，さらに列基本変形を繰り返して T_{ii} が単位行列を部分行列として含むように変形していき，次いでこの単位行列に対応する T の行から $T_{i1},\ldots,T_{i,i-1}$ 部分に残っている非零要素を列基本変形でさらに消去するという方針で，列基本変形のみで T を基本ブロック下三角行列とすることができる．したがって，定理に述べられている G が存在する． □

4. 条件数

列一次独立行列 C について，以下のように定義される数 χ_C，$\bar{\chi}_C$ を C の条件数と呼ぶ．

$$\chi_C \equiv \sup_{D\in\mathcal{D}} \|(C^TDC)^{-1}C^TD\|, \quad \bar{\chi}_C \equiv \sup_{D\in\mathcal{D}} \|C(C^TDC)^{-1}C^TD\|.$$

ここで $\mathcal{D}$ は対角成分が正であるような対角行列全体の集合である．これらの条件数が「有限」であることは，Dikin (1974) によって最初に示され，後に，Stewart (1989)，Todd (1990)，Ben-Tal and Teboulle (1990) らによって独立に別証明が与えられた．χ_C や $\bar{\chi}_C$ は，以下のような性質を持つ．

命題 4.1. C を k 列の列一次独立行列とする．この時条件数 χ_C，$\bar{\chi}_C$ は下記の性質を持つ．
(1) $\chi_C=\max\{\|G^{-1}\|:\ G$ は C の $k\times k$ 正則部分行列 $\}$.
(2) $\bar{\chi}_C=\max\{\|CG^{-1}\|:\ G$ は C の $k\times k$ 正則部分行列 $\}$.
(3) G が正則ならば，$\bar{\chi}_{CG}=\bar{\chi}_C$.
(4) χ_C，$\bar{\chi}_C$ は C の行置換と列置換に対して不変である．
(5) C が $k\times k$ 単位行列をその部分行列として含むとする．すると，C の任意の部分行列 $\tilde{C}$ について，$\|\tilde{C}\|\le\bar{\chi}_C$ が成立する．
(6) C が $k\times k$ 単位行列をその部分行列として含むとする．すると，C から列を選んで作られる部分行列 $\tilde{C}$ について，$\bar{\chi}_{\tilde{C}}\le\bar{\chi}_C$ が成立する．
(7) C が $k\times k$ 単位行列をその部分行列として含むならば $\chi_C\le\bar{\chi}_C$ が成立する．
(8) $\tilde{C}$ を C の行ベクトルからなる列一次独立な部分行列とする．この時 $\bar{\chi}_{\tilde{C}}\le\bar{\chi}_C$ が成立する．

証明． (1)については Forsgren (1996) の Corollary 2.2，(2)については Todd et al. (2001) の Theorem 1 を参照のこと．(3)，(4)は定義より自明．(5)については $\|\tilde{C}\|\le\|C\|$ であること，$k\times k$ 単位行列が(2)の $\bar{\chi}_C$ の特徴付けの最適化問題の許容解であるため，$\|C\|\le\bar{\chi}_C$ となることから従う．(6)については，$\bar{\chi}_{\tilde{C}}$ に対応する(2)の最適化問題を考えると，その許容解は必ず，$\bar{\chi}_C$ に対する(2)の最適化問題の許容解のある部分行列として現れることおよび行列のノルムは任意の部分行列のノルムよりも大きいことより従う(Monteiro and Tsuchiya, 2003, Proposition 2.2 (e)を参照のこと)．(7)については，(1)の特徴付けにおいて実際に G^{-1} が χ_C を達成するような G について，CG^{-1} を考えると，C の単位行列の部分に G^{-1} が現れるので成り立つ．(8)は，(2)および行列のノルムは任意の部分行列のノルムよりも大きいことから従う． □

計算複雑度の観点からは，χ_C の計算が NP 困難であることが Khachiyan (1995) において証

明されている．また，Vavasis and Ye (1996)においては，C が整数行列でその入力ビット数が L_C ならば，χ_C も $\bar{\chi}_C$ も $2^{\mathcal{O}(L_C)}$ で上から抑えられることが示されている．これらの条件数に関する研究としては，他に例えば Gonzaga and Lara (1997)などがある．

5. 2 層の場合

$\widetilde{T}$ を 2 層の基底ブロック下三角行列とする．この時，

$$\widetilde{T}=\begin{pmatrix}\widetilde{T}_{11} & 0\\ \widetilde{T}_{21} & \widetilde{T}_{22}\end{pmatrix}$$

と表現できる．また，$\widetilde{T}_{11}$ および $\widetilde{T}_{22}$ は列一次独立行列であり,(適当な行の入れ替えの後には)各々単位行列を部分行列として含む．また，$\widetilde{T}$ 自身も(適当な行の入れ替えの後には)単位行列を部分行列として含む．以下，$\widetilde{D}$ を各対角要素が正の対角行列とし，2 つの層に対応する対角行列を $\widetilde{D}_1$，$\widetilde{D}_2$ とする．以下，重み付き最小二乗解を与える行列 $(\widetilde{T}^T\widetilde{D}^2\widetilde{T})^{-1}\widetilde{T}^T\widetilde{D}^2$ について議論する．

ここで，$\widetilde{T}_{11}$ が適当な行の入れ替えののちには単位行列を部分行列として含むことを考えると，$\widetilde{T}_{21}=V\widetilde{T}_{11}$ と書ける．ここで，V はその列ベクトルが $\widetilde{T}_{21}$ の列あるいは 0 ベクトルからなる行列である．$\|V\|\le\|\widetilde{T}_{21}\|$ が成立することに注意しよう．すると，

$$\widetilde{T}=\begin{pmatrix}\widetilde{T}_{11} & 0\\ V\widetilde{T}_{11} & \widetilde{T}_{22}\end{pmatrix}$$

と書くことができる．したがって，

$$\widetilde{D}\widetilde{T}=\begin{pmatrix}\widetilde{D}_1\widetilde{T}_{11} & 0\\ \widetilde{D}_2\widetilde{T}_{21} & \widetilde{D}_2\widetilde{T}_{22}\end{pmatrix}=\begin{pmatrix}B & 0\\ WB & M\end{pmatrix}$$

と書ける．ここで，

$$B=\widetilde{D}_1\widetilde{T}_{11},\quad W=\widetilde{D}_2V\widetilde{D}_1^{-1},\quad \widetilde{T}_{21}=V\widetilde{T}_{11},\quad M=\widetilde{D}_2\widetilde{T}_{22}.$$

$\|V\|\le\|\widetilde{T}_{21}\|$ が成立することに注意すると，次の補題の証明は容易である．

補題 5.1. 以下の関係が成立する．
(1) $\|W\|\le\|\widetilde{D}_1^{-1}\|\|\widetilde{D}_2\|\|\widetilde{T}_{21}\|$.
(2) $\|\widetilde{D}_2^{-1}WB\|=\|\widetilde{T}_{21}\|$.
(3) $\|WB\|=\|\widetilde{D}_2\widetilde{T}_{21}\|\le\|\widetilde{D}_2\|\|\widetilde{T}_{21}\|$.
(4) $\|\widetilde{D}_2^{-1}W\|\le\|\widetilde{D}_1^{-1}\|\|\widetilde{T}_{21}\|$.

補題 5.2. 次の事実が成立する．
(1) $\widetilde{D}_1=d_1I$ ならば，$\|(B^TB)^{-1}B^T\widetilde{D}_1\|=\|(\widetilde{T}_{11}^T\widetilde{T}_{11})^{-1}\widetilde{T}_{11}^T\|\le 1$.
(2) $\|(M^TM)^{-1}M^T\widetilde{D}_2\|=\|(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\|\le\bar{\chi}_{\widetilde{T}}$.

証明. 最初の式については，$\widetilde{T}_{11}$ が(行の適当な並べ替えの後に)単位行列を部分行列として含むことおよび $\|(\widetilde{T}_{11}^T\widetilde{T}_{11})^{-1}\widetilde{T}_{11}^T\|=\|(\widetilde{T}_{11}^T\widetilde{T}_{11})^{-1}\|^{1/2}$ より明らか．2 番目については，

$$\|(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\|\le\chi_{\widetilde{T}_{22}}\le\bar{\chi}_{\widetilde{T}_{22}}\le\bar{\chi}_{\widetilde{T}}$$

より成立する．ここで，2 番目の不等式は命題 4.1（7），3 番目の不等式は命題 4.1（6），（8）よりしたがう．□

次の補題が重要である．

補題 5.3. $\widetilde{T}=[\widetilde{T}_1;\widetilde{T}_2]$ を 2 層からなる基底ブロック下三角行列とし，$\widetilde{D}_1=d_1 I$，$\widetilde{D}_2$ を各層 $\widetilde{T}_1,\widetilde{T}_2$ に対応する対角行列とする．この時，

$$(\widetilde{T}^T\widetilde{D}^2\widetilde{T})^{-1}\widetilde{T}^T\widetilde{D}^2=\begin{pmatrix}(\widetilde{T}_{11}^T\widetilde{T}_{11})^{-1}\widetilde{T}_{11}^T & 0\\ -(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{21}(\widetilde{T}_{11}^T\widetilde{T}_{11})^{-1}\widetilde{T}_{11}^T & (\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\end{pmatrix}+\begin{pmatrix}C_{11} & C_{12}\\ C_{21} & C_{22}\end{pmatrix} \tag{5.1}$$

と書ける．ここで $t=\|\widetilde{T}_{21}\|$，$\tau=\|\widetilde{D}_1^{-1}\|\|\widetilde{D}_2\|$ とおくと，

$$\begin{aligned}\|C_{11}\| &\le t^2\tau^2,\\ \|C_{12}\| &\le (1+t^2\tau^2)t\tau^2,\\ \|C_{21}\| &\le \|(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\|t^3\tau^2,\\ \|C_{22}\| &\le \|(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\|(2+t^2\tau^2)t^2\tau^2.\end{aligned}$$

証明.　まず，

$$\widetilde{T}^T\widetilde{D}^2\widetilde{T}\equiv\begin{pmatrix}F & G\\ G^T & H\end{pmatrix}=\begin{pmatrix}B^T(I+W^TW)B & B^TW^TM\\ M^TWB & M^TM\end{pmatrix}$$

となる．正則ブロック行列の逆行列の表現公式より，

$$(\widetilde{T}^T\widetilde{D}^2\widetilde{T})^{-1}=\begin{pmatrix}\widetilde{A} & \widetilde{B}^T\\ \widetilde{B} & \widetilde{C}\end{pmatrix}=\begin{pmatrix}L^{-1} & -L^{-1}GH^{-1}\\ -H^{-1}G^TL^{-1} & H^{-1}+H^{-1}G^TL^{-1}GH^{-1}\end{pmatrix}$$

を得る．ここで $L=F-GH^{-1}G^T$ である．

次いで，

$$\begin{aligned}\widetilde{A}&=L^{-1}=(F-GH^{-1}G^T)^{-1}\\ &=(B^T(I+W^TW)B-B^TW^TM(M^TM)^{-1}M^TWB)^{-1}\\ &=(B^T(I+W^TQ_MW)B)^{-1}=(B^TB+B^TW^TQ_MQ_MWB)^{-1}\\ &=(B^TB)^{-1}-(B^TB)^{-1}B^TW^TQ_M(I+Q_MWP_BW^TQ_M)^{-1}Q_MWB(B^TB)^{-1}\end{aligned} \tag{5.2}$$

と変形する．ここで，$P_B=B(B^TB)^{-1}B^T$，$Q_M=I-P_M$，$P_M=M(M^TM)^{-1}M^T$ とし，5 番目から 6 番目の等式に行くにあたり，Sherman-Morrison-Woodbury 公式を用いた．

また，

$$\widetilde{B}=-H^{-1}G^TL^{-1}=-(M^TM)^{-1}M^TWB\widetilde{A} \tag{5.3}$$

が成立する．

さらに，(5.2)を用いて，

$$\begin{aligned}\widetilde{C} &= H^{-1} + H^{-1}G^T L^{-1} G H^{-1} \\ &= (M^T M)^{-1} + (M^T M)^{-1} M^T W P_B \\ &\quad (I - P_B W^T Q_M (I + Q_M W P_B W^T Q_M)^{-1} Q_M W P_B) P_B W^T M^T (M^T M)^{-1} \\ &= (M^T M)^{-1} + (M^T M)^{-1} M^T W P_B (I + P_B W^T Q_M W P_B)^{-1} P_B W^T M (M^T M)^{-1}\end{aligned}$$

を得る．ここで Sherman-Morrison-Woodbury 公式を用いた．

これらの結果を用いて，

$$\begin{aligned}(\widetilde{T}^T \widetilde{D}^2 \widetilde{T})^{-1} \widetilde{T}^T \widetilde{D}^2 &= \begin{pmatrix} \widetilde{A} & \widetilde{B}^T \\ \widetilde{B} & \widetilde{C} \end{pmatrix} \begin{pmatrix} B^T \widetilde{D}_1 & B^T W^T \widetilde{D}_2 \\ 0 & M^T \widetilde{D}_2 \end{pmatrix} \\ &= \begin{pmatrix} \widetilde{A} B^T \widetilde{D}_1 & \widetilde{A} B^T W^T \widetilde{D}_2 + \widetilde{B}^T M^T \widetilde{D}_2 \\ \widetilde{B} B^T \widetilde{D}_1 & \widetilde{B} B^T W^T \widetilde{D}_2 + \widetilde{C} M^T \widetilde{D}_2 \end{pmatrix} \\ &= \begin{pmatrix} \widetilde{A} B^T \widetilde{D}_1 & \widetilde{A} B^T \widetilde{D}_1 \widetilde{D}_1^{-1} W^T \widetilde{D}_2 - \widetilde{A} B^T \widetilde{D}_1 \widetilde{D}_1^{-1} W^T M (M^T M)^{-1} M^T \widetilde{D}_2 \\ \widetilde{B} B^T \widetilde{D}_1 & \widetilde{B} B^T \widetilde{D}_1 \widetilde{D}_1^{-1} W^T \widetilde{D}_2 + \widetilde{C} M^T \widetilde{D}_2 \end{pmatrix} \\ &= \begin{pmatrix} \widetilde{A} B^T \widetilde{D}_1 & \widetilde{A} B^T \widetilde{D}_1 \widetilde{D}_1^{-1} W^T Q_M \widetilde{D}_2 \\ \widetilde{B} B^T \widetilde{D}_1 & \widetilde{B} B^T \widetilde{D}_1 \widetilde{D}_1^{-1} W^T \widetilde{D}_2 + \widetilde{C} M^T \widetilde{D}_2 \end{pmatrix} \equiv \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}\end{aligned}$$

を得る．ここで 3 番目の等号の第 12 成分について(5.3)を用いた．

以下順次，最右辺の各ブロック行列を評価する．まず，H_{11} については，(5.2)より

$$\begin{aligned}H_{11} &= (B^T B)^{-1} B^T \widetilde{D}_1 \\ &\quad - (B^T B)^{-1} B^T W^T Q_M (I + Q_M W P_B W^T Q_M)^{-1} Q_M W B (B^T B)^{-1} B^T \widetilde{D}_1\end{aligned}$$

を得る．右辺第 1 項は $(\widetilde{T}_{11}^T \widetilde{T}_{11})^{-1} \widetilde{T}_{11}^T$ となる．第 2 項については，$B = d_1 \widetilde{T}_{11}$ であり，$\widetilde{T}_{11}$ が(適当な行の入れ替えののちには)単位行列を部分行列として含むので，$\|(B^T B)^{-1}\| \le d_1^{-2} = \|\widetilde{D}_1^{-1}\|^2$ が成立すること，$\|WB\| \le \|\widetilde{D}_2\| \|\widetilde{T}_{21}\|$ であること(補題 5.1 (3))，補題 5.2 より，$\|(B^T B)^{-1} B^T \widetilde{D}_1\| = \|(\widetilde{T}_{11}^T \widetilde{T}_{11})^{-1} \widetilde{T}_{11}^T\| \le 1$ が成立すること，Q_M が射影行列であることなどより，そのノルムを

$$\begin{aligned}&\|(B^T B)^{-1} B^T W^T Q_M (I + Q_M W P_B W^T Q_M)^{-1} Q_M W B (B^T B)^{-1} B^T \widetilde{D}_1\| \\ &\le \|(B^T B)^{-1}\| \|B^T W^T\| \|Q_M\| \|(I + Q_M W P_B W^T Q_M)^{-1}\| \|Q_M\| \|WB\| \|(B^T B)^{-1} B^T \widetilde{D}_1\| \\ &\le \|\widetilde{T}_{21}\|^2 \|\widetilde{D}_1^{-1}\|^2 \|\widetilde{D}_2\|^2\end{aligned}$$

と上から押さえることができる．したがって，

$$\begin{aligned}H_{11} &= (B^T B)^{-1} B^T \widetilde{D}_1 + C_{11} = (\widetilde{T}_{11}^T \widetilde{T}_{11})^{-1} \widetilde{T}_{11}^T + C_{11}, \\ C_{11} &= -(B^T B)^{-1} B^T W^T Q_M (I + Q_M W P_B W^T Q_M)^{-1} Q_M W B (B^T B)^{-1} B^T \widetilde{D}_1, \\ \|C_{11}\| &\le \|\widetilde{T}_{21}\|^2 \|\widetilde{D}_1^{-1}\|^2 \|\widetilde{D}_2\|^2 = t^2 \tau^2\end{aligned} \tag{5.4}$$

を得る．

また，

$$C_{12} = H_{12} = H_{11} \widetilde{D}_1^{-1} W^T Q_M \widetilde{D}_2$$

なので，補題 5.2，補題 5.1 (1),(5.4) より，

$$\|C_{12}\| \leq (1+\|\widetilde{T}_{21}\|^2\|\widetilde{D}_1^{-1}\|^2\|\widetilde{D}_2^2\|)\|\widetilde{T}_{12}\|\|\widetilde{D}_1^{-1}\|^2\|\widetilde{D}_2\|^2 = (1+t^2\tau^2)t\tau^2$$

を得る．

一方, (5.3), (5.4) より,

$$\begin{aligned} H_{21} &= -(M^TM)^{-1}M^TWB\tilde{A}B^T\widetilde{D}_1 = -(M^TM)^{-1}M^TWBH_{11} \\ &= -(M^TM)^{-1}M^T\widetilde{D}_2\widetilde{D}_2^{-1}WB(B^TB)^{-1}B^T\widetilde{D}_1 - (M^TM)^{-1}M^T\widetilde{D}_2\widetilde{D}_2^{-1}WBC_{11} \\ &= -(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{21}(\widetilde{T}_{11}^T\widetilde{T}_{11})^{-1}\widetilde{T}_{11}^T + C_{21} \end{aligned}$$

となる．ここで

$$C_{21} = -(M^TM)^{-1}M^T\widetilde{D}_2\widetilde{D}_2^{-1}WBC_{11}$$

なので，補題 5.1 (2) と上の(5.4) より，

$$\begin{aligned} \|C_{21}\| &\leq \|(M^TM)^{-1}M^T\widetilde{D}_2\|\|\widetilde{T}_{21}\|\|\widetilde{T}_{21}\|^2\|\widetilde{D}_1^{-1}\|^2\|\widetilde{D}_2\|^2 \\ &= \|(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\|\|\widetilde{T}_{21}\|\|\widetilde{T}_{21}\|^2\|\widetilde{D}_1^{-1}\|^2\|\widetilde{D}_2\|^2 \\ &= \|(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\|t^3\tau^2 \end{aligned}$$

を得る．以下で使うのでさらに $\|H_{21}\|$ を評価する．補題 5.2 (1) より

$$\|(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{21}(\widetilde{T}_{11}^T\widetilde{T}_{11})^{-1}\widetilde{T}_{11}^T\| \leq \|(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\|\|\widetilde{T}_{21}\|$$

なので，上の $\|C_{21}\|$ に関する結果を合わせて，

$$\|H_{21}\| \leq \|(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\|t(1+t^2\tau^2)$$

となる．

最後に H_{22} を評価する．まず，$H_{22} = H_{21}\widetilde{D}_1^{-1}W^T\widetilde{D}_2 + \widetilde{C}M^T\widetilde{D}_2$ と書ける．この右辺第 2 項の一部が主要項となり，残りが誤差項となる．右辺第 1 項は，そのノルムを，補題 5.1 (1)を用いて

$$\begin{aligned} \|H_{21}\widetilde{D}_1^{-1}W^T\widetilde{D}_2\| &\leq \|H_{21}\|\|\widetilde{T}_{21}\|\|\widetilde{D}_1^{-1}\|^2\|\widetilde{D}_2\|^2 \\ &\leq \|(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\|(1+t^2\tau^2)t^2\tau^2 \end{aligned} \tag{5.5}$$

と評価できる．そして，第 2 項は

$$\begin{aligned} \widetilde{C}M^T\widetilde{D}_2 &= (M^TM)^{-1}M^T\widetilde{D}_2 \\ &\quad + (M^TM)^{-1}M^TWP_B(I+P_BW^TQ_MWP_B)^{-1}P_BW^TM(M^TM)^{-1}M^T\widetilde{D}_2 \\ &= (M^TM)^{-1}M^T\widetilde{D}_2 + (M^TM)^{-1}M^T\widetilde{D}_2\widetilde{D}_2^{-1}WP_B(I+P_BW^TQ_MWP_B)^{-1}P_BW^TP_M\widetilde{D}_2 \end{aligned}$$

となる．この右辺第 1 項は H_{22} の主要項 $(M^TM)^{-1}M\widetilde{D}_2 = (\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2$ となる．一方第 2 項のノルムは，補題 5.1 (1), (4)を用い，

$$\begin{aligned} &\|(M^TM)^{-1}M^T\widetilde{D}_2\widetilde{D}_2^{-1}WP_B(I+P_BW^TQ_MWP_B)^{-1}P_BW^TP_M^T\widetilde{D}_2\| \\ &\quad \leq \|(M^TM)^{-1}M^T\widetilde{D}_2\|\|\widetilde{T}_{21}\|^2\|\widetilde{D}_1^{-1}\|^2\|\widetilde{D}_2\|^2 \end{aligned} \tag{5.6}$$

と抑えられる.(5.5) と (5.6) より,

$$H_{22} = (M^TM)^{-1}M^T\widetilde{D}_2 + C_{22} = (\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2 + C_{22},$$

$$C_{22} = H_{21}\widetilde{D}_1^{-1}W^T\widetilde{D}_2 + (M^TM)^{-1}M^T\widetilde{D}_2\widetilde{D}_2^{-1}WP_B(I+P_BW^TQ_MWP_B)^{-1}P_BW^TP_M\widetilde{D}_2,$$
$$\|C_{22}\| \le \|(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\|(2+t^2\tau^2)t^2\tau^2$$

を得る． □

この補題より，以下の定理が成り立つ．

定理 5.4． $\widetilde{T}$ を 2 層からなる基底ブロック下三角行列とし，$\widetilde{D}_1 = d_1I$，$\widetilde{D}_2$ を対応する対角行列とする．この時，

$$(5.7)\quad (\widetilde{T}^T\widetilde{D}^2\widetilde{T})^{-1}\widetilde{T}\widetilde{D}^2 = \begin{pmatrix} (\widetilde{T}_{11}^T\widetilde{T}_{11})^{-1}\widetilde{T}_{11}^T & 0 \\ -(\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{21}(\widetilde{T}_{11}^T\widetilde{T}_{11})^{-1}\widetilde{T}_{11}^T & (\widetilde{T}_{22}^T\widetilde{D}_2^2\widetilde{T}_{22})^{-1}\widetilde{T}_{22}^T\widetilde{D}_2^2 \end{pmatrix} + C,$$
$$\|C\| \le \bar{\chi}_{\widetilde{T}}^4\tau^2 + 2(2+\bar{\chi}_{\widetilde{T}}^2\tau^2)\bar{\chi}_{\widetilde{T}}^3\tau^2$$

が成立する．ここで，$\tau = \|\widetilde{D}_1^{-1}\|\|\widetilde{D}_2\|$ である．

証明． 前補題の $\|C_{ij}\|(i,j=1,2)$ のノルムを評価する．まず，命題 4.1 (6)，(7)，(8) より $\|\widetilde{T}_{21}\| \le \bar{\chi}_{\widetilde{T}}$ が成り立つ．このことを考慮し，さらに補題 5.2 が成立することを用いると，

$$\|C_{11}\| \le \bar{\chi}_{\widetilde{T}}^2\tau^2,\ \|C_{12}\| \le (1+\bar{\chi}_{\widetilde{T}}^2\tau^2)\bar{\chi}_{\widetilde{T}}\tau^2,$$
$$\|C_{21}\| \le \bar{\chi}_{\widetilde{T}}^4\tau^2,\ \|C_{22}\| \le (2+\bar{\chi}_{\widetilde{T}}^2\tau^2)\bar{\chi}_{\widetilde{T}}^3\tau^2$$

が成立する．さらに $\bar{\chi}_{\widetilde{T}} \ge 1$ であることを用いてこれらの結果を適宜まとめて，上記の $\|C\|$ の評価を得る． □

(5.7) の右辺第 1 項が 2 層の場合の層別最小二乗解を与えるものであることを考えると，この系は，2 層の場合に，層別最小二乗解と重み付き最小二乗解の差のノルムが，

$$\bar{\chi}_{\widetilde{T}}^4\tau^2 + 2(2+\bar{\chi}_{\widetilde{T}}^2\tau^2)\bar{\chi}_{\widetilde{T}}^3\tau^2$$

に右辺ベクトルのノルムを掛けたもので押さえられることを示している．ここで，τ を十分に小さくした時に主要項となるのは $\bar{\chi}_{\widetilde{T}}^4\tau^2$ で，この項は，$(\widetilde{T}^T\widetilde{D}^2\widetilde{T})^{-1}\widetilde{T}\widetilde{D}^2$ の第 21 ブロックに由来するものであることを注意しておこう．

6．主定理の証明

本節では 3 節までの記法と同じものを用いる．以下，

$$S_W^{(k)} = (\bar{T}_kD_k^2\bar{T}_k^T)^{-1}\bar{T}_kD_k^2,$$
$$D_k = \mathrm{diag}(d_kI_{n_k}, d_{k+1}I_{n_{k+1}}, ..., d_NI_{n_N})$$

とおく．この時，次の命題が成立する．

命題 6.1． 以下の事実が成り立つ．
(1) $\chi_{\bar{T}_k} \le \bar{\chi}_{\bar{T}_k} \le \bar{\chi}_T = \bar{\chi}_A$ である．
(2) T の任意の部分行列 T' について，$\|T'\| \le \bar{\chi}_T = \bar{\chi}_A$.

証明. 命題 4.1 (5), (6), (7), (8) より成立する. □

補題 6.2. $\tau = \max[d_2/d_1, \dots, d_N/d_{N-1}]$ とする.

(6.1)
$$S_W^{(k)} - S_L^{(k)} = \begin{pmatrix} 0 & 0 \\ -(S_W^{(k+1)} - S_L^{(k+1)}) \begin{pmatrix} T_{k+1,k} \\ \vdots \\ T_{N,k} \end{pmatrix} (T_{kk}^T T_{kk})^{-1} T_{kk}^T & S_W^{(k+1)} - S_L^{(k+1)} \end{pmatrix} + C_k,$$

(6.2)
$$\|C_k\| \le \bar{\chi}_A^4 \tau^2 + 2\bar{\chi}_A^3 \tau^2 (2 + \bar{\chi}_A^2 \tau^2),$$

(6.3)
$$\|S_W^{(k)} - S_L^{(k)}\| \le (1 + \bar{\chi}_A) \|S_W^{(k+1)} - S_L^{(k+1)}\| + [\bar{\chi}_A^4 \tau^2 + 2\bar{\chi}_A^3 \tau^2 (2 + \bar{\chi}_A^2 \tau^2)]$$

が $k = 1, \dots, N-1$ について成立する. また, $S_W^{(N)} - S_L^{(N)} = 0$ である.

証明. 定理 5.4 を, $\widetilde{D}_1 = d_k I_{n_k}$, $\widetilde{D}_2 = \mathrm{diag}(d_{k+1} I_{n_{k+1}}, \dots, d_N I_N)$; $\widetilde{T}_{11} = T_{kk}$, $\widetilde{T}_{21} = [T_{k+1,k}; \cdots; T_{Nk}]$, $\widetilde{T}_{22} = \bar{T}_{k+1}$, $C = C_k$ として適用し, 命題 3.2 (i)の漸化式との差を取ることで漸化式(6.1)が得られる. $\|C_k\|$ の評価については, 定理 5.4 および命題 6.1 より従う. また, ノルムに関する漸化式(6.3)については,(6.1)において, $\|[T_{k+1,k}; \cdots; T_{N,k}]\| \le \|T\| \le \bar{\chi}_A$ が成立し, さらに, 補題 5.2 (1)より $\|(T_{kk}^T T_{kk})^{-1} T_{kk}\| \le 1$ が成立することと(6.2)を用いると容易に示すことができる. 最後に, $k = N$ の場合, 1 層しかないので定義より $S_W^{(N)} = S_L^{(N)} = 0$ であることは容易に従う. □

最後に主定理を証明する.

定理 6.3. $\tau = \max(d_2/d_1, \dots, d_N/d_{N-1})$ とする. この時, 像空間 $\mathrm{Im}(A)$ を基底ブロック下三角行列 T (の列の一次結合)で表現した場合の層別最小二乗解を求める行列 S_L と重み付き最小二乗解を求める行列 $S_W(d)$ の差について評価

(6.4)
$$\|S_W(d) - S_L\| \le [(1 + \bar{\chi}_A)^{N-1} - 1][\bar{\chi}_A^3 \tau^2 + 2\bar{\chi}_A^2 \tau^2 (2 + \bar{\chi}_A^2 \tau^2)]$$

が成立する. さらに, 像空間を元の行列 A で表現した場合, 層別最小二乗解を求める行列 $G^{-1} S_L$ と重み付き最小二乗解を求める行列 $G^{-1} S_W(d)$ の差については評価

(6.5)
$$\|G^{-1}(S_W(d) - S_L)\| \le \chi_A [(1 + \bar{\chi}_A)^{N-1} - 1][\bar{\chi}_A^3 \tau^2 + 2\bar{\chi}_A^2 \tau^2 (2 + \bar{\chi}_A^2 \tau^2)]$$

が成立する. 最後に, 残差空間で見た層別最小二乗解の変位を求める行列 TS_L と重み付き最小二乗解の変位を求める行列 $TS_W(d)$ の差は,

(6.6)
$$\|T(S_W(d) - S_L)\| \le [(1 + \bar{\chi}_A)^{N-1} - 1][\bar{\chi}_A^4 \tau^2 + 2\bar{\chi}_A^3 \tau^2 (2 + \bar{\chi}_A^2 \tau^2)]$$

と評価できる.

証明. 補題 6.2 の漸化式を繰り返し適用して等比級数の和の公式を適用して整理して(6.4)を得る.(6.5)は, $\|G^{-1}\| \le \chi_A$ であること, (6.6)は $\|T\| \le \bar{\chi}_A$ であることと(6.4)を用いると自明. □

7. まとめ

本論文では重み付き最小二乗法と層別最小二乗法の関係について考察した．層別最小二乗法は物理的なシステムのモデリングとも興味深い関係があることが示唆されている（笹川, 2005）．また，統計学や数理諸分野においてもいろいろな応用があると思われる．これらの話題について今後さらに研究を進めていきたい．

謝　辞

丁寧に査読し，数々の改善点を指摘していただいた査読者に感謝いたします．また，著者が研究者として駆け出しの時期に非常に多くの時間を割いて議論して下さり，統計数理の諸側面について幅広く教えていただいた田邉國士 統計数理研究所名誉教授（現 早稲田大学）に深く感謝いたします．本論文も多くの部分をその時期に田邉教授に教えていただいたことに負っています．

参 考 文 献

Ben-Tal, A. and Teboulle, M. A. (1990). A geometric property of the least squares solution of linear equations, *Linear Algebra and Applications*, **139**, 165–170.

Dikin, I. I. (1974). On the speed of an iterative process (in Russian), *Upravlyaemye Sistemi*, **12**, 54–60.

Forsgren, A. (1996). On linear least-squares problems with diagonally dominant weight matrices, *SIAM Journal on Matrix Analysis and Applications*, **17**, 763–788.

Gonzaga, C. C. and Lara, H. J. (1997). A note on properties of condition numbers, *Linear Algebra and Its Applications*, **261**, 269–273.

Khachiyan, L. (1995). On the complexity of approximating extremal determinants in matrices, *Journal of Complexity*, **11**, 138–153.

Megiddo, N., Mizuno, S. and Tsuchiya, T. (1998). A modified layered-step interior-point algorithm for linear programming, *Mathematical Programming*, **82**, 339–355.

Monteiro, R. D. C. and Tsuchiya, T. (2003). A variant of the Vavasis-Ye layered-step interior-point algorithm for linear programming, *SIAM Journal on Optimization*, **13**, 1054–1079.

笹川 卓（2005）. 2 次錐計画問題を用いた直流磁気シールドの最適化，博士論文，東京大学大学院情報理工学系研究科.

Stewart, G. W. (1989). On scaled projections and pseudoinverses, *Linear Algebra and Its Applications*, **112**, 189–193.

Todd, M. J. (1990). A Dantzig-Wolfe-like variant of Karmarkar's interior-point linear programming algorithm, *Operations Research*, **38**, 1006–1018.

Todd, M. J., Tunçel, L. and Ye, Y. (2001). Characterizations, bounds, and probabilistic analysis of two complexity measures for linear programming problems, *Mathematical Programming*, **90** (1), 59–70.

Tsuchiya, T. (1992). Global convergence property of the affine scaling method for primal degenerate linear programming problems, *Mathematics of Operations Research*, **17**, 527–557.

Vavasis, S. and Ye, Y. (1996). A primal-dual accelerated interior point method whose running time depends only on A, *Mathematical Programming*, **74**, 79–120.

404 Proceedings of the Institute of Statistical Mathematics Vol. 53, No. 2, 391–404 (2005)

Layered Least Squares Method
—Limit of Weighted Least Squares Method—

Takashi Tsuchiya

The Institute of Statistical Mathematics

The weighted least squares method is a simple generalization of the ordinary least squares method. This paper deals with an extreme case where the ratio among the weights in the weighted least squares problem diverges to infinity. It is known that a limiting solution exists and is well-defined. This limit of the weighted least squares method is referred to as the layered least squares method. In this paper, we analyze the behavior of the weighted least squares solution when the ratio among the weights tends to infinity, and establish bounds on the difference between the weighted least squares solution and the layered least squares solution in terms of the ratios and condition numbers of the matrix involved in the least squares problem.

Key words: Weighted least squares problems, layered least squares problems, condition numbers, interior-point algorithms.