

悪条件線形方程式の数値解析的研究

細田陽介

図・本館

悪条件線形方程式の数値解析的研究

細田陽介

名古屋大学図書	
和	1168017

はじめに

我々の目的は悪条件線形方程式 $Ax = b$ を解くことにある。通常、悪条件方程式とは、右辺のデータベクトル b に微小な摂動 b_Δ を加えたとき、解 x の変化 x_Δ が大きくなる恐れのある方程式を指す。作用素 A の悪条件性は $\|A\| \|A^\dagger\|$ を用いて定量的に表される。ここで A^\dagger は Moore-Penrose の一般逆作用素を指す。非退化核を持つ第1種フレドホルム積分方程式

$$\int_a^b K(s, t) f(t) dt = g(s), \quad s_{\min} \leq s \leq s_{\max}$$

は悪条件線形方程式の典型的な例である。

悪条件線形方程式の数値解法として、Tikhonov の方法に代表される正則化法が挙げられる。一般に正則化法は、ノイズに敏感な高周波成分を制御する方法であると言える。Tikhonov の正則化法はこれを正則化パラメータにより行なう。したがって、Tikhonov の正則化法を用いるときの本質的な問題は正則化パラメータの選択にある。

正則化パラメータの選択法としては、V. A. Morozov により提案された相変原理 (Discrepancy Principle, [24]) と、G. Wahba により提案された一般化交差検証法 (Generalized Cross-Validation, 略して GCV 法と呼ばれている [34]) が有名である。

相変原理は、計算結果の質と、入力データの質を一致させることを基礎としている。ここで、入力データの質はノイズレベルを、計算結果の質は残差ノルムを指す。これにより、事前にノイズレベルの上界が判っていればある意味において最良となる正則化パラメータを選択することができる。しかし、一般にノイズレベルは事前には不明である。

GCV 法の基本的な考え方は、与えられたデータ点の一部を取り除き、残りのデータ点を使って解を求め、それが取り除かれた部分を良く推定するようにパラメータを決定することにある。しかし、この手法は悪条件線形方程式の特殊性を考慮していないため、その有効性は以前より指摘されている [32]。

最近、P. C. Hansen により“L-カーブによる最適正則化法”が提唱された [15]。L-カーブとは Tikhonov の正則化法による近似解のノルムを縦軸に、その残差ノルムを横軸に取り、正則化パラメータを動かすことにより得られる図を指す。名前の由来は文字通り L 字型のカーブとなることにある。また、このカーブが L 字型となるのは悪条件線形方程式の場合に限る。

Tikhonov の正則化法を適用するとき、近似解のノルムもしくは残差ノルムが大きくなるような正則化パラメータを選択することは明らかに不都合である。この2つの量の中庸となるような正則化パラメータを選択すべきであると考えられる。この着想を具現化したものが L-カーブである。L-カーブはその中程にコーナーを持つ。そのコーナーを表現する正則化パラメータを選択せよというのが Hansen の主張である。

しかし、L-カーブはコーナー付近では正則化パラメータの動きに対してカーブの動きが極端に鈍くなる。そのため Hansen の主張する方法では正則化パラメータの選定は困難である。そこで我々は、正則化パラメータの選定基準として“L-カーブの曲率”を用い、これが最大となる正則化パラメータを選択する方法を提案した [16],[19]。L-カーブの曲率は行列の特異値分解を用いると正則化パラメータの関数として陽に表すことができる。そして、我々の手法の有効性は数値実験により実証された。

また、Morozov は、事前にノイズレベルを推定できれば、相変原理の意味において最適となる正則化パラメータを選択できることを示した。これより、統計的な立場からこの問題へのアプローチも考えられる。もし、統計的な手法を用いることにより、ノイズレベルを事前に推定することができるならば、相変原理の意味での最適の近似解が得られると考えられる。

ここで我々が用いた方法は、データ平滑化法として有効であると言われている AIC 最小化推定法 (Minimum AIC Estimation, 略して MAICE 法と呼ばれている [27]) と相変原理を組み合わせた MAICE-DP 法である

[17]. なお, AIC とは, 現実に観測されるデータを用いて統計的モデルの適切さを評価するために赤池弘次により導入された統計量であって, 次式のように定義される [2].

$$AIC = -2 \times \ln(\text{尤度の最大値}) + 2 \times (\text{自由なパラメータ数}).$$

パラメータ数は観測者が主観的, 経験的に決めるのが普通であるが, 情報量 AIC を最小とするモデルを最良として採用する方法を用いると, これを機械的に決めることができる. この AIC は情報量規準の推定値として広く使われている.

MAICE-DP 法は, まずノイズを含んだデータより回帰多項式と AIC を用いてノイズレベルを推定し, それを相変原理に適用することにより正則化パラメータを決定する方法である. この方法はノイズレベルの推定がうまく行なわれればかなり精度の良い正則化パラメータが得られるであろうと期待された. しかし, MAICE-DP 法は正則化パラメータ選定において, 数値的に不安定な非線形方程式を解かなければならないことがわかった. この非線形方程式は MAICE 法の許容誤差限界にですら敏感に反応してしまう. また, MAICE-DP 法はあくまでも相変原理の意味での最適の正則化パラメータを得ることを目的としているため, 相変原理が破綻をきたすような場合は無意味なものとなる. 我々の行なった数値実験より, データ点と解の近似に用いる基底関数の個数がかげ離れているような場合は相変原理の意味での最適の正則化パラメータでは誤差が最小とはならないことが判明した. これは, データ点が 100 個で A の実質的な階数が 30 程度の場合でも 100 個の基底関数を用いて解を近似しなければならないことを意味する. これは計算量の点から見ても無駄である. しかし, L-カーブの曲率による最適正則化法ではこのような場合でも 30 個前後の基底関数で高精度の近似解が得られた. これより, 悪条件線形方程式の正則化パラメータ選定法としては L-カーブの曲率による方法を用いるべきである.

Tikhonov の正則化法とは別に, 解空間を制限することにより高周波成分を切り捨てる正則化法も存在する. 通常, 解空間ならびにデータベクトル空間の基底として, 作用素 A の右, 左特異ベクトルを用いる. そして, データベクトル b の左特異ベクトルによるフーリエ係数と A の特異値との相関を考慮することにより, 解空間の次元を決定する. 我々はその空間において最良となる近似解を打ち切り最小 2 乗最小ノルム解と名付けた.

我々の行なった数値実験では, 特異値分解による打ち切り最小 2 乗最小ノルム解は, 悪条件線形方程式に対して良好な結果を示した. しかし, 特異値分解は直接法ではないので, 計算量が多いという欠点がある. また, 無駄な計算を伴っている. これに対して我々は, ピボッティング付修正グラム・シュミット法による QR 分解を 3 回行なうことにより得られる A の分解

$$A = URDV^T$$

による打ち切り最小 2 乗最小ノルム解を, 悪条件線形方程式の近似解として用いることを提案した [18]. ここで, U, V は列正規直交行列, R は悪条件でない上三角行列, D は対角行列である.

上の分解はピボッティングの効果により効率的に行なうことができる. また, これにより得られる打ち切り最小 2 乗最小ノルム解は, 特異値分解による打ち切り最小 2 乗最小ノルム解と同程度の近似解であることが数値実験により検証された.

本論文は以下のような構成になっている.

第 1 章では, 不適切問題を定義し, その中で我々が問題としている悪条件線形方程式を導出し, その数学的な意味付けを関数空間上の最小 2 乗最小ノルム問題の立場から記述する. 第 2 章では, 正則化法の収束証明, 収束速度などをデータにノイズが含まれている場合と含まれていない場合に分けて紹介する. 同様に, 代表的な正則化法である Tikhonov の正則化法についても考察する. 第 3 章では, 既存の正則化パラメータ選択法である Morozov の相変原理などを説明し, さらに, 我々の主張である L-カーブの曲率による最適正

則化法について言及する。また、統計的立場からの最適正則化法についての考察もこの章で行なう。第4章では、打ち切り最小2乗最小ノルム解を定義し、特異値分解による打ち切り最小2乗最小ノルム解と、我々が提案した直接法による打ち切り最小2乗最小ノルム解が悪条件線形方程式に対して同程度の有効性を持つことを数値実験により検証する。最後に、直接法により打ち切り最小2乗最小ノルム解を求めるプログラムを付録として掲載する。

目次

1 悪条件線形方程式	6
1.1 不適切問題とは	6
1.2 不適切問題の例	7
1.3 一般逆作用素	8
1.4 特異値分解	17
1.4.1 コンパクト自己共役作用素	17
1.4.2 特異値分解と悪条件線形方程式	21
2 正則化法	24
2.1 正則化法の一般論	24
2.2 ノイズを含んだデータに対する正則化法	29
2.3 Tikhonov の正則化法	32
3 正則化パラメータの選択	37
3.1 Morozov の相変原理	37
3.2 正則化法の離散化	42
3.3 AIC による正則化 —MAICE-DP 法—	45
3.3.1 AIC 最小化推定法	45
3.3.2 MAICE-DP 法のアルゴリズム	47
3.3.3 MAICE-DP 法の数値実験および考察	47
3.4 L-カーブによる正則化法	51
3.4.1 L-カーブ	51
3.4.2 L-カーブの曲率による正則化法	55
3.5 一般化交差検証法 —GCV 法—	58
3.6 まとめ	59
4 直接法による悪条件線形方程式の解法	61
4.1 打ち切り最小 2 乗最小ノルム解	61
4.2 特異値分解による打ち切り最小 2 乗最小ノルム解	62

4.3 QR 分解による打ち切り最小 2 乗最小ノルム解	64
4.4 アルゴリズム	65
4.5 数値実験およびその考察	66
4.6 まとめ	68
おわりに	69
参考文献	70
A 使用例の主プログラム	73
B 打ち切り最小 2 乗最小ノルム解のサブルーチン集	77

第 1 章

悪条件線形方程式

最近, 諸分野において逆問題が脚光を浴びている. 逆問題とは, 観測された現象からその原因あるいは現象を成立させている内部機構を決定する問題である. 順問題と比べると逆問題は格段の難しさを持っている. すなわち, 順問題においては, 通常適当な関数空間の中に与えられたデータに対して像が別の関数空間の中に一意に存在し, その対応が連続的になる. これに対し, 逆問題においては与えられたデータに対して解が一意に決まらないことがある. また, 一意的に決まる場合でも, データの微小な摂動に対して解が不安定な振る舞いをすることも多い. 1.1節で述べるように, J. Hadamard はこのような問題を不適切問題と名付けた. 不適切問題は長い間, 物理的に意味のない問題として, 物理数学の対象から除外されてきた. しかし, こうしたものも今では一般的な意味で適切な問題であることがわかったり, 解の中で最も合理的と思われるものを選択し, 数値計算によって実際にそれを求めることができるようになってきたりしている.

この章では, 先ず不適切問題の定義を与える. そして, 代表的な不適切問題である第 1 種フレドホルム積分方程式の性質を例をあげて説明する. 多くの線形不適切問題は, 悪条件線形方程式に帰着する. 1.4節では悪条件線形方程式の解析の有効な手段である特異値分解を準備する. また, 悪条件線形方程式の最小 2 乗最小ノルム問題の問題点もこの節で指摘する.

1.1 不適切問題とは

悪条件線形方程式は, 線形不適切問題 (linear ill-posed problem) より導かれる. 問題の適切性 (well-posedness) は J. Hadamard により最初に定式化された. 問題が適切であるとは, 次の 3 つの条件を満たしていることである.

- (1) 解が存在する.
- (2) 解が一意である.
- (3) 問題が安定である.

これらを数学的に定式化すると次のようになる. F, G を完備な距離空間, $d_F(\cdot, \cdot), d_G(\cdot, \cdot)$ を F, G での距離, T を F から G への作用素, $\mathcal{D}(T), \mathcal{R}(T)$ をそれぞれ T の定義域, 値域とする. 方程式 $Tf = g, g \in G$ の解 $f \in F$ を求める問題において,

- (1') $\mathcal{R}(T) = G$.
- (2') $f_1, f_2 \in \mathcal{D}(T)$ に対し, $Tf_1 = Tf_2$ ならば, $f_1 = f_2$.

(3') 任意の $\varepsilon > 0$ に対して, $\delta = \delta(\varepsilon) > 0$ が存在して, $d_G(Tf_1, Tf_2) < \delta$ ならば $d_F(f_1, f_2) < \varepsilon$ となる.

(1'), (2') の条件は T の逆作用素 T^{-1} の存在を表す. (3') は T^{-1} の連続性を意味する. 不適切問題とは, 適切でない問題を指す.

多くの不適切問題は第1種フレドホルム積分方程式

$$\int_{I_t} K(s, t)f(t)dt = g(s), \quad s \in I_s \quad (1.1)$$

に帰着する. ここで, K は積分核である.

上の第1種フレドホルム積分方程式について, 適切性の3つの条件を考察しよう. もし, K が連続で, f が可積分であれば, g もまた連続となる. そのため, g として不連続な関数を与えると, 上式の方程式を満たすような可積分関数 f は存在しない. これより, 第1種フレドホルム積分方程式は, 可積分関数 g に対し, 解の存在すら明らかでない方程式であると言える.

解の一意性が明らかではない例として, 方程式

$$\int_0^\pi s \sin t f(t) dt = s$$

が挙げられる. 積分核は $K(s, t) = s \sin t$ である. この方程式を満たす解は $f(t) = 1/2$ である. しかし,

$$f_n(t) = \frac{1}{2} + \sin nt, \quad n = 1, 2, \dots$$

も解となる.

次に, 最も重要な要素である安定性について検証する. 今, $K(\cdot, \cdot)$ を任意の2乗可積分な関数とする. このとき, リーマン・ルベグの定理より,

$$\int_0^\pi K(s, t) \sin nt dt \rightarrow 0, \quad (n \rightarrow \infty)$$

が成り立つ. これより, n を十分大きくとれば,

$$\tilde{g}(s) = g(s) + C \int_0^\pi K(s, t) \sin nt dt$$

は, $g(s)$ に僅かに摂動を加えたデータであると言える. しかし, もし C を大きくとれば, 積分方程式 $\int_0^\pi K(s, t)f(t)dt = \tilde{g}(s)$ の解 $f(t) + C \sin nt$ は, $f(t)$ とは著しく異なる関数となる. これは, 第1種フレドホルム積分方程式の不安定性を示す典型的な例である.

1.2 不適切問題の例

不適切問題は数学, 工学を問わず, さまざまな分野で現れる. それらの多くは第1種フレドホルム積分方程式として定式化される. いくつかの例として, 以下のものが挙げられる.

1. 逆ラプラス変換
2. 計算機トモグラフィー (CT : Computed Tomography)
3. 解析関数の解析接続
4. 画像処理

5. 逆熱伝導方程式

これらの中で、逆熱伝導方程式を取り上げ、第1種フレドホルム積分方程式に定式化してみよう [28]. 方程式系

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2}, & (0 < x < \pi, 0 < t < \infty) \\ u|_{t=0} &= u_0(x), & (0 < x < \pi) \\ u|_{x=\pi} &= u|_{x=0} = 0, & (0 < t < \infty) \end{aligned} \quad (1.2)$$

を考える. 以下, 領域 Ω に対し $C^m(\Omega)$ は Ω 上 m 回連続微分可能な関数全体, $C^m(\bar{\Omega})$ は $C^m(\Omega)$ の元で m 回までの微係数が $\bar{\Omega}$ 上に連続的に拡張される関数全体を表し, $C_0(\bar{\Omega})$ は境界で 0 となる $C(\bar{\Omega})$ の元全体を示すことにする.

この方程式系において初期値 $u_0 \in C_0[0, \pi]$ が未知量であり, 時刻 $t = 1$ における温度分布

$$g(x) = u|_{t=1}, \quad (0 < x < \pi)$$

が既知量であったとしよう. フーリエ級数を用いれば, 方程式 (1.2) の解 $u(x, t) \in C^0([0, \pi] \times [0, \infty)) \cap C^2([0, \pi] \times (0, \infty))$ は

$$u(x, t) = \sum_{n=1}^{\infty} a_n e^{-n^2 t} \sin nx$$

と展開できる. ただし

$$a_n = \frac{2}{\pi} \int_0^\pi u_0(x) \sin nx dx$$

はフーリエ係数である. したがって, この問題は

$$K(x, y) = \frac{2}{\pi} \sum_{n=1}^{\infty} e^{-n^2} \sin nx \sin ny$$

を積分核とする第1種フレドホルム積分方程式

$$\int_0^\pi K(x, y) u_0(y) dy = g(x)$$

に帰着できる.

1.3 一般逆作用素

今, $\{\varphi_i\}, \{\psi_i\}$ をそれぞれ $L_2(I_t), L_2(I_s)$ の任意の完全正規直交基底とする. 積分核 $K(s, t) \in L_2(I_s \times I_t)$ は

$$K(s, t) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \alpha_{ij} \psi_i(s) \varphi_j(t) \quad (1.3)$$

と表すことができる. もし, 上式の右辺の級数が有限級数ならば K は退化核であるという. そうでないとき, K は非退化核であるという [6].

ここで,

$$a_{ij} = \int_{I_s} \left(\psi_i(s) \int_{I_t} K(s, t) \varphi_j(t) dt \right) ds \quad (1.4)$$

を ij 成分とする行列を A ,

$$b_i = \int_{I_s} \psi_i(s) g(s) ds$$

を第 i 成分とするベクトルを b とすると, 第 1 種フレドホルム積分方程式 (1.1) は, 点列空間 l_2 上で定義された線形方程式

$$Ax = b \quad (1.5)$$

に変換することができる. このとき, 解ベクトル x の要素は, 解関数 f の $\{\varphi_j\}$ によるフーリエ係数となる. また, K が退化核であるとき, A が有限次元行列となるのは定義より明らかである. 以下 A は l_2 から l_2 への有界線形作用素とする.

まず, 方程式 (1.5) の最小 2 乗問題について考察する. $Ax = b$ の最小 2 乗解 \bar{x}_0 は次の定理により特徴付けられる. なお, $\bar{}$ は集合 の閉包を意味する.

定理 1.3.1 [8] A を l_2 から l_2 への有界線形作用素, P を l_2 から $\overline{\mathcal{R}(A)}$ への直交射影とする. このとき, $\bar{x}_0 \in l_2$ に対して, 次の 3 条件は同値である.

- (i) $A\bar{x}_0 = Pb$
- (ii) 任意の $x \in l_2$ に対して, $\|A\bar{x}_0 - b\| \leq \|Ax - b\|$
- (iii) $A^T A\bar{x}_0 = A^T b$

[証明] 先ず (i) から (ii) を示す. ピタゴラスの定理と $Pb - b \in \overline{\mathcal{R}(A)}^\perp$ より, 任意の $x \in l_2$ について

$$\begin{aligned} \|Ax - b\|^2 &= \|Ax - Pb\|^2 + \|Pb - b\|^2 \\ &= \|Ax - Pb\|^2 + \|A\bar{x}_0 - b\|^2 \\ &\geq \|A\bar{x}_0 - b\|^2 \end{aligned}$$

が成り立つ.

次に, (ii) から (iii) を示す. $Pb \in \overline{\mathcal{R}(A)}$ ゆえ, $\lim_n Ax_n = Pb$ なる列 $\{x_n\} \subset l_2$ が存在し, 仮定より

$$\|b - Pb\|^2 = \lim_n \|b - Ax_n\|^2 \geq \|b - A\bar{x}_0\|^2$$

がいえる. これより,

$$\begin{aligned} \|A\bar{x}_0 - b\|^2 &= \|A\bar{x}_0 - Pb\|^2 + \|Pb - b\|^2 \\ &\geq \|A\bar{x}_0 - Pb\|^2 + \|b - A\bar{x}_0\|^2 \end{aligned}$$

となり, $A\bar{x}_0 = Pb$ が示される. ゆえに,

$$A\bar{x}_0 - b = Pb - b \in \overline{\mathcal{R}(A)}^\perp = \mathcal{N}(A^T)$$

となるから, $A^T(A\bar{x}_0 - b) = 0$, つまり, (iii) が示せた.

最後に (iii) から (i) を示す. 今, $A^T A\bar{x}_0 = A^T b$ であるから $A\bar{x}_0 - b \in \overline{\mathcal{R}(A)}^\perp$ である. これより,

$$P(A\bar{x}_0 - b) = A\bar{x}_0 - Pb = 0$$

となる. ■

与えられた $b \in l_2$ に対し, 上の条件 (i)-(iii) を満たすベクトル \bar{x}_0 全体を \mathcal{C}_b と書くことにすれば, 次の定理が得られる.

定理 1.3.2 [28]

(1) $\mathcal{C}_b \subset l_2$ は閉凸集合である.

(2) 次の3条件は同値である.

(i) $\mathcal{C}_b \neq \emptyset$.

(ii) $Pb \in \mathcal{R}(A)$.

(iii) $b \in \mathcal{R}(A) + \mathcal{R}(A)^\perp$.

これより, $\mathcal{C}_b \neq \emptyset$ ならば, 方程式 $Ax = b$ の最小2乗最小ノルム問題

$$\inf\{\|x\| : x \in \mathcal{C}_b\}$$

に一意解が存在することがわかる. しかし, $\mathcal{R}(A)$ が閉部分空間でないならば, $Pb \notin \mathcal{R}(A)$ なる $b \in l_2$ が存在する. このような b に対しては上の最小2乗最小ノルム問題は解を持たない. 線形方程式 $Ax = b$ の最小2乗最小ノルム解 x_0 は, 次に定義する一般逆作用素により表される.

定義 1.3.1 作用素 A は l_2 から l_2 への有界線形作用素. $b \in \mathcal{D}(A^\dagger) = \mathcal{R}(A) + \mathcal{R}(A)^\perp$, x_0 を方程式 $Ax = b$ の最小2乗最小ノルム解とする. このとき, $A^\dagger b = x_0$ で定義される作用素 A^\dagger を A の一般逆作用素という.

一般逆作用素 A^\dagger の性質を以下に述べる. 先ず, $\mathcal{R}(A)$ が閉部分空間であるとき, A^\dagger は有界線形作用素となることを示す. そのために二つの補題を与える.

補題 1.3.1 [8] A は l_2 から l_2 への有界線形作用素. $\mathcal{R}(A)$ は閉部分空間とする. このとき, 任意の $x \in \mathcal{N}(A)^\perp$ に対して

$$\|Ax\| \geq C\|x\|$$

なる $C > 0$ が存在する.

[証明] 作用素 $A|_{\mathcal{N}(A)^\perp}$ を A の $\mathcal{N}(A)^\perp$ への制限とすると, $A|_{\mathcal{N}(A)^\perp}$ は完備な空間 $\mathcal{N}(A)^\perp$ から完備な空間 $\mathcal{R}(A)$ への1対1な上への写像となるから, $(A|_{\mathcal{N}(A)^\perp})^{-1}$ は有界線形作用素となる. ここで

$$C = \frac{1}{\|(A|_{\mathcal{N}(A)^\perp})^{-1}\|}$$

とおくと, 任意の $x \in \mathcal{N}(A)^\perp - \{0\}$ について

$$\frac{\|Ax\|}{\|x\|} \geq \frac{1}{\|(A|_{\mathcal{N}(A)^\perp})^{-1}\|} \quad \text{for } \forall x \in \mathcal{N}(A)^\perp - \{0\}$$

であることからこの補題が得られる. ただし, 作用素 A のノルム $\|A\|$ は

$$\|A\| \stackrel{d}{=} \sup_x \frac{\|Ax\|}{\|x\|}$$

により定義される. ■

補題 1.3.2 [8] 上の補題と同様の仮定の下で,

$$\mathcal{R}(A^\dagger) = \mathcal{R}(A^T) = \mathcal{R}(A^\dagger A)$$

である.

[証明] 先ず, $\mathcal{R}(A^\dagger) = \mathcal{R}(A^T)$ を示す. $\mathcal{R}(A^T) = \mathcal{N}(A)^\perp$ より, 任意の $b \in \mathcal{D}(A^\dagger)$ について, $A^\dagger b \in \mathcal{N}(A)^\perp$ を示せばよい. 今, $A^\dagger b$ は二つのベクトル $u_1 \in \mathcal{N}(A)^\perp, u_2 \in \mathcal{N}(A)$ を用いて

$$A^\dagger b = u_1 + u_2 \in \mathcal{N}(A)^\perp + \mathcal{N}(A)$$

と表されると仮定する. このとき u_1 は

$$Au_1 = A(u_1 + u_2) = AA^\dagger b = Pb$$

であるから, 方程式 $Ax = b$ の最小 2 乗解となる. ここで, P は l_2 から $\mathcal{R}(A)$ への直交射影である. そして, $u_2 \neq 0$ とすると,

$$\|u_1\|^2 < \|u_1 + u_2\|^2 = \|A^\dagger b\|^2$$

となり, $A^\dagger b$ がノルム最小の最小 2 乗解であることと矛盾する. これより, $u_2 = 0$, つまり, $A^\dagger b = u_1 \in \mathcal{N}(A)^\perp$ が示せた.

次に逆を示す. 任意の $u \in \mathcal{N}(A)^\perp$ に対し, $b = Au$ とする. ベクトル u が方程式 $Ax = b$ の最小 2 乗最小ノルム解であること, つまり, $u = A^\dagger b$ を示せばよい. ベクトル u が $Ax = b$ の最小 2 乗解であることは

$$Au = PAu = Pb$$

より明らかである. ここで, v をそれ以外の任意の最小 2 乗解とすると,

$$Av = Pb = Au$$

より $\bar{u} \stackrel{d}{=} u - v \in \mathcal{N}(A)$ となる. これより,

$$\|v\|^2 = \|u\|^2 + \|\bar{u}\|^2 \geq \|u\|^2$$

であるから, u は $Ax = b$ の最小 2 乗最小ノルム解となる.

また, 任意の $b \in l_2$ について,

$$A^\dagger b = A^\dagger Pb \in \mathcal{R}(A^\dagger A)$$

となるから, $\mathcal{R}(A^\dagger) = \mathcal{R}(A^\dagger A)$ が示せた. ■

定理 1.3.3 [8] A を l_2 から l_2 への有界線形作用素, $\mathcal{R}(A)$ は閉部分空間であるとする. このとき A^\dagger は有界線形作用素となる.

[証明] 線形性は明らかであるから, 有界性のみを示す. 任意の $b \in l_2$ について, 補題 1.3.2 より $A^\dagger b \in \mathcal{R}(A^\dagger) = \mathcal{N}(A)^\perp$ である. ゆえに, 補題 1.3.1 より

$$\|AA^\dagger b\| \geq C\|A^\dagger b\|$$

なる $C > 0$ が存在する. そして, $AA^\dagger b = Pb$ となることから,

$$\|b\| \geq \|Pb\| \geq C\|A^\dagger b\|$$

となり, A^\dagger の有界性が証明された. ■

次に, A^\dagger が有界作用素となるのは $\mathcal{R}(A)$ が閉部分空間となることに限ることを示す.

定理 1.3.4 [8] 作用素 A を l_2 から l_2 への有界線形作用素とする. A^\dagger が有界線形作用素となるための必要十分条件は, $\mathcal{R}(A)$ が閉部分空間となることである.

[証明] $\mathcal{R}(A)$ が閉部分空間のとき A^\dagger が有界線形作用素となるのは定理 1.3.3 で示した。ここでは $\mathcal{R}(A)$ は閉部分空間ではないと仮定する。 P を l_2 から $\overline{\mathcal{R}(A)}$ への直交射影とすると、任意の $b \in \mathcal{D}(A^\dagger)$ について

$$AA^\dagger b = Pb$$

となる。 A^\dagger を有界作用素とすると、 $\overline{\mathcal{D}(A^\dagger)} = l_2$ 上で定義された有界作用素 \hat{A}^\dagger へと一意的に拡張できる。しかし、 $b \in \overline{\mathcal{R}(A)} - \mathcal{R}(A)$ に対して $AA^\dagger b \neq Pb$ である。 ■

以上のことより、線形作用素 A の一般逆行列 A^\dagger が有界となるのは、 $\mathcal{R}(A)$ が閉部分空間のときで、また、そのときに限ることがわかった。

次に、コンパクト作用素の一般逆作用素について考察する。第 1 種フレドホルム型積分作用素

$$(Tf)(s) \stackrel{d}{=} \int_{I_1} K(s,t)f(t)dt, \quad s \in I, \quad (1.6)$$

は $K(s,t) \in L_2(I_s \times I_t)$ のときコンパクト作用素となる。上式を離散化することにより得られる作用素 A も同様である。それを示すために、先ず、コンパクト集合の定義を与える。

定義 1.3.2 位相空間 X の部分集合 \mathcal{U} が、次の性質を持つとき、コンパクト集合であるという。

[X の開集合の族 $\{\mathcal{G}_\nu\}$ があって、 $\cup_\nu \mathcal{G}_\nu \supset \mathcal{U}$ ならば、適当な有限個の ν_1, \dots, ν_n をとって $\cup_{i=1}^n \mathcal{G}_{\nu_i} \supset \mathcal{U}$ とできる。すなわち、 \mathcal{U} の開被覆は有限部分被覆を持つ。

$\overline{\mathcal{U}}$ がコンパクト集合であるとき、 \mathcal{U} を相対コンパクトという。

定義 1.3.3 位相空間 X の部分集合 \mathcal{U} が次の性質を持つとき、点列コンパクトであるという。

[\mathcal{U} の元の列はつねに \mathcal{U} の元に収束する部分列を持つ。

X が距離空間のときはコンパクトと点列コンパクトは一致する [22]。我々は主にヒルベルト空間で議論を進めるため、以後コンパクト集合の定義として、点列コンパクトを採用する。

次に示す補題は、Riesz の補題として知られている。

補題 1.3.3 [22] X をノルム空間、 M を X の閉真部分空間とする。このとき、任意の $\varepsilon > 0$ に対して、 $\|x\| = 1$ かつ、

$$d(x, M) \geq 1 - \varepsilon$$

を満たす $x \in X$ が存在する。ここで、 $d(x, M) \stackrel{d}{=} \sup_{y \in M} \|x - y\|$ である。

[証明] $x_0 \in X - M$ をとり、 $d_0 = d(x_0, M)$ とおけば、 $d_0 > 0$ である。ゆえに、 $y_0 \in M$ で

$$\|x_0 - y_0\| \leq (1 + \varepsilon)d_0$$

となるものが存在する。 $\|x_0 - y_0\| = r$ とし

$$x = \frac{x_0 - y_0}{r}$$

とおけば、 $\|x\| = 1$ であり、また任意の $y \in M$ に対して

$$\begin{aligned} \|x - y\| &= \left\| \frac{x_0 - y_0}{r} - y \right\| \\ &= \frac{1}{r} \|x_0 - (y_0 + ry)\| \\ &\geq \frac{d_0}{r} \geq \frac{1}{1 + \varepsilon} \geq 1 - \varepsilon \end{aligned}$$

となり、 $d(x, M) \geq 1 - \varepsilon$ が成り立つ。 ■

定理 1.3.5 [22] 無限次元ノルム空間 X において, 閉球 $\{x \in X : \|x\| \leq r\} (r > 0)$ はコンパクトではない.

[証明] 先ず, $\|x_1\| = 1$ なる $x_1 \in X$ をとる. $M_1 = \text{span}\{x_1\}$ と考えて, 上の Riesz の補題を適用すれば, $\|x_2\| = 1$ かつ $d(x_2, M_1) \geq 1/2$ となる $x_2 \in X$ がとれる. 次に, $M_2 = \text{span}\{x_1, x_2\}$ と考えて, $\|x_3\| = 1$ かつ $d(x_3, M_2) \geq 1/2$ となる $x_3 \in X$ がとれる. これを繰り返して X の元の列 $\{x_n\}$ を作れば, すべての n について $\|x_n\| = 1$ であり, また $m > n$ のとき

$$\|x_m - x_n\| \geq d(x_m, M_{m-1}), \quad M_{m-1} = \text{span}\{x_1, \dots, x_{m-1}\}$$

である. これより, $r > 0$ のとき $\{rx_n\}$ は閉球 $\{x \in X : \|x\| \leq r\}$ に属する元の列で, そのいかなる部分集合も収束しない. よって, この閉球はコンパクトではない. ■

コンパクト作用素は次のように定義される.

定義 1.3.4 X, Y をノルム空間とし, X から Y への線形作用素 T を考える. 容易にわかるように, T が有界であるための必要十分条件は, X の有界集合 \mathcal{U} の像 $T\mathcal{U}$ がつねに Y で有界となることである. これに対し, 有界集合 \mathcal{U} の像 $T\mathcal{U}$ がつねに Y で相対コンパクトであるとき, T はコンパクト作用素であるという.

ノルム空間 X の部分集合 \mathcal{U} が相対コンパクトであれば, \mathcal{U} は有界集合である [22]. これより, コンパクト作用素は必ず有界作用素である. しかし, この逆は一般には成立しない. 例えば, 無限次元ノルム空間 X における恒等作用素 I は有界であるが, 定理 1.3.5 よりコンパクトではない.

線形作用素 A が有界で, $\dim \mathcal{R}(A) < \infty$ であるとき, A を有限階の作用素 (または退化作用素) と呼ぶ. 有限次元ノルム空間の有界な部分集合は相対コンパクトになる [22] ことから, 有限階の作用素はつねにコンパクト作用素となる.

定理 1.3.6 [22] $\{A_n\}$ を l_2 から l_2 へのコンパクト作用素列とする. このとき, $\|A - A_n\| \rightarrow 0 (n \rightarrow \infty)$ ならば, A もまたコンパクト作用素となる.

[証明] 対角線論法を用いて証明する. $\{A_n\}$ はコンパクト作用素だから, 定義より有界列 $\{x_n\}$ に対して, 部分列が存在して $\{A_1 x_n^{(1)}\}$ は収束する. さらに, $\{x_n^{(1)}\}$ の部分列 $\{x_n^{(2)}\}$ が存在して $\{A_2 x_n^{(2)}\}$ は収束する. これを繰り返して, 一般に $\{x_n^{(m)}\}$ は $\{x_n^{(m-1)}\}$ の部分列で $\{A_m x_n^{(m)}\}$ が収束するようにとる. そこで, $y_n = x_n^{(n)}$ とおけば, 明らかに $\{y_n\}$ は $\{x_n\}$ の部分列であり, すべての A_m に対して $\{A_m y_n\}$ は収束する.

任意の $\varepsilon > 0$ に対して, $\|A_m - A\| < \varepsilon$ となるような m を固定する. $\{A_m y_n\}$ が収束するから, 番号 n_0 を十分大きくとれば

$$k, j \geq n_0 \implies \|A_m y_k - A_m y_j\| < \varepsilon.$$

これより, $p, q \geq n_0$ ならば, $c = \sup_n \|x_n\|$ とおいて

$$\begin{aligned} \|A y_k - A y_j\| &\leq \|A y_k - A_m y_k\| + \|A_m y_k - A_m y_j\| + \|A_m y_j - A y_j\| \\ &< \varepsilon (\|y_k\| + 1 + \|y_j\|) \\ &\leq (2c + 1)\varepsilon. \end{aligned}$$

よって, $\{A y_n\}$ はコーシー列であるから, l_2 の完備性より収束する. よって A はコンパクトである. ■

今, $K(s, t)$ を (1.3) により与えられる非退化核とし, 退化核 $K_n(s, t)$ を次式のように定義する.

$$K_n(s, t) \stackrel{d}{=} \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} \psi_i(s) \varphi_j(t).$$

作用素 A_n は退化核 $K_n(s, t)$ を持つ第1種フレドホルム型積分作用素

$$(T_n f)(s) \stackrel{d}{=} \int_{I_t} K_n(s, t) f(t) dt$$

を (1.4) により離散化されたものとする。このとき、 A_n は有限階で、明らかに

$$\|A - A_n\| \rightarrow 0, \quad (n \rightarrow \infty)$$

である。これより、(1.4) により得られた作用素 A はコンパクト作用素となることが示された。

次に、コンパクト作用素の固有値に関する重要な定理を述べる。

定理 1.3.7 [25] A を l_2 からそれ自身へのコンパクト作用素とすれば、 A の固有値は高々加算個の点からなり、それは 0 以外には集積しない。また、 λ を A の非零の固有値とすれば、 λ に属する固有空間の次元は有限である。

[証明] はじめに、 $\{\lambda_k\}_{k=1}^n$ を相異なる有限個の固有値とし、 \mathbf{x}_k を λ_k に対応する固有ベクトルとすれば、 $\{\mathbf{x}_j\}_{j=1}^n$ は 1 次独立であることを証明する。

$\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{m-1}\}$ は 1 次独立であるが、 $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ は 1 次独立でなかったとすると、

$$\mathbf{x}_m = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_{m-1} \mathbf{x}_{m-1}$$

と表される。 $A\mathbf{x}_m = \lambda_m \mathbf{x}_m$ を用いると、

$$0 = \alpha_1(\lambda_1 - \lambda_m)\mathbf{x}_1 + \alpha_2(\lambda_2 - \lambda_m)\mathbf{x}_2 + \dots + \alpha_{m-1}(\lambda_{m-1} - \lambda_m)\mathbf{x}_{m-1}$$

となる。ところが、 $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{m-1}\}$ は独立だから

$$\alpha_1 = \alpha_2 = \dots = \alpha_{m-1} = 0$$

となる。すると $\mathbf{x}_m = 0$ となり矛盾となる。よって $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ は 1 次独立である。

定理の前半部分を証明するためには、任意の $\varepsilon > 0$ に対して $|\lambda| \geq \varepsilon$ となるような固有値 λ は高々有限個しかないことをいえば十分である。そこで、ある $\varepsilon > 0$ に対して $|\lambda_n| \geq \varepsilon$ となるような相異なる固有値 λ_n が無限個存在したと仮定する。そうすると、それらの固有値に対応する固有ベクトル $\{\mathbf{x}_n\}$ は 1 次独立になる。 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ から張られる線形部分空間を

$$M_n \stackrel{d}{=} \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$$

と書くと、 M_{n-1} は M_n の閉真部分空間となる。 M_n は有限次元だから、補題 1.3.3 より、 $\|\mathbf{u}_n\| = 1$ で、 M_{n-1} のどの \mathbf{x} に対しても $\|\mathbf{u}_n - \mathbf{x}\| \geq 1/2$ であるような $\mathbf{u}_n \in M_n$ が存在する。 \mathbf{u}_n は $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ の 1 次結合で表されるから、 $A\mathbf{u}_n$ は $\{A\mathbf{x}_1, A\mathbf{x}_2, \dots, A\mathbf{x}_n\}$ 、すなわち $\{\lambda_1 \mathbf{x}_1, \lambda_2 \mathbf{x}_2, \dots, \lambda_n \mathbf{x}_n\}$ の 1 次結合で表され、結局、 $A\mathbf{u}_n$ は $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ の 1 次結合で表される。ところが、 $\mathbf{x} = \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \dots + \beta_n \mathbf{x}_n$ と表される点 \mathbf{x} に対しては

$$(\lambda_n - A)\mathbf{x} = \beta_1(\lambda_n - \lambda_1)\mathbf{x}_1 + \dots + \beta_{n-1}(\lambda_n - \lambda_{n-1})\mathbf{x}_{n-1}$$

が成り立つから $(\lambda_n - A)\mathbf{u}_n \in M_{n-1}$ となり、したがって、 $1 \leq m < n$ ならば

$$z = (\lambda_n - A)\mathbf{u}_n + A\mathbf{u}_m \in M_{n-1}$$

となる. そうすると

$$\begin{aligned} Au_n - Au_m &= \lambda_n u_n - (\lambda_n u_n - Au_n + Au_m) \\ &= \lambda_n \left(u_n - \frac{z}{\lambda_n} \right) \end{aligned}$$

だから

$$\|Au_n - Au_m\| \geq \frac{|\lambda_n|}{2} \geq \frac{\epsilon}{2}$$

となる. ところが $\|u_n\| = 1$ だから $\{u_n\}$ は有界列で, A がコンパクト作用素だから, その像 $\{Au_n\}$ は収束列を含む. これは上式より明らかに不可能である.

次に λ に属する固有空間は有限次元であることを示す. λ に属する固有空間を M とすれば, M は

$$M = \{x \in l_2 : Ax = \lambda x\}$$

と表せる. そこで, $\frac{1}{\lambda}A$ を M に制限した作用素は明らかに M における恒等作用素で, しかもコンパクトである. よって, M においては有界集合は必ず相対コンパクトであるから, 定理 1.3.5 より M は有限次元である. ■

コンパクト作用素の一般逆作用素は次のような性質を持つ.

定理 1.3.8 [20] A を l_2 から l_2 へのコンパクト作用素とする. このとき, $\mathcal{R}(A)$ が閉部分空間となるのは, $\mathcal{R}(A)$ が有限次元のときで, また, そのときに限る.

[証明] A はコンパクト作用素で, $\mathcal{R}(A)$ は閉部分空間であるとする. このとき, 定理 1.3.4 より, A^\dagger は $\mathcal{D}(A^\dagger) = l_2$ で有界となり, AA^\dagger もコンパクト作用素となる. P を l_2 から $\mathcal{R}(A)$ への直交射影とすると, $AA^\dagger = P$ となる. そして, AA^\dagger の $\mathcal{R}(A)$ への制限を $AA^\dagger|_{\mathcal{R}(A)}$ とすると,

$$AA^\dagger|_{\mathcal{R}(A)} = P|_{\mathcal{R}(A)} = I|_{\mathcal{R}(A)}$$

となり, $\mathcal{R}(A)$ から $\mathcal{R}(A)$ への恒等写像 $I|_{\mathcal{R}(A)}$ はコンパクト作用素とならなければならない. しかし, 定理 1.3.7 より, 恒等写像がコンパクト作用素となるのは値域が有限次元のときに限る. これより, A^\dagger が有界ならば $\mathcal{R}(A)$ が有限次元となることが示された.

逆は, 任意の有限次元部分空間が閉部分空間となることから明らかである. ■

上の定理より, 第 1 種フレドホルム積分方程式を離散化した作用素 A の一般逆作用素 A^\dagger が有界となるのは, $\mathcal{R}(A)$ が有限次元のとき, つまり, $K(s, t)$ が退化核のときに限ることが示された.

ちなみに, 第 1 種フレドホルム型積分作用素 (1.6) は $C(I_t)$ から $C(I_s)$ へのコンパクト作用素とみなすこともできる. バナッハ空間 $C(I_t)$ はノルム

$$\|f\| \stackrel{d}{=} \sup_{t \in I_t} |f(t)|, \quad f \in C(I_t)$$

において完備である. $C(I_s)$ も同様である.

(1.6) により定義される, バナッハ空間 $C(I_t)$ から $C(I_s)$ への線形作用素 T がコンパクト作用素であることを証明するには, 無限次元関数空間の部分集合がコンパクトになるための条件を示す必要がある. そのため, 関数族の同程度有界と同程度連続を次に定義する.

定義 1.3.5 関数族 $\{f_\mu(t) : \mu \in \Lambda\}$ を集合 $\mathcal{U} \subseteq \mathbf{R}$ で定義された実数値関数の族とする. Λ は添数集合である. ある $M > 0$ が存在して, すべての $\mu \in \Lambda$ に対して \mathcal{U} 上で $|f_\mu(t)| \leq M$ となるとき, この族は同程度有界であ

るという。つぎに t_0 を \mathcal{U} の点としよう。任意の $\varepsilon > 0$ に対してある $\delta > 0$ が存在して、 $t \in (t_0 - \delta, t_0 + \delta) \cap \mathcal{U}$ ならば $|f_\mu(t) - f_\mu(t_0)| < \varepsilon$ がすべての $\mu \in \Lambda$ に対して成り立つとき、族 $\{f_\mu(t)\}$ は t_0 において同程度連続であるという。集合 \mathcal{U} の各点で同程度連続のとき、 \mathcal{U} で同程度連続であるという。

各 $f_\mu(t)$ が連続関数であれば、 $\varepsilon > 0$ に対して $|f_\mu(t) - f_\mu(t')| < \varepsilon$ が $|t - t'| < \delta$ ならば成り立つように $\delta > 0$ はとれるわけであるが、この δ がすべての $\mu \in \Lambda$ に対して共通にとれるというのが同程度連続ということの意味である。

次の定理は Ascoli-Arzelà の定理として知られている。

定理 1.3.9 [25], [36] $\{f_\mu(t)\}$ を有界閉区間 $[a, b]$ で同程度有界、同程度連続な関数族とすれば、 $[a, b]$ 上で一様収束するような関数列がその族から選び出せる。すなわち、同程度有界で同程度連続な関数族は $C[a, b]$ 内でコンパクトである。

[証明] 仮定より、 $\{f_\mu(t)\}$ は $[a, b]$ 上で同程度連続であるから、任意の $\varepsilon > 0$ に対して、ある $\delta = \delta(\varepsilon)$ が存在して、 $t, t' \in [a, b]$, $|t - t'| < \delta$ ならば $|f_\mu(t) - f_\mu(t')| < \varepsilon$ が成り立つ。

今、 $[a, b]$ に含まれる有理点の集合を $\{r_n\}_{n=1}^\infty$ とする。 $\{f_\mu(t)\}$ が $[a, b]$ で同程度有界だから、 $\{f_\mu(r_1)\}$ は有界集合となる。したがって、それは収束する数列を含む。それを $\{f_{n_1}(r_1)\}$ とする。次に、 $\{f_{n_1}(r_2)\}$ を考えれば、それも有界数列だから、収束する部分数列 $\{f_{n_2}(r_2)\}$ を含む。同様に $\{f_{n_2}(r_3)\}$ は収束する部分数列を含む。この操作を続けた後、対角線にくる関数列 $\{f_{n_n}(t)\}$ に注目すれば、 t がどの有理数 r_j であっても $\{f_{n_n}(r_j)\}$ は収束していることがわかる。この $f_{n_n}(t)$ を改めて $f_n(t)$ と書く。

次に、 $\{f_n(t)\}$ が $[a, b]$ 上で一様収束することを示す。 $\{f_n(t)\}$ は $\{f_\mu(t)\}$ の部分列であるから同程度連続である。したがって、任意の $\varepsilon > 0$ に対して、 $t', t'' \in [a, b]$, $|t' - t''| < \delta$ ならば $|f_n(t') - f_n(t'')| < \varepsilon/3$ となるように、すべての n に対して共通な $\delta > 0$ がとれる。有理数の集合 $\{r_n\}$ は $[a, b]$ 内の稠密集合だから、 $t \in [a, b]$ の任意の点とすると $|t - r| < \delta$ であるような $r \in \{r_n\}$ は必ず存在する。これより、

$$\begin{aligned} |f_n(t) - f_m(t)| &\leq |f_n(t) - f_n(r)| + |f_n(r) - f_m(r)| + |f_m(r) - f_m(t)| \\ &\leq |f_n(r) - f_m(r)| + \frac{2}{3}\varepsilon \end{aligned}$$

となる。 $\{f_n(t)\}$ は有理点上では収束しているから、ある番号 N が存在して、 $n, m > N$ ならば $|f_n(r) - f_m(r)| < \varepsilon/3$ となるから、

$$|f_n(t) - f_m(t)| < \varepsilon$$

がいえた。すなわち、 $\{f_n(t)\}$ は $[a, b]$ 上で一様収束している。■

上の定理を用いて、第 1 種フレドホルム型積分作用素がコンパクト作用素となることを証明する。

定理 1.3.10 [25] (1.6) により定義された作用素 T はコンパクト作用素となる。

[証明] $F = C(I_t)$, $G = C(I_s)$, $K(s, t)$ は $I_s \times I_t$ で連続であるとする。また、 \mathfrak{B} を F の有界集合、 $f \in F$ について $g = Tf$ とする。 \mathfrak{B} は有界集合だから、ある $M > 0$ が存在して $f \in \mathfrak{B}$ ならば $\sup_{a \leq t \leq b} |f(t)| \leq M$ である。また、 $K(s, t) \in C(I_s \times I_t)$ だから、一様連続である。したがって、任意の $\varepsilon > 0$ に対して $|s - s'| < \delta$ かつ $|t - t'| < \delta$ ならば $|K(s, t) - K(s', t')| < \varepsilon$ となるような $\delta > 0$ がとれる。そうすると、

$$\begin{aligned} |g(s) - g(s')| &\leq \int_a^b |K(s, t) - K(s', t)| |f(t)| dt \\ &\leq \varepsilon M(b - a) \end{aligned}$$

となるから、これは $\{g(s) : g = Tf, f \in \mathfrak{B}\}$ が同程度連続であることを示す。よって、定理 1.3.9 より、 T はコンパクト作用素となる。■

1.4 特異値分解

前節で我々は一般逆行列の様々な性質を見てきた。この節では、最小2乗最小ノルム解の具体的な表現法として最も一般的である特異値分解について考察する。まずはじめに、特異値分解の理論の基礎となる、コンパクト自己共役作用素の性質について述べる。

1.4.1 コンパクト自己共役作用素

n 次元空間での対称行列 A に対しては、 A の直交する n 個の固有ベクトル e_1, e_2, \dots, e_n を座標軸にとれば A は対角行列に標準化されることは周知の事実である。これを無限次元空間に拡張することを考える。

ここでの目的はコンパクト自己共役作用素 A の非零の固有値 $\{\lambda_n\}$ に属する固有ベクトル系 $\{x_n\}$ が $\mathcal{R}(A)$ の完全正規直交系となり、任意の $x \in \mathcal{D}(A)$ について、 Ax が

$$Ax = \sum_k \lambda_k(x, x_k)x_k$$

の形に展開できることを示すことである。

まず、自己共役作用素の定義を与える。

定義 1.4.1 l_2 から l_2 への線形作用素 A が $A = A^T$ であるとき、自己共役作用素 (またはエルミット作用素) であるという。

零作用素 O 、恒等作用素 I は明らかに自己共役作用素である。

自己共役作用素の固有ベクトルは、有限次元の対称行列の固有ベクトルと同様に、次のような性質を持つ。

定理 1.4.1 [25] A を自己共役作用素とすると、 A の相異なる固有値に属する固有ベクトルは直交する。

[証明] 今、 λ_1, λ_2 を A の相異なる固有値、 x_1, x_2 をそれぞれに属する固有ベクトルとすると

$$\lambda_1(x_1, x_2) = (\lambda_1 x_1, x_2) = (Ax_1, x_2) = (x_1, Ax_2) = (x_1, \lambda_2 x_2) = \lambda_2(x_1, x_2)$$

ゆえに $(\lambda_1 - \lambda_2)(x_1, x_2) = 0$ となり、 $\lambda_1 \neq \lambda_2$ より、 x_1 と x_2 は直交する。■

次に、自己共役作用素の正值性を定義する。

定義 1.4.2 有界線形自己共役作用素 A が任意の $x \in l_2$ に対して

$$(Ax, x) \geq 0$$

を満たすとき、 A を正值自己共役作用素という。また、二つの自己共役作用素 A, B について、 $A - B$ が正值であるとき、すなわち、任意の $x \in l_2$ について

$$(Ax, x) \geq (Bx, x)$$

であるとき、 $A \geq B$ と書く。

定理 1.4.2 [22] A を任意の有界線形作用素とすると、 $A^T A$ と AA^T はともに正值自己共役作用素で、

$$\|A^T A\| = \|AA^T\| = \|A\|^2$$

を満たす。

[証明] 任意の $\boldsymbol{x} \in \mathcal{D}(A)$ に対して

$$(A^T A \boldsymbol{x}, \boldsymbol{x}) = (A \boldsymbol{x}, A \boldsymbol{x}) = \|A \boldsymbol{x}\|^2 \geq 0$$

であるから $A^T A$ は正値自己共役作用素である。さらに、

$$\|A \boldsymbol{x}\|^2 = (A^T A \boldsymbol{x}, \boldsymbol{x}) \leq \|A^T A \boldsymbol{x}\| \|\boldsymbol{x}\| \leq \|A^T A\| \|\boldsymbol{x}\|^2$$

であるから $\|A\|^2 \leq \|A^T A\|$ である。一方

$$\|A^T A\| \leq \|A^T\| \|A\| = \|A\|^2.$$

よって、 $\|A^T A\| = \|A\|^2$ となる。 AA^T についても同様である。■

定理 1.4.3 [22] 自己共役作用素 A が $A \geq 0$ であるとき、任意の $\boldsymbol{x}, \boldsymbol{y} \in l_2$ に対して

$$(A \boldsymbol{x}, \boldsymbol{y})^2 \leq (A \boldsymbol{x}, \boldsymbol{x})(A \boldsymbol{y}, \boldsymbol{y})$$

である。

[証明] $(A \boldsymbol{x}, \boldsymbol{y}) = \alpha$ とおき、 $\alpha \neq 0$ のとき証明すればよい。今、仮定より $A \geq 0$ であるから、任意の実数 t に対して

$$\begin{aligned} (A(t\boldsymbol{x} + \alpha\boldsymbol{y}), t\boldsymbol{x} + \alpha\boldsymbol{y}) &= t^2(A\boldsymbol{x}, \boldsymbol{x}) + t\alpha(A\boldsymbol{x}, \boldsymbol{y}) + t\alpha(A\boldsymbol{y}, \boldsymbol{x}) + \alpha^2(A\boldsymbol{y}, \boldsymbol{y}) \\ &= t^2(A\boldsymbol{x}, \boldsymbol{x}) + 2t\alpha^2 + \alpha^2(A\boldsymbol{y}, \boldsymbol{y}) \\ &\geq 0 \end{aligned}$$

が成り立つ。したがって、上式を t の 2 次式とみたときの判別式は

$$\alpha^4 - \alpha^2(A\boldsymbol{x}, \boldsymbol{x})(A\boldsymbol{y}, \boldsymbol{y}) \leq 0$$

でなければならない。 $\alpha \neq 0$ より定理が証明された。■

定理 1.4.4 [22] 自己共役作用素 A が $0 \leq A \leq I$ であるならば、すべての自然数 n に対して $A^{n+1} \leq A^n$ である。

[証明] 仮定より、任意の $\boldsymbol{x} \in \mathcal{D}(A)$ に対して

$$0 \leq (AA\boldsymbol{x}, A\boldsymbol{x}) \leq (A\boldsymbol{x}, A\boldsymbol{x})$$

であるから、定理 1.4.3 の \boldsymbol{y} を $A\boldsymbol{x}$ とおくことにより、

$$(A\boldsymbol{x}, A\boldsymbol{x})^2 \leq (A\boldsymbol{x}, \boldsymbol{x})(A(A\boldsymbol{x}), A\boldsymbol{x}) \leq (A\boldsymbol{x}, \boldsymbol{x})(A\boldsymbol{x}, A\boldsymbol{x})$$

となる。これより

$$(A^2\boldsymbol{x}, \boldsymbol{x}) = (A\boldsymbol{x}, \boldsymbol{x}) \leq (A\boldsymbol{x}, \boldsymbol{x})$$

すなわち、 $A^2 \leq A$ が成り立つ。よって、定理は $n=1$ のとき成り立ち、また $A^0 = 0$ として $n=0$ でも成り立っている。

$n \geq 2$ のときは、 $n-2$ で成り立っているとすれば、

$$(A^{n+1}\boldsymbol{x}, \boldsymbol{x}) = (A^{n-1}(A\boldsymbol{x}), A\boldsymbol{x}) \leq (A^{n-2}(A\boldsymbol{x}), A\boldsymbol{x}) = (A^n\boldsymbol{x}, \boldsymbol{x})$$

となり、帰納法によりすべての n で成り立つ。■

定理 1.4.5 [22] A を有界な自己共役作用素とする. このとき

$$M_A \stackrel{d}{=} \sup_{\|x\| \leq 1} (Ax, x), \quad m_A \stackrel{d}{=} \inf_{\|x\| \leq 1} (Ax, x)$$

とおけば, 次の式が成り立つ.

$$m_A I \leq A \leq M_A I \quad (1.7)$$

$$\|A\| = \max\{M_A, |m_A|\} = \sup_{\|x\| \leq 1} |(Ax, x)| \quad (1.8)$$

[証明] 任意の $x \neq 0$ に対して, $y = x/\|x\|$ とおけば, $\|y\| = 1$ であるから

$$(Ax, x) = \|x\|^2 (Ay, y) \leq M_A \|x\|^2 = (M_A I x, x).$$

よって, $A \leq M_A I$ である. $m_A I \leq A$ も同様である.

次に, もし $m_A < 0$ であるならば,

$$|m_A| = -m_A = \sup_{\|x\|=1} (-(Ax, x))$$

であるから (1.8) の第 2 項と第 3 項は一致する. この値を c とおく. $-cI \leq A \leq cI$ であるから, $c = 0$ のときは $A = O$ で $\|A\| = c$ が成り立つ. $c > 0$ のときは, $O \leq A + cI \leq 2cI$ より, $S = (2c)^{-1}(A + cI)$ とおけば $O \leq S \leq I$ である. よって, 定理 1.4.4 より

$$2c(A + cI) = 4c^2 S \geq 4c^2 S^2 = (A + cI)^2 = A^2 + 2cA + c^2 I$$

となる. したがって, $A^2 \leq c^2 I$ であるから, 任意の $x \in l_2$ に対して

$$\|Ax\|^2 = (Ax, Ax) = (A^2 x, x) \leq c^2 \|x\|^2$$

となる. これより, $\|A\| \leq c$ がいえた. 一方, $\|x\| \leq 1$ のとき

$$|(Ax, x)| \leq \|Ax\| \|x\| \leq \|A\| \|x\|^2 \leq \|A\|$$

であるから, $c \leq \|A\|$ となる. よって, (1.8) が証明された. ■

定理 1.4.6 [25] A をコンパクト自己共役作用素で零作用素でないものとする. そうすると, $\|A\|$ または $-\|A\|$ は A の固有値であり, それに属する固有ベクトルで $|(Ax, x)| = \|A\|$, $\|x\| = 1$ となるようなものが存在する.

[証明] 定理 1.4.5 から, $\|x_n\| = 1$ で (Ax_n, x_n) が $\|A\|$ または $-\|A\|$ に収束するような $\{x_n\}$ が存在することがわかる. 収束する値. すなわち $\|A\|$ か $-\|A\|$ かのどちらか一方を λ と書くと, λ は実数だから

$$\begin{aligned} 0 \leq \|Ax_n - \lambda x_n\|^2 &= \|Ax_n\|^2 - 2\lambda(Ax_n, x_n) + \lambda^2 \|x_n\|^2 \\ &\leq \|A\|^2 - 2\lambda(Ax_n, x_n) + \lambda^2 \rightarrow 0 \end{aligned}$$

となり, $\|Ax_n - \lambda x_n\| \rightarrow 0$ がいえる. 仮定より, A はコンパクトだから, $\{Ax_n\}$ は収束する部分列を含む. それを $\{Ax_{n_i}\}$ とすると, $\lambda \neq 0$ より $\{x_{n_i}\}$ も収束しなければならない. その収束先を $x_n \rightarrow x$ とすると, $\|x\| = 1$ となる. また, $Ax_{n_i} \rightarrow Ax$ だから, $Ax = \lambda x$ となる. これは x が λ に属する固有ベクトルであることを示している. $|(Ax, x)| = \|A\|$ は明らかである. ■

定理 1.4.7 [25] A をコンパクト自己共役作用素とすれば,

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq 0$$

という固有値の列とそれぞれに属する固有ベクトルの正規直交系 $\{\mathbf{x}_k\}$ が存在して, 任意の $\mathbf{x} \in \mathcal{D}(A)$ に対して

$$A\mathbf{x} = \sum_k (A\mathbf{x}, \mathbf{x}_k)\mathbf{x}_k = \sum_k \lambda_k (\mathbf{x}, \mathbf{x}_k)\mathbf{x}_k \quad (1.9)$$

と展開できる. A の 0 でない固有値は列 $\{\lambda_k\}$ の中に必ず現れ, どの固有値 λ_i に属する固有空間も有限次元で, その次元は λ_i が列 $\{\lambda_k\}$ の中に何回現れるかという回数と一致する.

[証明] 定理 1.4.6 で得られた固有値を λ_1 , 固有ベクトルを \mathbf{x}_1 とする. $H_1 = l_2, H_2 = \{\mathbf{x} \in H_1 : (\mathbf{x}, \mathbf{x}_1) = 0\}$ とおく. $\mathbf{x} \in H_2$ ならば

$$(A\mathbf{x}, \mathbf{x}_1) = (\mathbf{x}, A\mathbf{x}_1) = (\mathbf{x}, \lambda_1\mathbf{x}_1) = \lambda_1(\mathbf{x}, \mathbf{x}_1) = 0$$

だから $AH_2 \subseteq H_2$ である.

A の H_2 への制限 $A|_{H_2}$ はコンパクト自己共役作用素であるから, もしそれが零作用素でなければ, 定理 1.4.6 を用いて, $A\mathbf{x}_2 = \lambda_2\mathbf{x}_2, \|\mathbf{x}_2\| = 1$ なる固有値 λ_2 , 固有ベクトル \mathbf{x}_2 が得られる. このとき, $|\lambda_2|$ は $A|_{H_2}$ のノルムに等しいから $|\lambda_2| \leq |\lambda_1|$ である. これを繰り返すと固有値 $\lambda_1, \lambda_2, \dots$ とそれぞれに属する固有ベクトル $\mathbf{x}_1, \mathbf{x}_2, \dots$ が得られ,

$$|\lambda_1| \geq |\lambda_2| \geq \dots$$

となる. このとき, ある n が存在して A の H_{n+1} への制限が零作用素になるならば,

$$A\mathbf{x} = \sum_{k=1}^n (\mathbf{x}, \mathbf{x}_k)A\mathbf{x}_k = \sum_{k=1}^n \lambda_k (\mathbf{x}, \mathbf{x}_k)\mathbf{x}_k$$

と表される. これは任意の $\mathbf{x} \in l_2$ に対して

$$\mathbf{y}_n = \mathbf{x} - \sum_{k=1}^n (\mathbf{x}, \mathbf{x}_k)\mathbf{x}_k$$

とおけば, $i = 1, 2, \dots, n$ に対して $(\mathbf{y}_n, \mathbf{x}_i) = 0$ であるから $\mathbf{y}_n \in H_{n+1}$ となり, $A\mathbf{y}_n = 0$ より示される. これは有限次元対称行列の対角化に他ならない.

さて, どんな n に対しても $A|_{H_{n+1}}$ が零作用素とならない場合を考える. 今, A はコンパクト作用素であるから, 定理 1.3.7 より, $\lambda_n \rightarrow 0 (n \rightarrow \infty)$ である. 上と同様に $\mathbf{y}_n = \mathbf{x} - \sum_{k=1}^n (\mathbf{x}, \mathbf{x}_k)\mathbf{x}_k$ とおくと, $\{\mathbf{x}_k\}$ の直交性から

$$\|\mathbf{y}_n\|^2 = \|\mathbf{x}\|^2 - \sum_{k=1}^n (\mathbf{x}, \mathbf{x}_k)^2 \leq \|\mathbf{x}\|^2$$

がいえ. また, $\mathbf{y}_n \in H_{n+1}$ であり, $|\lambda_{n+1}|$ は $A|_{H_{n+1}}$ のノルムに等しいから

$$\|A\mathbf{y}_n\| \leq |\lambda_{n+1}| \|\mathbf{y}_n\| \leq |\lambda_{n+1}| \|\mathbf{x}\|$$

である. これより, $A\mathbf{y}_n \rightarrow 0$ がいえた. ゆえに

$$A\mathbf{x} = \sum_{k=1}^{\infty} (A\mathbf{x}, \mathbf{x}_k)\mathbf{x}_k = \sum_{k=1}^{\infty} \lambda_k (\mathbf{x}, \mathbf{x}_k)\mathbf{x}_k$$

となる.

次に、0でない固有値 λ で上に述べたどの λ_k とも一致しないものがあつたとする。 λ に属する固有ベクトル \boldsymbol{x} はどの \boldsymbol{x}_k とも直交しているから上式より $A\boldsymbol{x} = 0$ とならなければならない。そうすると、 $0 = A\boldsymbol{x} = \lambda\boldsymbol{x}$ となって $\boldsymbol{x} = 0$ となり矛盾となる。したがって、0でない固有値はすべて上式の $\{\lambda_k\}$ のどれかになっている。

一方、定理1.3.7より、固有値は高々可算個で $\lambda_n \rightarrow 0$ だから、どの固有値も列 $\{\lambda_k\}$ の中に無限回現れるということ、すなわちその固有値に属する固有空間が無限次元であることはあり得ない。■

次に、上の定理により得られたコンパクト自己共役作用素 A の固有ベクトル系 $\{\boldsymbol{x}_k\}$ が $\overline{\mathcal{R}(A)}$ の完全正規直交系であることを示す。

定理 1.4.8 [25] A はコンパクト自己共役作用素、 $\{\lambda_n\}, \{\boldsymbol{x}_n\}$ は定理1.4.7で定められたものとする。このとき、

$$M \stackrel{d}{=} \overline{\text{span}(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots)}$$

とすると、 $M^\perp = \mathcal{N}(A)$ である。したがって、0が A の固有値でないときそのときに限って $\{\boldsymbol{x}_n\}$ は l_2 の完全正規直交系となる。

[証明] $\boldsymbol{x} \perp \boldsymbol{x}_k (k = 1, 2, \dots)$ ならば(1.9)より $\boldsymbol{x} \in \mathcal{N}(A)$ である。ゆえに $M^\perp \subseteq \mathcal{N}(A)$ である。逆に $\boldsymbol{x} \in \mathcal{N}(A)$ ならば

$$(\boldsymbol{x}, \boldsymbol{x}_k) = \frac{1}{\lambda_k} (\boldsymbol{x}, A\boldsymbol{x}_k) = \frac{1}{\lambda_k} (A\boldsymbol{x}, \boldsymbol{x}_k) = 0$$

より、 $\mathcal{N}(A) \subseteq M^\perp$ である。これで $M^\perp = \mathcal{N}(A)$ がいえた。

正規直交系 $\{\boldsymbol{x}_k\}$ が l_2 で完全系であるのは $M^\perp = \{0\}$ のときそのときだけであり、それは $\mathcal{N}(A) = \{0\}$ 、すなわち0が A の固有値ではないことを意味している。■

1.4.2 特異値分解と悪条件線形方程式

作用素 A を l_2 から l_2 へのコンパクト作用素、 $A^T A$ の固有値を $\{\lambda_n\}$ 、それに属する固有ベクトルを $\{\boldsymbol{v}_n\}$ とする。このとき、定理1.4.2より $A^T A$ は正値自己共役作用素だから、 $\lambda_k \geq 0 (k = 1, 2, \dots)$ である。今、固有値は

$$\lambda_1 \geq \lambda_2 \geq \dots \geq 0$$

であるとし、非零の λ_k に対してベクトル $\{\boldsymbol{u}_k\}$ を

$$\boldsymbol{u}_k \stackrel{d}{=} \frac{1}{\sqrt{\lambda_k}} A\boldsymbol{v}_k$$

と定義する。すると、ベクトル $\{\boldsymbol{u}_k\}$ は

$$AA^T \boldsymbol{u}_k = \frac{1}{\sqrt{\lambda_k}} AA^T A\boldsymbol{v}_k = \sqrt{\lambda_k} A\boldsymbol{v}_k = \lambda_k \boldsymbol{u}_k$$

より、 AA^T の固有値 $\{\lambda_k\}$ に属する固有ベクトルとなる。

定義 1.4.3 $\sigma_k \stackrel{d}{=} \sqrt{\lambda_k}$ を作用素 A の特異値、 $\{\boldsymbol{u}_k\}, \{\boldsymbol{v}_k\}$ をそれぞれ $\{\sigma_k\}$ に属した左、右特異ベクトルという。

特異ベクトル $\{\boldsymbol{u}_k\}, \{\boldsymbol{v}_k\}$ は、定理1.4.8より、それぞれ $\overline{\mathcal{R}(A)} = \mathcal{N}(A^T)^\perp$ 、 $\overline{\mathcal{R}(A^T)} = \mathcal{N}(A)^\perp$ の完全正規直交基底となる。

定理 1.4.9 コンパクト作用素 A は

$$U \stackrel{d}{=} [u_1, u_2, \dots], \quad V \stackrel{d}{=} [v_1, v_2, \dots], \quad \Sigma \stackrel{d}{=} \text{diag}(\sigma_1, \sigma_2, \dots)$$

を用いて

$$A = U\Sigma V^T \tag{1.10}$$

と分解することができる。

[証明] 任意の $x \in \mathcal{D}(A)$ について, $Ax = U\Sigma V^T x$ を示せばよい. $\{v_k\}$ は $\mathcal{N}(A)^\perp$ の完全正規直交基底であるから, x は

$$x = \sum_i c_i v_i + \bar{x}, \quad \bar{x} \in \mathcal{N}(A)$$

と表すことができる. したがって,

$$Ax = \sum_i c_i A v_i = \sum_i c_i \sigma_i u_i$$

となる. 一方

$$U\Sigma V^T x = \sum_i c_i U\Sigma V^T v_i = \sum_i c_i \sigma_i u_i$$

となり, $Ax = U\Sigma V^T x$ が示せた. ■

上の定理の分解 (1.10) を A の特異値分解という. 以後, 定義 1.4.3 を満たす $\{\sigma_k; u_k, v_k\}$ を A の特異値分解と呼ぶ.

この特異値分解を用いると, 方程式 $Ax = b$ の最小 2 乗最小ノルム解 $x_0 = A^\dagger b$ を陽に表すことができる.

定理 1.4.10 [10] A をコンパクト作用素, $b \in \mathcal{D}(A^\dagger) = \mathcal{R}(A) + \mathcal{R}(A)^\perp$ とする. このとき, 方程式 $Ax = b$ の最小 2 乗最小ノルム解は

$$x_0 = \sum_i \frac{(b, u_i)}{\sigma_i} v_i \tag{1.11}$$

と表すことができる.

[証明] ベクトル x_0 が \mathcal{C}_b に属し, なおかつ $\mathcal{N}(A)^\perp$ に属していることを示せばよい.

x_0 が方程式 $A^T A x = A^T b$ を満たしていることは, 分解 (1.10) を用いると容易に確かめることができる. また, $\{v_k\}$ は $\mathcal{N}(A)^\perp$ の完全正規直交基底であるから, $x_0 \in \mathcal{N}(A)^\perp$ は明らかである. ■

上の定理より, A の特異値分解が求められれば, 方程式 $Ax = b$ の最小 2 乗最小ノルム解が求められることがわかった. これについては第 4 章で詳しく述べる.

さて, x_0 のノルムについて考察しよう. x_0 のノルムは $\{v\}$ の正規直交性から

$$\|x_0\|^2 = \sum_i \left(\frac{(b, u_i)}{\sigma_i} \right)^2$$

と表すことができる. しかし, A が無限階のコンパクト作用素のとき, σ_i は 0 に集積するため, 上式の右辺の無限級数の和が常に有限値をとるとは限らない. もし, 有限値でないならば $x_0 \notin l_2$ となり, $Ax = b$ は l_2 に解を持たない.

次の定理は, 方程式 $Ax = b$ が l_2 内に解を持つための必要十分条件を表す. この定理は Picard の条件として知られている.

定理 1.4.11 [10] 方程式 $Ax = b$ が l_2 に解を持つための必要十分条件は、 $b \in \mathcal{N}(A^T)^\perp$ 、かつ

$$\sum_i \frac{|(b, u_i)|^2}{\sigma_i^2} < \infty$$

を満たすことである。

第1種フレドホルム積分方程式において、積分核 $K(s, t)$ が退化核であれば A は有限階となり、問題は不適切であるが安定である。しかも、 $b \in \mathcal{N}(A^T)^\perp$ ならば必ず解を持つ。しかし、 $K(s, t)$ が非退化核のときは σ_k は 0 に集積し、 $x_0 \in l_2$ は自明ではない。我々はこのような方程式を悪条件線形方程式と定義する。

定義 1.4.4 $\{\sigma_i\}$ を作用素 A の特異値とする。すべての特異値が正で、かつ

$$\lim_{k \rightarrow \infty} \sigma_k = 0 \tag{1.12}$$

を満たしているとき、線形方程式 $Ax = b$ は悪条件であるという。

以上のことを踏まえた上で、我々の問題を明記する。我々の目的は非退化核 $K(s, t)$ を持つ第1種フレドホルム積分方程式

$$\int_{I_1} K(s, t) f(t) dt = g(s), \quad s \in I_2$$

を解くことにある。この方程式は $L_2(I_1), L_2(I_2)$ に適当な正規直交基底を導入することにより、線形方程式

$$Ax = b, \quad b \in l_2$$

に変換できる。このとき、作用素 A はコンパクト作用素となり、 A の一般逆作用素 A^\dagger は非有界となる。それゆえ、方程式 $Ax = b$ は悪条件となる。

もし、 A の特異値分解 $\{\sigma_k; u_k, v_k\}$ が得られ、定理 1.4.11 の条件を満たしているならば、上の方程式は数学的には可解で、最小 2 乗最小ノルム解の意味で唯一解を持つ。

しかし、 A の特異値分解を計算機を用いて精度良く求めることは困難である。特に、微小な特異値に属する特異ベクトルほど難しい [12]。また、悪条件であるため、微小な摂動にすら敏感に反応してしまう。

我々は悪条件線形方程式に対して、正則化法と、直接法による解法、二つのアプローチを試みる。

第 2 章

正則化法

作用素 A がコンパクトで、なおかつ無限階であるとき、一般逆作用素 A^\dagger は有界とはならない。これは、方程式 $Ax = b$ の右辺に微小な摂動を加えただけで、解が大きく変化してしまう恐れがあることを意味している。この現象の本質は、 A の特異値が 0 に集積することにある。

一般にデータは観測誤差などのノイズが含まれていることが多い。このとき、微小な特異値により、ノイズが拡大され、解は全く無意味なものとなる。正則化法はこのような不適切性を緩和することを目的に考案された。

正則化法に要求される性質は、ノイズを 0 に収束させたとき、近似解が $b \in \mathcal{D}(A^\dagger)$ ならば $A^\dagger b$ に収束し、そうでなければ発散することである。通常 Tikhonov の正則化法と呼ばれている手法は、この問題を解決し、不適切問題に実際的な意味を与えることに成功している。

この章では、先ず 2.1 節で正則化法の一般論を述べる。次に、2.2 節ではデータがノイズを含んでいる場合の正則化法の性質を紹介する。2.3 節では、代表的な正則化法である Tikhonov の正則化法を導入し、その収束性などを考察する。

2.1 正則化法の一般論

第 1 章で述べたように、第 1 種フレドホルム積分方程式

$$\int_{I_1} K(s, t) f(t) dt = g(s), \quad s \in I_2$$

は適当な離散化を施すことにより線形方程式

$$Ax = b, \quad b \in l_2$$

に変換できる。このとき、作用素 A はコンパクトとなり、上の方程式は悪条件となる。我々の目的は悪条件線形方程式を解くことにある。

1.3 節で述べたように、 A がコンパクト作用素のとき、 A^\dagger が有界となるのは、 $\mathcal{R}(A)$ が有限次元の場合に限る。ここで、我々は適当な有界作用素 R_λ により A^\dagger を近似することを考える。ここでの近似とは、任意の $b \in \mathcal{D}(A^\dagger)$ について、

$$R_\lambda b \rightarrow A^\dagger b, \quad (\lambda \rightarrow 0)$$

を意味する。

以後、我々は $A^T A$ を \tilde{A} で表し、 $\mathbf{x}_0 \stackrel{\text{d}}{=} A^\dagger \mathbf{b}$ とする。定理 1.3.1 の (iii) より、 \mathbf{x}_0 は

$$\tilde{A} \mathbf{x}_0 = A^T \mathbf{b}$$

を満たす。これより、もし \tilde{A} が可逆であれば、 $\mathbf{x}_0 = \tilde{A}^{-1} A^T \mathbf{b}$ となる。たとえ \tilde{A} が可逆でなくとも、 \mathbf{x}_0 を

$$R_\lambda(\tilde{A}) A^T \mathbf{b}, \quad \lambda > 0$$

の形で近似することを考える。ここで、 $R_\lambda(t)$ は区間 $[0, \|A\|^2]$ 上で定義された連続関数で、作用素 $R_\lambda(A)$ は一般に次のように定義される。

定義 2.1.1 A をコンパクト自己共役作用素、 $\{\lambda_k\}$ を A の固有値、 \mathbf{x}_k をそれに属した固有ベクトルとする。区間 $[0, \|A\|^2]$ 上で定義された実数値連続関数 $f(t)$ に対して、作用素 $f(A)$ を

$$f(A) \mathbf{x} \stackrel{\text{d}}{=} \sum_i f(\lambda_i) (\mathbf{x}, \mathbf{x}_i) \mathbf{x}_i \quad (2.1)$$

と定義する。

次の定理はスペクトル写像定理と呼ばれている。

定理 2.1.1 [10] 作用素 A のスペクトル $\sigma(A)$ を

$$\sigma(A) \stackrel{\text{d}}{=} \{\lambda \in \mathbb{C} : \|(A - \lambda I)^{-1}\| = \infty\}$$

と定義する。 A を l_2 から l_2 への有界自己共役作用素、 f を $\sigma(A)$ を含む区間上で定義された実数値連続関数とする。このとき

$$\sigma(f(A)) = f(\sigma(A))$$

が成り立つ。

作用素 $R_\lambda(\tilde{A}) A^T$ により A^\dagger を近似するために任意の $\mathbf{b} \in \mathcal{D}(A^\dagger)$ について

$$\lim_{\lambda \rightarrow 0} R_\lambda(\tilde{A}) A^T \mathbf{b} = A^\dagger \mathbf{b}$$

を保証する条件を $R_\lambda(t)$ に課さなければならない。ここで我々は以下のことを仮定する。

仮定 2.1.1 (i) 任意の $t > 0$ について

$$R_\lambda(t) \rightarrow \frac{1}{t}, \quad (\lambda \rightarrow 0).$$

(ii) $|t R_\lambda(t)|$ は区間 $[0, \|A\|^2]$ で一様有界。

ここで、 $R_\lambda(t)$ は正則化関数と呼ばれている。この仮定の下で、次の定理により、収束が保証される。

定理 2.1.2 [9] $\{R_\lambda\}_{\lambda > 0}$ は仮定 2.1.1 を満たしている、区間 $[0, \|A\|^2]$ 上で定義された実数値連続関数とする。このとき、任意の $\mathbf{b} \in \mathcal{D}(A^\dagger)$ に対し、

$$R_\lambda(\tilde{A}) A^T \mathbf{b} \rightarrow A^\dagger \mathbf{b}, \quad (\lambda \rightarrow 0)$$

が成り立つ。

[証明] 任意の多項式 p について

$$p(A^T A)A^T = A^T p(AA^T)$$

が成り立つことは自明である。ワイエルシュトラスの近似定理により、 $[0, \|A\|^2]$ 区間上で定義された連続関数に対しても同様の結果が成り立つ。これより、 $\hat{A} \triangleq AA^T$ とすると

$$R_\lambda(\hat{A})A^T b = A^T R_\lambda(\hat{A})b \quad (2.2)$$

となり、 $R_\lambda(\hat{A})A^T b \in \mathcal{R}(A^T)$ である。

A の特異値分解を用いると、

$$\begin{aligned} R_\lambda(\hat{A})A^T b &= \sum_{i=1}^{\infty} R_\lambda(\sigma_i^2)(A^T b, v_i)v_i \\ &= \sum_{i=1}^{\infty} \sigma_i R_\lambda(\sigma_i^2)(b, u_i)v_i \\ &= \sum_{i=1}^{\infty} \sigma_i^2 R_\lambda(\sigma_i^2) \frac{(b, u_i)}{\sigma_i} v_i \end{aligned}$$

となり、定理の仮定より、上式は $\lambda \rightarrow 0$ のとき

$$x_c = \sum_{i=1}^{\infty} \frac{(b, u_i)}{\sigma_i} v_i$$

に収束する。■

また、データ b が $\mathcal{D}(A^\dagger)$ に入っていない場合は次の定理が成り立つ。

定理 2.1.3 [10] もし、 $b \notin \mathcal{D}(A^\dagger)$ ならば、 $\{R_{\lambda_n}(\hat{A})A^T b\}$ が弱収束するような列 $\lambda_n \rightarrow 0$ は存在しない。

[証明] 弱収束とは、内積による収束を意味する。すなわち、 x_n が $x \in l_2$ に弱収束するのを $x_n \xrightarrow{w} x$ と表すとすると、

$$x_n \xrightarrow{w} x \iff (x_n, y) \rightarrow (x, y), \quad (\forall y \in l_2)$$

である。

P を l_2 から $\overline{\mathcal{R}(A)} = \mathcal{N}(A^T)^\perp = \mathcal{N}(\hat{A})^\perp$ への直交射影とする。 A がコンパクト作用素であるための必要十分条件は、 $x_n \xrightarrow{w} x$ ならば $Ax_n \rightarrow Ax$ であることから [22]、もし、ある $x \in l_2$ に対して

$$R_{\lambda_n}(\hat{A})A^T P b = R_{\lambda_n}(\hat{A})A^T b \xrightarrow{w} z, \quad (n \rightarrow \infty)$$

ならば、(2.2) より、

$$\hat{A} R_{\lambda_n}(\hat{A}) P b \rightarrow A z, \quad (n \rightarrow \infty)$$

となる。しかし、仮定より

$$\hat{A} R_{\lambda_n}(\hat{A}) P b \rightarrow P b, \quad (n \rightarrow \infty)$$

であるから、 $P b = A z$ 、つまり、 $b \in \mathcal{D}(A^\dagger)$ となる。■

ヒルベルト空間内の任意の有界列は、弱収束する部分列を持つ。これより、直ちに次の系が得られる。

系 2.1.1 [10] もし、 $b \notin \mathcal{D}(A^\dagger)$ ならば、 $\lim_{\lambda \rightarrow 0} \|R_\lambda(\hat{A})A^T b\| = \infty$ である。

上の定理 2.1.2 と系 2.1.1 より, 関数 $R_\lambda(t)$ が仮定 2.1.1 を満たしているとき, 正則化法による近似解 $R_\lambda(\tilde{A})A^T b$ は, $b \in \mathcal{D}(A^\dagger)$ ならば解に収束し, そうでなければ発散することがわかる.

次に, 収束の速度について考察する. 以下簡単のため,

$$x_\lambda \stackrel{d}{=} R_\lambda(\tilde{A})A^T b, \quad (\lambda > 0)$$

とする.

定理 2.1.2 と定理 2.1.3 より, x_λ が x_0 に収束するための必要十分条件は, $b \in \mathcal{D}(A^\dagger)$ である. P を l_2 から $\overline{\mathcal{R}(A)}$ への直交射影とすると, この条件は, $Pb \in \mathcal{R}(A)$ と同じである. 今, P_S を閉部分空間 S への直交射影とすると,

$$\mathcal{R}(A) = \mathcal{R}(AP_{\mathcal{N}(A)^\perp})$$

が成り立つ. そして, 定理 1.4.7 より

$$P_{\mathcal{N}(A)^\perp} x = P_{\mathcal{N}(\tilde{A})^\perp} x = \lim_{\nu \rightarrow 0^+} \tilde{A}^\nu x$$

である. なお, 上式の二つめの等号は式 (2.1) において $f(t) = t^\nu$ とすることにより得られる. 我々は, $Pb \in \mathcal{R}(A)$ の条件を, ある $\nu > 0$ について, $Pb \in \mathcal{R}(A\tilde{A}^\nu)$ というより少し強い条件に置き換える.

また, 仮定 2.1.1 の (ii) の条件を, 任意の $t \in [0, \|A\|^2]$ に対して

$$t^\nu |1 - tR_\lambda(t)| \leq \omega(\lambda, \nu) \quad (2.3)$$

とする. ここで, $\omega(\lambda, \nu)$ は収束速度関数と呼ばれるもので, 任意の $\nu > 0$ について

$$\omega(\lambda, \nu) \rightarrow 0, \quad (\lambda \rightarrow 0)$$

を満たしているものとする.

補題 2.1.1 [10] 任意の $\nu \in [0, \|A\|^2]$ について

$$\mathcal{R}(\tilde{A}^\nu) \subseteq \mathcal{N}(A)^\perp.$$

が成り立つ.

[証明] $\{\lambda_k\}$ を \tilde{A} の非零固有値, $\{w_k\}$ をそれに属する固有ベクトルとすると, 定理 1.4.7 より

$$\tilde{A}^\nu x = \sum_k \lambda_k^\nu(x, w_k) w_k$$

と書ける. そして,

$$w_k = \frac{1}{\lambda_k} \tilde{A} w_k = A^T \frac{1}{\lambda_k} A w_k \in \mathcal{R}(A^T)$$

より, $\tilde{A}^\nu x \in \overline{\mathcal{R}(A^T)} = \mathcal{N}(A)^\perp$ である. ■

定理 2.1.4 [10] ある $\nu > 0$ と, ある $w \in l_2$ について, $Pb = A\tilde{A}^\nu w$ ならば,

$$\|x_0 - x_\lambda\| \leq \omega(\lambda, \nu) \|w\|$$

が成り立つ.

[証明] x_0 は仮定より

$$Ax_0 = Pb = A\tilde{A}^\nu w$$

であるから, $x_0 - \tilde{A}^\nu w \in \mathcal{N}(A)$ となる. しかし, 補題 2.1.1より, $x_0, \tilde{A}^\nu w \in \mathcal{N}(A)^\perp$, つまり $x_0 - \tilde{A}^\nu w \in \mathcal{N}(A) \cap \mathcal{N}(A)^\perp = \{0\}$ となり, $x_0 = \tilde{A}^\nu w$ である. そして

$$\begin{aligned} x_\lambda &= R_\lambda(\tilde{A})A^T b = R_\lambda(\tilde{A})A^T Pb \\ &= R_\lambda(\tilde{A})\tilde{A}x_0 = R_\lambda(\tilde{A})\tilde{A}^{\nu+1}w \end{aligned}$$

となり, これより定理 2.1.1を用いると

$$\begin{aligned} \|x_0 - x_\lambda\| &= \|\tilde{A}^\nu(I - R_\lambda(\tilde{A})\tilde{A})w\| \\ &\leq \omega(\lambda, \nu)\|w\| \end{aligned}$$

がいえる. ■

上の証明の中で, $Pb = A\tilde{A}^\nu w$ という条件が, $x_0 = \tilde{A}^\nu w$ を意味することを示した. 逆に, $x_0 = \tilde{A}^\nu w$ とすると, 定理 1.3.1の (iii) より, $Pb = Ax_0 = A\tilde{A}^\nu w$ がいえる. したがって, $Pb \in \mathcal{R}(A\tilde{A}^\nu)$ というデータ b への条件は, $x_0 \in \mathcal{R}(\tilde{A}^\nu)$ という解についての条件と同値となる.

次に我々は, ある $\nu \geq 1$ について, $Pb \in \mathcal{R}(\hat{A}^\nu)$ のときの収束速度を考察する. 反復法による正則化法の収束速度を調べるときは, このような条件を課した方が便利である. この条件は, $Pb \in \mathcal{R}(A\hat{A}^{\nu-1}A^T)$ と同値であるから, 定理 2.1.4より, $\nu > 1$ ならばただちに $\omega(\lambda, \nu - 1)$ という収束速度が得られる. しかし, 実は, これよりも速い収束速度が得られる. 先ず, 補題を1つ与える.

補題 2.1.2 [10] ある $\nu \geq 1$, ある $w \in l_2$ について, $Pb = \hat{A}^\nu w$ ならば

$$\|x_0 - x_\lambda\|^2 \leq \omega(\lambda, \nu - 1)\|A(x_0 - x_\lambda)\|\|w\|$$

が成り立つ.

[証明] 前定理の証明より,

$$x_0 = A^T \hat{A}^{\nu-1} w, \quad x_\lambda = A^T R_\lambda(\hat{A}) \hat{A}^\nu w$$

が得られた. これより

$$x_0 - x_\lambda = A^T(I - R_\lambda(\hat{A})\hat{A})\hat{A}^{\nu-1}w$$

となり

$$\begin{aligned} \|x_0 - x_\lambda\|^2 &= (x_0 - x_\lambda, A^T(I - R_\lambda(\hat{A})\hat{A})\hat{A}^{\nu-1}w) \\ &= (A(x_0 - x_\lambda), (I - R_\lambda(\hat{A})\hat{A})\hat{A}^{\nu-1}w) \\ &\leq \omega(\lambda, \nu - 1)\|A(x_0 - x_\lambda)\|\|w\| \end{aligned}$$

が得られる. ■

定理 2.1.5 [10] ある $\nu \geq 1$, ある $w \in l_2$ に対し, $Pb = \hat{A}^\nu w$ ならば,

$$\|x_0 - x_\lambda\| \leq \sqrt{\omega(\lambda, \nu - 1)\omega(\lambda, \nu)}\|w\|$$

である.

[証明] 前定理の証明より

$$\mathbf{x}_0 - \mathbf{x}_\lambda = A^T(I - R_\lambda(\hat{A})\hat{A})\hat{A}^{\nu-1}\mathbf{w}$$

が示された。上式に \hat{A} を作用させると

$$\hat{A}(\mathbf{x}_0 - \mathbf{x}_\lambda) = A^T\hat{A}^\nu(I - R_\lambda(\hat{A})\hat{A})\mathbf{w}$$

となり

$$\begin{aligned} \|A(\mathbf{x}_0 - \mathbf{x}_\lambda)\|^2 &= (\hat{A}(\mathbf{x}_0 - \mathbf{x}_\lambda), \mathbf{x}_0 - \mathbf{x}_\lambda) \\ &= (\hat{A}^\nu(I - R_\lambda(\hat{A})\hat{A})\mathbf{w}, A(\mathbf{x}_0 - \mathbf{x}_\lambda)) \\ &\leq \omega(\lambda, \nu)\|A(\mathbf{x}_0 - \mathbf{x}_\lambda)\|\|\mathbf{w}\| \end{aligned}$$

が得られる。したがって

$$\|A(\mathbf{x}_0 - \mathbf{x}_\lambda)\| \leq \omega(\lambda, \nu)\|\mathbf{w}\|$$

である。上式と前補題より

$$\|\mathbf{x}_0 - \mathbf{x}_\lambda\|^2 \leq \omega(\lambda, \nu - 1)\omega(\lambda, \nu)\|\mathbf{w}\|^2$$

となり、定理が証明された。■

2.2 ノイズを含んだデータに対する正則化法

前節では、データがノイズを含まない場合の正則化法について考察した。しかし、実際には、データに観測誤差などのノイズが含まれている場合が多い。

今、ノイズを含むデータ b^δ は、

$$\|b - b^\delta\| \leq \delta$$

を満たしているものとする。そして、そのデータを用いて、前節と同様の形式

$$\mathbf{x}_\lambda^\delta \stackrel{d}{=} R_\lambda(\hat{A})A^T b^\delta \tag{2.4}$$

により方程式 $Ax = b$ の最小2乗最小ノルム解 \mathbf{x}_0 の近似を行なう。このとき、もし、ある意味においてノイズレベル δ に応じて正則化パラメータ $\lambda = \lambda(\delta)$ を選ぶことにより、

$$\lim_{\delta \rightarrow 0} R_{\lambda(\delta)}(\hat{A})A^T b^\delta = \mathbf{x}_0$$

となれば、近似(2.4)は正則であるという。正則化法は、正則化関数 R_λ の選択だけでなく、正則化パラメータ $\lambda(\delta)$ の選択法にも依存する。

以下、仮定2.1.1に対応して、

$$\sup_{\lambda > 0} \max_{0 \leq t \leq \|A\|^2} |tR_\lambda(t)| = C^2 \tag{2.5}$$

なる定数 C が存在すると仮定する。また、関数 $r(\lambda)$ を

$$r(\lambda) \stackrel{d}{=} \max\{|R_\lambda(t)| : t \in [0, \|A\|^2]\} \tag{2.6}$$

と定義する。

補題 2.2.1 [10]

$$\|A(\mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta)\| \leq \delta C^2.$$

[証明]

$$\tilde{A}(\mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta) = \tilde{A}R_\lambda(\tilde{A})A^T(\mathbf{b} - \mathbf{b}^\delta)$$

であるから、仮定 (2.5) より

$$\begin{aligned}\|A(\mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta)\|^2 &= (\tilde{A}(\mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta), \mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta) \\ &= (\tilde{A}R_\lambda(\tilde{A})A^T(\mathbf{b} - \mathbf{b}^\delta), \mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta) \\ &= (\hat{A}R_\lambda(\hat{A})(\mathbf{b} - \mathbf{b}^\delta), A(\mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta)) \\ &\leq \delta C^2 \|A(\mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta)\|\end{aligned}$$

となる. ■

補題 2.2.2 [10]

$$\|\mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta\| \leq \delta C \sqrt{r(\lambda)}.$$

[証明]

$$\mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta = A^T R_\lambda(\hat{A})(\mathbf{b} - \mathbf{b}^\delta)$$

より, (2.6) と上の補題を用いると,

$$\begin{aligned}\|\mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta\|^2 &= (\mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta, A^T R_\lambda(\hat{A})(\mathbf{b} - \mathbf{b}^\delta)) \\ &= (A(\mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta), R_\lambda(\hat{A})(\mathbf{b} - \mathbf{b}^\delta)) \\ &\leq \delta^2 C^2 r(\lambda)\end{aligned}$$

が成り立つ. ■

我々は, 正則化パラメータ選択関数 $\lambda(\delta)$ は $\lambda(0) = 0$ の下で

$$\lambda : [0, \infty) \rightarrow [0, \infty)$$

を満たす単調な連続関数であると仮定する. 次の定理は $\mathbf{x}_{\lambda(\delta)}^\delta$ が \mathbf{x}_0 に強収束するための十分条件を示している.

定理 2.2.1 [10] $\mathbf{b} \in \mathcal{D}(A^\dagger)$, $\delta \rightarrow 0$ について $\lambda(\delta) \rightarrow 0$, $\delta^2 r(\lambda(\delta)) \rightarrow 0$ とする. このとき

$$\mathbf{x}_{\lambda(\delta)}^\delta \rightarrow \mathbf{x}_0, \quad (\delta \rightarrow 0)$$

である.

[証明] 補題 2.2.2 より,

$$\begin{aligned}\|\mathbf{x}_0 - \mathbf{x}_{\lambda(\delta)}^\delta\| &\leq \|\mathbf{x}_0 - \mathbf{x}_{\lambda(\delta)}\| + \|\mathbf{x}_{\lambda(\delta)} - \mathbf{x}_{\lambda(\delta)}^\delta\| \\ &\leq \|\mathbf{x}_0 - \mathbf{x}_{\lambda(\delta)}\| + \delta \sqrt{r(\lambda(\delta))} C\end{aligned}$$

である. 定理 2.1.2 と仮定 $\lim_{\delta \rightarrow 0} \delta \sqrt{r(\lambda(\delta))} = 0$ より, $\mathbf{b} \in \mathcal{D}(A^\dagger)$ ならば $\mathbf{x}_{\lambda(\delta)}^\delta \rightarrow \mathbf{x}_0$ が示された. ■

次に, H. W. Engl [7] によって導入された弱正則性について考察する. 近似解の族 $\{\mathbf{x}_{\lambda(\delta)}^\delta\}$ が弱正則であるとは, $\delta_n \rightarrow 0$ となる任意の列 $\{\delta_n\}$ について

$$\mathbf{x}_{\lambda(\delta_n)}^{\delta_n} \xrightarrow{w} \mathbf{x}_0$$

が成り立つことを意味する. 同様に, 強収束するときは強正則であるという.

上の定理で, $\delta r(\lambda(\delta)) \rightarrow 0$ が強正則であるための十分条件であることを示した. 次に我々は

$$\limsup_{\delta \rightarrow 0} \delta^2 r(\lambda(\delta)) < \infty \quad (2.7)$$

が弱正則の十分条件であることを示す.

定理 2.2.2 [10] $b \in \mathcal{D}(A^\dagger)$ で, (2.7) を満たしているとする. このとき, $\{x_{\lambda(\delta)}^\delta\}$ は弱正則である.

[証明] $\{\delta_n\}$ を $\delta_n \rightarrow 0$ を満たす列とする. 今, 仮定より $\lambda(\delta_n) \rightarrow 0$ であるから, 定理 2.1.2 より

$$x_{\lambda(\delta_n)} \rightarrow x_0, \quad (n \rightarrow \infty)$$

である. また,

$$x_0 - x_{\lambda(\delta_n)}^{\delta_n} = x_0 - x_{\lambda(\delta_n)} + x_{\lambda(\delta_n)} + x_{\lambda(\delta_n)}^{\delta_n}$$

と書き表されるから, 定理を証明するには

$$x_{\lambda(\delta_n)} + x_{\lambda(\delta_n)}^{\delta_n} \xrightarrow{w} 0, \quad (n \rightarrow \infty) \quad (2.8)$$

を示せば十分である.

補題 2.2.2 より

$$\|x_{\lambda(\delta_n)} + x_{\lambda(\delta_n)}^{\delta_n}\|^2 \leq C^2 \delta_n^2 r(\lambda(\delta_n))$$

である. そして, 仮定 (2.7) から, $\{x_{\lambda(\delta_n)} + x_{\lambda(\delta_n)}^{\delta_n}\}$ は有界列であることがわかる. パナツハ・ステインハウスの定理より, (2.8) を証明するには, l_2 の稠密な部分集合に属する任意の元 z について

$$(z, x_{\lambda(\delta_n)} + x_{\lambda(\delta_n)}^{\delta_n}) \rightarrow 0, \quad (n \rightarrow \infty) \quad (2.9)$$

を示せばよい. したがって, l_2 の稠密な部分空間

$$\mathcal{D}(A^\dagger) = \mathcal{R}(A^T) + \mathcal{R}(A^T)^\perp = \mathcal{R}(A^T) + \mathcal{N}(A)$$

について (2.9) を検証する.

$z \in \mathcal{N}(A)$ のときは

$$\begin{aligned} x_{\lambda(\delta_n)} + x_{\lambda(\delta_n)}^{\delta_n} &= R_{\lambda(\delta_n)}(\tilde{A})A^T(b - b^{\delta_n}) \\ &= A^T R_{\lambda(\delta_n)}(\hat{A})(b - b^{\delta_n}) \end{aligned}$$

より明らかである. $z \in \mathcal{R}(A^T)$ のとき, $z = A^T w$ とすると

$$\begin{aligned} (z, x_{\lambda(\delta_n)} + x_{\lambda(\delta_n)}^{\delta_n}) &= (w, \hat{A} R_{\lambda(\delta_n)}(\hat{A})(b - b^{\delta_n})) \\ &\leq \|w\| C^2 \delta_n \rightarrow 0, \quad (n \rightarrow \infty) \end{aligned}$$

となり, 任意の $z \in \mathcal{D}(A^\dagger)$ について (2.9) が示された. ■

Engl は同様に, 弱発散, 強発散の概念も導入した [7]. 次の二つの定理において, 我々は正則化パラメータ選択関数 $\lambda: [0, \infty) \rightarrow [0, \infty)$ は $\lambda(0) = 0$ を満たす狭義単調増加関数であると仮定する. 先ず, 条件

$$\lim_{\delta \rightarrow 0} \delta^2 \lambda(\delta) R_{\lambda(\delta)}(\lambda(\delta))^2 = \infty \quad (2.10)$$

が強発散の十分条件であることを示す.

定理 2.2.3 [7] A は無限階で, $b \in \mathcal{D}(A^\dagger)$, $\lambda(\delta)$ が上の条件 (2.10) を満たしていれば, $\{x_{\lambda(\delta_n)}^{\delta_n}\}$ が弱収束しない $\delta_n \rightarrow 0$ なる $\{\delta_n\}$, $\|b - b^{\delta_n}\| \leq \delta_n$ なるベクトル b^{δ_n} が存在する. このとき $\{x_{\lambda(\delta)}^{\delta}\}$ は強発散であるという.

[証明] $\{\sigma_k; u_k, v_k\}$ を A の特異値分解, $\mu_k = \sigma_k^2$ とする. 今, A は有限階ではないので, $\mu_k \rightarrow 0 (n \rightarrow \infty)$ である. 正則化パラメータ選択関数 λ についての条件より, $\lambda(\delta_n) = \mu_n, \delta_n \rightarrow 0$ なる列 $\{\delta_n\}$ が存在する. 今, $b^{\delta_n} = b + \delta_n v_n$ とすると, 定理 2.1.2 より,

$$x_{\lambda(\delta_n)} = R_{\lambda(\delta_n)}(\tilde{A})A^T b \rightarrow x_0, \quad (n \rightarrow \infty) \quad (2.11)$$

であるから

$$\begin{aligned} x_{\lambda(\delta_n)} - x_{\lambda(\delta_n)}^{\delta_n} &= R_{\lambda(\delta_n)}(\tilde{A})A^T (b - b^{\delta_n}) \\ &= \delta_n \sigma_n R_{\lambda(\delta_n)}(\mu_n) u_n \end{aligned}$$

となる. したがって, 仮定 (2.10) より

$$\begin{aligned} \|x_{\lambda(\delta_n)} - x_{\lambda(\delta_n)}^{\delta_n}\|^2 &= \delta_n^2 \mu_n R_{\lambda(\delta_n)}(\mu_n)^2 \\ &= \delta_n^2 \lambda(\delta_n) R_{\lambda(\delta_n)}(\lambda(\delta_n))^2 \rightarrow \infty \end{aligned}$$

となる. (2.11) より

$$\|x_{\lambda(\delta_n)}^{\delta_n}\| \rightarrow \infty, \quad (n \rightarrow \infty)$$

となり, ゆえに $\{x_{\lambda(\delta_n)}^{\delta_n}\}$ は弱収束しない. ■

次に, 条件 (2.10) を

$$\liminf_{\delta \rightarrow 0} \delta^2 \lambda(\delta) R_{\lambda(\delta)}(\lambda(\delta))^2 > 0 \quad (2.12)$$

と置き換える.

定理 2.2.4 [7] A は無限階, $b \in \mathcal{D}(A^\dagger)$ で, 仮定 (2.12) を満たしていれば, $\{x_{\lambda(\delta_n)}^{\delta_n}\}$ が x_0 に強収束しないような $\delta_n \rightarrow 0$ なる列 $\{\delta_n\}$ と $\|b - b^{\delta_n}\| \leq \delta_n$ なるベクトル b^{δ_n} が存在する. このとき $\{x_{\lambda(\delta)}^{\delta}\}$ は弱発散であるという.

[証明] 前定理の証明と同様の理論を展開することにより, 十分大きな n に対して,

$$\|x_{\lambda(\delta_n)} - x_{\lambda(\delta_n)}^{\delta_n}\|^2 = \delta^2 \lambda(\delta_n) R_{\lambda(\delta_n)}(\lambda(\delta_n))^2 \geq C$$

を満たす $C > 0$ が存在することが仮定 (2.12) より導かれる. しかし,

$$x_{\lambda(\delta_n)} \rightarrow x_0, \quad (n \rightarrow \infty)$$

であることから, $\{x_{\lambda(\delta_n)}^{\delta_n}\}$ が x_0 に強収束することは不可能である. ■

2.3 Tikhonov の正則化法

第1種フレドホルム積分方程式

$$Tf \stackrel{d}{=} \int_a^b K(s, t) f(t) dt = g(s)$$

を解く難しさは、一般逆作用素 T^{-1} の不安定性に起因する。1.1節で示したように、解 f の擾動は、積分作用素により遮られて、 g にはほとんど反映されない。

A. N. Tikhonov はそれらの振動を切り捨て、解を正則化する方法を提案した [30]。それは、汎関数

$$\int_a^b (Tf(s) - g(s))^2 ds + \lambda \int_a^b [p(s)f(s)^2 + q(s)f''(s)^2] ds \quad (2.13)$$

を最小とする関数 f を近似解とする方法である。ここで、 p と q は非負で $p+q$ が恒等的には 0 にならない関数で、 λ は正のパラメータである。

この手法は $p(s) \equiv 1, q(s) \equiv 0$ とおくことにより、汎関数

$$\hat{F}_\lambda(\mathbf{x}) \stackrel{d}{=} \|A\mathbf{x} - \mathbf{b}\|^2 + \lambda\|\mathbf{x}\|^2 \quad (2.14)$$

を最小化する問題と見ることができる。また、適当な関数空間を設定することにより、一般の形式 (2.13) も上式の形式に変換することが可能である。ここでは、(2.14) を最小化する問題を Tikhonov の正則化法と呼ぶ。

この 2 次汎関数 \hat{F}_λ が一意の最小点 \mathbf{x}_λ を持つことは明らかである。その最小点は、関数 $f_\lambda(t)$ を任意の $\mathbf{w} \in l_2$ に対して

$$f_\lambda(t) \stackrel{d}{=} \hat{F}_\lambda(\mathbf{x}_\lambda + t\mathbf{w})$$

と定義すると、 $f'_\lambda(0) = 0$ により特徴付けられる。そして、

$$f'_\lambda(0) = 2(A^T A\mathbf{x}_\lambda - A^T \mathbf{b} + \lambda\mathbf{x}_\lambda, \mathbf{w})$$

より、 \mathbf{x}_λ は方程式

$$A^T A\mathbf{x}_\lambda + \lambda\mathbf{x}_\lambda = A^T \mathbf{b} \quad (2.15)$$

の解として表される。

Tikhonov の正則化法は、2.1節の正則化関数 $R_\lambda(t)$ として

$$R_\lambda(t) = \frac{1}{t + \lambda}$$

を用いた正則化法である。また、条件 (2.3) を満たすように収束速度関数を、

$$\omega(\lambda, \nu) = \lambda^\nu, \quad 0 < \nu \leq 1$$

とすると、定理 2.1.4 の系として、次の系が得られる。

系 2.3.1 [24] もし、ある $\nu \in (0, 1]$ について、 $\mathbf{x}_0 \in \mathcal{R}(\hat{A}^\nu)$ ならば、

$$\|\mathbf{x}_0 - \mathbf{x}_\lambda\| = O(\lambda^\nu)$$

である。

同様に、定理 2.1.5 で $\nu = 1$ とおくことにより、次の系が得られる。

系 2.3.2 [24] もし、 $\mathbf{x}_0 \in \mathcal{R}(A^T)$ ならば、

$$\|\mathbf{x}_0 - \mathbf{x}_\lambda\| = O(\lambda^{\frac{1}{2}})$$

である。

系 2.3.1 より, $\nu = 1$ のとき, $\mathbf{x}_0 \in \mathcal{R}(\tilde{A})$ ならば, 収束速度 $O(\lambda)$ が得られることがわかる. 逆に, $O(\lambda)$ が最良の収束速度であることが次の二つの定理より得られる. その前に, 定理 1.4.11 の系として, 次の系を与える.

系 2.3.3 $\mathbf{b} \in \mathcal{N}(A^T)^\perp$ とする. このとき, $\nu \geq 0$ に対し, $\mathbf{b} \in \mathcal{R}(A\tilde{A}^\nu)$ と,

$$\sum_{i=1}^{\infty} \frac{|(\mathbf{b}, \mathbf{u}_i)|^2}{\sigma_i^{4\nu+2}} < \infty$$

は同値である.

[証明] 作用素 $A\tilde{A}^\nu$ の特異値分解が $\{\sigma_k^{2\nu+1}; \mathbf{u}_k, \mathbf{v}_k\}$ となることから導かれる. ■

定理 2.3.1 [10] もし, $\mathbf{b} \in \mathcal{D}(A^\dagger)$ で, $\|\mathbf{x}_0 - \mathbf{x}_\lambda\| = o(\lambda)$ ならば, $\mathbf{x}_0 = 0$ で, 任意の $\lambda > 0$ に対して, $\mathbf{x}_\lambda = 0$ である.

[証明] $\mathbf{x}_0 \neq 0$ とし, $\mathbf{e}_\lambda \stackrel{d}{=} \mathbf{x}_0 - \mathbf{x}_\lambda$ とおく.

$$\begin{aligned} (\tilde{A} + \lambda I)\mathbf{e}_\lambda &= A^T \mathbf{b} - (\tilde{A} + \lambda I)\mathbf{x}_0 \\ &= -\lambda \mathbf{x}_0 \end{aligned}$$

より

$$\lambda \|\mathbf{x}_0\| \leq (\|A\|^2 + \lambda) \|\mathbf{e}_\lambda\|$$

である. 今, 仮定より上式の右辺は $o(\lambda)$ であるから, $\mathbf{x}_0 = 0$ となる. また, P を l_2 から $\overline{\mathcal{R}(A)}$ への直交射影とすると,

$$0 = A\mathbf{x}_0 = P\mathbf{b}$$

となる. これより, 任意の λ について

$$\mathbf{x}_\lambda = (\tilde{A} + \lambda I)^{-1} A^T \mathbf{b} = (\tilde{A} + \lambda I)^{-1} A^T P\mathbf{b} = 0$$

となる. ■

定理 2.3.2 [10] A がコンパクト作用素, $\mathbf{b} \in \mathcal{D}(A^\dagger)$ で, $\|\mathbf{x}_0 - \mathbf{x}_\lambda\| = O(\lambda)$ ならば, $\mathbf{x}_0 \in \mathcal{R}(\tilde{A})$ である.

[証明] A の特異値分解を用いると, 定理 1.4.10 より \mathbf{x}_0 は

$$\mathbf{x}_0 = \sum_{i=1}^{\infty} \frac{(P\mathbf{b}, \mathbf{u}_i)}{\sigma_i} \mathbf{v}_i$$

と表せる. 同様に \mathbf{x}_λ も

$$\mathbf{x}_\lambda = (\tilde{A} + \lambda I)^{-1} A^T P\mathbf{b} = \sum_{i=1}^{\infty} \frac{\sigma_i}{\sigma_i^2 + \lambda} (P\mathbf{b}, \mathbf{u}_i) \mathbf{v}_i$$

となる. これより, 誤差の 2 乗は

$$\begin{aligned} \|\mathbf{x}_0 - \mathbf{x}_\lambda\|^2 &= \left\| \sum_{i=1}^{\infty} \left(\frac{1}{\sigma_i} - \frac{\sigma_i}{\sigma_i^2 + \lambda} \right) (P\mathbf{b}, \mathbf{u}_i) \mathbf{v}_i \right\|^2 \\ &= \lambda^2 \sum_{i=1}^{\infty} \left(\frac{1}{\sigma_i(\sigma_i^2 + \lambda)} \right)^2 |(P\mathbf{b}, \mathbf{u}_i)|^2 \end{aligned}$$

となる. 今, 仮定より $\|\mathbf{x}_0 - \mathbf{x}_\lambda\|^2 = O(\lambda^2)$ であるから

$$\sum_{i=1}^{\infty} \left(\frac{1}{\sigma_i(\sigma_i^2 + \lambda)} \right)^2 |(P\mathbf{b}, \mathbf{u}_i)|^2 < \infty, \quad (\lambda \rightarrow 0)$$

でなければならない. これは,

$$\sum_{i=1}^{\infty} \frac{|(P\mathbf{b}, \mathbf{u}_i)|^2}{\sigma_i^6} < \infty$$

を意味している. よって, 系 2.3.3 より, $P\mathbf{b} \in \mathcal{R}(AA^T A)$ となり, これより, $\mathbf{x}_0 \in \mathcal{R}(\hat{A})$ が示せた. ■

次に, データ \mathbf{b} にノイズが含まれている場合, つまり,

$$\mathbf{x}_\lambda^\delta = (\hat{A} + \lambda I)^{-1} A^T \mathbf{b}^\delta$$

のときを考察する. ここで, \mathbf{b}^δ は $\|\mathbf{b} - \mathbf{b}^\delta\| \leq \delta$ を満たしているとする. また, $R_\lambda(t) = (t + \lambda)^{-1}$ より, (2.5) の C は $C = 1$. (2.6) の $r(\lambda)$ は $r(\lambda) = 1/\lambda$ とおける. 補題 2.2.2 より,

$$\begin{aligned} \|\mathbf{x}_0 - \mathbf{x}_\lambda^\delta\| &\leq \|\mathbf{x}_0 - \mathbf{x}_\lambda\| + \|\mathbf{x}_\lambda - \mathbf{x}_\lambda^\delta\| \\ &\leq \|\mathbf{x}_0 - \mathbf{x}_\lambda\| + \frac{\delta}{\sqrt{\lambda}} \end{aligned}$$

となり, 系 2.3.1 と系 2.3.2 から次の二つの系が得られる.

系 2.3.4 [10] ある $\nu \in (0, 1]$ に対し, $\mathbf{x}_0 \in \mathcal{R}(\hat{A}^\nu)$ で, $\lambda = C'\delta^{\frac{2\nu}{2\nu+1}}$ ならば,

$$\|\mathbf{x}_0 - \mathbf{x}_\lambda^\delta\| = O(\delta^{\frac{2\nu}{2\nu+1}})$$

である.

系 2.3.5 [10] $\mathbf{x}_0 \in \mathcal{R}(A^T)$ で, ある $C' > 0$ に対し, $\lambda = C'\delta$ ならば,

$$\|\mathbf{x}_0 - \mathbf{x}_\lambda^\delta\| = O(\sqrt{\delta})$$

である.

これらの系は, $\mathbf{x}_0 \in \mathcal{R}(\hat{A})$ で, $\lambda = C'\delta^{\frac{2}{3}}$ のとき, $O(\delta^{\frac{2}{3}})$ が得られることを示している. ノイズが含まれていない場合と同様に, この収束がもっとも速いことが次の二つの定理 2.3.3 および定理 2.3.4 より得られる.

補題 2.3.1 [10] $\mathbf{b} \in \mathcal{D}(A^\dagger)$ と仮定する. このとき, $\mathbf{x}_0 \neq 0$ ならば

$$\lambda(\delta) = O(\|\mathbf{x}_0 - \mathbf{x}_{\lambda(\delta)}^\delta\|) + O(\delta), \quad (\delta \rightarrow 0)$$

である.

[証明]

$$(\hat{A} + \lambda(\delta)I)(\mathbf{x}_0 - \mathbf{x}_{\lambda(\delta)}^\delta) = \lambda(\delta)\mathbf{x}_0 + A^T(\mathbf{b} - \mathbf{b}^\delta)$$

より,

$$\lambda(\delta)\|\mathbf{x}_0\| \leq \delta\|A^T\| + O(\|\mathbf{x}_0 - \mathbf{x}_{\lambda(\delta)}^\delta\|)$$

となる. ■

定理 2.3.3 [10] A は有限階ではない作用素で, b^δ とは無関係に $\|x_o - x_\lambda^\delta\| = o(\delta^{2/3})$ とすると, $x_o = 0$ である.

[証明] 今, A は無限階だから, $\sigma_n \rightarrow 0$ ($n \rightarrow \infty$) である. $\delta_n = \sigma_n^3$, $b^{\delta_n} = b + \delta_n u_n$ とすると,

$$\begin{aligned} x_{\lambda(\delta_n)}^{\delta_n} - x_o &= x_{\lambda(\delta_n)} - x_o + x_{\lambda(\delta_n)}^{\delta_n} - x_{\lambda(\delta_n)} \\ &= x_{\lambda(\delta_n)} - x_o + \delta_n (\tilde{A} + \lambda(\delta_n)I)^{-1} A^T u_n \end{aligned}$$

となり, これより

$$\begin{aligned} \|x_{\lambda(\delta_n)}^{\delta_n} - x_o\|^2 &= \|x_{\lambda(\delta_n)} - x_o\|^2 + \frac{2\delta_n \sigma_n}{\sigma_n^2 + \lambda(\delta_n)} (x_{\lambda(\delta_n)} - x_o, v_n) + \left(\frac{\delta_n \sigma_n}{\sigma_n^2 + \lambda(\delta_n)} \right)^2 \\ &\geq \frac{2\delta_n^{4/3}}{\delta_n^{2/3} + \lambda(\delta_n)} (x_{\lambda(\delta_n)} - x_o, v_n) + \frac{\delta_n^{8/3}}{(\delta_n^{2/3} + \lambda(\delta_n))^2} \end{aligned}$$

となる. 上式の両辺を $\delta_n^{4/3}$ で割ると,

$$\frac{\|x_{\lambda(\delta_n)}^{\delta_n} - x_o\|^2}{\delta_n^{4/3}} \geq \frac{2}{\delta_n^{2/3} + \lambda(\delta_n)} (x_{\lambda(\delta_n)} - x_o, v_n) + \frac{\delta_n^{4/3}}{(\delta_n^{2/3} + \lambda(\delta_n))^2}$$

が得られ, $x_o \neq 0$ とすると, 上の補題と定理の仮定より, $\lambda(\delta_n) \delta_n^{-2/3} \rightarrow 0$ ($\delta_n \rightarrow 0$) となり,

$$0 \geq 2 \limsup_{\delta_n \rightarrow 0} \frac{(x_{\lambda(\delta_n)} - x_o, v_n)}{\delta_n^{2/3} + \lambda(\delta_n)} + 1$$

とならなければならない. しかし, 今, 仮定より $\|b - b^{\delta_n}\| \leq \delta_n$ を満たすどんな b^{δ_n} についても $\|x_{\lambda(\delta_n)}^{\delta_n} - x_o\| = o(\delta^{2/3})$ であるから, $b = b^{\delta_n}$ とおくと, $\|x_{\lambda(\delta_n)} - x_o\| = o(\delta^{2/3})$ が得られ, 上式は $0 \geq 1$ となる. これは矛盾である. ■

定理 2.3.4 [10] ある $C' > 0$ に対し, $\lambda = C' \delta^{2/3}$, $\|b - b^\delta\| \leq \delta$ を満たすどんな b^δ に対しても $\|x_o - x_\lambda^\delta\| = O(\delta^{2/3})$ とすると, $x_o \in \mathcal{R}(\tilde{A})$ である.

[証明] $\|b\| = 1$, $b^\delta = (1 + \delta)b$ とする.

$$x_o - x_{\lambda(\delta)}^\delta = \sum_{i=1}^{\infty} \left\{ \frac{(b, u_i)}{\sigma_i} - \frac{\sigma_i}{\sigma_i^2 + \lambda(\delta)} (b^\delta, u_i) \right\} v_i$$

より,

$$\|x_o - x_{\lambda(\delta)}^\delta\|^2 = \sum_{i=1}^{\infty} \left\{ \frac{\lambda(\delta) - \delta \sigma_i^2}{\sigma_i(\sigma_i^2 + \lambda(\delta))} \right\}^2 |(b, u_i)|^2$$

となる. したがって, ある $M > 0$ が存在して, 任意の N について

$$\begin{aligned} M &\geq \frac{\|x_o - x_{\lambda(\delta)}^\delta\|^2}{\delta^{4/3}} = \sum_{i=1}^{\infty} \left\{ \frac{C' - \delta^{1/3} \sigma_i^2}{\sigma_i(\sigma_i^2 + \lambda(\delta))} \right\}^2 |(b, u_i)|^2 \\ &\geq \sum_{i=1}^N \left\{ \frac{C' - \delta^{1/3} \sigma_i^2}{\sigma_i(\sigma_i^2 + \lambda(\delta))} \right\}^2 |(b, u_i)|^2 \end{aligned}$$

である. ここで $\delta \rightarrow 0$ とすることにより

$$\sum_{i=1}^N C'^2 \frac{|(b, u_i)|^2}{\sigma_i^6} \leq M$$

そして, 系 2.3.3 より $b \in \mathcal{R}(A\tilde{A})$, つまり, $x_o \in \mathcal{R}(\tilde{A})$ となる. ■

第 3 章

正則化パラメータの選択

Tikhonov に代表される正則化法を用いるときの本質的な問題は正則化パラメータの選択にある。正則化パラメータ λ_0 が最適であるとは、

$$\|x_0 - x_{\lambda_0}^{\delta}\| = \min_{0 \leq \lambda < \infty} \|x_0 - x_{\lambda}^{\delta}\|$$

を満たすことをいう。

代表的な正則化パラメータ選択法として V. A. Morozov により提案された相変原理と, G. Wahba により導入された一般化交差検証法 (GCV 法) がある。この二つの手法は、期せずして東西両陣営の数学者により提唱された。

Morozov はノイズの上界が事前に判れば、誤差の上界となるように定義された関数を最小化することが可能であることを示した。これより、もし統計的検定の手法を用いることによりノイズレベルを推定できれば、正則化パラメータを決定することができると考えられる。我々は、最も有効であるデータ平滑化法の一つである AIC 最小化推定法を用いて上のアイデアの検証を行なった。

また、最近 P. C. Hansen により L-カーブによる最適正則化法が発表された。この発想自体は新しいものではなく、以前より用いられていたものであるが、悪条件線形方程式に適用され、命名したのは Hansen が最初である。この曲線はその名のごとく L 字型となる。これは悪条件線形方程式の特性である。この特性を利用するのがこの手法の特徴である。

しかし、Hansen が主張する方法ではパラメータの選択は困難である。そこで、我々は L-カーブの曲率が最大となるパラメータを選択する方法を提案する。行列の特異値分解を用いると L-カーブの曲率は正則化パラメータの陽な関数として表されるため、パラメータの決定は容易となる。

相変原理と GCV 法の概要はそれぞれ 3.1 節と 3.5 節で紹介する。統計的検定による最適正則化法の有効性は 3.3 節で検証する。3.4 節では L-カーブによる正則化法を述べる。我々の主張である L-カーブの曲率による正則化法は 3.4.2 節で、数値実験も含めて解説する。

3.1 Morozov の相変原理

正則化パラメータの選択法として、V. A. Morozov により提案された相変原理 (Discrepancy Principle) が有名である [24]。この原理は、計算結果の質と、入力データの質を一致させることを基礎としている。

今、データ b の近似として b^{δ} が与えられ、これより

$$Ax = b, \quad b \in \mathcal{R}(A)$$

を解くことを考える. ここで, A はコンパクト作用素, 近似データ b^δ は,

$$\|b - b^\delta\| \leq \delta < \|b^\delta\| \quad (3.1)$$

を満たしているとする. ここで b についての条件は $b \in \mathcal{D}(A^\dagger)$ ではなく, $b \in \mathcal{R}(A)$ であることに注意されたい. なお, 条件 (3.1) の $\delta < \|b^\delta\|$ は深刻な制約ではない. これは, ノイズが入力データよりも小さいことを意味しているにすぎない. もし, そうでなければ, どんな解析法を用いても, 解を近似することは不可能である.

Morozov の相変原理は, 仮定 (3.1) の下で, 判定式

$$\|Ax_\lambda^\delta - b^\delta\| = \delta \quad (3.2)$$

を満たす $\lambda = \lambda(\delta)$ を正則化パラメータとして用いる方法である. これは, 次のように解釈することも可能である.

データに関する情報 (3.1) から, 方程式 $Ax = b$ の最小 2 乗最小ノルム解を求めるとき, 集合

$$\mathcal{D}_\delta \stackrel{\text{d}}{=} \{x : \|Ax - b^\delta\| \leq \delta\}$$

の中のノルム最小のベクトルを近似解として採用することは, 合理的な方法である. ここで, 集合 \mathcal{D}_δ は閉凸集合であるから, ノルム最小なベクトル x_δ は一意に存在する. 次に, この x_δ が (3.2) を満たしていることを示す. 今, $\|Ax_\delta - b^\delta\| < \delta$ とすると, $0 < t < 1$ で 1 に十分近い t をとれば, $\|Atx_\delta - b^\delta\| < \delta$ となる. しかし, このとき $\|tx_\delta\| < \|x_\delta\|$ であるから, x_δ がノルム最小なベクトルであるという仮定に反する. これより, x_δ は (3.2) を満たす.

次の定理は, 判定式 (3.2) を満たすパラメータは唯一つであることを示している.

定理 3.1.1 [10] b と b^δ は (3.1) を満たしているとする. このとき, λ の関数 $\|Ax_\lambda^\delta - b^\delta\|$ は連続で狭義単調増加関数であり, その値域に δ を含む.

[証明] x_λ^δ は A の特異値分解を用いると,

$$x_\lambda^\delta = (\hat{A} + \lambda I)^{-1} A^T b = \sum_{i=1}^{\infty} \frac{\sigma_i}{\sigma_i^2 + \lambda} (b^\delta, u_i) v_i$$

と表せることから, 残差は

$$\begin{aligned} Ax_\lambda^\delta - b^\delta &= A(\hat{A} + \lambda I)^{-1} b^\delta - b^\delta \\ &= \sum_{i=1}^{\infty} \frac{-\lambda}{\sigma_i^2 + \lambda} (b^\delta, u_i) u_i - Qb^\delta \end{aligned}$$

となる. ここで, Q は l_2 から $\mathcal{N}(A^T) = \mathcal{R}(A)^\perp$ への直交射影である. これより, 残差ノルムの 2 乗は

$$\|Ax_\lambda^\delta - b^\delta\|^2 = \sum_{i=1}^{\infty} \left(\frac{\lambda}{\sigma_i^2 + \lambda} \right)^2 |(b^\delta, u_i)|^2 + \|Qb^\delta\|^2$$

と表せる. 上式より, λ の関数 $\|Ax_\lambda^\delta - b^\delta\|$ は明らかに狭義単調増加連続関数である. また, $b \in \mathcal{R}(A)$ より,

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \|Ax_\lambda^\delta - b^\delta\|^2 &= \|(I - Q)b^\delta\|^2 + \|Qb^\delta\|^2 = \|b^\delta\|^2 > \delta^2 \\ \lim_{\lambda \rightarrow 0} \|Ax_\lambda^\delta - b^\delta\|^2 &= \|Qb^\delta\|^2 = \|Q(b - b^\delta)\|^2 \leq \|b - b^\delta\|^2 \leq \delta^2 \end{aligned}$$

となり、その値域に δ を含むことがわかる。■

次に、相変原理(3.2)により選択したパラメータの意義を考察する。2乗誤差 $\|\mathbf{x}_o - \mathbf{x}_\lambda^\delta\|^2$ ができるだけ小さい近似解 $\mathbf{x}_\lambda^\delta$ を求めるのが我々の目的である。

$$\begin{aligned}\|\mathbf{x}_o - \mathbf{x}_\lambda^\delta\|^2 &= \|\mathbf{x}_\lambda^\delta\|^2 - \frac{2}{\lambda} \left(A^T (A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta), \mathbf{x}_o \right) + \|\mathbf{x}_o\|^2 \\ &= \|\mathbf{x}_\lambda^\delta\|^2 - \frac{2}{\lambda} (A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta, \mathbf{b}) + \|\mathbf{x}_o\|^2 \\ &= \|\mathbf{x}_\lambda^\delta\|^2 + \frac{2}{\lambda} (A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta, \mathbf{b}^\delta) + \frac{2}{\lambda} (A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta, \mathbf{b}^\delta - \mathbf{b}) + \|\mathbf{x}_o\|^2 \\ &\leq \|\mathbf{x}_\lambda^\delta\|^2 - \frac{2}{\lambda} (A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta, \mathbf{b}^\delta) + \frac{2\delta}{\lambda} \|A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta\| + \|\mathbf{x}_o\|^2\end{aligned}$$

より、

$$\hat{E}(\lambda; \mathbf{b}^\delta) \triangleq \|\mathbf{x}_\lambda^\delta\|^2 - \frac{2}{\lambda} (A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta, \mathbf{b}^\delta) + \frac{2\delta}{\lambda} \|A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta\| + \|\mathbf{x}_o\|^2$$

と定義すると、関数 $\hat{E}(\lambda; \mathbf{b}^\delta)$ は2乗誤差の上界の一つとなる。ここで、この $\hat{E}(\lambda; \mathbf{b}^\delta)$ は常に2乗誤差よりも大というわけではない。もし、 \mathbf{b} を AA^T のノルム1の固有ベクトル、対応する固有値を $\sigma > 0$ 、 $\mathbf{b}^\delta = (1 + \delta)\mathbf{b}$ とすると、

$$\|\mathbf{x}_o - \mathbf{x}_\lambda^\delta\|^2 = \hat{E}(\lambda; \mathbf{b}^\delta)$$

となり、等号が成立する例が存在する。

相変原理は、 $\hat{E}(\lambda; \mathbf{b}^\delta)$ を最小とするパラメータを選択していることが次の定理よりわかる。

定理 3.1.2 [10] \mathbf{b} と \mathbf{b}^δ は(3.1)を満たしているとする。このとき、 $\hat{E}(\lambda; \mathbf{b}^\delta)$ が最小となるのは、(3.2)のときで、また、そのときに限る。

[証明] 先ず、任意の $\lambda > 0$ について $\|A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta\| > 0$ を示す。今、ある $\lambda > 0$ に対し、 $A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta = 0$ とする。

$$\begin{aligned}A^T(A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta) &= \hat{A}(\hat{A} + \lambda I)^{-1}A^T\mathbf{b}^\delta - A^T\mathbf{b}^\delta \\ &= (\hat{A}(\hat{A} + \lambda I)^{-1} - I)A^T\mathbf{b}^\delta \\ &= ((\hat{A} + \lambda I)(\hat{A} + \lambda I)^{-1} - \lambda(\hat{A} + \lambda I)^{-1} - I)A^T\mathbf{b}^\delta \\ &= -\lambda(\hat{A} + \lambda I)^{-1}A^T\mathbf{b}^\delta \\ &= -\lambda\mathbf{x}_\lambda^\delta\end{aligned}\tag{3.3}$$

より、このとき $\mathbf{x}_\lambda^\delta = 0$ となる。これは $\mathbf{b}^\delta \in \mathcal{N}(A^T) = \mathcal{R}(A)^\perp$ を意味する。仮定(3.1)と $\mathbf{b} \in \mathcal{R}(A)$ より

$$\delta^2 \geq \|\mathbf{b} - \mathbf{b}^\delta\|^2 = \|\mathbf{b}\|^2 + \|\mathbf{b}^\delta\|^2 > \|\mathbf{b}\|^2 + \delta^2$$

となり矛盾である。

次に、 $\hat{E}(\lambda; \mathbf{b}^\delta)$ を λ で微分すると

$$\frac{d}{d\lambda} \hat{E}(\lambda; \mathbf{b}^\delta) = 2 \left(1 - \frac{\delta}{\|A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta\|} \right) \frac{\|(\hat{A} + \lambda I)^{-1/2}(A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta)\|^2}{\lambda}$$

となる。ここで

$$\frac{\|(\hat{A} + \lambda I)^{-1/2}(A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta)\|^2}{\lambda}$$

は上と同様の理由により常に正である。そして、定理 3.1.1 より λ の関数 $\|A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta\|$ は狭義単調増加関数であることから

$$\begin{aligned} \|A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta\| < \delta & \quad \text{のとき} & \quad \frac{d}{d\lambda} \tilde{E}(\lambda; \mathbf{b}^\delta) < 0 \\ \|A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta\| > \delta & \quad \text{のとき} & \quad \frac{d}{d\lambda} \tilde{E}(\lambda; \mathbf{b}^\delta) > 0 \end{aligned}$$

となり、定理が示せた。■

相変原理の正則性は、次の定理により示される。

定理 3.1.3 [10] もし、 \mathbf{b} と \mathbf{b}^δ が (3.1) を、 $\lambda = \lambda(\delta)$ が (3.2) を満たしているならば、

$$\mathbf{x}_{\lambda(\delta)}^\delta \rightarrow \mathbf{x}_o, \quad (\delta \rightarrow 0)$$

である

[証明] $\mathbf{x}_{\lambda(\delta)}^\delta$ は汎関数

$$F_{\lambda(\delta)}(\mathbf{x}; \mathbf{b}^\delta) \triangleq \|A\mathbf{x} - \mathbf{b}^\delta\|^2 + \lambda(\delta)\|\mathbf{x}\|^2$$

の最小点であることから、

$$\begin{aligned} \|A\mathbf{x}_{\lambda(\delta)}^\delta - \mathbf{b}^\delta\|^2 + \lambda(\delta)\|\mathbf{x}_{\lambda(\delta)}^\delta\|^2 &= F_{\lambda(\delta)}(\mathbf{x}_{\lambda(\delta)}^\delta; \mathbf{b}^\delta) \\ &\leq F_{\lambda(\delta)}(\mathbf{x}_o; \mathbf{b}^\delta) \\ &= \|\mathbf{b} - \mathbf{b}^\delta\|^2 + \lambda(\delta)\|\mathbf{x}_o\|^2 \\ &\leq \delta^2 + \lambda(\delta)\|\mathbf{x}_o\|^2 \end{aligned}$$

が得られる。これより、 $\|A\mathbf{x}_{\lambda(\delta)}^\delta - \mathbf{b}^\delta\| = \delta$ であるから、すべての $\delta > 0$ について、

$$\|\mathbf{x}_{\lambda(\delta)}^\delta\| \leq \|\mathbf{x}_o\| \tag{3.4}$$

がいえ。したがって、ヒルベルト空間内の任意の有界列は、弱収束する部分列を含むから、ある部分列 $\{\delta_n\}$ があって、ある $\mathbf{y} \in l_2$ について

$$\mathbf{x}_{\lambda(\delta_n)}^{\delta_n} \rightharpoonup \mathbf{y}$$

である。今、 A はコンパクト作用素であるから、

$$A\mathbf{x}_{\lambda(\delta_n)}^{\delta_n} \rightarrow A\mathbf{y}$$

となる。そして、(3.2) より

$$\|A\mathbf{x}_{\lambda(\delta_n)}^{\delta_n} - \mathbf{b}^{\delta_n}\| = \delta_n \rightarrow 0$$

(3.1) より $\mathbf{b}^{\delta_n} \rightarrow \mathbf{b}$ であるから、 $A\mathbf{y} = \mathbf{b}$ となる。また、 $\{\mathbf{x}_{\lambda(\delta_n)}^{\delta_n}\} \subset \mathcal{R}(A^T)$ より、 $\mathbf{y} \in \overline{\mathcal{R}(A^T)} = \mathcal{N}(A)^\perp$ であるから、結局、 $\mathbf{y} = \mathbf{x}_o$ がいえ。

次に、強収束することを示す。バナッハ・スタインハウスの定理を用いると

$$\|\mathbf{x}_o\| \leq \liminf_{n \rightarrow \infty} \|\mathbf{x}_{\lambda(\delta_n)}^{\delta_n}\|$$

が得られる。また、(3.4) より、

$$\limsup_{n \rightarrow \infty} \|\mathbf{x}_{\lambda(\delta_n)}^{\delta_n}\| \leq \|\mathbf{x}_o\|$$

であるから,

$$\|\mathbf{x}_o\| \leq \liminf_{n \rightarrow \infty} \|\mathbf{x}_{\lambda(\delta_n)}^{\delta_n}\| \leq \limsup_{n \rightarrow \infty} \|\mathbf{x}_{\lambda(\delta_n)}^{\delta_n}\| \leq \|\mathbf{x}_o\|$$

となり,

$$\|\mathbf{x}_{\lambda(\delta_n)}^{\delta_n}\| \rightarrow \|\mathbf{x}_o\|$$

である. ヒルベルト空間において, $\mathbf{x}_{\lambda(\delta_n)}^{\delta_n} \xrightarrow{w} \mathbf{x}_o$ かつ $\|\mathbf{x}_{\lambda(\delta_n)}^{\delta_n}\| \rightarrow \|\mathbf{x}_o\|$ ならば, $\mathbf{x}_{\lambda(\delta_n)}^{\delta_n} \rightarrow \mathbf{x}_o$ であるから [22], $\mathbf{x}_{\lambda(\delta_n)}^{\delta_n} \rightarrow \mathbf{x}_o$ が示された.

列 $\{\mathbf{x}_{\lambda(\delta)}^{\delta}\}$ が \mathbf{x}_o に収束しないならば, 任意の $\varepsilon > 0$ について \mathbf{x}_o の ε -近傍に入らない無限部分列をとることができる. しかし, その部分列に上の論法を適用すると矛盾となる. ゆえに, $\{\mathbf{x}_{\lambda(\delta)}^{\delta}\}$ は \mathbf{x}_o に収束する. ■

V. A. Vinokurov は相変原理を用いるときの, 正則化パラメータの上界を事前に与える定理を示している.

定理 3.1.4 [33] パラメータ $\lambda = \lambda(\delta)$ が (3.2) を満たしているならば,

$$\lambda(\delta) \leq \frac{\delta \|A\|^2}{\|\mathbf{b}^{\delta}\| - \delta}$$

である.

[証明] (3.2) と (3.3) より,

$$\begin{aligned} \|\mathbf{b}^{\delta}\| - \delta &= \|\mathbf{b}^{\delta}\| - \|A\mathbf{x}_{\lambda(\delta)}^{\delta} - \mathbf{b}^{\delta}\| \\ &\leq \|A\mathbf{x}_{\lambda(\delta)}^{\delta}\| \\ &\leq \frac{\|A\|^2 \|A\mathbf{x}_{\lambda(\delta)}^{\delta} - \mathbf{b}^{\delta}\|}{\lambda(\delta)} \\ &= \frac{\|A\|^2 \delta}{\lambda(\delta)} \end{aligned}$$

となる. ■

次に相変原理の収束速度を与える定理を紹介する. 今, 方程式 $A\mathbf{x} = \mathbf{b}$ の最小2乗最小ノルム解 \mathbf{x}_o は $\mathcal{R}(A^T)$ に属しているものとする.

定理 3.1.5 [10] もし, $\mathbf{x}_o \in \mathcal{R}(A^T)$ ならば, $\|\mathbf{x}_o - \mathbf{x}_{\lambda(\delta)}^{\delta}\| = O(\sqrt{\delta})$ である.

[証明] ある $\mathbf{w} \in l_2$ について, $\mathbf{x}_o = A^T \mathbf{w}$ とすると, (3.1), (3.2), (3.4) より

$$\begin{aligned} \|\mathbf{x}_{\lambda(\delta)}^{\delta} - \mathbf{x}_o\|^2 &= \|\mathbf{x}_{\lambda(\delta)}^{\delta}\|^2 - 2(\mathbf{x}_{\lambda(\delta)}^{\delta}, \mathbf{x}_o) + \|\mathbf{x}_o\|^2 \\ &\leq 2\{\|\mathbf{x}_o\|^2 - (\mathbf{x}_{\lambda(\delta)}^{\delta}, \mathbf{x}_o)\} \\ &= 2(\mathbf{x}_o - \mathbf{x}_{\lambda(\delta)}^{\delta}, A^T \mathbf{w}) \\ &= 2(A\mathbf{x}_o - A\mathbf{x}_{\lambda(\delta)}^{\delta}, \mathbf{w}) \\ &= 2(\mathbf{b} - \mathbf{b}^{\delta}, \mathbf{w}) + 2(\mathbf{b}^{\delta} - A\mathbf{x}_{\lambda(\delta)}^{\delta}, \mathbf{w}) \\ &\leq 4\delta \|\mathbf{w}\| \end{aligned}$$

となる. ■

作用素 A が有限階であれば, 上の収束速度は $O(\delta)$ に改善できる. これは以下により示すことができる. 作用素 A が有限階であるとき, 定理 1.3.4, 定理 1.3.8 より, A^\dagger は有界作用素となる. また, $\mathbf{x}_o, \mathbf{x}_{\lambda(\delta)}^{\delta} \in N(A)^\perp$

より,

$$\begin{aligned} \mathbf{x}_{\lambda(\delta)}^\delta - \mathbf{x}_0 &= A^\dagger A \mathbf{x}_{\lambda(\delta)}^\delta - A^\dagger A \mathbf{x}_0 \\ &= A^\dagger (A \mathbf{x}_{\lambda(\delta)}^\delta - \mathbf{b}) \end{aligned}$$

となる. これより,

$$\begin{aligned} \|\mathbf{x}_{\lambda(\delta)}^\delta - \mathbf{x}_0\| &\leq \|A^\dagger\| \{\|A \mathbf{x}_{\lambda(\delta)}^\delta - \mathbf{b}^\delta\| + \|\mathbf{b}^\delta - \mathbf{b}\|\} \\ &\leq 2\|A^\dagger\|\delta \end{aligned}$$

が得られる.

次の定理は, A が有限階でない場合は $O(\sqrt{\delta})$ が最速の収束速度であることを示している.

定理 3.1.6 [10] 条件 (3.1) を満たす任意の $\mathbf{b}, \mathbf{b}^\delta$ に対し, $\|\mathbf{x}_{\lambda(\delta)}^\delta - \mathbf{x}_0\| = o(\sqrt{\delta})$ ならば, A は有限階の作用素である.

[証明] A を無限階の作用素, $\mathbf{b} = \mathbf{u}_1$, $\mathbf{b}^{\delta_n} = \mathbf{u}_1 + \delta_n \mathbf{u}_n$ と仮定する. このとき, 明らかに $\mathbf{x}_0 = \mathbf{v}_1/\sigma_1$ で, $\mathbf{b}, \mathbf{b}^{\delta_n}$ は

$$\|\mathbf{b} - \mathbf{b}^{\delta_n}\| = \delta_n < \|\mathbf{b}^{\delta_n}\|$$

を満たす. また,

$$\begin{aligned} \mathbf{x}_{\lambda(\delta_n)}^{\delta_n} &= (\tilde{A} + \lambda(\delta_n)I)^{-1} A^T \mathbf{u}_1 + \frac{\mathbf{v}_1}{\sigma_1} + \delta_n (\tilde{A} + \lambda(\delta_n)I)^{-1} A^T \mathbf{u}_n \\ &= \frac{-\lambda(\delta_n)}{\sigma_1(\sigma_1^2 + \lambda(\delta_n))} \mathbf{v}_1 + \frac{\delta_n \sigma_n}{\sigma_n^2 + \lambda(\delta_n)} \mathbf{v}_n \end{aligned}$$

となるから, $\delta_n = \sigma_n^2$ とおくと,

$$\|\mathbf{x}_{\lambda(\delta_n)}^{\delta_n} - \mathbf{x}_0\|^2 \geq \left(\frac{\delta_n \sigma_n}{\sigma_n^2 + \lambda(\delta_n)} \right)^2 = \left(\frac{\sqrt{\delta_n}}{1 + \lambda(\delta_n)/\delta_n} \right)^2$$

となる. 今, 仮定より,

$$\frac{\sqrt{\delta_n}}{1 + \lambda(\delta_n)/\delta_n} = o(\sqrt{\delta_n})$$

であるから, これは $\lambda(\delta_n)/\delta_n \rightarrow \infty$ を意味している. しかし, 定理 3.1.4 より

$$\frac{\lambda(\delta_n)}{\delta_n} \leq \frac{\|A\|^2}{\sqrt{1 + \delta_n^2} - \delta_n}$$

となり, 上式の右辺は $\|A\|^2$ に収束する. これより矛盾となる. ■

3.2 正則化法の離散化

この節では積分方程式の有限次元方程式による近似と正則化法について考察する.

第1種フレドホルム積分方程式

$$\int_{I_1} K(s, t) f(t) dt = g(s), \quad s \in L_s$$

に適切な基底関数 $\{\varphi_i\}$, $\{\psi_i\}$ を導入することにより, l_2 上の線形方程式

$$Ax = b \quad (3.5)$$

が得られたとする. 1.3節で定めたように, A の ij 成分と b の第 i 成分はそれぞれ

$$a_{ij} = \int_{I_s} \left(\psi_i(s) \int_{I_t} K(s,t) \varphi_j(t) dt \right) ds \quad (3.6)$$

$$b_i = \int_{I_s} \psi_i(s) g(s) ds \quad (3.7)$$

である. 先ず, 方程式 (3.5) を有限次元の方程式で近似することを考える.

今, 作用素 $A^{[N]}$, $A^{<N>}$ をそれぞれ

$$A^{<N>} = (a_{ij}^{<N>}), \quad a_{ij}^{<N>} = \begin{cases} a_{ij}, & 1 \leq i, j \leq N, \\ 0, & \text{それ以外} \end{cases}$$

$$A^{[N]} = A - A^{<N>}$$

と定義する. 同様に, ベクトル $b^{[N]}$, $b^{<N>}$ をそれぞれ

$$b^{<N>} = (b_1^{<N>}, b_2^{<N>}, \dots)^T \in l_2, \quad b_i^{<N>} = \begin{cases} b_i, & 1 \leq i \leq N, \\ 0, & N < i, \end{cases}$$

$$b^{[N]} = b - b^{<N>}$$

とする. また, $A^{(N)} = (a_{ij}^{(N)})$, $b^{(N)} = (b_1^{(N)}, b_2^{(N)}, \dots)^T \in l_2$ を

$$a_{ij}^{(N)} = \begin{cases} (3.6) \text{ の右辺を数値積分により近似したもの,} & 1 \leq i, j \leq N, \\ 0, & \text{それ以外,} \end{cases}$$

$$b_i^{(N)} = \begin{cases} (3.7) \text{ の右辺を数値積分により近似したもの,} & 1 \leq i \leq N, \\ 0, & N < i \end{cases}$$

とおき, 丸め誤差 ε について

$$\|A^{(N)} - A^{<N>}\| = O(\varepsilon)$$

$$\|b^{(N)} - b^{<N>}\| = O(\varepsilon)$$

であると仮定する. これは数値積分による誤差が丸め誤差程度であることを意味する. さらに, 十分大きな N をとって

$$\|A^{[N]}\| = O(\varepsilon), \quad \|b^{[N]}\| = O(\varepsilon) \quad (3.8)$$

であると仮定すると, 作用素 A , ベクトル b はそれぞれ $A^{(N)}$, $b^{(N)}$ により

$$\|A^{(N)} - A\| = \|A^{(N)} - (A^{<N>} + A^{[N]})\| \leq \|A^{(N)} - A^{<N>}\| + \|A^{[N]}\| = O(\varepsilon)$$

$$\|b^{(N)} - b\| = \|b^{(N)} - (b^{<N>} + b^{[N]})\| \leq \|b^{(N)} - b^{<N>}\| + \|b^{[N]}\| = O(\varepsilon)$$

と近似できることがわかる. このとき, 重要なことは, 基底関数と数値積分則の選択である. 基底関数は (3.8) において, N を小さくできることが望ましい. これは, 数値積分にも影響を及ぼす. そして, 数値積分則には, 精度が高く, 計算効率に良い手法が適していることはいうまでもない.

次に, 方程式

$$A^{(N)} x^{(N)} = b^{(N)} \quad (3.9)$$

に Tikhonov の正則化法を適用したときの解について考察する. 上の方程式に Tikhonov の正則化法を適用すると

$$(\tilde{A}^{(N)} + \lambda I) \mathbf{x}_\lambda^{(N)} \tilde{A}^{(N)T} \mathbf{b}^{(N)}, \quad \tilde{A}^{(N)} \stackrel{d}{=} A^{(N)T} A^{(N)} \quad (3.10)$$

となる. この方程式は本質的に N 元連立 1 次方程式である. ここで, $(\tilde{A} + \lambda I)$ の条件数を κ で表すことにすると, $\mathbf{x}_\lambda^{(N)}$ と \mathbf{x}_λ の関係

$$\frac{\|\mathbf{x}_\lambda^{(N)} - \mathbf{x}_\lambda\|}{\|\mathbf{x}_\lambda\|} \leq \frac{\kappa}{1 - \kappa \|\Delta \tilde{A}\| / \|\tilde{A} + \lambda I\|} \left(\frac{\|\Delta \tilde{A}\|}{\|\tilde{A} + \lambda I\|} + \frac{\|A^{(N)T} \mathbf{b}^{(N)} - A^T \mathbf{b}\|}{\|A^T \mathbf{b}\|} \right) \quad (3.11)$$

が得られる [35]. ただし, $\Delta \tilde{A} \stackrel{d}{=} \tilde{A}^{(N)} - \tilde{A}$ である. 条件数 κ は

$$\kappa = \|\tilde{A} + \lambda I\| \|(\tilde{A} + \lambda I)^{-1}\| \approx \frac{\sigma_1^2 + \lambda}{\lambda}$$

であるから,

$$\begin{aligned} \|\Delta \tilde{A}\| &= \varepsilon_1 = O(\varepsilon) \\ \|A^{(N)T} \mathbf{b}^{(N)} - A^T \mathbf{b}\| &= \varepsilon_2 = O(\varepsilon) \end{aligned}$$

である. $\|A^T \mathbf{b}\| = O(1)$. $\sigma_1^2 \gg \lambda \gg \varepsilon$ と仮定すると, (3.11) の右辺は

$$\begin{aligned} \frac{\kappa}{1 - \kappa \|\Delta \tilde{A}\| / \|\tilde{A} + \lambda I\|} \left(\frac{\|\Delta \tilde{A}\|}{\|\tilde{A} + \lambda I\|} + \frac{\|A^{(N)T} \mathbf{b}^{(N)} - A^T \mathbf{b}\|}{\|A^T \mathbf{b}\|} \right) &\approx \frac{\sigma_1^2}{\lambda - \varepsilon_1} \left(\frac{\varepsilon_1}{\sigma_1^2} + \varepsilon_2 \right) \\ &\approx \frac{\varepsilon_1}{\lambda} + \frac{\sigma_1^2 \varepsilon_2}{\lambda} \end{aligned}$$

となる. これは正則化パラメータ λ が ε よりも十分に大きければ, (3.10) の解 $\mathbf{x}_\lambda^{(N)}$ は意味のあるものであることを示している.

次に, $A^{(N)}$ の特異値と A の特異値との関係について考察する. P. C. Hansen は $A^{<N>}$ の特異値と A の特異値との関係について以下のことを示した [12].

定理 3.2.1 [12] $\{\sigma_i^{<N>}\}_{i=1}^N$ を $A^{<N>}$ の特異値とすると, 次の関係が成り立つ.

$$\sigma_i^{<N>} \leq \sigma_i^{<N+1>} \leq \sigma_i, \quad i = 1, \dots, N.$$

定理 3.2.2 [12] $\delta_N \stackrel{d}{=} \|A^{<N>} - A\|$ とすると,

$$0 \leq \sigma_i - \sigma_i^{<N>} \leq \delta_N, \quad i = 1, \dots, N$$

が成り立つ.

定理 3.2.3 [12]

$$\sum_{i=1}^N (\sigma_i - \sigma_i^{<N>}) \leq \delta_N^2.$$

定理 3.2.4 [12]

$$\sigma_i^{<N>} \leq \sigma_i \leq \left((\sigma_i^{<N>})^2 + \delta_N^2 \right)^{\frac{1}{2}}, \quad i = 1, \dots, N.$$

これらの結果は $A^{(N)}$ が数値積分により精度良く $A^{<N>}$ を近似していて, なおかつ δ_N が十分に小さければ, $A^{(N)}$ の特異値は A の特異値を良く近似していることを示している.

以後この章において我々は方程式 (3.9) を $A\mathbf{x} = \mathbf{b}$ と見なし議論を進める.

3.3 AIC による正則化 —MAICE-DP 法—

3.5節で述べる PRESS 法, GCV 法は統計的検定を基礎としている. 本節でも同様に, 統計的な立場から正則化パラメータを推定する方法について考察する [17].

3.1節で述べたように, Morozov は事前にノイズの上界が判れば, 常に誤差の上界となるように定義された関数 $\hat{E}(\lambda; b^\delta)$ を残差により最小化できることを示した. これより, もし, 統計的手法を用いてノイズレベルを推定できれば, 正則化パラメータの選択ができるであろうと予測される. 以下, 情報量 AIC (Akaike Information Criterion, [1],[2],[3]) を用いたデータ平滑化法, AIC 最小化推定法 (Minimum AIC Estimation, MAICE 法, [27]) と, 相変原理を用いたパラメータ推定法について考察する. 我々はこの手法を MAICE-DP 法と命名する.

3.3.1 AIC 最小化推定法

ノイズのレベルを推定する問題への一つのアプローチとして回帰多項式を利用する方法がある. しかし, このままではあてはめるべき多項式の次数がわからないという問題点が残る. もし, 次数を低く抑えると, 回帰多項式は真のデータから偏る. 逆に次数の高いものを許容すると, 回帰多項式はノイズの影響を過度に受けることになる. この問題に対して田辺國士は情報量 AIC を用いて, この多項式の次数を決定する MAICE 法を提案した [27].

情報量 AIC は, 現実に観測されるデータを用いて統計的モデルの適切さを評価するために赤池弘次により導入された統計量である. それは, 次式のように定義される [2].

$$AIC = -2 \times \ln(\text{尤度の最大値}) + 2 \times (\text{自由なパラメータの数}). \quad (3.12)$$

ここで, パラメータ数とは, モデル内において自由に変化させられるものの数である. この AIC は情報量規準の推定値として広く使われ, また, その有効性がさまざまな分野から報告されている.

本節では, 問題を次のように定式化する. m 個の観測点 $\{s_i\}_{i=1}^m$ 上で観測されたデータ $\{b_i^\delta\}_{i=1}^m$ は, 平均が 0 で, 未知分散が δ_0^2 のノイズ $\{\varepsilon_i\}_{i=1}^m$ を含んでいるものとする. すなわち,

$$b_i^\delta = g(s_i) + \varepsilon_i.$$

これより, 第 1 種フレドホルム積分方程式

$$\int_a^b K(s, t) f(t) dt = g(s), \quad s_{\min} \leq s \leq s_{\max}$$

を満たす可積分関数 $f(t)$ を求めることが我々の問題である.

まず, データ $\{b_i^\delta\}$ から回帰多項式を用いて $g(s)$ を推測する MAICE 法の概要を紹介する.

回帰多項式 $h_q(s)$ の形は一般に

$$h_q(s) = \sum_{i=0}^q c_i \varphi_i(s)$$

と表される. ここで, $\varphi_i(s)$ は i 次の多項式である. $g(s)$ の推定に q 次の多項式 $h_q(s)$ を用いると, このモデルに対応する $h_q(\cdot)$ と δ_0^2 の対数尤度関数 $\ln L(b^\delta; h_q, \delta_0^2)$ は

$$-2 \ln L(b^\delta; h_q, \delta_0^2) = m \ln(2\pi\delta_0^2) + \frac{\|b^\delta - G_q c_q\|^2}{\delta_0^2} \quad (3.13)$$

で与えられる。ただし、

$$G_q \stackrel{\text{d}}{=} \begin{pmatrix} \varphi_0(s_1) & \cdots & \varphi_q(s_1) \\ \vdots & \ddots & \vdots \\ \varphi_0(s_m) & \cdots & \varphi_q(s_m) \end{pmatrix} \in \mathbf{R}^{m \times (q+1)}, \quad \mathbf{c}_q \stackrel{\text{d}}{=} \begin{pmatrix} c_0 \\ \vdots \\ c_q \end{pmatrix} \in \mathbf{R}^{q+1} \quad (3.14)$$

とする。したがって、次数 q を固定したときの係数 \mathbf{c}_q の最尤推定値は (3.13) の右辺の第 2 項の分子を最小とする最小 2 乗解

$$\mathbf{c}'_q \stackrel{\text{d}}{=} G_q^{\dagger} \mathbf{b}^{\delta} = (\bar{c}_0^{(q)}, \dots, \bar{c}_q^{(q)})^T \quad (3.15)$$

となる。このときの $g(\cdot)$ と δ^2 の最尤推定値はそれぞれ

$$\bar{g}_q(s) \stackrel{\text{d}}{=} \sum_{i=0}^q \bar{c}_i^{(q)} \varphi_i(s),$$

$$\bar{\delta}_q^2 \stackrel{\text{d}}{=} \frac{\bar{R}_q^2}{m}$$

となる。ここで、 \bar{R}_q^2 は残差平方和

$$\bar{R}_q^2 \stackrel{\text{d}}{=} \|\mathbf{b}^{\delta} - G_q \mathbf{c}'_q\|^2 = \|(I - G_q G_q^{\dagger}) \mathbf{b}^{\delta}\|^2 \quad (3.16)$$

を表す。

上述したように問題は回帰多項式の次数 q の決定にある。次数が決まれば係数は最尤推定法で求めればよいのであるが、残差平方和 \bar{R}_q^2 は q が大きくなるにつれ単調に減少するから、 q を選ぶ基準とはなり得ない。一般に q は観測者が \bar{R}_q^2 の動きをにらみながら、経験的、主観的に選ぶのが普通である。しかし、MAICE 法を用いると、これを客観的に決めることができる。

次数 q のモデルを採用したときの統計量 AIC を $\text{AIC}(q; \mathbf{b}^{\delta})$ とすると、(3.12) の第 1 項は (3.13) に $\bar{\delta}_q^2$ を代入した

$$m \ln \bar{\delta}_q^2 + m \ln 2\pi\epsilon$$

となる。そして、自由なパラメータ数は多項式の係数と δ であるから $q+2$ となり、結局

$$\begin{aligned} \text{AIC}(q; \mathbf{b}^{\delta}) &= m \ln \bar{\delta}_q^2 + 2q + m \ln 2\pi\epsilon + 4 \\ &= m \ln \bar{R}_q^2 + 2q + m \ln \frac{2\pi\epsilon^2}{m} + 4 \end{aligned} \quad (3.17)$$

となる。上式を最小とする q を \bar{q} とすると、

$$\bar{g}_{\bar{q}}(s) = \sum_{i=0}^{\bar{q}} \bar{c}_i^{(\bar{q})} \varphi_i(s)$$

が MAICE 法が選択する $g(s)$ の推定関数となる。そして、データベクトル \mathbf{b}^{δ} に含まれるノイズの推定レベルは

$$\left(\sum_{i=1}^m (b_i^{\delta} - \bar{g}_{\bar{q}}(s_i))^2 \right)^{\frac{1}{2}} \quad (3.18)$$

となる。これは残差平方和 $\bar{R}_{\bar{q}}^2$ の平方根に他ならない。

ただし、注意すべきこととして、この方法が漸近理論に基づいていることを考慮すると、たとえば、

$$\frac{q+1}{m} < \frac{1}{10} \quad (3.19)$$

となる範囲で用いるべきであると田辺は指摘している [27].

3.3.2 MAICE-DP 法のアゴリズム

以上の算法を次にまとめる. 先ず, 積分核 $K(s, t)$ と m 個の標本点上でのデータ $\{b_i^\delta\}$ を与える.

[Step 1] $L_2[a, b]$ 上の n 個の基底関数 $\{\psi_j(t)\}$ を与え, 適当な数値積分則を用いて,

$$\int_a^b K(s_i, t)\psi_j(t)dt \quad (3.20)$$

を ij 成分とする行列 A を計算する.

[Step 2] 行列 A の特異値分解

$$\begin{aligned} A &= U\Sigma V^T \\ U &= (\mathbf{u}_1, \dots, \mathbf{u}_n) \in \mathbf{R}^{m \times n}, \quad V = (\mathbf{v}_1, \dots, \mathbf{v}_n) \in \mathbf{R}^{n \times n} \\ \Sigma &= \text{diag}(\sigma_1, \dots, \sigma_n) \end{aligned}$$

を求める. ベクトル $\{\mathbf{u}_i\}$, $\{\mathbf{v}_i\}$ は正規直交基底である.

[Step 3] 区間 $[s_{\min}, s_{\max}]$ 上の適当な多項式 $\{\varphi_i(s)\}$ を与え, $0 \leq q \leq \frac{m}{10}$ の範囲で, 行列 G_q と情報量 $\text{AIC}(q; b^\delta)$ を計算する. $\text{AIC}(q; b^\delta)$ が最小となる q を \bar{q} とすると, このときの残差平方和 $\bar{R}_{\bar{q}}^2$ が MAICE 法によるノイズの 2 乗ノルムの推定値である. それを δ^2 とする. なお, 最小 2 乗解 (3.15) は不必要である.

[Step 4] Tikhonov の正則化法による近似解 $\mathbf{x}_\lambda^\delta$ の残差の 2 乗ノルムを

$$r(\lambda; b^\delta)^2 = \|A\mathbf{x}_\lambda^\delta - b^\delta\|^2$$

とすると, 方程式

$$r(\lambda; b^\delta) = \bar{\delta}^2 \quad (3.21)$$

の解が MAICE-DP 法が選択した正則化パラメータである. 上の方程式は行列の特異値分解を用いると,

$$\sum_{i=1}^n \left(\frac{\lambda}{\sigma_i^2 + \lambda} (\mathbf{u}_i, b^\delta) \right)^2 = \bar{\delta}^2$$

と λ の陽な非線形方程式として表せる. 関数 $r(\lambda; b^\delta)$ は狭義単調増加関数であるから [10], この方程式の解は一意である.

[Step 5] 方程式 (3.21) の解を λ_q とすると,

$$\bar{f}_{\lambda_q}(t) \stackrel{\text{d}}{=} \sum_{i=1}^n \bar{x}_i \psi_i(t)$$

が MAICE-DP 法による第 1 種フレドホルム積分方程式の近似解である. ここで, $\mathbf{x}_{\lambda_q}^\delta \stackrel{\text{d}}{=} (\bar{x}_1, \dots, \bar{x}_n)^T$ である.

3.3.3 MAICE-DP 法の数値実験および考察

次に, MAICE-DP 法の実用性を数値実験により検証する. 数値実験は第 1 種フレドホルム積分方程式

$$\int_0^1 \sqrt{s^2 + t^2} f(t) dt = \frac{(1 + s^2)^{\frac{3}{2}} - s^3}{3}, \quad 0 \leq s \leq 1$$

に対して以下の手順で行なった。

右辺のデータ関数の定義域である $[0, 1]$ 区間に 100 個の一樣乱数を用いて標本点を探った。ノイズは平均が 0 で標準偏差が 1.0×10^{-7} の正規乱数を用いた。解関数空間の基底関数 $\{\psi_j(t)\}$ およびデータの平滑化にはルジャンドル多項式を採用した。その理由は、ガウス・ルジャンドル数値積分則を用いることにより (3.20) の数値積分が容易に、高精度で行なえるからである。残差平方和 (3.16) の計算にはライブラリ [37] の LEQLSD を、行列の特異値分解と非線形方程式 (3.21) の解法にはライブラリ [38] の DASVD1 と DTSD1 を用いた。計算は全て倍精度演算で行なった。

まず、解関数空間の基底関数として 0 次から 99 次までのルジャンドル多項式を用いて数値実験を行なった。このとき、行列 A のサイズは 100×100 となる。その結果を表 1 に示す。

表 1: ノイズベクトル 1. に対する実験結果.

q	AIC	Est Noise	Parameter	Error
0	-0.37770E+02	0.81152E+00	0.38150E+01	0.19323E+00
1	-0.37715E+03	0.14723E+00	0.12640E+01	0.11940E+00
2	-0.73859E+03	0.23922E-01	0.36257E+00	0.30437E-01
3	-0.11368E+04	0.32348E-02	0.12312E+00	0.14288E-01
4	-0.16215E+04	0.28377E-03	0.30732E-01	0.44238E-02
5	-0.23867E+04	0.61224E-05	0.36435E-02	0.91680E-03
6	-0.25153E+04	0.31875E-05	0.24891E-02	0.71509E-03
7	-0.27385E+04	0.10336E-05	0.85365E-03	0.33435E-03
8	-0.27610E+04	0.91420E-06	0.32386E-03	0.11075E-03
9	-0.27602E+04	0.90902E-06	0.27465E-03	0.17600E-03
10	-0.27599E+04	0.90107E-06	0.20436E-03	0.36813E-03

ここで、 q はデータの平滑化に用いた回帰多項式の次数を表す。AIC は $AIC(q; b^{\delta})$ から定数項を除いたもの、Est Noise はノイズレベルの推定値、Parameter は正則化パラメータの平方根、解の平方根、Error は $\|x_0 - x_{\lambda_0}^{\delta}\|$ である。

この表から $q = 8$ のとき AIC が最小となり、そのときに誤差が最小となることがわかる。なお、この実験に用いたノイズベクトルを以下の数値実験での便宜上ノイズベクトル 1. とする。

表 1 では回帰多項式の次数毎に誤差を表したが、次の図 1 は、パラメータを連続的に動かしたときの誤差の曲線の中で、MAICE-DP 法が選択した正則化パラメータが、どのような位置にあるかを図示したものである。

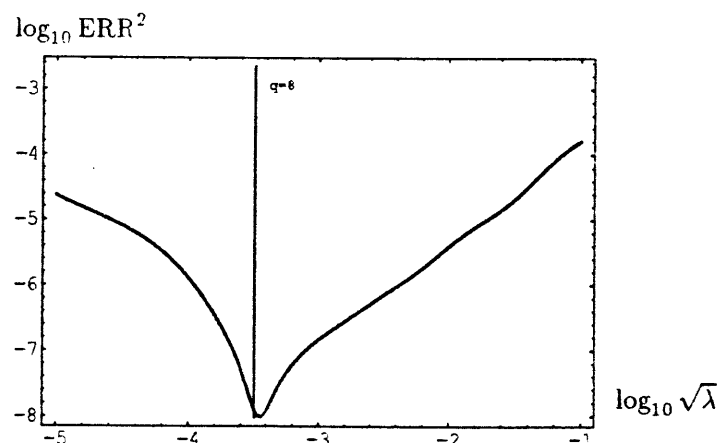


図 1: MAICE-DP 法による正則化パラメータ

図中で $q = 8$ と書かれている直線が MAICE-DP 法が選択したパラメータの位置を表している。この図を見る限り、MAICE-DP 法はかなり精度の良い正則化パラメータの推定に成功しているといえる。

次に、ノイズベクトル1.と同じノルムの、二つの異なるノイズベクトル2., 3.を用いての数値実験を行った。その結果がそれぞれ表2と表3である。

表2: ノイズベクトル2.に対する実験結果.

q	AIC	Est Noise	Parameter	Error
7	-0.27387E+04	0.10324E-05	0.74514E-03	0.57932E-03
8	-0.27615E+04	0.91212E-06	0.17651E-04	0.30937E-02
9	-0.27598E+04	0.91083E-06	0.14885E-04	0.42792E-02
10	-0.27578E+04	0.91071E-06	0.14649E-04	0.43921E-02

表3: ノイズベクトル3.に対する実験結果.

q	AIC	Est Noise	Parameter	Error
7	-0.27380E+04	0.10360E-05	0.64393E-03	0.38688E-03
8	-0.27575E+04	0.93071E-06	0.15292E-05	0.27655E-01
9	-0.27555E+04	0.93064E-06	0.14882E-05	0.28563E-01
10	-0.27536E+04	0.92991E-06	0.10680E-05	0.41027E-01

表2, 表3ともに $q = 8$ のとき AIC が最小になっているが, 表1とは違い, そのときに誤差最小とはならない。これらの表からは, MAICE-DP 法は最適の正則化パラメータを推定しているとは言い難い。これについては, 図2から次のように考察される。

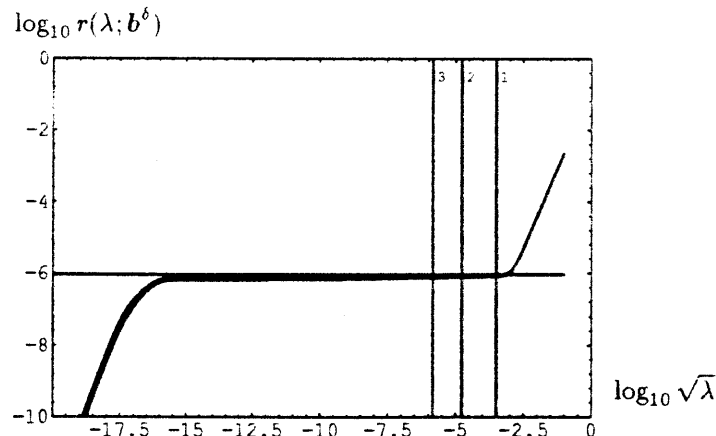


図2: ノイズベクトル1., 2., 3. に対する MAICE-DP 法が選択した正則化パラメータと推定ノイズレベルと残差ノルム。

図中の1, 2, 3と番号付けられた直線が, それぞれノイズベクトル1., 2., 3. に対して MAICE-DP 法が選択した正則化パラメータの位置を表している。上述したように, $r(\lambda; b^\delta)$ は狭義単調増加関数である。しかし, あるところから増加が鈍り, ほぼ平坦になっていることがわかる。そして, 平坦となるのは大体ノイズレベル付近である。この現象はこの数値例だけではなく, これ以外の悪条件線形方程式においても観測される。

問題は, この残差ノルムが平坦になるところにある。この範囲内ではパラメータの動きに対して残差ノルムの変化は鈍いが, 誤差は鋭く変化している。MAICE-DP 法は図2中の横線である推定ノイズレベルと残差ノルムのカーブとの交点を正則化パラメータとして採用する方法である。上で実験を行なった3つのノイズに対する MAICE 法による推定ノイズレベルはそれぞれ 0.914×10^{-6} , 0.912×10^{-6} , 0.931×10^{-6}

である。この差は MAICE 法の許容誤差範囲内であると言える。事実、図 2 では 3 本の横軸はほとんど 1 本に見える。しかし、残差ノルムの曲線がほぼ平坦であるため、これほどの差ですら無視できず、正規化パラメータは 2 桁近い差が生じている。さらに、誤差の曲線もパラメータに敏感であるため、大きく違っているのである。以上のことから考えると、ノイズベクトル 1. に対する実験は、偶然成功した例であると考えべきである。なお、図中の繁雑を避けるため、それぞれのノイズベクトルに対する誤差曲線は割愛したが、ほぼ同様の振る舞いであることを付け加えておく。

次の図 3 は行列 A の特異値 σ_i を大きい順に並べて常用対数の尺度で表示したものである。

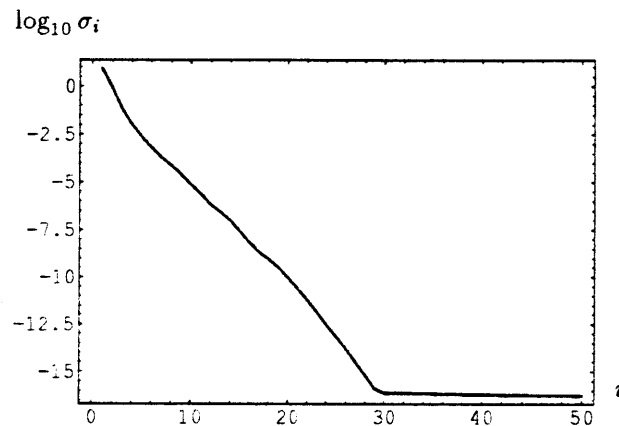


図 3: 行列 A の特異値.

特異値は 30 までは順調に 0 に収束していくが、それ以降は丸め誤差レベルで頭打ちとなり横ばいとなる。理論上、非退化核を持つ積分作用その特異値は、0 に集積する。しかし、倍精度演算であるため、丸め誤差以下の特異値はそのレベルに丸められてしまいます。これより、丸め誤差レベルにある特異値は信頼性が薄く、この場合であれば、解関数空間の基底関数は 30 個程度で十分である。

そこで、40 個のルジャンドル多項式 (0~39 次) を用いて同様の実験を行なった。その結果が表 4 である。

表 4: 基底関数 40 個での実験結果.

q	AIC	Est Noise	Parameter	Error
7	-0.27385E+04	0.10336E-05	0.12054E-02	0.43683E-03
8	-0.27610E+04	0.91420E-06	0.10776E-02	0.40295E-03
9	-0.27602E+04	0.90902E-06	0.10715E-02	0.40127E-03
10	-0.27599E+04	0.90107E-06	0.10621E-02	0.39865E-03

この表で、AIC と Est Noise の項は、ノイズ、標本点、データは同じものを用いているので、表 1 と同じである。しかし、表 1 では $q=8$ のときに誤差が最小となったが、ここでは最小とならない。さらに、表中では省略したが、 q を 10 よりも大きくすると、誤差は減少する。

これは、図 4 より、次のように考察することができる。図 4 は解関数空間の基底関数の個数を $n = 20, 40, 60, 100$ としたときの相変原理の誤差の上界 $\hat{E}(\lambda; b^b)$ と (E1), 誤差 (E2) を図示したものである。縦線は E1 と E2 が最小となるパラメータの位置を示している。横線はノイズレベルを表している。この図からわかるように、 $\hat{E}(\lambda; b^b)$ が最小となる λ と、誤差が最小となる λ は、基底関数の個数が少ないほど離れたものとなる。MAICE-DP 法は相変原理を基にした手法であるため、相変原理が選択した正規化パラメータと真の最適の正規化パラメータとがかけ離れている場合は無意味なものとなる。そのため、 $n = 40$ のときは表 4 のように芳しくない結果となったのである。

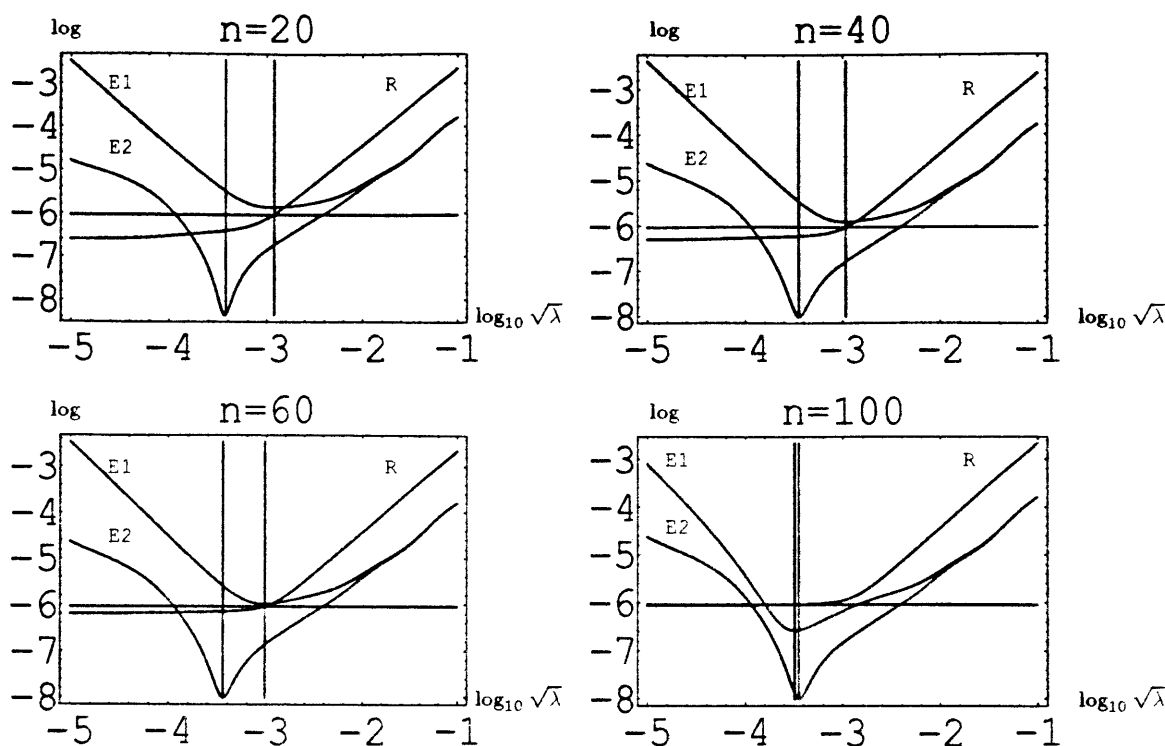


図 4: 誤差推定関数 $\hat{E}(\lambda; b^{\delta})$ と誤差の関係.

上でも述べたように、MAICE 法はノイズレベルを高精度で推定している。これより、MAICE-DP 法は相変原理の意味においての最適の正則化パラメータの選択には成功している。しかし、この手法はパラメータの選定のために不安定な非線形方程式 (3.21) を解かなければならないという問題点が残る。次の節で我々は、MAICE-DP 法よりも有効で、なおかつ安定である“L-カーブによる最適正則化法”を提示する。

3.4 L-カーブによる正則化法

3.4.1 L-カーブ

正則化法による近似解 x_{λ} 、もしくは x_{λ}^{δ} の情報を的確に表示させる方法としては、 x_{λ}^{δ} のノルムと、残差 $Ax_{\lambda}^{\delta} - b^{\delta}$ のノルムを対にして表示する方法が考えられる。データ b が $\mathcal{R}(A)$ に属しているとき、 $\|Ax_{\lambda}^{\delta} - b^{\delta}\|$ は小さいが、 $\|x_{\lambda}^{\delta}\|$ が大きい。または、 $\|x_{\lambda}^{\delta}\|$ は適度な大きさであるが、 $\|Ax_{\lambda}^{\delta} - b^{\delta}\|$ が大きいような x_{λ}^{δ} は、近似解として適当でないのは明らかである。そのため、上のような方法で、二つの量の妥協をとるアイデアが考えられる。この節では、この着想より正則化パラメータを得る方法について考察する。

この表示法を実際に、最初に用いたのは [21] である。また、類似の方法が [23] にも見られる。最近では、[31],[11],[13] などである。この図形に“L-カーブ”と名付けたのは P. C. Hansen[15] である。

なお、以下において、 A は第 1 種フレドホルム積分方程式 (1.1) に適当な数値積分を施すことにより得られた N 行 N 列の行列であるとする。同様に、 b, b^{δ} も N 次元ベクトルに離散化されたものとする。ここで、数値積分による離散化誤差は、計算機イプシロンに比べ、小さくなるようにあらかじめ分点数 N を十分大きくとるものとする。なお、数値積分については 4.5 節で述べる。

定義 3.4.1 [16],[15]

$$r(\lambda; \mathbf{b}^\delta) \stackrel{\text{d}}{=} \|A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta\|, \quad x(\lambda; \mathbf{b}^\delta) \stackrel{\text{d}}{=} \|\mathbf{x}_\lambda^\delta\|$$

とすると, L-カーブとは, $r(\lambda; \mathbf{b}^\delta)$ を x 座標, $x(\lambda; \mathbf{b}^\delta)$ を y 座標として, 正則化パラメータ λ を $[0, \infty)$ の範囲で動かすことにより得られる点 $(r(\lambda; \mathbf{b}^\delta), x(\lambda; \mathbf{b}^\delta))$ の集合である.

L-カーブは次のような性質を持つ.

定理 3.4.1 [16]

$$\mathbf{b}_o^\delta \stackrel{\text{d}}{=} \sum_{i=1}^N (\mathbf{u}_i, \mathbf{b}^\delta) \mathbf{u}_i, \quad \mathbf{x}_o^\delta \stackrel{\text{d}}{=} \sum_{i=1}^N \frac{(\mathbf{u}_i, \mathbf{b}^\delta)}{\sigma_i} \mathbf{v}_i$$

とすると,

$$\begin{cases} r(\lambda; \mathbf{b}^\delta) \rightarrow \|\mathbf{b}_o^\delta\|, \\ x(\lambda; \mathbf{b}^\delta) \rightarrow 0, \end{cases} \quad (\lambda \rightarrow \infty),$$

$$\begin{cases} r(\lambda; \mathbf{b}^\delta) \rightarrow 0, \\ x(\lambda; \mathbf{b}^\delta) \rightarrow \|\mathbf{x}_o^\delta\|, \end{cases} \quad (\lambda \rightarrow 0)$$

である.

[証明] $A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta, \mathbf{x}_\lambda^\delta$ はそれぞれ特異値分解を用いて

$$A\mathbf{x}_\lambda^\delta - \mathbf{b}^\delta = \sum_{i=1}^N \frac{\lambda}{\sigma_i^2 + \lambda} (\mathbf{b}^\delta, \mathbf{u}_i) \mathbf{u}_i$$

$$\mathbf{x}_\lambda^\delta = \sum_{i=1}^N \frac{\sigma_i}{\sigma_i^2 + \lambda} (\mathbf{b}^\delta, \mathbf{u}_i) \mathbf{v}_i$$

と表すことができる. これより明らかである. ■

定理 3.4.2 [16] $\lambda \in (0, \infty)$ において, $x(\lambda; \mathbf{b}^\delta)$ は $r(\lambda; \mathbf{b}^\delta)$ の単調減少関数である.

[証明] 前定理の証明より,

$$r(\lambda; \mathbf{b}^\delta) = \left(\sum_{i=1}^N \left(\frac{\lambda}{\sigma_i^2 + \lambda} \right)^2 (\mathbf{b}^\delta, \mathbf{u}_i)^2 \right)^{1/2}$$

$$x(\lambda; \mathbf{b}^\delta) = \left(\sum_{i=1}^N \left(\frac{\sigma_i}{\sigma_i^2 + \lambda} \right)^2 (\mathbf{b}^\delta, \mathbf{u}_i)^2 \right)^{1/2}$$

が得られる. これらを λ で微分すると,

$$\frac{d}{d\lambda} r(\lambda; \mathbf{b}^\delta) = \frac{\sum_{i=1}^N \frac{\sigma_i^2 \lambda}{(\sigma_i^2 + \lambda)^3} (\mathbf{b}^\delta, \mathbf{u}_i)^2}{r(\lambda; \mathbf{b}^\delta)}$$

$$\frac{d}{d\lambda} x(\lambda; \mathbf{b}^\delta) = \frac{\sum_{i=1}^N \frac{-\sigma_i^2}{(\sigma_i^2 + \lambda)^3} (\mathbf{b}^\delta, \mathbf{u}_i)^2}{x(\lambda; \mathbf{b}^\delta)}$$

となる. 今, $\lambda > 0$ であるから,

$$\frac{d}{d\lambda} r(\lambda; \mathbf{b}^\delta) > 0, \quad \frac{d}{d\lambda} x(\lambda; \mathbf{b}^\delta) < 0$$

となり, これより

$$\frac{dx(\lambda; b^\delta)}{dr(\lambda; b^\delta)} = \frac{dx(\lambda; b^\delta)}{d\lambda} \frac{d\lambda}{dr(\lambda; b^\delta)} < 0$$

が得られる. ■

ただし, 定理 3.4.1において, $\|x_0^\delta\|$ は $N \rightarrow \infty$ で発散する量である.

L-カーブの大きな特徴は, 曲線の途中に“コーナー”を持つことである. 悪条件線形方程式でなければ, このコーナーは現れない. 名前の由来もこれに依存する.

特徴 [16],[15] ベクトル Δb を摂動ベクトルとする. すなわち, $b^\delta = b + \Delta b$ とし, 次を仮定する.

1. 右辺ベクトル b の特異ベクトルによるフーリエ係数の絶対値 $|(u_i, b)|$ は, 特異値 σ_i よりも速く 0 に収束する.
2. 摂動ベクトル Δb は, 平均が 0, 標準偏差が δ_0 の正規乱数ベクトルである.
3. Δb のノルムは, $\|\Delta b\| < \|b\|$ を満たす.

このとき, L-カーブ $(r(\lambda; b^\delta), x(\lambda; b^\delta))$ はコーナーを持つ. そのコーナーは, 大体 $(\|\Delta b\|, \|x_0\|)$ の付近である.

ここで, 上の 3 つの条件のうち, 仮定 1. は, 定理 1.4.11 に対比させて, 離散的 Picard 条件 [14] と呼ばれている.

上の特徴に正確な証明を与えることはできない. しかし, この現象は次のように説明することができる.

今, ベクトル x_λ^δ を二つのベクトル

$$x_\lambda \triangleq \sum_{i=1}^N \frac{\sigma_i}{\sigma_i^2 + \lambda} (u_i, b) v_i, \quad x_\lambda^\Delta \triangleq \sum_{i=1}^N \frac{\sigma_i}{\sigma_i^2 + \lambda} (u_i, \Delta b) v_i$$

により, $x_\lambda^\delta = x_\lambda + x_\lambda^\Delta$ とする. 離散的 Picard 条件より, σ_i は $|(u_i, b)|$ よりも速く 0 に収束する. しかし, $|(u_i, \Delta b)/\sigma_i|$ は σ_i が 0 に収束していくに従って増大していく. ゆえに, $\lambda \ll 1$ では, $\|x_\lambda\| \ll \|x_\lambda^\Delta\|$ となる. これは, x_λ^δ は x_λ^Δ の項が支配していることを意味する.

λ を増加させると, x_λ^Δ の項に影響が出て, 曲線は降下を始める. さらに λ を増加させると x_λ の項にも影響が出始める. このとき, x_λ^δ を支配する項が x_λ^Δ から x_λ へと変わり, 曲線は水平となる. そして, $(\|\Delta b\|, 0)$ へと収束していく.

つまり, L-カーブは曲線の中ほどに“コーナー”を持ち, そのコーナーは x_λ^δ を支配する項が x_λ^Δ から x_λ へと変わる点で, $\|\Delta b\| = \delta$ とすると, $(\delta, \|x_0\|)$ 付近にある.

定理 3.1.2 を用いるならば, $r(\lambda; b^\delta) = \delta$ となるとき, 誤差の上界 $\hat{E}(\lambda; b^\delta)$ が最小となる. これより, コーナーを表現する λ は, 相変原理の意味での最適の正則化パラメータといえる.

ここで, 注目すべきことは, 問題を解くにあたって, ノイズベクトルのノルムに関する情報を必要としないことである.

次に, 実際の第 1 種フレドホルム積分方程式に上の L-カーブを適用する. 図 5 は, 積分方程式

$$\int_0^1 e^{st} f(t) dt = \frac{e^{s+1} - 1}{s + 1}, \quad 0 \leq s \leq 1 \quad (3.22)$$

に対する結果である. 上式の解は $f(t) = e^t$ で, [5] より引用した. 摂動ベクトルは平均が 0, 標準偏差が 1.0×10^{-5} の正規乱数を用いた. 離散化は 50 点ガウス・ルジャンドル数値積分を用いた. このときの摂

動ベクトルのノルムは $\delta = 0.66 \times 10^{-4}$, $\|\mathbf{x}_0\| = 1.78$ である。数値計算は倍精度演算, 特異値分解と正規乱数の生成はライブラリ [38] を用いた。

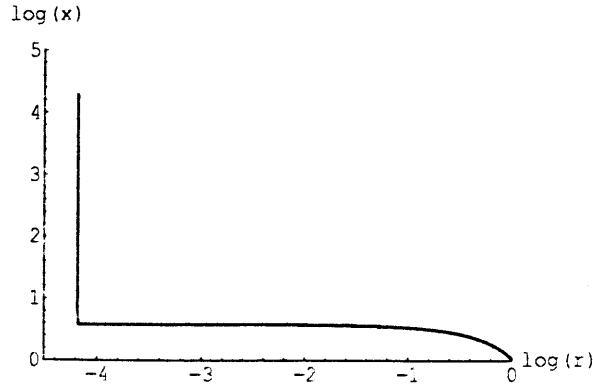


図 5: 積分方程式 (3.22) の L-カーブ。

図は横軸, 縦軸とも常用対数の尺度である。 λ は最大特異値の 2 乗 σ_1^2 から, 計算機イプシロン $\epsilon = 1.0 \times 10^{-15}$ の範囲で動かした。曲線が $(\delta, \|\mathbf{x}_0\|)$ 付近でコーナーを持つ様子がわかる。

図 6 は, 第 1 種フレドホルム積分方程式

$$\int_0^1 \sqrt{s^2 + t^2} f(t) dt = \frac{(1 + s^2)^{\frac{3}{2}} - s^3}{3}, \quad 0 \leq s \leq 1 \quad (3.23)$$

に適用した結果である。この方程式の解は $f(t) = t$, (3.22) 同様に [5] より引用した。 $\delta = 0.73 \times 10^{-4}$, $\|\mathbf{x}_0\| = 0.58$ である。

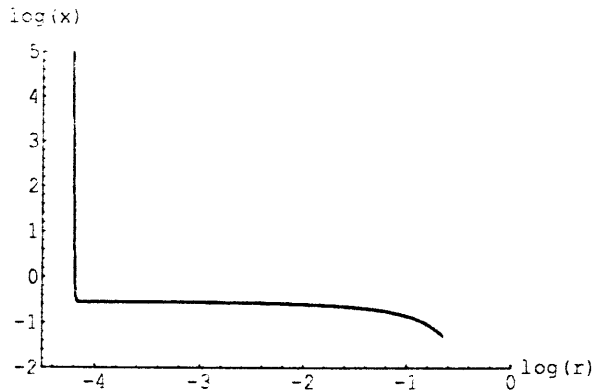


図 6: 積分方程式 (3.23) の L-カーブ。

次に, 積分方程式の特異値と右辺のデータ関数の特異関数によるフーリエ係数が正確にわかっている第 1 種フレドホルム積分方程式

$$\begin{aligned} \frac{ab}{2} \int_{-1}^1 \left(\frac{\sin(\pi(s+t))}{a^2 - 2ab \cos(\pi(s+t)) + b^2} + \frac{\sin(\pi(s-t))}{a^2 - 2ab \cos(\pi(s-t)) + b^2} \right) f(t) dt \\ = \frac{b \sin \pi s}{1 - 2b \cos \pi s + b^2}, \quad -1 \leq s \leq 1 \end{aligned} \quad (3.24)$$

対する同様の結果を図 7 に示す。上式の解は

$$f(t) = \frac{a - \cos \pi t - a^2}{1 - 2a \cos \pi t + a^2}$$

であって、この k 番目の特異値及び右辺のフーリエ係数はそれぞれ $(b/a)^k, b^k$ である。数値実験では $a = 0.2, b = 0.05$ と設定した。

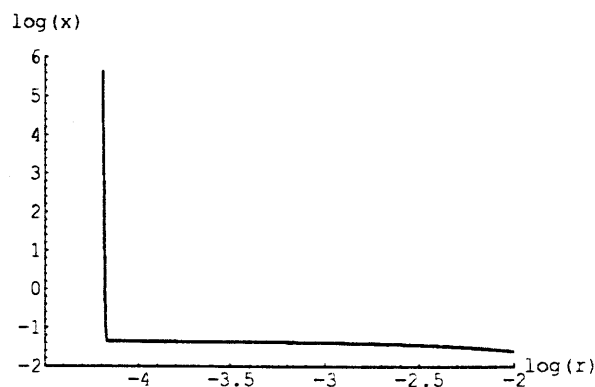


図 7: 積分方程式 (3.24) の L-カーブ.

最後に、退化核を持つ第 1 種フレドホルム積分方程式

$$\int_0^1 (s-t)^2 f(t) dt = \frac{85}{3} s^2 - \frac{85}{3} s + \frac{143}{15}, \quad 0 \leq s \leq 1 \quad (3.25)$$

への同様の結果を図 8 に示す。

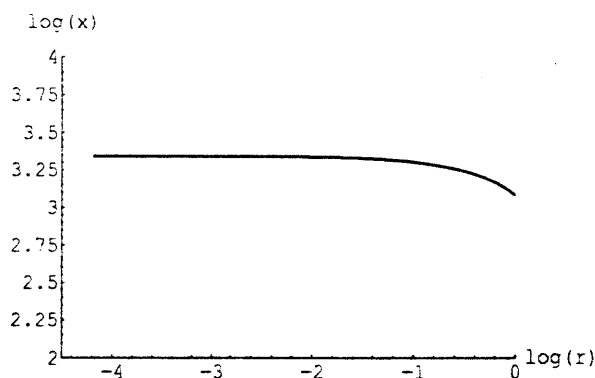


図 8: 積分方程式 (3.25) の L-カーブ.

これは悪条件線形方程式ではないので、曲線にコーナーが現れない。

3.4.2 L-カーブの曲率による正則化法

次に、 λ に関する L-カーブの動きを検証する。図 9 の左図は積分方程式 3.24 に対し、

$$\lambda_n = 10^{-n}, \quad n = 1, 2, \dots, 15$$

の点で L-カーブを図示したものである。右図は各点での誤差

$$E(\lambda; b^\delta) \stackrel{d}{=} \|\mathbf{x}_0 - \mathbf{x}_\lambda^\delta\| \quad (3.26)$$

を表す。

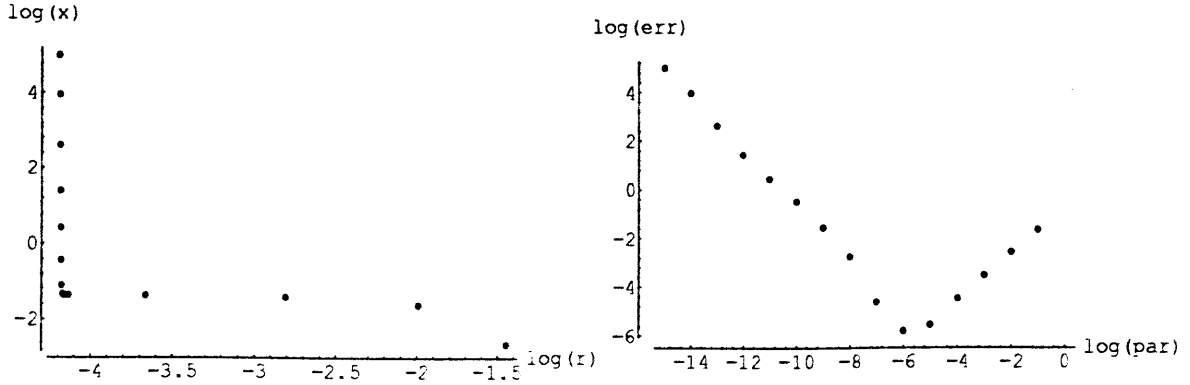


図 9: L-カーブの動きと、誤差.

この図から観測できるように、コーナー付近では λ の動きに比べて、曲線の動きは鈍くなる。しかし、コーナー付近でL-カーブが停滞している間の、 x_λ^* の有効桁数は2桁から6桁と幅広い。これでは、L-カーブのコーナーを表現する λ というだけではパラメータの限定は困難である。

そこで、我々は、L-カーブの曲率が最大となる λ を正則化パラメータとして選択することを提案する [16].

ベクトル値関数 $y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}$ の t における曲率 $\kappa(t)$ は、

$$\kappa(t) = \frac{|y_1'(t)y_2''(t) - y_1''(t)y_2'(t)|}{(y_1'(t)^2 + y_2'(t)^2)^{\frac{3}{2}}}$$

と表せる。これより、L-カーブの λ における曲率は、特異値分解を用いることにより、

$$\kappa(\lambda) = \frac{1}{\left(r(\lambda; \mathbf{b}^\delta)^2 + \lambda^2 x(\lambda; \mathbf{b}^\delta)^2\right)^{\frac{3}{2}}} \times \left| \frac{r(\lambda; \mathbf{b}^\delta)^2 x(\lambda; \mathbf{b}^\delta)^2 (\Sigma_1(\lambda) + 3\lambda \Sigma_2(\lambda))}{\Sigma_3(\lambda)^2} - \lambda \left(\lambda r(\lambda; \mathbf{b}^\delta)^2 + x(\lambda; \mathbf{b}^\delta)^2 \right) \right| \quad (3.27)$$

と、 λ の陽関数として表すことができる。ここで、

$$\begin{aligned} \Sigma_1(\lambda) &\stackrel{\text{d}}{=} \sum_{i=1}^N \frac{\sigma_i^2(\sigma_i^2 - 2\lambda)}{(\sigma_i^2 + \lambda)^4} (\mathbf{u}_i, \mathbf{b}^\delta)^2 \\ \Sigma_2(\lambda) &\stackrel{\text{d}}{=} \sum_{i=1}^N \frac{\sigma_i^2}{(\sigma_i^2 + \lambda)^4} (\mathbf{u}_i, \mathbf{b}^\delta)^2 \\ \Sigma_3(\lambda) &\stackrel{\text{d}}{=} \sum_{i=1}^N \frac{\sigma_i^2}{(\sigma_i^2 + \lambda)^3} (\mathbf{u}_i, \mathbf{b}^\delta)^2 \end{aligned}$$

である。また、L-カーブは常用対数の尺度で用いるが、曲率は不変である。

図 10は積分方程式 (3.22) に対する L-カーブの曲率と誤差 (3.26) を図示したものである。

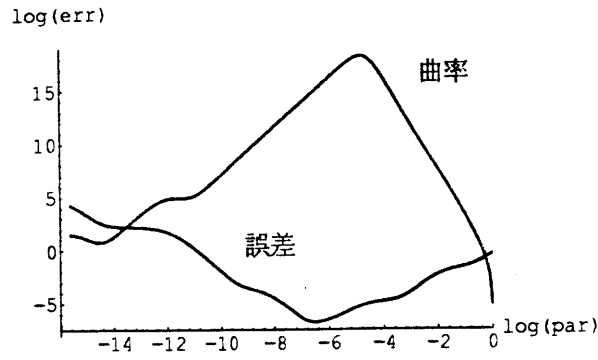


図 10: 積分方程式 (3.22) の L-カーブの曲率と誤差.

これより, 有効桁数 5 桁程度の近似解が得られることがわかる.

図 11 と図 12 は, それぞれ積分方程式 (3.23) と (3.24) の同様の結果である.

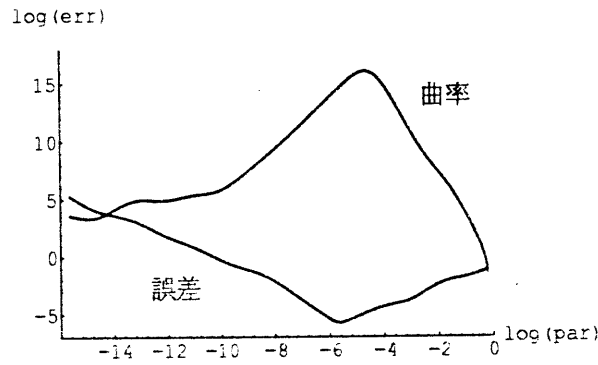


図 11: 積分方程式 (3.23) の L-カーブの曲率と誤差.

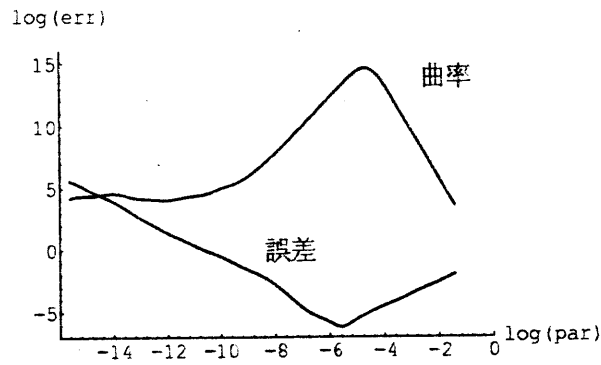


図 12: 積分方程式 (3.24) の L-カーブの曲率と誤差.

これらも, 有効桁数 5 桁程度の近似解が得られている.

最後の例として, MAICE-DP 法の数値実験に用いた問題に対する結果を示す. MAICE-DP 法ではデータ点が 100 個で基底関数が 40 個のときの実験結果は芳しいものではなかったが, L-カーブの曲率による最適正則化法では良好な結果が得られた.

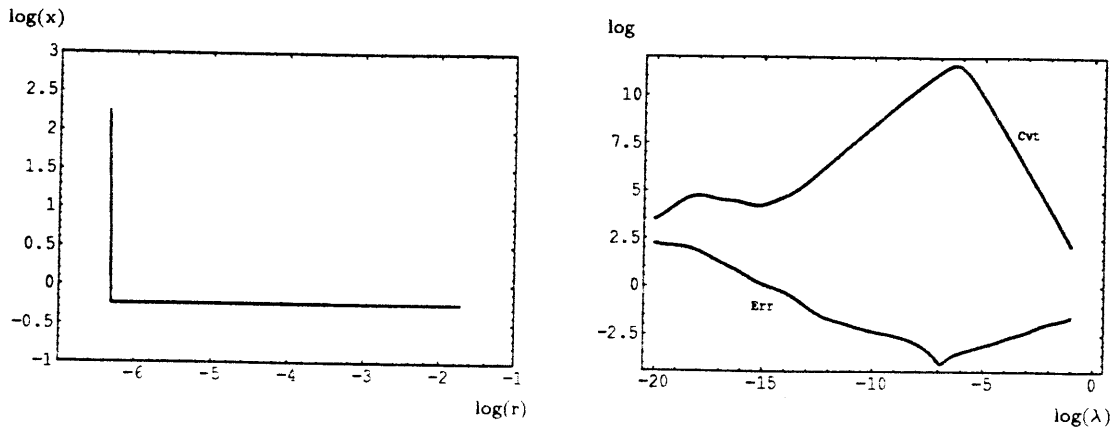


図 13: MAICE-DP 法の数値実験 (解の基底関数 $n = 40$) に用いた問題に対する L-カーブと曲率と誤差.

3.5 一般化交差検証法 —GCV 法—

Morozov の相変原理は、正則化パラメータの選択法として、旧東欧諸国の数学者を中心に支持を集めた。一方、西側諸国では G. Wahba により提唱された一般化交差検証法 (Generalized Cross-Validation [34], 以下 GCV 法) が有名である。この手法は、D. M. Allen により提案された交差検証法 ((Ordinary) Cross-Validation, または、Prediction Sum of Squares, 略して PRESS 法とも呼ばれている [4]) を基にしている。

PRESS 法の基本的な考え方は、「与えられたデータ点の一部を取り除き (通常 1 点), 残りのデータ点を使って解を求め, それが除かれた部分をよく推定するようにパラメータを決定する」というものである。このアイデアを実際に数式で表現すると以下ようになる。今, $b^{\delta, (k)}$ を b^{δ} の第 k 成分を 0 に置き換えたベクトル, すなわち,

$$b^{\delta, (k)} \stackrel{d}{=} (b_1^{\delta}, \dots, b_{k-1}^{\delta}, 0, b_{k+1}^{\delta}, \dots, b_N^{\delta})^T \in \mathbb{R}^N, \quad b = (b_1^{\delta}, \dots, b_N^{\delta})^T \in \mathbb{R}^N$$

データ $b^{\delta, (k)}$ に対する Tikhonov の正則化法による近似解を

$$x_{\lambda}^{\delta, (k)} \stackrel{d}{=} (A^T A + \lambda I)^{-1} A^T b^{\delta, (k)} \in \mathbb{R}^N$$

とする。もし、正則化パラメータ λ の選択が妥当であるならば、ベクトル $Ax_{\lambda}^{\delta, (k)}$ の第 k 成分 $[Ax_{\lambda}^{\delta, (k)}]_k$ は b_k^{δ} のよい推定となることが予想される。すなわち、残差 $([Ax_{\lambda}^{\delta, (k)}]_k - b_k^{\delta})^2$ をすべてのデータ点 b_k^{δ} , $k = 1, \dots, N$ で求め、その総和が最小となるような λ を、最適の λ として採用するのが PRESS 法である。つまり、パラメータ選択のための判定式は

$$P(\lambda; b^{\delta}) \stackrel{d}{=} \frac{1}{N} \sum_{k=1}^N ([Ax_{\lambda}^{\delta, (k)}]_k - b_k^{\delta})^2 \quad (3.28)$$

と表わされる。上式は、二つの行列

$$\hat{A}_{\lambda} \stackrel{d}{=} A(A^T A + \lambda I)^{-1} A^T = (\hat{a}_{ij})$$

$$\hat{B}_{\lambda} \stackrel{d}{=} \text{diag}\left(\frac{1}{1 - \hat{a}_{11}}, \dots, \frac{1}{1 - \hat{a}_{NN}}\right)$$

を定義することにより,

$$P(\lambda; b^{\delta}) = \frac{1}{N} \|\hat{B}_{\lambda}(I - \hat{A}_{\lambda})b^{\delta}\| \quad (3.29)$$

とも表すことができる。

しかし、このように導入された PRESS 法には、二つの弱点がある。一つは計算量の多さである。もう一つは A が対角行列に近い場合の弱さである。極端な例として A が対角行列の場合を考える。このとき、 \hat{A}_λ も対角行列となり、判定式は $P(\lambda; b^\delta) = \frac{1}{N} \sum_{k=1}^N (b_k^\delta)^2$ となる。これは最小点が存在しない。

これらの原因は PRESS 法が直交変換によって不変ではないことに依存する。そして、この要請から考案された手法が GCV 法である。すなわち、GCV 法は PRESS 法の直交変換不変版であるといえる。この一般化は、 A の特異値分解

$$A = U \Sigma V^T$$

と、フーリエ行列

$$W = (w_{ij}), \quad w_{ij} = \frac{1}{\sqrt{N}} e^{\frac{2\pi ijk}{N}}$$

によって実現される。これにより、PRESS 法のパラメータ判定式 (3.29) は、

$$V(\lambda; b^\delta) \triangleq \frac{\frac{1}{N} \|(I - \hat{A}_\lambda) b^\delta\|^2}{\left(\frac{1}{N} \text{Trace}(I - \hat{A}_\lambda)\right)^2} \quad (3.30)$$

と変換される。上式を最小とする λ を正則化パラメータとして採用する手法が GCV 法である。

GCV 法の詳細、収束証明は参考文献 [34],[10]などを参照されたい。

しかし、この手法の悪条件線形方程式に対する有効性への疑問ははやくから指摘されている [32]。GCV 法を MAICE-DP 法で数値実験に用いた問題 ($m = n = 100$, データのノイズはノイズベクトル 1.) に適用した結果を図 14 に示す。

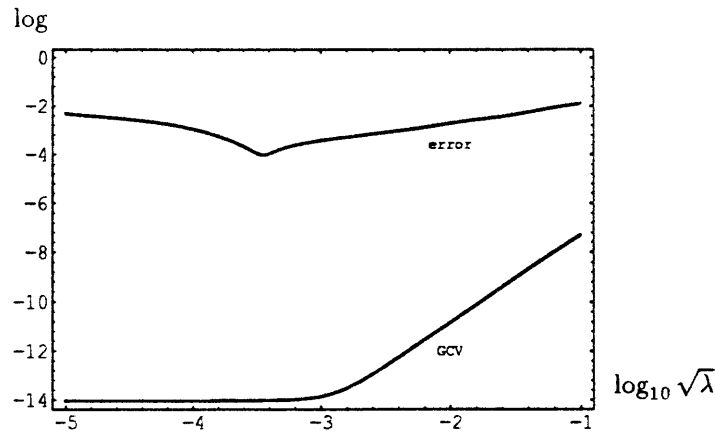


図 14: 方程式 (3.23) に対する GCV 法の結果。

図中の GCV は (3.30) の $V(\lambda; b^\delta)$ を表す。図が表すように、 $V(\lambda; b^\delta)$ はあるところからはほぼ平坦となり、パラメータを選定することは困難である。

3.6 まとめ

Tikhonov の正則化法を悪条件線形方程式に適用するときの本質的な問題は正則化パラメータの選択にある。この正則化パラメータの選択法としては、相変原理と GCV 法が有名である。我々は MAICE-DP 法と L-カーブの曲率による最適正則化法を提案した。MAICE-DP 法は相変原理の意味での最適正則化には

成功している。しかし、相変原理がうまくいかない場合は無意味な結果となる。また、不安定な非線形方程式を解く必要があるという問題点も残る。

L-カーブの曲率による最適正則化法は悪条件線形方程式に対して良好な結果が得られた。また、相変原理がうまくいかない場合にも有効であることが数値実験により確認された。これより、正則化パラメータの選択法としてはL-カーブの曲率による最適正則化法を用いるべきであるというのが我々の結論である。

第 4 章

直接法による悪条件線形方程式の解法

この章では右辺のデータにノイズを含んでいない悪条件線形方程式

$$Ax = b, \quad b \in l_2$$

の解法について考察する。第 1 章で述べたように、一般にこの方程式の解は l_2 空間に存在するとは限らない。そのため我々は打ち切り最小 2 乗最小ノルム解を導入する。

通常、悪条件線形方程式の解法としては特異値分解が用いられる。我々も、先ず特異値分解を用いて打ち切り最小 2 乗最小ノルム解による悪条件線形方程式の数値解法の検証を行なった。その結果は良好であった。

しかし、特異値分解は固有値問題として定式化されるため、直接解法では求めることができず、計算量が多くなるという欠点を持っている。我々の目的は直接解法により特異値分解と同程度の近似解を打ち切り最小 2 乗最小ノルム解として求めることである。

行列の QR 分解による線形方程式の解法は古典的な数値計算法の一つである。また、ランク落ちした線形方程式の最小 2 乗最小ノルム問題に対しても、ピボティング付修正グラム・シュミット法などを用いることにより、1 回の QR 分解で最小 2 乗解が、2 回の QR 分解で最小 2 乗最小ノルム解が得られる [26]。しかし、1, 2 回の QR 分解では打ち切り最小 2 乗最小ノルム解を得ることはできない。我々は、悪条件線形方程式の解法として、3 回の QR 分解による打ち切り最小 2 乗最小ノルム解を求める方法を提案する。この方法により、特異値分解による打ち切り最小 2 乗最小ノルム解と同程度の精度の近似解が得られることが数値実験により確認された。

次の 4.1 節で一般的な打ち切り最小 2 乗最小ノルム解の定義を与える。4.2 節では特異値分解による打ち切り最小 2 乗最小ノルム解について述べ、その有効性を数値実験により検証する。4.3 節では、修正グラム・シュミット法を用いた 3 回の QR 分解を基礎とした悪条件線形方程式の直接法による解法を提示する。4.4 節で、その具体的なアルゴリズムを、4.5 節で、その有効性を数値実験により実証する。

4.1 打ち切り最小 2 乗最小ノルム解

線形方程式 $Ax = b$ の係数行列 A が次の形に分解されたとする。

$$A = URDV^T, \quad D \stackrel{d}{=} \text{diag}(d_1, d_2, \dots), \quad d_1 \geq d_2 \geq \dots > 0. \quad (4.1)$$

ここで U, V は正規直交行列、 R は悪条件でない上三角行列、 D は対角行列である。このとき、 $Ax = b$ の打ち切り最小 2 乗最小ノルム解を次のように定義する。

定義 4.1.1 [18] $\varepsilon > 0$ を与え, $\mathbf{c} \stackrel{d}{=} U^T \mathbf{b} = (c_1, c_2, \dots)^T$ とし, $\sum_{i>n} c_i^2 < \varepsilon^2$ を満たす最小の $n = n(\varepsilon)$ に対し,

$$A_n \stackrel{d}{=} URD_n V^T, \quad D_n \stackrel{d}{=} \text{diag}(d_1, \dots, d_n, 0, \dots)$$

を係数とする方程式 $A_n \mathbf{x} = \mathbf{b}$ の最小 2 乗最小ノルム解を $A \mathbf{x} = \mathbf{b}$ の打ち切り最小 2 乗最小ノルム解という.

定義からわかるように, 打ち切り最小 2 乗最小ノルム解は分解 (4.1), \mathbf{b} と ε に依存する. また, ε に応じて n を切ったとき, R が上三角行列であることから, 解空間の部分空間の次元 n が決まる. その部分空間上での最小 2 乗最小ノルム解が逆代入と D_n^{-1} による変換だけで求められる.

打ち切り最小 2 乗最小ノルム解は, 制限された解空間の中では最良の解である. もし, ε として計算機イブシロンをとれば, その解は与えられた演算桁数の下では, ほぼ限界の精度に達していることになる.

4.2 特異値分解による打ち切り最小 2 乗最小ノルム解

行列 A の分解 (4.1) において, 最も一般的であるのは行列の特異値分解 (1.10) である. この節では特異値分解を用いての悪条件線形方程式の最小 2 乗最小ノルム解の解法を述べる. 先ず, 残差ベクトルのノルムを評価する. U, V が正規直交行列であることから,

$$\begin{aligned} \|A \mathbf{x} - \mathbf{b}\| &= \|U \Sigma V^T \mathbf{x} - \mathbf{b}\| \\ &= \|\Sigma \mathbf{z} - \mathbf{c}\| \end{aligned}$$

となる. ただし,

$$\mathbf{z} \stackrel{d}{=} V^T \mathbf{x} = (z_1, z_2, \dots)^T, \quad \mathbf{c} \stackrel{d}{=} U^T \mathbf{b} = (b_1, b_2, \dots)^T$$

である. したがって, 方程式 $A \mathbf{x} = \mathbf{b}$ は一般性を失うことなく単独の方程式を単に列記したものとなる.

$$\sigma_i z_i = c_i, \quad i = 1, 2, \dots$$

次に, 特異値分解による打ち切り最小 2 乗最小ノルム解について述べる. 行列 A の特異値分解 (1.10) が得られたとする. このとき, 特異値分解による打ち切り最小 2 乗最小ノルム解 $\mathbf{x}_n^{(\text{SVD})}$ は, 定義 4.1.1 にしたがって, $R = I, D = \Sigma$, 残差の許容誤差限界 $\varepsilon = \varepsilon_b$ を与えることにより, \mathbf{c} を $n = n(\varepsilon_b)$ で打ち切り,

$$\mathbf{x}_n^{(\text{SVD})} = \sum_{i=1}^n \frac{c_i}{\sigma_i} \mathbf{v}_i \quad (4.2)$$

として求めることができる. 近似解 $\mathbf{x}_n^{(\text{SVD})}$ の誤差は

$$E_{(\text{SVD})}(n) \stackrel{d}{=} \sum_{i>n} \left(\frac{c_i^2}{\sigma_i^2} \right)^{\frac{1}{2}} \quad (4.3)$$

で表される. この級数の収束が速ければ, 事実上 (4.2) の右辺のフーリエ係数の最後の 2 項の絶対値の和

$$\tilde{E}_{(\text{SVD})}(n) \stackrel{d}{=} \left| \frac{c_{n-1}}{\sigma_{n-1}} \right| + \left| \frac{c_n}{\sigma_n} \right| \quad (4.4)$$

により誤差を推定することができる. 級数 (4.3) の収束が遅ければ, 悪条件線形方程式の数値解法はいつそう困難となる.

数値計算では, 計算機イブシロン ε_μ に対して, ε_μ 以下となる特異値は 0 に置き換える. このとき, もし $\sigma_{m+1} < \varepsilon_\mu$ なる最小の $m = m(\varepsilon_\mu)$ について $\sum_{i>m} c_i^2 > \varepsilon_\mu^2$ ならば, ノルムが $1/\varepsilon_\mu$ 以下の数値解は存在しない. この意味で l_2 には解は存在しない.

悪条件線形方程式に対する、特異値分解による打ち切り最小2乗最小ノルム解の有効性を数値実験で検証する。まず、第1種フレドホルム積分方程式(3.22)を50点ガウス・ルジャンドル数値積分を用いて離散化して解く。このとき得られた特異値分解による打ち切り最小2乗最小ノルム解の計算結果を図15に示す。数値計算はすべて倍精度演算、すなわち $\epsilon_\mu = 1.0 \times 10^{-15}$ で行なった。特異値分解はライブラリ[38]のDASVD1を用いた。

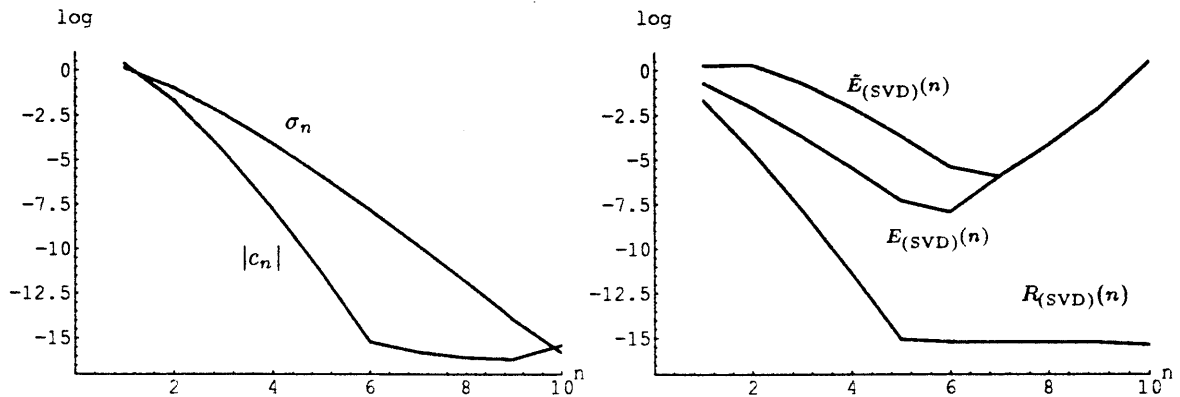


図15: 方程式(3.22)に対する特異値分解による打ち切り最小2乗最小ノルム解の数値結果.

図15の左図は離散化した行列の特異値と左特異ベクトル $\{u_i\}$ によるデータベクトル b のフーリエ係数の絶対値を常用対数の尺度でプロットしたものである。どちらも丸め誤差レベルまでは順調に小さくなる様子が観測できる。これから、離散化した行列の実質的な次元は10程度であるといえる。

右図は(4.2)の n を順次増やしたときの誤差 $E_{(SVD)}(n)$ 、推定誤差 $\hat{E}_{(SVD)}(n)$ と残差 $R_{(SVD)}(n)$ の振る舞いを示したものである。ここで $E_{(SVD)}(n)$ と $R_{(SVD)}(n)$ は、それぞれ

$$E_{(SVD)}(n) \stackrel{d}{=} \|\mathbf{x}_o - \mathbf{x}_n^{(SVD)}\|$$

$$R_{(SVD)}(n) \stackrel{d}{=} \left(\sum_{i=n+1}^m c_i^2 \right)^{\frac{1}{2}}$$

を表す。ただし、 \mathbf{x}_o は解 $f(t) = e^t$ を離散化したベクトル、 m は計算機イプシロン ϵ_μ 以上の特異値の個数である。有効桁数7桁程度の近似解が得られていることがわかる。

次に、積分方程式(3.24)に対する同様の結果を図16に示す。

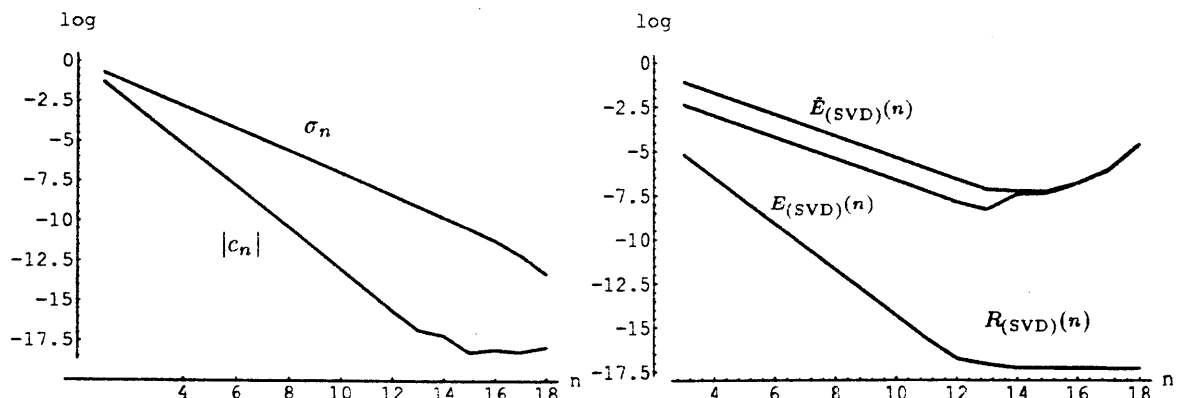


図16: 方程式(3.24)に対する特異値分解による打ち切り最小2乗最小ノルム解の数値結果.

この例においても有効桁数7桁程度の近似解が得られている。いずれも特異値分解による打ち切り最小2乗最小ノルム解が、悪条件線形方程式に対して有効であることを示している。

4.3 QR 分解による打ち切り最小2乗最小ノルム解

前節で述べたように、特異値分解による打ち切り最小2乗最小ノルム解は、悪条件線形方程式に対して有効である。しかし、特異値分解は直接法ではないので計算量が多いという欠点がある。また、無駄な計算も伴っている。方程式(3.22)を離散化した行列の実質的な階数は10程度であるが、この情報は特異値分解を行なう前には不明である。つまり、この例では10個の特異値とそれに対応した特異ベクトル(u_i と v_i)を得るために50個の特異値と特異ベクトルを求めている。それ以外の40個の特異値と特異ベクトルは全て無駄となった。

以下、ピボット付修正グラム・シュミット法を用いたQR分解による悪条件線形方程式の直接法による解法を提案する。特異値分解との対応を明確にするために、本節においても U, V, c 等の記号を用いるが、意味は異なることを注意しておく。

行列 A の特異値はすべて正であるとする。 A の列ベクトルにピボット付修正グラム・シュミット法を適用することにより

$$AP = Q_1DS, \quad D \stackrel{d}{=} \text{diag}(d_1, d_2, \dots) \quad (4.5)$$

の分解が得られたとする。ただし、 Q_1 は列ベクトルが正規直交系の行列、 $S = (s_{ij})$ は上三角行列、そして、ピボット付修正の効果により

$$\begin{aligned} d_1 &\geq d_2 \geq \dots > 0, \\ s_{ij} &= 1, \\ |s_{ij}| &\leq 1, \quad j > i \end{aligned} \quad (4.6)$$

が成り立つ。以後、簡単のため A の列ベクトルのピボット付修正は事前に行なわれているものとして P を省略する。

ここで、 S の条件数を考察する。 S が有限次元行列のとき、その評価が[26]にある。しかしながら、 S が無次元行列のとき、その評価は発散してしまう。ここでは、典型的な場合として、 $s_{ij}, j > i$ が区間 $[-1, 1]$ 上の一様乱数の場合を考察する。このとき、 $S^T S$ の非対角要素は、 s_{ij} の統計的独立性から、対角要素に比べて無視できる程度となり、 S の条件数は有限である。したがって、(4.5)の分解において A の悪条件性(1.12)は、対角行列 D だけに反映されると考えてよい。

次に、 S^T をQR分解すると、 $S^T = Q_2L^T$ 、すなわち

$$A = Q_1MDQ_2^T, \quad M \stackrel{d}{=} DLD^{-1}$$

が得られる。 Q_2 は列ベクトルが正規直交系の行列、 L は下三角行列である。 L の条件数は、 Q_2 の正規直交性から、 S の条件数と同じである。さらに、 D で L を相似変換すると、

$$A = Q_1MDQ_2^T, \quad M \stackrel{d}{=} DLD^{-1} \quad (4.7)$$

となり、 D の対角要素の単調性(4.6)より下三角行列 M の対角成分は非対角成分に対して相対的に大きくなる。

対角行列 D による相似変換が L の対角要素をどれほど強調しているかを示すために次の指標を定義する。

$$d_L(k) \stackrel{d}{=} \frac{1}{m-k} \sum_{j=1}^{m-k} |l_{k+j,j}|, \quad L = (l_{ij}),$$

$$d_M(k) \stackrel{d}{=} \frac{1}{m-k} \sum_{j=1}^{m-k} |m_{k+j,j}|, \quad M = (m_{ij}).$$

ここで m は行列 L, M の次数である. 積分方程式 (3.22) と (3.24) の行列に関する上式の指標を求めたのが 図 17 の左図と右図である.

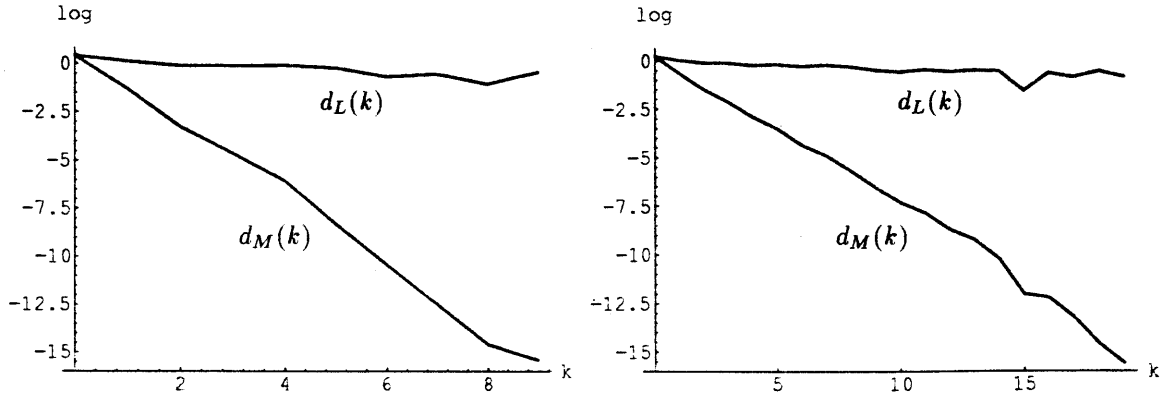


図 17: 行列 L と M の対角優位性の比較.

M が L に比べて対角要素が優位である様子がわかる.

さらに, M に QR 分解を施すことにより, $M = Q_3 R$ が得られる. ここで, Q_3 は列ベクトルが正規直交系の行列, R は正則な上三角行列である. 以上, 3 回の QR 分解を用いて, A の分解

$$A = Q_1 Q_3 R D Q_2^T$$

が得られる. ここで, $U \stackrel{d}{=} Q_1 Q_3$, $V \stackrel{d}{=} Q_2$ とおくことにより,

$$A = U R D V^T \tag{4.8}$$

となる.

本題に立ち返って, $Ax = b$ を解くことを考える. 行列 A の分解 (4.8) が得られれば, 悪条件線形方程式 $Ax = b$ の打ち切り最小 2 乗最小ノルム解 $x_n^{(QR)}$ は, 定義 4.1.1 にしたがって, 残差の許容誤差限界 $\epsilon = \epsilon_\mu$ を与えることにより容易に求めることができる.

4.4 アルゴリズム

前節で述べた分解 (4.8) を用いて, 悪条件線形方程式 $Ax = b$ の解法について述べる. 先ず, A, b , 計算機イプシロン $\epsilon_\mu > 0$ と残差の許容誤差限界 $\epsilon_b \geq \epsilon_\mu$ を与える.

[Step 1] A の列ベクトルに対し, ピボット列のノルムが ϵ_μ 以下になるまでピボットイング付修正グラム・シュミット法を施すことにより, A の QR 分解

$$AP = Q_1 D S + O(\epsilon_\mu)$$

を求める. ただし, P はピボットイング行列である. $m = m(\epsilon_\mu)$ 列でピボット列のノルムが ϵ_μ 以下になったとすると, Q_1 は列数が m の縦長の列正規直交行列, S は行数が m の横長の台形型行列, D は m 次対角行列となる.

[Step 2] ベクトル $\Delta b = b - Q_1 Q_1^T b$ を計算する. このとき, $\|\Delta b\|_2 > \varepsilon_\mu$ ならば元の方程式は l_2 に解を持たない. 計算停止.

[Step 3] S^T の QR 分解を行なう. すなわち, $S = LQ_2^T$ を求める. ここで, L は m 次の正則な下三角行列である.

[Step 4] D による L の相似変換 $M = DLD^{-1}$ を計算し, 引き続き M の QR 分解 $M = Q_3 R$ を求める. ここで, R は m 次の正則な上三角行列である.

[Step 5] ベクトル $c = Q_3^T Q_1^T b = (c_1, \dots, c_m)^T$ を求め, 残差の許容誤差限界 ε_b に基づき,

$$\sum_{i>n}^m c_i^2 < \varepsilon_b^2$$

を満たす最小の $n = n(\varepsilon_b)$ を決定する. 上式を満たす $n < m$ が存在しないときは $n = m$ とする.

[Step 6] $c_n = (c_1, \dots, c_n, 0, \dots, 0)^T \in \mathbb{R}^m$ とし, $Ry_n = c_n$ を逆代入で解く.

[Step 7] $D_n Q_2^T x = y_n = (y_1, \dots, y_n, 0, \dots, 0)^T$ を x について解く.

$$\xi_i = \begin{cases} \frac{y_i}{d_i}, & i = 1, 2, \dots, n \\ 0, & i > n \end{cases}$$

$$x = Q_2(\xi_1, \xi_2, \dots, \xi_n, 0, \dots)^T$$

最後に x にピボット行行列 P の逆変換を行えば近似解 $x_n^{(QR)}$ が得られる. このときの残差の2乗の上界は $\|\Delta b\|^2 + \varepsilon_b^2$ である. 解の誤差は

$$\hat{E}_{(QR)}(n) \triangleq \left| \frac{c_{n-1}}{d_{n-1}} \right| + \left| \frac{c_n}{d_n} \right| \quad (4.9)$$

で推定する.

計算法としては, A の最終列に b を加えた行列に修正グラム・シュミット法を施すと効率が良い. もちろん, b の列はピボット行の対象から外さなければならない. これにより, [Step 2] の $Q_1 b$ と Δb が同時に計算できる.

4.5 数値実験およびその考察

第1種フレドホルム積分方程式

$$\int_a^b K(s, t) f(t) dt = g(s), \quad s_{\min} \leq s \leq s_{\max}$$

の離散化法を述べる. 一般性を失うことなく, 区間 $[a, b]$ 及び $[s_{\min}, s_{\max}]$ は区間 $[-1, 1]$ にとる.

区間 $[-1, 1]$ 上の N 点ガウス・ルジャンドル数値積分則の分点を $\{x_1, \dots, x_N\}$, 重みを $\{w_1, \dots, w_N\}$ とする. 今,

$$K \stackrel{d}{=} (K(x_i, x_j)) \in \mathbb{R}^{N \times N}$$

$$W \stackrel{d}{=} \text{diag}(\sqrt{w_1}, \dots, \sqrt{w_N}) \in \mathbb{R}^{N \times N}$$

$$g \stackrel{d}{=} (g(x_1), \dots, g(x_N))^T \in \mathbb{R}^N$$

を定義すると、未知関数 f に対する積分方程式の残差ノルムのは、 \mathbf{R}^N における 2 乗ノルム $\|\cdot\|$ を用いて

$$\left\| \int_{-1}^1 K(s,t)f(t)dt - g(s) \right\|_{L_2} \approx \left\| W^{\frac{1}{2}} \hat{K} W^{\frac{1}{2}} f - g \right\|, \quad f \stackrel{d}{=} (f(x_1), \dots, f(x_N))^T$$

で近似できる。ここで、数値積分による離散化誤差は、計算機イプシロン ε_μ に比べ、小さくなるようにあらかじめ分点数 N を十分大きくとるものとする。したがって、積分方程式の最小 2 乗最小ノルム問題は、線形方程式

$$Ax = b, \quad A \stackrel{d}{=} W^{\frac{1}{2}} \hat{K} W^{\frac{1}{2}}, \quad b \stackrel{d}{=} W^{\frac{1}{2}} g$$

の最小 2 乗最小ノルム問題に帰着する。このときの解は $x = W^{\frac{1}{2}} f$ である。

方程式 (3.22) に対する計算結果を図 18 に示す。

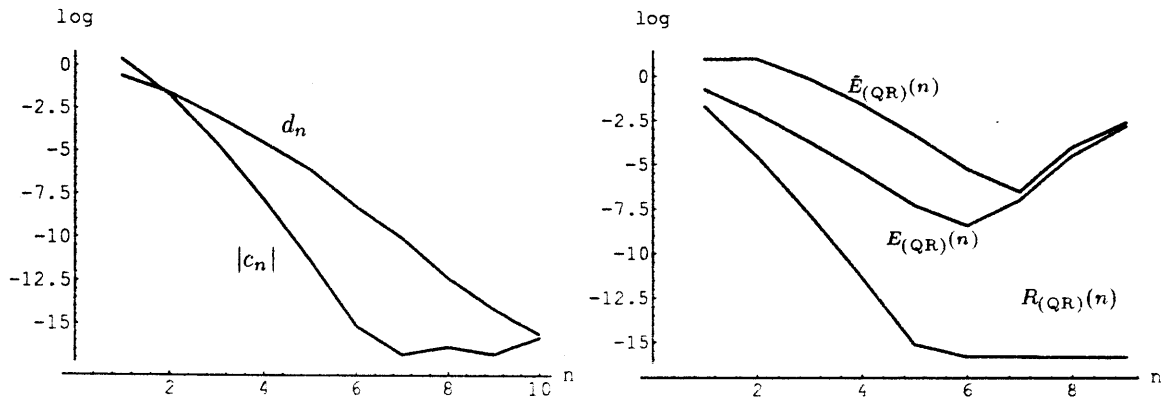


図 18: 方程式 (3.22) に対する分解 (4.8) による打ち切り最小 2 乗最小ノルム解の数値結果。

左図の d_n , c_n は、それぞれ

$$D = \text{diag}(d_1, \dots, d_m) \quad A = Q_1 Q_3 R D Q_3^T, \\ c = Q_3^T Q_1^T b = (c_1, \dots, c_m)^T$$

であって、それらの振る舞いは図 15 の左図に酷似している。これは A の悪条件性が対角行列 D によく反映されていることを意味している。

右図の図中の $E_{(QR)}(n)$, $\hat{E}_{(QR)}(n)$, $R_{(QR)}(n)$ は、それぞれ誤差、推定誤差 (4.9), 残差ノルムを表す。ここでの誤差、残差ノルムは

$$E_{(QR)}(n) \stackrel{d}{=} \|x_o - x_n^{(QR)}\| \quad (4.10)$$

$$R_{(QR)} \stackrel{d}{=} \left(\sum_{i>n} c_i^2 \right)^{\frac{1}{2}}, \quad (4.11)$$

である。今、 $\varepsilon_b = 1.0 \times 10^{-14}$ とすることにより、有効桁数 7 桁程度の近似解が得られることがわかる。これは特異値分解による打ち切り最小 2 乗最小ノルム解と同程度の精度である。

方程式 (3.24) に対する計算結果を図 19 に示す。

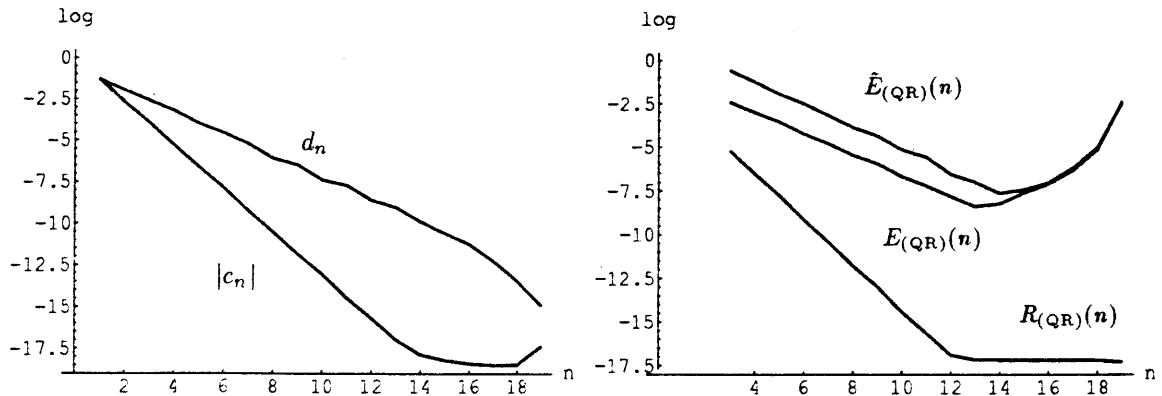


図 19: 方程式 (3.24) に対する分解 (4.8) による打ち切り最小 2 乗最小ノルム解の数値結果.

この場合も特異値分解による打ち切り最小 2 乗最小ノルム解と同程度の精度の計算結果を得た.

4.6 まとめ

打ち切り最小 2 乗最小ノルム解は、解空間を制限し、その空間内での最良の近似解である。本方法を用いるときの問題は基底ベクトル系の選択にある。基底ベクトル系として A の特異ベクトル系を用いることは合理的である。事実、数値実験においても、特異値分解による打ち切り最小 2 乗最小ノルム解は、与えられた演算桁数の下で、ほぼ限界の精度が得られた。しかし、特異値分解は固有値問題として取り扱われるため、計算量が多いという欠点を持つ。

これに対して我々は、直接法による打ち切り最小 2 乗最小ノルム解を提示した。この直接法はピボティング付を含む 3 回の QR 分解からなる。特異値分解による方法と比較し、本方法は同程度の近似解が得られることが数値実験より検証された。

おわりに

第1種フレドホルム積分方程式の数値解法において、典型的な悪条件線形方程式 $Ax = b$ が現れる。我々は、 A の条件数 $\|A\| \|A^\dagger\| = \infty$ の線形方程式を取り扱った。

Tikhonov の正則化法における正則化パラメータ選択問題について、既存の方法の問題点を指摘し、L-カーブの曲率による最適正則化法を提案した。この手法の有効性を数値実験により実証した。

また、データに AIC 最小化推定法を適用することによりノイズレベルを推定し、それをもとに相変原理を用いて正則化パラメータを選定する MAICE-DP 法についてその有効性を考察した。MAICE-DP 法は相変原理の意味においての最適の正則化パラメータの選定には成功しているといえる。しかし、相変原理がうまくいかない例においては、やはり良好な結果を得ることができなかった。これに対して、L-カーブの曲率による最適正則化法はこのような例においても高精度の近似解が得られる。これより、悪条件線形方程式の正則化パラメータ選定法には L-カーブの曲率による最適正則化法を用いるべきであるというのが我々の結論である。

また、我々は悪条件線形方程式に対して打ち切り最小2乗最小ノルム解を定義し、それを求める直接法を提示した。この直接法の主要な部分は A に対する3回の QR 分解からなる。特異値分解による方法と比較し、本方法は精度においても遜色がない。与えられた演算桁数の下で、ほぼ限界の精度が得られることを示した。

謝辞

本研究の遂行にあたり、多大なる御指導を賜りました鳥居達生名古屋大学教授に深甚なる感謝の意を表明します。

本論文の作成にあたり、有益な御助言をいただいた細江繁幸名古屋大学教授、杉浦洋名古屋大学助教授に厚くお礼申し上げます。

また、正則化法について御指導いただいた北川高嗣筑波大学助教授に深く感謝します。

参考文献

- [1] Akaike, H., *A new look at the statistical model identification*, IEEE Transactions on Automatic Control, AC-19, 6(1974), 716-730.
- [2] 赤池弘次, 情報量規準 AIC とは何か, 数理科学, 153(1976), 5-11.
- [3] 赤池弘次, モデルによってデータを測る, 数理科学, 213(1981), 7-10.
- [4] Allen, D. M., *The relationship between variable selection and data augmentation and a method for prediction*, Technometrics, 16(1974), 125-127.
- [5] Baker, C. T. H., Fox, L., Mayers, D. F. and Wright, K., *Numerical solution of Fredholm Integral Equations of first kind*. Comput. J., 7(1964), 141-148.
- [6] Cochran, J. A., "The Analysis of Linear Integral Equations", McGraw-Hill, New York, 1972.
- [7] Engl, H. W., *Necessary and sufficient conditions for convergence of regularization methods for solving linear operator equations of the first kind*, Numer. Func. Anal. & Optimiz., 3(1981), 201-222.
- [8] Groetsch, C. W., "Generalized Inverses of linear Operators", Marcel Dekker, New York and Basel, 1977.
- [9] Groetsch, C. W., *On a class of regularization methods*, Boll. Un. Mat. Ital., Ser. 17-b(1980), 1411-1419.
- [10] Groetsch, C. W., "The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind", Pitman, Boston, London and Melbourne, 1984.
- [11] Hansen, P. C., *Solution of ill-posed problems by means of truncated SVD*, in Numerical Mathematics, Singapore 1988, R. P. Agarwal, Y. M. Chow, and S. J. Wilson, Eds., ISNM 86, Birkhäuser, Basel, Switzerland, 1988, 179-192.
- [12] Hansen, P. C., *Computation of the singular value expansion*, Computing, 40(1988), 185-199.
- [13] Hansen, P. C., *Truncated SVD solutions to discrete ill-posed problems with ill-determined numerical rank*, SIAM J. Sci. Statist. Comput., 11(1990), 503-518.
- [14] Hansen, P. C., *The discrete Picard condition for discrete ill-posed problems*, BIT, 30(1990), 658-672.
- [15] Hansen, P. C., *Analysis of discrete ill-posed problems by means of L-curve*, SIAM Review, 34(1992), 561-580.

- [16] 細田陽介, 北川高嗣, L-カーブによる不適切問題の最適正則化について, 日本応用数学会論文誌, 2(1992), 55-68.
- [17] 細田陽介, 北川高嗣, 不適切問題に対する MAICE-DP 法による最適正則化について, 日本応用数学会論文誌, 3(1993), 47-59.
- [18] 細田陽介, 鳥居達生, 悪条件線形方程式に対する一つの直接解法 — 打ち切り最小2乗最小ノルム解 —, 日本応用数学会論文誌に投稿中.
- [19] Kitagawa, T. and Hosoda, Y., *New approaches to the optimal regularization*, in IUTAM SYMPOSIUM TOKYO/JAPAN 1993, Inverse Problems in Engineering Mechanics, M. Tanaka and H. D. Bui, Eds., Springer-Verlag, Berlin Heidelberg, 1993, 21-26.
- [20] Kammerer, W. J. and Nashed, M. Z., *Iterative methods for best approximate solutions of linear integral equations of the first and second kinds*, J. Math. Anal. Appl., 40(1972), 547-573.
- [21] Lawson, C. L. and Hanson, R. J., "Solving Least Squares Problems", Prentice-Hall, Englewood Cliffs, NJ, 1974, Chap. 25 and 26.
- [22] 前田周一郎, "関数解析", 森北出版, 東京, 1974.
- [23] Miller, K., *Least squares methods for ill-posed problems with a prescribed bound*, SIAM J. Math. Anal., 1(1970), 52-74.
- [24] Morozov, V. A., *The error principle in the solution of operational equations by the regularization methods*, USSR Comput. Math. and Math. Phys., 8(1968), 63-87.
- [25] 村上温夫, "関数解析", 朝倉書店, 東京, 1980.
- [26] 中川 徹, 小柳義夫, 最小2乗法による実験データ解析, 東京大学出版会, 東京, 1982, 67-72.
- [27] 田辺國士, 統計的データの誤差の処理, 数値計算における誤差 — 一松 信・戸川隼人編, 共立出版, 東京, 1975, 113-125.
- [28] 田中 博, 岡部政之, 鈴木 貴, "岩波講座 応用数学 逆問題", 岩波書店, 東京, 1993.
- [29] Taylor, A. E., "Introduction to Functional Analysis", Wiley, New York, 1958.
- [30] Tikhonov, A. N., *Solution of incorrectly formulated problems and the regularization method*, Soviet Math., 4(1963), 1035-1038.
- [31] Tikhonov, A. N., *On the problems with approximately specified information*, in Ill-Posed Problems in the Natural Sciences, A. N. Tikhonov and A. V. Goncharsky, Eds., MIR, Moscow, 1987, 13-20.
- [32] Varah, J. M., *Pitfalls in the numerical solution of linear ill-posed problems*, SIAM J. Sci. Stat. Comput., 4(1983), 164-176.
- [33] Vinokurov, V. A., *Two notes on the choice of regularization parameter*, USSR Comput. Math. and Math. Phys., 12(No.2) (1972), 249-253.
- [34] Wahba, G., *Practical approximate solutions to linear operator equations when the data are noisy*, SIAM J. Numer. Anal., 14(1977), 651-667.

- [35] 山本哲朗, “数値解析入門”, サイエンス社, 東京, 1976.
- [36] Yosida, K., “Functional Analysis”, Springer-Verlag, Berlin Heidelberg New York, 1974.
- [37] NUMPAC ライブラリ・プログラム利用の手引き, 二宮市三監修, 名古屋大学大型計算機センター.
- [38] 科学用サブルーチンライブラリ SSL II 使用手引書, 富士通株式会社, 1987.

付録 A

使用例の主プログラム

第1種フレドホルム積分方程式

$$\int_0^1 e^{st} f(t) dt = \frac{e^{s+1} - 1}{s+1}, \quad 0 \leq s \leq 1$$

をガウス・ルジャンドル数値積分則を用いて離散化し、修正グラム・シュミット法によるQR分解を3回行なうことにより打ち切り最小2乗最小ノルム解を求めるプログラムを以下に示す。なお、上の積分方程式の厳密解は $f(t) = e^t$ で、サンプルプログラムではガウス・ルジャンドル数値積分則は30点を用いた。

サンプルプログラム

```
IMPLICIT REAL*8 (A-H,O-Z)
PARAMETER (NN=30)
DIMENSION A(NN,NN),R(NN,NN),U(NN,NN),V(NN,NN),WK(NN,NN+1)
DIMENSION PT(NN),WT(NN),DATA(NN),D(NN),NP(NN),APF(NN),WW(NN)
EXTERNAL FKER,GD
C
C 数値積分則の分点を求めるための Newton 法の収束判定定数と、QR 分解の
C  ピボットイングのしきい値は EPS = 1.0D-15 を用いる。
C
EPS=1.0D-15
DA=0.0D0
DB=1.0D0
C
N=NN
C
C [-1,1] 区間のガウス・ルジャンドル数値積分則の分点と重率を求める。
C
CALL TGLEGD(N,PT,WT,EPS,ICON)
IF(ICON.NE.0) WRITE(*,1001) ICON
1001 FORMAT('ERROR OCCURED IN TGLEGD ICON=',I5)
C
C [-1,1] 区間のガウス・ルジャンドル数値積分則の分点と重率を
```

```

C      [DA,DB] 区間に変換する.
C
      DO 10 I=1,N
          PT(I)=((DB-DA)*PT(I)+DA+DB)*0.5DO
          WT(I)=(DB-DA)*0.5DO*WT(I)
10     CONTINUE
C
C      積分作用素とデータ関数を離散化する.
C
      CALL DCKLD1(A,NN,N,N,DATA,PT,WT,PT,WT,FKER,GD)
C
C      積分作用素を離散化して得られた行列 A を分解する.
C      同時にデータベクトルも展開する.
C
      CALL QR3DCP(A,NN,N,N,U,V,R,D,DATA,NP,EPS,RS,NK,WK,ICON)
      IF(ICON.NE.0) WRITE(*,1002) ICON
1002  FORMAT('ERROR OCCURED IN QR3DCP ICON=',I5)
C
C      データベクトルのフーリエ係数を打ち切るための定数を
C      EPX = 1.0D-13 と定める.
C
      EPX=1.0D-13
      CALL SOLVE1(R,NN,V,D,N,NK,DATA,APF,NP,K,EST,EPX,WW)
C
C      サブルーチン SOLVE1 により求められた近似解 APF を,
C      数値積分則の重み WT で割ることにより標本点上での近似解の
C      値が求まる.
C
      DO 20 I=1,N
          APF(I)=APF(I)/SQRT(WT(I))
20     CONTINUE
C
      ERR=0.0DO
      DO 30 I=1,N
          ERR=ERR+(EXF(PT(I))-APF(I))*(EXF(PT(I))-APF(I))
30     CONTINUE
      ERR=SQRT(ERR)
      RX=0.0DO
      DO 40 I=K+1,NK
          RX=RX+DATA(I)*DATA(I)
40     CONTINUE
      RX=SQRT(RX)
C
      WRITE(*,1003)

```

```

1003 FORMAT(10X,' 標本点',16X,' 近似解',14X,' 厳密解',/)
      WRITE(*,1004) (PT(I),APF(I),EXF(PT(I)),I=N,1,-1)
      WRITE(*,1005) NK,K,RS,RX,ERR,EST
1004 FORMAT(F20.15,2F20.10)
1005 FORMAT(/,'A の数値的階数',10X,I3,
&          /,' 解近似に用いた基底の個数',I3,
&          /,'A の値域と B との距離',D10.3,
&          /,' 近似解の残差ノルム',3X,D10.3,
&          /,' 誤差',17X,D10.3,
&          /,' 推定誤差',13X,D10.3)
C
      STOP
      END
C
      FUNCTION FKER(S,T)
      IMPLICIT REAL*8 (A-H,O-Z)
      FKER=EXP(S*T)
      RETURN
      END
C
      FUNCTION GD(S)
      IMPLICIT REAL*8 (A-H,O-Z)
      GD=(EXP(S+1.0D0)-1.0D0)/(S+1.0D0)
      RETURN
      END
C
      FUNCTION EXF(T)
      IMPLICIT REAL*8 (A-H,O-Z)
      EXF=EXP(T)
      RETURN
      END

```

実行結果

標本点	近似解	厳密解
0.001553257962675	1.0015546265	1.0015544649
0.008165938360126	1.0081995037	1.0081993706
0.019989067515846	1.0201902748	1.0201901868
0.036899976285363	1.0375892675	1.0375892321
0.058719732103974	1.0604779667	1.0604779812
0.085217118808616	1.0889534216	1.0889534733
0.116111283947587	1.1231207805	1.1231208505
0.151074752603342	1.1630835311	1.1630835984
0.189736908505379	1.2089314501	1.2089314962
0.231687925928990	1.2607262139	1.2607262275
0.276483115230955	1.3184847113	1.3184846902
0.323647637234561	1.3821602475	1.3821601995
0.372681536916055	1.4516220380	1.4516219782
0.423065043195708	1.5266336435	1.5266335899
0.474264078722341	1.6068312920	1.6068312610
0.525735921277659	1.6917033498	1.6917033509
0.576934956804292	1.7805724935	1.7805725267
0.627318463083945	1.8725823862	1.8725824418
0.676352362765439	1.9666907971	1.9666908578
0.723516884769045	2.0616710893	2.0616711355
0.768312074071010	2.1561237879	2.1561238033
0.810263091494621	2.2484994929	2.2484994700
0.848925247396658	2.3371337180	2.3371336612
0.883888716052413	2.4202933380	2.4202932634
0.914782861191384	2.4962332821	2.4962332139
0.941280267896026	2.5632610151	2.5632609793
0.963100023714637	2.6198053397	2.6198053568
0.980010932484154	2.6644852922	2.6644853712
0.991834061639874	2.6961747561	2.6961748914
0.998446742037325	2.7140627410	2.7140629129
A の数値的階数 9		
解近似に用いた基底の個数 5		
A の値域と B との距離 0.312D-15		
近似解の残差ノルム 0.841D-15		
誤差 0.397D-06		
推定誤差 0.411D-03		

数値的階数は1回目のピボットリング付き修正グラム・シュミット法が打ち切られた回数を指す。行列 A の実質的な値域の次元は9程度であることを示す。解近似に用いた基底の個数は、絶対値がEPS 以上であるデータベクトルのフーリエ係数の個数を表す。この例では近似解の構成に5個の基底ベクトルが用いられた。A の値域と B との距離は上により求められた A の実質的な値域とデータベクトル b との距離(2-ノルム)を表す。もし、この値がピボットリングのしきい値EPS より大きければ方程式 $Ax = b$ に解はない。

近似解の残差ノルム、誤差はそれぞれ式(4.10)、(4.11)を指す。推定誤差は式(4.9)により定められた量である。

付録 B

打ち切り最小 2 乗最小ノルム解のサブ ルーチン集

```
C
C *****
C 修正グラム・シュミット法による QR 分解を 3 回行なうことにより
C 行列 A を
C
C           T
C   A = U R D V
C
C の形式に展開する.
C
C --引数の説明--
C   CALL QR3DCP(A,LA,M,N,U,V,R,D,B,NP,EPS,NK,WK,ICON)
C
C   A ... 入力 実数型 2 次元配列.
C           分解する行列 A. 内容は保存されない.
C
C   LA ... 入力 整数型.
C           A の配列宣言における第 1 添字の値. ( LA >= M ).
C
C   M ... 入力 整数型.
C           A の行数.
C
C   N ... 入力 整数型.
C           A の列数.
C
C   U ... 出力 実数型 2 次元配列.
C           データ空間の正規直交基底が列ベクトルに入る.
C           配列宣言の第 1 添字の値は LA でなければならない.
C
C   V ... 出力 実数型 2 次元配列.
```

```

C          解空間の正規直交基底が行ベクトルに入る。
C          配列宣言の第1添字の値は LA でなければならない。
C
C          R ... 出力 実数型 2次元配列。
C          上三角行列 R。
C          配列宣言の第1添字の値は LA でなければならない。
C
C          D ... 出力 実数型 1次元配列, 大きさ N。
C          対角行列 D の対角成分。
C
C          B ... 入出力 実数型 1次元配列, 大きさ M。
C          データベクトル B を入力, U の列ベクトルによる
C          フーリエ係数が出力される。
C
C          NP ... 出力 整数型 1次元配列, 大きさ N。
C          ピボットイングによる置換行列。
C
C          EPS .. 入力 実数型。
C          ピボットイングのしきい値。
C
C          RS ... 出力 実数型。
C          行列 A の値域とベクトル B との距離 (2-ノルム)。
C
C          NK ... 出力 整数型。
C          A の数値的階数。
C
C          WK ... 作業領域 実数型 2次元配列。
C          列数は N+1。
C          配列宣言の第1添字の値は LA でなければならない。
C
C          ICON 出力 整数型。
C          ICON = 0 : 正常終了。
C          ICON = K : 2 回目の QR 分解において K 列目の直交化で
C                   ノルムが EPS 以下となった。
C                   処理は打ち切る。
C          ICON = -K : 3 回目の QR 分解において K 列目の直交化で
C                   ノルムが EPS 以下となった。
C                   処理は打ち切る。
C
C          *****
C          SUBROUTINE QR3DCP(A, LA, M, N, U, V, R, D, B, NP, EPS, RS, NK, WK, ICON)
C          IMPLICIT REAL*8 (A-H, O-Z)
C          DIMENSION A(LA, N), U(LA, N), V(LA, N), R(LA, N),
C          &          D(N), B(M), NP(N), WK(LA, N+1)

```



```

      ICON=0
C
      DO 10 J=1,N
        NP(J)=J
10     CONTINUE
C
      DO 110 K=1,N
        QMAX=0.0DO
        JMAX=K
        DO 30 J=K,N
          JP=NP(J)
          TMP=0.0DO
          DO 20 I=1,M
            TMP=TMP+A(I,JP)*A(I,JP)
20         CONTINUE
          IF(TMP.GT.QMAX) THEN
            QMAX=TMP
            JMAX=J
          ENDIF
30        CONTINUE
        QMAX=SQRT(QMAX)
        IF(K.EQ.1) ANRM=QMAX
        D(K)=QMAX
        IF(QMAX/ANRM.LE.EPS) THEN
          NK=K-1
          GOTO 120
        ENDIF
        IF(JMAX.NE.K) THEN
          JP=NP(JMAX)
          NP(JMAX)=NP(K)
          NP(K)=JP
          DO 40 I=1,K-1
            TMP=V(I,JMAX)
            V(I,JMAX)=V(I,K)
            V(I,K)=TMP
40         CONTINUE
          ENDIF
          V(K,K)=1.DO
          KP=NP(K)
          DO 50 I=1,M
            A(I,KP)=A(I,KP)/QMAX
50         CONTINUE
          DO 80 J=K+1,N
            JP=NP(J)

```

```

        TMP=0.0DO
        DO 60 I=1,M
            TMP=TMP+A(I,JP)*A(I,KP)
60     CONTINUE
        V(K,J)=TMP/QMAX
        DO 70 I=1,M
            A(I,JP)=A(I,JP)-TMP*A(I,KP)
70     CONTINUE
80     CONTINUE
        TMP=0.0DO
        DO 90 I=1,M
            TMP=TMP+B(I)*A(I,KP)
90     CONTINUE
        WK(K,1)=TMP
        DO 100 I=1,M
            B(I)=B(I)-TMP*A(I,KP)
100    CONTINUE
110    CONTINUE
C
        NK=N
120    CONTINUE
C
        RS=0.0DO
        DO 130 I=1,M
            RS=RS+B(I)*B(I)
130    CONTINUE
        RS=SQRT(RS)
C
        DO 190 K=1,NK
            KP=K+1
            QMAX=0.0DO
            DO 140 J=1,N
                QMAX=QMAX+V(K,J)*V(K,J)
140    CONTINUE
            QMAX=SQRT(QMAX)
            IF(QMAX.LT.EPS) THEN
                ICON=K
                RETURN
            ENDIF
            WK(K,KP)=QMAX
            DO 150 J=1,N
                V(K,J)=V(K,J)/QMAX
150    CONTINUE
            DO 180 I=KP,NK

```

```

        TMP=0.000
        DO 160 J=1,N
            TMP=TMP+V(K,J)*V(I,J)
160     CONTINUE
        WK(I,KP)=TMP
        DO 170 J=1,N
            V(I,J)=V(I,J)-TMP*V(K,J)
170     CONTINUE
180     CONTINUE
190 CONTINUE
C
        DO 210 I=1,NK
            TMP=D(I)
            DO 200 J=2,I+1
                WK(I,J)=TMP*WK(I,J)/D(J-1)
200     CONTINUE
210 CONTINUE
C
        DO 290 K=1,NK
            KP=K+1
            QMAX=0.000
            DO 220 I=1,NK
                QMAX=QMAX+WK(I,KP)*WK(I,KP)
220     CONTINUE
            QMAX=SQRT(QMAX)
            IF(QMAX.LE.EPS) THEN
                ICON=-K
                RETURN
            ENDIF
            R(K,K)=QMAX
            DO 230 I=1,NK
                WK(I,KP)=WK(I,KP)/QMAX
230     CONTINUE
            DO 260 J=KP,NK
                JP=J+1
                TMP=0.000
                DO 240 I=1,NK
                    TMP=TMP+WK(I,KP)*WK(I,JP)
240     CONTINUE
                R(K,J)=TMP
                DO 250 I=1,NK
                    WK(I,JP)=WK(I,JP)-TMP*WK(I,KP)
250     CONTINUE
260 CONTINUE

```

```

      TMP=0.0D0
      DO 270 I=1,NK
        TMP=TMP+WK(I,KP)*WK(I,1)
270    CONTINUE
      B(K)=TMP
      DO 280 I=1,NK
        WK(I,1)=WK(I,1)-TMP*WK(I,KP)
280    CONTINUE
290  CONTINUE
C
      DO 320 I=1,M
        DO 310 J=1,NK
          JP=J+1
          TMP=0.0D0
          DO 300 K=1,NK
            TMP=TMP+A(I,NP(K))*WK(K,JP)
300    CONTINUE
          U(I,J)=TMP
310    CONTINUE
320  CONTINUE
C
      RETURN
      END
C
C *****
C サブルーチン QR3DCP で得られた行列 A の分解より,  $Ax = b$  の
C 近似解を求める.
C
C --引数の説明--
C      CALL SOLVE1(R,LR,V,D,N,NK,B,X,NP,K,EST,EPS,WW)
C
C      R ... 入力 実数型 2 次元配列.
C           QR3DCP により得られた上三角行列 R.
C
C      LR ... 入力 整数型.
C            R の配列宣言における第 1 添字の値. ( LR >= N ).
C
C      V ... 入力 実数型 2 次元配列.
C           QR3DCP により得られた正規直交行列 V.
C           配列宣言の第 1 添字の値は LR でなければならない.
C
C      D ... 入力 実数型 1 次元配列. 大きさ N.
C           QR3DCP により得られた対角行列 D の対角成分.
C

```

```

C      N ... 入力 整数型.
C          方程式の解ベクトルの次元.
C
C      NK ... 入力 整数型.
C          QR3DCP により得られた A の数値的階数.
C
C      CC ... 入力 実数型 1 次元配列. 大きさ N.
C          QR3DCP により得られたデータベクトル B の
C          フーリエ係数.
C
C      X ... 出力 実数型 1 次元配列. 大きさ N.
C          Ax = b の近似解.
C
C      NP ... 入力 整数型 1 次元配列. 大きさ N.
C          QR3DCP により得られたピボットイングによる
C          置換行列.
C
C      K ... 出力 整数型.
C          解近似に用いた基底ベクトルの数.
C
C      EST .. 出力 実数型.
C          近似解 X の推定誤差.
C
C      EPX .. 入力 実数型.
C          解近似に用いるフーリエ係数のしきい値.
C          QR3DCP で用いた EPS 以上である必要がある.
C
C      WW ... 作業領域 実数型 1 次元配列. 大きさ N.
C      *****
C      SUBROUTINE SOLVE1(R,LR,V,D,N,NK,CC,X,NP,K,EST,EPX,WW)
C      IMPLICIT REAL*8 (A-H,O-Z)
C      DIMENSION R(LR,N),V(LR,N),D(N),CC(N),X(N),NP(N),WW(N)
C-----
C      K=0
10     K=K+1
        IF((ABS(CC(K)).GT.EPX).AND.(K.LE.NK)) GOTO 10
        K=K-1
C
C      IF(K.EQ.0) THEN
C          DO 20 J=1,N
C              X(J)=0.0D0
20     CONTINUE
        EST=0.0D0
        RETURN

```

```

      ENDIF
C
      X(K)=CC(K)/R(K,K)
      DO 40 I=K-1,1,-1
          TMP=0.0DO
          DO 30 J=K,I+1,-1
              TMP=TMP+R(I,J)*X(J)
30          CONTINUE
          X(I)=(CC(I)-TMP)/R(I,I)
40          CONTINUE
C
      DO 60 I=1,N
          TMP=0.0DO
          DO 50 J=1,K
              TMP=TMP+V(J,I)*X(J)/D(J)
50          CONTINUE
          WW(I)=TMP
60          CONTINUE
C
      DO 70 J=1,N
          X(NP(J))=WW(J)
70          CONTINUE
C
      IF(K.EQ.1) THEN
          EST=ABS(CC(1))/D(1)
      ELSE
          EST=ABS(CC(K-1))/D(K-1)+ABS(CC(K))/D(K)
      ENDIF
C
      RETURN
      END
C
C *****
C 第1種フレドホルム型積分作用素とデータ関数を離散化する.
C
C --引数の説明--
C      CALL DCKLD1(A,LA,M,N,B,SX,SW,TX,TW,FK,G)
C
C      A ... 出力 実数型 2次元配列.
C
C      LA ... 入力 整数型.
C              Aの配列宣言における第1添字の値. ( LA >= M ).
C
C      M ... 入力 整数型.

```

```

C           A の行数.
C
C   N ... 入力  整数型.
C           A の列数.
C
C   B ... 出力  実数型 1 次元配列, 大きさ M.
C           データ関数を離散化したベクトルが出力される.
C
C   SX ... 入力  実数型 1 次元配列, 大きさ M.
C           データ関数の定義域の M 点ガウス・ルジャンドル数値
C           積分則の分点.
C
C   SW ... 入力  実数型 1 次元配列, 大きさ M.
C           データ関数の定義域の M 点ガウス・ルジャンドル数値
C           積分則の重率.
C
C   TX ... 入力  実数型 1 次元配列, 大きさ N.
C           解関数の定義域の N 点ガウス・ルジャンドル数値
C           積分則の分点.
C
C   TW ... 入力  実数型 1 次元配列, 大きさ N.
C           解関数の定義域の N 点ガウス・ルジャンドル数値
C           積分則の重率.
C
C   FK ... 入力  実数型関数副プログラム.
C           第 1 種フレドホルム型積分作用素の積分核.
C           使用者が積分変数だけの 2 変数の関数副プログラム
C           として用意しなければならない.
C
C   G ... 入力  実数型関数副プログラム.
C           積分方程式のデータ関数.
C           使用者が積分変数だけの 1 変数の関数副プログラム
C           として用意しなければならない.
C
C *****
SUBROUTINE DCKLD1(A,LA,M,N,B,SX,SW,TX,TW,FK,G)
IMPLICIT REAL*8 (A-H,O-Z)
DIMENSION A(LA,N),B(N),SX(M),SW(M),TX(N),TW(N)
-----C-----
DO 20 I=1,M
  TMP=SQRT(SW(I))
  DO 10 J=1,N
    A(I,J)=TMP*FK(SX(I),TX(J))*SQRT(TW(J))
10  CONTINUE
  B(I)=TMP*G(SX(I))

```

```

20 CONTINUE
RETURN
END

C
C *****
C [-1,1] 区間のガウス・ルジャンドル数値積分則の分点と重率を与える.
C このサブルーチンは二宮市三による作成で, NUMPAC よりの転載である.
C
C --引数の説明--
C CALL TGLEGD(ND,X,W,EP,ICON)
C
C ND ... 入力 整数型
C 分点数 n. ( 2 <= n <= 50 ).
C
C X ... 出力 実数型 1 次元配列, 大きさ N.
C 分点 xk の値が出力される. ( k=1,2,...,N).
C
C W ... 出力 実数型 1 次元配列, 大きさ N.
C 重率 wk の値が出力される. ( k=1,2,...,N).
C
C EP ... 入力 実数型.
C 分点 xk を求めるときの Newton 法における収束判定定数.
C
C ICON 出力 整数型.
C ICON = 0 : 正常終了.
C ICON = 10000 : EPS が小さすぎるので EPS = 1.0D-15 に
C 引き上げて計算を行なった.
C ICON = 30000 : 入力引数 ND の制限が破られた.
C *****
C SUBROUTINE TGLEGD(ND,X,W,EP ,ICON)
C IMPLICIT REAL*8 (A-H,O-Z)
C DIMENSION X(ND),W(ND)
C LEGENDRE-GAUSS
C IF(ND.LT.1) GO TO 40
C IF(ND.GT.50)GO TO 40
C ICON=0
C EPS=DMAX1(1.0D-15,EP)
C IF(EPS.GT.EP) ICON=10000
C NDM1=ND-1
C N=(ND+1)/2
C FND=ND
C DO 30 K=1,N
C ISW=ISW+1
C XK=DSIN(DFLOAT(ND+1-K-K)/(FND+0.5)*1.570796)

```



```

10 PO=1.0
   P1=XK
   DO 20 I=1,NDM1
     PN=(DFLOAT(I+I+1)*XK*P1-DFLOAT(I)*PO)/DFLOAT(I+1)
     PO=P1
20  P1=PN
     XK21=(1.-XK)*(1.+XK)
     PN=P1*XK21/(PO-XK*P1)/FND
     XK=XK-PN
     IF(DABS(PN).GT.EPS) GO TO 10
     ISW=ISW+1
     IF(ISW.LT.2) GO TO 10
     X(K)= XK
     W(K)=2.*XK21/((FND*PO)**2)
     J=ND-K+1
     X(J)=-XK
30  W(J)=W(K)
     RETURN
40  ICON=30000
     RETURN
50  X(1)=0.0
     W(1)=2.0
     RETURN
     END

```