

Rank Revealing QR 分解による Tikhonov の正則化法

仲田 晋 (Susumu Nakata) 北川 高嗣 (Takashi Kinami) 細田 陽介 (Yousuke Hosoda)
筑波大学理工学研究科 筑波大学電子・情報工学系 富山県立大学工学部

1 はじめに

我々の目的は悪条件線形方程式の高速な数値解法の構築にある。一般に方程式

$$(1.1) \quad Ax = b, \quad A \in \mathbb{R}^{m \times n}$$

が悪条件であるとは、右辺データベクトル b に微小な摂動 Δb を加えたとき、解 x の変化 Δx が大きくなる方程式を意味する。

悪条件線形方程式の代表的な数値解法として Tikhonov の正則化法 [3],[4] がある。通常 Tikhonov の正則化法は係数行列の特異値分解を用いて解を求める。しかし、特異値分解は計算量が多いという欠点を持つため、大規模な問題に対しては困難が伴うこととなる。

本論文では特異値分解の代りに Rank Revealing QR 分解 [1],[2] を用いた正則化法のアルゴリズムを提案する。Rank Revealing QR 分解は特異値分解に比べ高速であり、特に A の数値的階数が小さい場合は少ない計算量で分解できる。

Tikhonov の正則化法は、ある正則化パラメータ λ に対して汎関数

$$(1.2) \quad F_\lambda(x) = \|Ax - b\|_2^2 + \lambda^2 \|Sx\|_2^2$$

を最小化するベクトル x_λ を悪条件線形方程式の近似解とする方法である。ここで S はスタビライザー行列と呼ばれるもので、 $S = I_n$ としたときを Tikhonov の正則化法の標準形と呼ぶ。

従って Tikhonov の正則化法では正則化パラメータ λ の選択が問題となる。最適正則化パラメータの推定法として代表的なものに一般化交差検証法 (Generalized Cross Validation, 以下 GCV 法) [7] がある。この方法ではさまざまな λ についてその最適化規準関数の値を求め、その値が最小となる λ を最適正則化パラメータとする方法である。特異値分解による Tikhonov の正則化法の場合、一つの λ に対する最適化基準関数の計算量は $O(k)$ である。ここで k は係数行列の数値的階数を表す。

本論文で提案する Rank Revealing QR 分解を用いた方法では、特異値分解よりも計算量の少ない 2 回の QR 分解で計算可能であり、最適化基準関数は特異値分解と同様 $O(k)$ で計算可能である。

2 特異値分解による Tikhonov の正則化法

本節では悪条件線形方程式

$$(2.1) \quad Ax = b, \quad A \in \mathbb{R}^{m \times n}$$

に対する，特異値分解による Tikhonov の正則化法について説明する．係数行列 A の条件数 $\text{cond}(A)$ は

$$\text{cond}(A) \stackrel{d}{=} \|A\| \|A^\dagger\|$$

で表すことにする．またこれは係数行列の特異値を用いて次のように表すことができる [6]．

$$\text{cond}(A) = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)}$$

ただし， $\sigma_{\max}(A), \sigma_{\min}(A)$ はそれぞれ A の最大，非零な最小特異値を表す．条件数が大きい場合，特に計算機イプシロンの逆数 $1/\varepsilon$ と同程度もしくはそれ以上であるとき，この方程式は悪条件であるという．

Tikhonov の正則化法の標準形

$$(2.2) \quad F_\lambda(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \lambda^2 \|\mathbf{x}\|_2^2$$

を適用したときの正則化解の特異値分解を用いた計算法について述べる．ただし，上式を最小化するベクトル $\mathbf{x}_\lambda \in \mathbf{R}^n$ は

$$(2.3) \quad \mathbf{x}_\lambda = (\mathbf{A}^T \mathbf{A} + \lambda^2 \mathbf{I}_n)^{-1} \mathbf{A}^T \mathbf{b}$$

で与えられる．

係数行列 A の特異値分解

$$(2.4) \quad \mathbf{A} = \mathbf{U} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \mathbf{V}^T$$

$$\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_m] \in \mathbf{R}^{m \times m}, \quad \mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_n] \in \mathbf{R}^{n \times n}$$

$$\mathbf{U}^T \mathbf{U} = \mathbf{U} \mathbf{U}^T = \mathbf{I}_m, \quad \mathbf{V}^T \mathbf{V} = \mathbf{V} \mathbf{V}^T = \mathbf{I}_n$$

$$\Sigma = \text{diag}(\sigma_1, \dots, \sigma_k) \in \mathbf{R}^{k \times k} \quad \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \in \mathbf{R}^{m \times n}$$

が与えられているとする．ここで k はしきい値 μ に対する数値的階数を表す．

このとき，Tikhonov の正則化法の解 $\mathbf{x}_\lambda^{(S)}$ は

$$(2.5) \quad \mathbf{x}_\lambda^{(S)} = \mathbf{V} \begin{bmatrix} (\Sigma^2 + \lambda^2 \mathbf{I}_k)^{-1} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \mathbf{U}^T \mathbf{b}$$

$$(2.6) \quad = \sum_{i=1}^k \left(\frac{\sigma_i}{\sigma_i^2 + \lambda^2} \mathbf{u}_i^T \mathbf{b} \right) \mathbf{v}_i$$

となる．ただし， $\Sigma^2 = \text{diag}(\sigma_1^2, \dots, \sigma_k^2)$ である．

これより，係数行列の特異値分解と右辺ベクトル $\mathbf{b}$ のフーリエ係数 $\mathbf{u}_i^T \mathbf{b}$ があらかじめ求められているならば，一つの λ に対して $nk + 2k$ 回の乗除算で正則化解を計算することができる．

3 QR 分解による正則化法

本節では, Rank Revealing QR 分解を用いた正則化法について定式化する.

まず, Rank Revealing QR 分解の定義は以下のように与えられる.

定義 3.1 [1],[5]

行列 $A \in \mathbf{R}^{m \times n}$ の QR 分解

$$(3.1) \quad \begin{aligned} A\Pi &= QR \\ &= Q \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix} \end{aligned}$$

において, 条件

$$(3.2) \quad \begin{aligned} \text{cond}(R_{11}) &\approx \sigma_1/\sigma_k \\ \sigma_k &\gg \sigma_{k+1} = O(\|R_{22}\|_2) \end{aligned}$$

を満たすような置換行列 Π が存在するとき, この QR 分解を Rank Revealing QR 分解という. ただし, Q は $m \times m$ の正規直交行列, R_{11} は $k \times k$ の上三角行列である.

ここで任意の行列に対して Rank Revealing QR 分解が存在する [5].

本論文では Rank Revealing QR 分解における置換行列は Π , 通常のピボッティング付 QR 分解における置換行列は P で表すことにする.

次に, Rank Revealing QR 分解による正則化法について述べる. 悪条件線形方程式 (2.1) が与えられているとする. 数値的階数決定のためのしきい値 μ を与え, 係数行列 A の行ベクトルに対して Rank Revealing QR 分解を施し, 以下のように分解する.

$$(3.3) \quad \Pi A = \begin{bmatrix} \hat{L} & 0 \end{bmatrix} \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^T + \Delta A$$

ここで, $\hat{L} \in \mathbf{R}^{m \times k}$ は対角成分が 1 で, それ以外の成分の絶対値が 1 以下の下三角行列, $D \in \mathbf{R}^{k \times k}$ は対角行列で, その成分 $d_1, d_2, \dots, d_k$ はピボッティングの効果により

$$(3.4) \quad |d_1| \geq |d_2| \geq \dots \geq |d_k| > 0$$

の順に並んでいる. $V = [v_1, \dots, v_n] \in \mathbf{R}^{n \times n}$ は正規直交行列, Π は行置換行列, $k \leq \min(m, n)$ は μ に依存した A の数値的階数,

$$\begin{bmatrix} \hat{L} & 0 \end{bmatrix} \in \mathbf{R}^{m \times n}, \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \in \mathbf{R}^{n \times n}$$

である. また ΔA はしきい値 μ による QR 分解の誤差で, $\|\Delta A\|_2 = O(\mu)$ である.

式 (3.3) の両辺に Π^{-1} を左から乗ずることにより

$$\begin{aligned} A &= [\Pi^{-1}\hat{L} \ 0] \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^T + \Delta A \\ &= [L \ 0] \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^T + \Delta A, \quad L = \Pi^{-1}\hat{L} \end{aligned}$$

となる。以下

$$A = A_\mu = [L \ 0] \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^T$$

とおく。

次に、 $[L \ 0]$ の列に対して通常の QR 分解をほどこすことにより

$$[L \ 0] = U \begin{bmatrix} \hat{R} & 0 \\ 0 & 0 \end{bmatrix}$$

が得られる。ただし、 $\hat{R} \in \mathbf{R}^{k \times k}$ は上三角行列、 $U = [\mathbf{u}_1, \dots, \mathbf{u}_m] \in \mathbf{R}^{m \times m}$ は正規直交行列、

$$\begin{bmatrix} \hat{R} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbf{R}^{m \times n}$$

である。以上の分解より、 A の分解

$$(3.5) \quad A = U \begin{bmatrix} \hat{R} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^T + \Delta A$$

が得られる。このとき、上三角行列 $\hat{R}$ は Rank Revealing QR 分解の効果により悪条件とはならないことに注意する。また、 U, V の最初の k 個の列ベクトル $\{\mathbf{u}_i\}_{i=1}^k, \{\mathbf{v}_i\}_{i=1}^k$ はそれぞれ A の値域 $R(A)$ 、ゼロ空間の直交補空間 $N(A)^\perp$ の正規直交基底となる。

上式の分解において、上三角行列 $\hat{R}$ に対角行列 D による相似変換

$$R = D^{-1}\hat{R}D$$

を行うことにより、分解 (3.5) は

$$(3.6) \quad A = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} R & 0 \\ 0 & 0 \end{bmatrix} V^T + \Delta A$$

となる。ただし、上式においては

$$\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \in \mathbf{R}^{m \times m}, \quad \begin{bmatrix} R & 0 \\ 0 & 0 \end{bmatrix} \in \mathbf{R}^{m \times n}$$

である。これにより D の対角成分の単調性 (3.4) から, $\hat{R}$ の対角成分が強調され, R はさらに良条件となる。

係数行列の分解 (3.6) を用いて方程式 $Ax = b$ に対する Tikhonov の正則化法の汎関数 (1.2) を次のように定式化する。

$$(3.7) \quad F_\lambda(x) = \|Ax - b\|_2^2 + \lambda^2 \|Sx\|_2^2, \quad S = \begin{bmatrix} R & 0 \\ 0 & 0 \end{bmatrix} V^T = RV_k^T.$$

ただし, $V_k = [v_1, \dots, v_k] \in \mathbf{R}^{n \times k}$ である。しかし, $N(A) = N(S) \neq \{0\}$ であるため, 汎関数 (3.7) を最小化するベクトル $x \in \mathbf{R}^n$ は一意には定まらない。そこで我々は汎関数 (3.7) を最小化するベクトル $x \in N(A)^\perp$ を考える。すると汎関数 (3.7) を最小化するベクトル x_λ は

$$x_\lambda = (A^T A + \lambda^2 S^T S)^\dagger A^T b$$

として与えられる。これは分解 (3.6) を用いると

$$(3.8) \quad x_\lambda = V \begin{bmatrix} R^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} (D^2 + \lambda^2 I_k)^{-1} D & 0 \\ 0 & 0 \end{bmatrix} U^T b$$

$$(3.9) \quad = V_k R^{-1} (D^2 + \lambda^2 I_k)^{-1} D U_k^T b$$

と書き表すことができる。ただし, $U_k = [u_1, \dots, u_k] \in \mathbf{R}^{m \times k}$, $D^2 = \text{diag}(d_1^2, \dots, d_k^2)$ である。

上式の正則化解は, 分解 (3.6) と $U_k^T b$ があらかじめ計算されていれば, $nk + 2k + k(k+1)/2$ 回の乗除算で求めることができる。これは特異値分解による正則化解 (2.6) を求めるための乗除算の回数よりも R による後退代入分だけ多い。もし, 数値的階数 k が m, n よりも格段に小さい場合, この増加分による影響は小さい。

4 GCV 法による最適正則化パラメータの推定

本節では, 正則化法を適用するときの最適の正則化パラメータを推定する方法の1つである GCV 法について説明する。

GCV 法は最も有名な最適正則化パラメータ推定法の一つであり, 最適化基準関数が最大となる λ を最適正則パラメータとして採用する方法である。GCV 法の最適化規準関数は

$$(4.1) \quad \text{GCV}(\lambda) = \frac{\|Ax_\lambda - b\|_2^2}{(\text{trace}(I_m - AA_\lambda^-))^2}$$

として表される [7]。ここで A_λ^- は式 (2.3) の場合,

$$\begin{aligned} A_\lambda^- &= (A^T A + \lambda^2 I)^{-1} A^T \\ &= V^{(S)} \begin{bmatrix} (\Sigma^2 + \lambda^2 I_k)^{-1} \Sigma & 0 \\ 0 & 0 \end{bmatrix} U^{(S)T} \end{aligned}$$

となる。ここで $U^{(S)}, V^{(S)}$ はそれぞれ特異値分解 (2.4) で得られた正規直交行列 U, V を表す。これより、 $U_k^{(S)} \in \mathbf{R}^{m \times n}$ を $U^{(S)}$ の最初の k 列のベクトルから成る行列、

$$\Delta \mathbf{b}^{(S)} = \mathbf{b} - U_k^{(S)} U_k^{(S)T} \mathbf{b}, \quad \mathbf{c}^{(S)} = U_k^{(S)T} \mathbf{b} = (c_1^{(S)}, \dots, c_k^{(S)})^T$$

とすれば、GCV 法の最適化規準関数は

$$(4.2) \quad \text{GCV}^{(S)}(\lambda) = \frac{\sum_{i=1}^k \left(\frac{\lambda^2}{\sigma_i^2 + \lambda^2} c_i^{(S)} \right)^2 + \|\Delta \mathbf{b}^{(S)}\|_2^2}{\left(\sum_{i=1}^k \frac{\lambda^2}{\sigma_i^2 + \lambda^2} + (m - k) \right)^2}$$

となる。

次に方程式 $A\mathbf{x} = \mathbf{b}$ に QR 分解による正則化法 (3.7) を適用したときの GCV 法の最適化規準関数を与える。このとき正則化行列は

$$A_{\lambda}^{-} = V^{(Q)} \begin{bmatrix} R^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} (D^2 + \lambda^2 I_k)^{-1} D & 0 \\ 0 & 0 \end{bmatrix} U^{(Q)T}$$

となる。ここで $U^{(Q)}, V^{(Q)}$ はそれぞれ分解 (3.6) における正規直交行列 U, V を表す。特異値分解による Tikhonov の正則化法のとときと同様に、 $U_k^{(Q)}$ を $U^{(Q)}$ の最初の k 列のベクトルから成る行列、

$$\Delta \mathbf{b}^{(Q)} = \mathbf{b} - U_k^{(Q)} U_k^{(Q)T} \mathbf{b}, \quad \mathbf{c}^{(Q)} = U_k^{(Q)T} \mathbf{b} = (c_1^{(Q)}, \dots, c_k^{(Q)})^T$$

とすれば、GCV 法の最適化規準関数は

$$(4.3) \quad \text{GCV}^{(Q)}(\lambda) = \frac{\sum_{i=1}^k \left(\frac{\lambda^2}{d_i^2 + \lambda^2} c_i^{(Q)} \right)^2 + \|\Delta \mathbf{b}^{(Q)}\|_2^2}{\left(\sum_{i=1}^k \frac{\lambda^2}{d_i^2 + \lambda^2} + (m - k) \right)^2}$$

となる。

式を見れば明らかなように、式 (4.2) と式 (4.3) は同じ形式となってる。これより、どちらの規準関数も一つの λ に対して $O(k)$ 回の乗除算で求めることができることがわかる。

5 数値実験と考察

本節では QR 分解による正則化法の実用性を数値実験を通して考察する。

数値実験は第一種フレドホルム積分方程式

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

を用いて行った。この積分方程式の真の解は $f(t) = t$ である。データ点 $\{s_i\}$ として $[0, 1]$ 区間上の 100 個の一樣乱数を、数値積分は 100 点ガウス・ルジャンドル則を用い、線形方程式 $A\mathbf{x} = \mathbf{b}$ の形に離散

化した。離散化したあと、右辺ベクトル $\mathbf{b}$ にノイズとして平均が 0 で標準偏差が 1.0×10^{-6} の正規乱数から成るベクトル $\Delta \mathbf{n}$ を加えた。このとき、 $\|\Delta \mathbf{n}\|_2 = 1.0129 \times 10^{-5}$ であった。数値的階数決定のためのしきい値は $\mu = 1.0 \times 10^{-8}$ を用いた。

計算機は HP Vectra XU(Pentium Pro 200MHz) を用い、計算はすべて倍精度実数演算で行った。特異値分解 (2.4) と 2 回の QR 分解による分解 (3.6) を得るためにかかった実行時間を表 1 に示す。

Table 1. CPU-time of both decompositions.

	SVD(2.4)	QR2(3.6)
CPU-time(sec)	0.84	0.07

この問題の場合、行列の分解に要する時間は QR 分解を用いた方が 10 倍以上速い。

特異値 σ_i と右辺のフーリエ係数 $\{c_i^{(S)}\}$ 、分解 (3.6) における対角行列 D の対角成分 $\{d_i\}$ と右辺のフーリエ係数 $\{c_i^{(Q)}\}$ を常用対数尺度でそれぞれ図示する。なお、以下の図はすべて左側が特異値分解を用いた場合 (SVD)、右側が QR 分解を用いた場合 (QR2) の結果を表している。

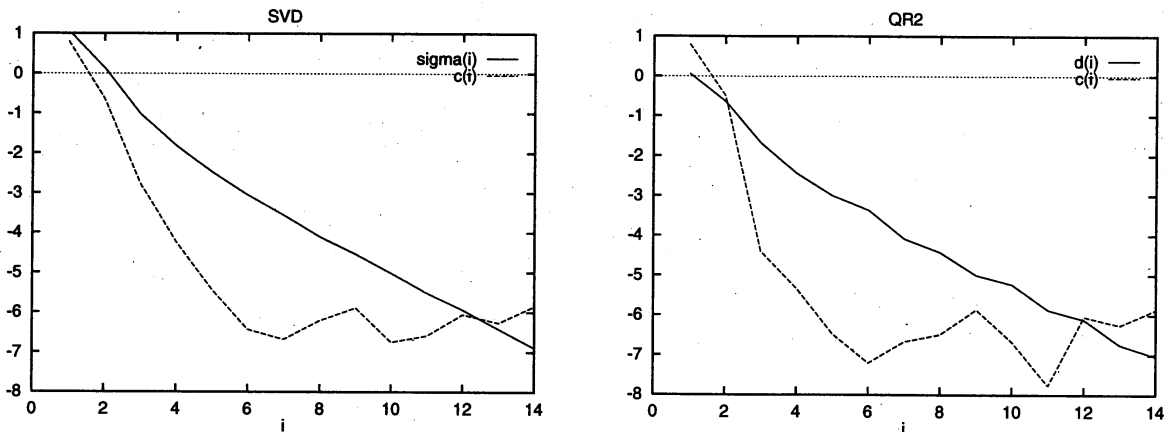


Fig. 1. Singular values and Fourier coefficients for SVD (2.4), diagonal elements of D and Fourier coefficients for decomposition (3.6)

図から σ_i と $|d_i|$, $c_i^{(S)}$ と $c_i^{(Q)}$ がそれぞれよく似た振舞をしていることがわかる。

特異値分解と QR 分解それぞれの場合の GCV 最適化規準関数 (4.2),(4.3) と真の解ベクトルを $\mathbf{x}_0$ としたときの誤差関数

$$(5.1) \quad E^{(S)}(\lambda) = \|\mathbf{x}_\lambda^{(S)} - \mathbf{x}_0\|_2, \quad E^{(Q)}(\lambda) = \|\mathbf{x}_\lambda^{(Q)} - \mathbf{x}_0\|_2,$$

の関係を図 2 に示す。ただし、横軸、縦軸は常用対数尺度で表してある。図からはわかり難いが GCV 規準関数は特異値分解による Tikhonov の正則化法の場合は $\log_{10} \lambda = -3.101$ で、2 回 QR 分解による正則化法の場合は $\log_{10} \lambda = -3.488$ で最小値をとっている。これより GCV 法を用いればどちらの方法でも 3 桁程度の近似解が得られることがわかる。

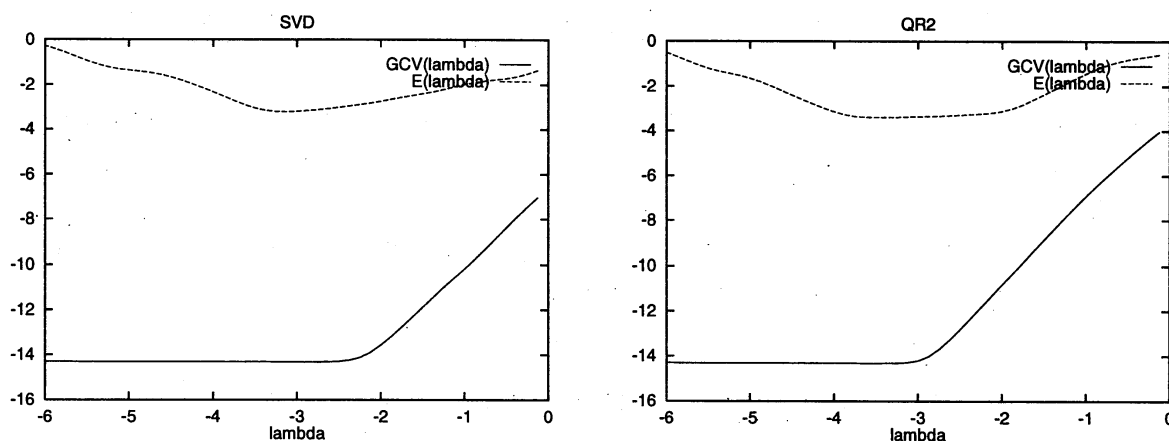


Fig. 2. GCV functions (4.2),(4.3) and errors of both methods.

6 まとめ

我々は悪条件線形方程式に対する数値計算法として, Rank Revealing QR 分解を含む 2 回の QR 分解から得られる正則化法を提案した. 本方法は Rank Revealing QR 分解の性質である数値的階数の決定により数値的に安定で, 特異値分解に比べてより高速で, なおかつ特異値分解による Tikhonov の正則化法を用いるときの最適化規準が同じ計算量で導入できることがわかった. 数値実験においても特異値分解による Tikhonov の正則化法と同程度の近似解が高速に得られることがわかった.

参考文献

- [1] Chan, T.F., *Rank Revealing QR-factorizations*, Linear Algebra Apl., **88/89**(1987), 67-82.
- [2] Chan, T.F., and Hansen, P.C., *Some applications of the rank revealing QR factorization*, SIAM J.Sci.Stat.Comput., **13**(1992), 727-741..
- [3] Groetsch, C.W., *Inverse Problems in the Mathematical Sciences*, Vieweg,Braunschweig, 1993.
- [4] Groetsch, C.W., *The theory of Tikhonov regularization for Fredholm equations of the first kind*, Pitman,Boston,London and Melbourne,1984.
- [5] Y.P.Hong and C.-T.Pan, *Rank Revealing QR factorization and the Singular Value Decomposition*, Mathematics of Computation,**58**(1992), 213-232.
- [6] 中川徹, 小柳義夫, 最小二乗法による実験データ解析, 東京大学出版会, 東京, 1982.
- [7] Wahba, G., *Generalized cross-validation as a method for choosing a good ridge parameter*, TECHNOMETRICS, **21**(1979), 215-223.