Vol.2013-HPC-139 No.1
2013/5/29

# 限量記号消去法を用いた回帰モデルの予測力向上

折居茂夫[†1]　　山本義郎[†2]

並列処理時間の回帰モデルの予測力向上のため，モデルパラメータに正負の条件を付けた連立残差不等式を解いてモデルパラメータを決定する方法を提案する．連立残差不等式は，数式処理のアルゴリズムの一つ限量記号消去法(QE)を用いて解く．まず全ての残差式を満足する最小残差値を QE で求める．次にこの最小残差値に対するモデルパラメータ値を QE で決定する．モデルパラメータに正負の条件を付けてこれを決定すると，この最小残差値より大きな残差を生み出す基底関数のモデルパラメータが零となる．これにより過剰適合が軽減してモデルの予測力が向上する．またパラメータ値が零の基底関数を削除することが予測力向上につながる．これらのことを簡単なデータのモデル化と，並列処理時間のモデル化により例示する．

# Improving Predictive Power of Regression Models Using Quantifier Elimination

SHIGEO ORII[†1]　　YAMAMOTO YOSHIRO[†2]

We propose a novel method improving predictive power of regression models for parallel processing time. The method solves simultaneous inequalities for residual to obtain model parameters restricted in the positive or negative range. Quantifier elimination (QE), a symbolic computation algorithm is used to solve the simultaneous inequalities for residual. First a minimum value of the residual that satisfies the simultaneous inequalities is obtained with QE. Second model parameters are obtained based on the minimum residual. Solving the simultaneous inequalities added a positive or negative condition to model parameters, zero value becomes the solution for model parameters whose bases functions create large residuals than the minimum residual. The zero value reduces over fitting phenomena and improves prediction power of regression models. Removing basis functions based on the zero value also improve predictive power of regression models. These are illustrated by modeling with simple data and with timing data of a parallel processing.

## 1. はじめに

並列処理時間の回帰モデルのモデルパラメータを最小二乗法により決定することが広く行われている．最小二乗法を適用すると，モデルパラメータ決定に用いたデータを内挿する良いモデルを構築できる．一方データの外挿領域の予測をする場合，実際に測定した値と比べてみると，予測が大きく外れていることをしばしば経験する．この過剰適合状態の原因としては，用いた基底関数がデータの振舞いをうまく記述できない，基底関数過多，使用したデータが外挿領域で優勢となる振舞いの情報をあまり含んでいない，情報に比べて大きなノイズの存在等が考えられる．この中の基底関数過多では，余分な基底関数を駆使しノイズに起因する残差までもが削減される．このとき余分な基底関数の中に冪乗等の急激に増大するものがあると，データを外挿した領域での予測値は実際とは大きく異なることになる．そこでプロセッサの増加に対して滑らかになるように正規化したデータを用いてこのような過剰適合状態を軽減する方法を提案してきた[1, 2].

今回我々は過剰適合の原因の一つ基底関数過多に着目し，予測に不要な基底関数をモデル構築時点に見つける方法を研究した．このために必要な情報源として，最小二乗法が平方と和を取る前の残差に着目した．予測に不要な基底関数は，残差の一部を増加するよう働くと考えられる．そこでまず与えられた全データの残差に対する連立不等式を作り，連立不等式が成り立つ最小残差を求める．次にその最小残差に対する基底関数の係数即ちモデルパラメータを決定する．この計算においてモデルパラメータに正負の条件を付けておくと，最小残差より大きな残差を生じる基底関数のモデルパラメータは零となるというのが基本的アイディアである．

連立残差不等式は，数式処理のアルゴリズムの一つ限量記号消去法(QE)を用いて解く．QE を用いると連立不等式に対する厳密解が得られ，求めた最小残差も厳密解となる．従ってモデルパラメータが零になる基底関数も厳密に特定することができる．

本論文では過剰適合の原因となる予測に不適当な基底関数のモデルパラメータを零値として得ることにより回帰モデルの予測力向上を図る方法を提案する．またモデルパラメータが零値となった不要な基底関数をモデルから削除し，回帰モデルの予測力向上を図ることを提案する．

2 章は本論文で扱う課題について述べる．3 章で課題を解決する基本的アイディアを示す．4 章でこのアイディアを QE で解く方法を示す． 5 章では過剰適用が軽減され予測力が向上した例と，不要な基底関数を削除して並列処理の時間モデルの予測力が向上した例を示す． 6 章でまとめを行う．

†1 GDE00740@nifty.com
†2 東海大学　理学部

## 2. 予測モデル構築における課題

ここでは最小二乗法(LSM)で予測モデルを作る時の課題について述べる．$N$ 個のデータ $y_i$ を $M$ 個の基底関数 $f_k(x)$ で表された回帰モデル $a_0+\Sigma a_k f_k(x)$を考える．係数 $a_k$ はモデルパラメータである．LSM では式（1）の残差平方和 $E$ を最小にして $a_k$ を決定する．

$$E=\sum_{i=1}^{N}\left(y_i-a_0-\sum_{k=1}^{M}a_k\cdot f_k(x_i)\right)^2 \tag{1}$$

$E$ を最小にするため両辺をモデルパラメータで $\partial E/\partial a_k=-2\cdot\Sigma f_k(x_i)\cdot[y_i-a_0-\Sigma a_k\cdot f_k(x_i)]$のように微分して整理すると式（2）を得る．

$$\sum_{i=1}^{N}f_k(x_i)\cdot a_0+\sum_{i=1}^{N}f_k(x_i)\cdot f_1\cdot a_1+\cdots+\sum_{i=1}^{N}f_k(x_i)\cdot f_M\cdot a_M=\sum_{i=1}^{N}f_k(x_i)\cdot y_i \tag{2}$$

ここで 1 番目の基底関数$f_1$ がデータ $y_i$ をモデル化するのに必要ない基底関数の場合を考え，そのような $y_i$ を式(3)と仮定する．

$$y_i=a_0'+a_2'\cdot f_2(x_i)+\cdots+a_M'\cdot f_M(x_i)+d_i \tag{3}$$

ここに $d_i$ は式(3)の基底関数で表しきれないデータからの差異で，測定誤差等を総合したものである．この式(3)を式(2)に代入して整理すると式(4)を得る．

$$\sum_{i=1}^{N}f_k(x_i)\cdot(a_0-a_0')+\sum_{i=1}^{N}f_k(x_i)\cdot f_1(x_i)\cdot a_1+\sum_{i=1}^{N}f_k(x_i)\cdot f_2(x_i)\cdot(a_2-a_2')+\cdots+\sum_{i=1}^{N}f_k(x_i)\cdot f_M(x_i)\cdot(a_M-a_M')=\sum_{i=1}^{N}f_k(x_i)\cdot d_i \tag{4}$$

式(4)からわかるように $a_1$ は $d_i$ が零でない場合は零とはならない．このように LSM では $f_1$ のような不要な基底関数があるとそれも $E$を最小化するために使わる．このような不要な基底関数が入った過剰適合状態で $f_1$ が $x$ の冪乗等の急激に増加する関数であれば，データ $y_i$ の外挿領域においてその項が優勢となり，結果モデルの予測力が低下する．従ってモデル構築時にこのような不要な基底関数を見つけて削除することが，回帰モデルの予測力向上につながる課題の一つと考える．

## 3. 基本的アイディア

過剰適合を引き起こす不要な基底関数を，モデル構築時に見つけ出すためには，それを行うための情報が必要となる．本研究ではこの情報源として式(5)の各データに対する残差 $e_i$ に着目した．

$$e_i=y_i-a_0-\sum_{k=1}^{M}a_k\cdot f_k(x_i) \tag{5}$$

全ての残差の情報を考慮するため，残差 $e_i$ の最大値以上のある残差 $e$ を用いて式(5)を式(6)のように連立残差不等式で表す．

$$e\geq\left|y_i-a_0-\sum_{k=1}^{M}a_k\cdot f_k(x_i)\right| \tag{6}$$

ここで式(6)を満足する $e$ の最小値を求める．$e$ の最小値かつ $e_i$ の最大値であるこれを，今後本論文では $e_{max}$ と記述する．$e_{max}$ は全ての $e_i$ を基にして決定されているので，ある $i$ の範囲の $e_i$ だけを減少し，他の範囲では増加するような基底関数は $e_{max}$ を増すように働くと考えられ，そのような基底関数のモデルパラメータは零になると考えられる．

このような状態を具体的に把握するため，式(3)を式(6)に代入し，全ての $i$ で成り立つ $e$ の最小値 $e_{max}$ を求めて代入すると式(7)を得る．

$$e_{max}\geq\left|\begin{matrix}(a_0'-a_0)-a_1\cdot f_1(x_i)+\\(a_2'-a_2)\cdot f_2(x_i)+\cdots+\\(a_M'-a_M)\cdot f_M(x_i)+d_i\end{matrix}\right| \tag{7}$$

$-a_1\cdot f_1(x_i)+d_i$ の項が無ければ $e_{max}$ は零なので，この項が $e_{max}$ を決定している．従って $e_{max}$ は，$\mathrm{Max}(|d_i|)$か或いはそれより小さい値となる可能性を持つ$|-a_1\cdot f_1(x_i)+d_i|$と考えられる．ここでもしモデルパラメータ $a_1$ が正負どちらの値をも取れるとすると，$f_1(x_i)$と $d_i$ の和で $\mathrm{Max}(d_i)$より小さい値となる可能性があり，その場合 $a_1\neq 0$ となる．一方 $a_1\geq 0$ あるいは $a_1\leq 0$ とすると，$\mathrm{Max}(d_i)<|-a_1\cdot f_1(x_i)+d_i|$にならないために $a_1=0$ であることが必要な場合が生じる．

このようにモデルパラメータに正負の条件を付けて $e_{max}$ を求め，それを基にモデルパラメータを決定し，$e_{max}$ より大きな残差を生じる基底関数モデルのモデルパラメータが零になるよう仕向けることが，提案する方法の基本的アイディアである．

## 4. QE による $e_{max}$ とモデルパラメータの決定

与えられた $N$ 個のデータ $y_i$ と $M$ 個の基底関数 $f_k(x)$を用いて $e_{max}$ とモデルパラメータを決定するために，数式処理のアルゴリズムの一つ限定記号消去法(QE)を用いる[3, 4]. QE を適用するため式(6)を論理式として表わすと，式(8)となる．ベクトル $\mathbf{x}_i$ は変数が複数であることを意味する．

$$\bigwedge_{i=1}^{N}\left(-e\leq y_i-a_0-\sum_{k=1}^{M}a_k\cdot f_k(\mathbf{x}_i)\leq e\right) \tag{8}$$

ここで QE を用い式(8)の $a_k$ を全て消去すると $e$ を変数とする関数 $F(e)$を得ることができる．３章で述べたようにこ

の $F(e)$の最小値が $e_{\max}$ となる．これらの表記を式(9)に示す．

$$\exists a_0 \exists a_1 \cdots \exists a_M \left( \bigwedge_{i=1}^{N} -e \le y_i - a_0 - \sum_{k=1}^{M} a_k \cdot f_k(\mathbf{x}_i) \le e \right) \overset{\text{QE}}{\Rightarrow} F(e) \qquad (9)$$
$$e_{\max} = Min(F(e))$$

次に，得られた $e_{\max}$ を用いてモデルパラメータ $a_k$ を QE オペレーションで決定する．式(10)は得られた $e_{\max}$ を基に $a_0$以外のモデルパラメータを消去して $a_0$を決定する論理式である．

$$\exists a_1 \cdots \exists a_M \left( \bigwedge_{i=1}^{N} -e_{\max} \le y_i - a_0 - \sum_{k=1}^{M} a_k \cdot f_k(\mathbf{x}_i) \le e_{\max} \right) \overset{\text{QE}}{\Rightarrow} F(a_0) \qquad (10)$$

QE オペレーションに，3 章の基本的アイディアで述べたパラメータが正か負である条件を加えるには，論理式にこれらの条件を追加すればよい．例えば $a_0 \ge 0$ かつ $a_1 \ge 0$ かつ $a_2 \ge 0$ のような条件付は，式(9)で $M=2$ とすると下記のようになる．

$$\exists a_0 \exists a_1 \exists a_2 \left( \begin{array}{l} \bigwedge_{i=1}^{N} \left( -e \le y_i - a_0 - a_1 \cdot f_1(\mathbf{x}_i) - a_2 \cdot f_2(\mathbf{x}_i) \le e \right) \\ \wedge a_0 \ge 0 \wedge a_1 \ge 0 \wedge a_2 \ge 0 \end{array} \right) \overset{\text{QE}}{\Rightarrow} F(e)$$

## 5. 適用事例

QE によりモデルの予測力が向上する 3 つの事例を示す．比較のため最小二乗法(LSM)で決定した結果を列挙する．QE オペレーションを行うツールとして Mathematica [5]のバージョン 8 の Reduce 関数を，LSM のツールとして同 FindFit を用いた．

### 5.1 $y=x^2$ のモデル化とモデル予測力向上

QE により過剰適合の原因となる基底関数のモデルパラメータが零になるかを確かめるため，データを $y=x^2$ より作成した．基底関数に多項式 $f(x)= a_1 \cdot x+a_2 \cdot x^2+a_3 \cdot x^3+a_4 \cdot x^4+a_5 \cdot x^5$ 用いた．データは，$x_i$=(1.1, 1.25, 1.4, 1.55, 1.7, 1.85)に対する $y$ を標準偏差 $\sigma$ =0 の時の $y_i$ とし，$y$ に $\sigma$ =0.01 のばらつきを持たせた $y'_i$ の 2 ケースを用意いた．モデルの予測力を確かめるため， $x_i$ の $x$ の最大値(～2）の 10 倍の $x$ =0～20 間のモデル値と $y$ を比較した．

この問題を式(8)のような論理式で記述すると式(11-1)となる．この式に式(9)のようにQEオペレーションすると $e_{\max}$ =0 を得る．次に式(10)のように QE オペレーションを行いモデルパラメータを決定した．

$$\bigwedge_{i=1}^{6} \left( -e \le y_i - a_0 + \sum_{k=1}^{5} a_k \cdot f_k(x_i) \le e \right) \qquad (11\text{-}1)$$

表 1-1 は標準偏差 $\sigma$ =0 の場合である．QE で求めたモデルパラメータ値は $a_2$=1 でそれ以外の不要な基底関数のそれらは零であり，正確な予測ができるモデルが構築できることがわかる．LSM で決定したパラメータもほぼ同様である．

表 1-1　モデルパラメータ（$\sigma$ =0）

| | $a_o$ | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ |
|---|---|---|---|---|---|---|
| *QE* | 0 | 0 | 1 | 0 | 0 | 0 |
| *LSM* | 7.83 $\times 10^{-13}$ | -2.38 $\times 10^{-12}$ | 1 | -1.63 $\times 10^{-12}$ | 4.55 $\times 10^{-13}$ | -4.85 $\times 10^{-14}$ |

表 1-1 のモデルパラメータを使って x=0～20 までの予測値を $y=x^2$ と比較したものを図 1-1 に示す．図は $\sigma$ =0 の場合，データの範囲を超えた領域で，QE と LSM どちらもよい予測を与え，外挿に対応できることがわかる．

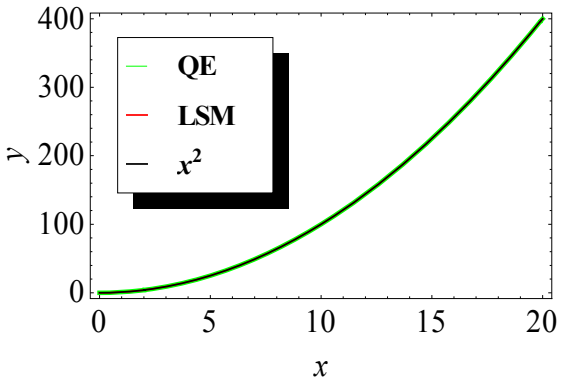


図 1-1　回帰モデルによる予測の確認（$\sigma$ =0）

次に $\sigma$ =0.01 の揺らぎを与えて作ったデータ $y'_i$ に対するモデルパラメータを表 1-2 に示す．有効数字 3 桁で QE と LSM の値は同じである． しかし全てのパラメータが大きな値となり，$y=x^2$ とは異なった過剰適合状態でモデル化されていることがわかる．

表 1-2　モデルパラメータ（$\sigma$ =0.01, $e_{\max}$ =0）

| | $a_o$ | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ |
|---|---|---|---|---|---|---|
| *QE* | -107. | 376. | -522. | 362. | -124. | 16.9 |
| *LSM* | -107. | 376. | -522. | 362. | -124. | 16.9 |

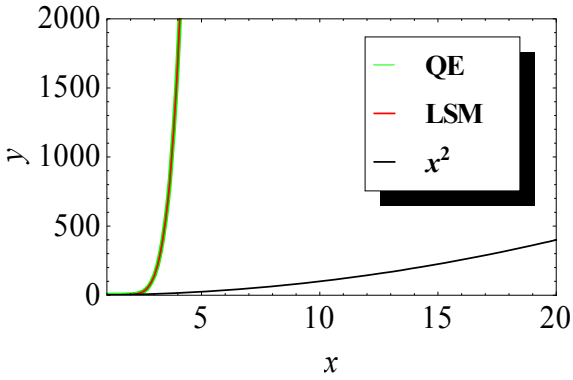


図 1-2　回帰モデルによる予測の確認（$\sigma$ =0.01）

図 1-2 は，このようなモデルパラメータでもデータの x 区間 1.1～1.8 の間の内挿には十分な精度である一方，デ

ータ範囲外では予測ができないことを示す．

更に，上記と同じ σ =0.01 の揺らぎを与えたデータに対し，3 章で述べたように $a_i \geq 0$ ($i$=0,.., $M$)の条件を入れ QE と LSM を適用した．論理式を式(11-2)に示す．

$$\bigwedge_{i=1}^{6}\left(-e \leq (y_i' - a_0 - \sum_{k=1}^{5} a_k \cdot f_k(x_i)) \leq e\right) \quad (11\text{-}2)$$
$$\wedge a_0 \geq 0 \wedge a_1 \geq 0 \wedge a_2 \geq 0 \wedge a_3 \geq 0 \wedge a_4 \geq 0 \wedge a_5 \geq 0$$

表 1-3 は，QE によるモデル化では $a_0, a_1, a_4, a_5$ が零となり，過剰適合が改善されていることがわかる．その結果，図 1-3 に示すようにデータの範囲を大きく外れると，LSM では $y$ とモデルの乖離が大きくなるが，QE はよい予測を与え，外挿に対応できることがわかる．

表 1-3　モデルパラメータ（σ =0.01, $a_i \geq 0$）

| | $a_o$ | $a_1$ | $a_2$ | $a_3(10^{-3})$ | $a_4$ | $a_5$ |
|---|---|---|---|---|---|---|
| *QE* | 0 | 0 | 0.989 | 5.62 | 0 | 0 |
| *LSM* | 1.15 ×10$^{-2}$ | 1.88 ×10$^{-2}$ | 0.956 | 11.1 | 1.79 ×10$^{-3}$ | 5.41 ×10$^{-4}$ |

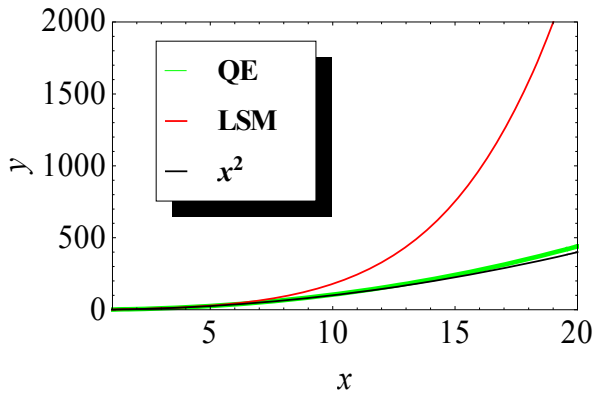


図 1-3　回帰モデルによる予測の確認（σ =0.01, $a_i \geq 0$）

一方 LSM でも不要な基底関数のパラメータ値が表 1-2 に比べて小さくなったが，予測値は一致しない．

結果，QE でモデルパラメータが正であるいは零である条件を加えた場合，予測に不要な 4 つのモデルパラメータが零となって過剰適合が緩和され，データ $x$ の約 10 倍の領域を予測できるようになったことを確認できた．

### 5.2 $y=x^2$ -Log(x)のモデル化とモデル予測力向上

5.1 節では 1 つの関数から作ったデータに対するモデル化だったが，現実には複数の現象の重ね合わせをモデル化する場合が想定される．QE により，このような場合に対しても過剰適合が緩和されて予測力向上が図れるかを調べるため，2 つの関数 $x^2$ と Log(x)を用いて $y=x^2$ -Log(x)のデータを作り，それに対するモデルを構築した．基底関数の集合としては，6.1 節と同程度の難しさを意識して $f(x)= a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3 + a_4 \cdot x^4 + a_5 \cdot x^5 + a_6 \cdot \mathrm{Log}(x)$を用いた．データは 6.1 節と同じ $x_i$=(1.1, 1.25, 1.4, 1.55, 1.7, 1.85)，$y_i$ は $x_i$ に対する σ =0 と σ =0.01 の時の $y=x^2$ -Log(x)値を用いた．

モデル化の結果，σ =0 の場合は 5.1 節と同様 QE と LSM で構築したモデルは，$x$=20 の $y$ を予測できた．σ =0.01 のばらつきを持たせた場合，$a_i$($i$=1,.., $M$)に正負の条件を付けないときは，モデルは 5.1 節と同様に QE 法も LSM 法も $x$=20 の $y$ を予測できなかった．

一方 $a_6 \leq 0$ と条件付けすると，表 2 に示すように，QE によるモデルは過剰適合が緩和される．一方 LMS によるモデルは全てのモデルパラメータが非零となり過剰適合が見られる．

表 2　モデルパラメータ（σ =0.01）

| | $a_o$ | $a_1$ | $a_2$ | $a_3$ |
|---|---|---|---|---|
| *QE* | 0 | 0 | 1.00 | 0.00376 |
| *LSM* | 0.0273 | 0.0332 | 0.921 | 0.0162 |

| | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|
| *QE* | 0 | 0 | -1.04 |
| *LSM* | 0.00142 | 0.000282 | -0.905 |

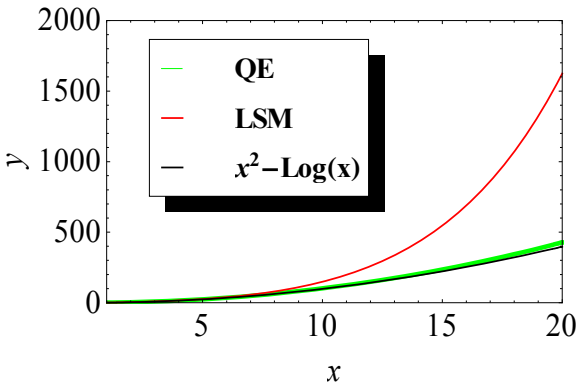


図 2　回帰モデルによる予測の確認（σ =0.01, $a_i \leq 0$）

モデルを使った予測結果を図 2 に示す．QE によるモデルは，データの範囲を大きく外れてもよい予測を与えるが，LSM では $y$ とモデルの乖離が大きくなる．この結果は，複数の現象が重なったデータに対しても，QE では過剰適合を緩和して予測力を向上できる場合があることを示す．

### 5.3 並列処理時間モデルの構築

これまでは関数から簡単なデータを作ってそれらのモデルを構築し，QE では過剰適合が緩和され予測力が向上することを確認したが，ここでは分子動力学プログラム[6]を並列計算機 IBM SP2 で実行した時の処理時間モデルを使い，QE による過剰適合緩和と予測力向上を確認する．

### (1) 基底関数の抽出

モデル化のための基底関数をプログラムから抽出した．モデル化したプログラムは，ファンデルワールス力と重力の影響下で時間発展するアルゴン原子の巨視的振舞いをシミュレーションする分子動力学プログラムである．

並列計算されるこのようなシミュレーションプログラムの処理時間をモデル化する場合，処理時間の大部分を費やすカーネルを見つけてその部分の処理時間を式化し，それらをアッセンブルしてプログラムの処理時間モデルを構築する方法がある．そこで時間全体の 99%の時間を占めるサブルーチンの処理時間を式化してこれらのサブルーチンをモデル化した．具体的には，これらのサブルーチンのカーネル(do ループ，MPI 通信部)の処理時間を，問題の大きさ $n$ の関数であるループ回転数に比例するものとし，次にその並列処理時間はプロセッサ数 $p$ に反比例するとして式化したものを用いた [7]．時間全体の 99%の時間を占めるサブルーチンは 5 つあり，そのうち 3 つのサブルーチンが並列化されている．

これらの処理時間の総和 $T_{MD}$ に対し，変数 $p$ と $n$ で式をまとめると，プログラム全体の処理時間を式(12)のように，8 つのモデルパラメータを持つ 2 変数 $p$, $n$ のモデルとして記述することができる．ここではこれらのモデルパラメータは全て零か正の値と置くことにする．

$$T_{MD}(p,n) = a_0 + a_1 \cdot n^2 + a_2 \cdot \frac{n^2}{p} + a_3 \cdot n + a_4 \cdot \frac{n}{p} + a_5 \cdot n \cdot p + \frac{a_6}{p} + a_7 \cdot p \qquad (12)$$

### (2) QE によるモデルパラメータの決定

式(12)とその条件を式(8)のように論理式で記述すると式(13)となる．データ $y_i$ は，$p$={2, 4, 6, 8,10}と，$n$={7200, 12800, 20000, 39200}の組み合わせで 20 通りの測定した実行時間からランダムに 7 点を 5 回抽出して用いた．

得られたモデルパラメータを抽出ケース毎に表 3-1 に示す．表はこのようなモデル化においても，モデルパラメータが零のものがあり，過剰適合が緩和されることを示す．

$$\begin{aligned} &\bigwedge_{i=1}^{N}\left(-e \le (y_i - a_0 - \sum_{k=1}^{7} a_k \cdot T_{MD_k}(x_i)) \le e\right) \\ &\wedge a_0 \ge 0 \wedge a_1 \ge 0 \wedge a_2 \ge 0 \wedge a_3 \ge 0 \\ &\wedge a_4 \ge 0 \wedge a_5 \ge 0 \wedge a_6 \ge 0 \wedge a_7 \ge 0 \end{aligned} \qquad (13)$$

表 3-1 は，抽出データ No.1 のモデルパラメータが範囲を持って得られたことを示す．このように QE を用いると結果が不等式となる場合がある．このような幅があるモデルパラメータを除外し，図 3 に No.2 から 5 までのモデルパラメータを用いて予測した例を示す．

表 3-1　QE により決定した式(12)のモデルパラメータ

| No | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $a_0$ | ≥0<br>≤6.94 | 0 | 64.0 | 0 | 22.0 |
| $a_1$ | ≥0<br>≤$3.33\cdot10^{-7}$ | 0 | $8.88\cdot10^{-8}$ | 0 | $4.80\cdot10^{-7}$ |
| $a_2$ | ≥$3.76\cdot10^{-6}$<br>≤$4.43\cdot10^{-6}$ | $4.26\cdot10^{-6}$ | $5.39\cdot10^{-6}$ | $2.74\cdot10^{-6}$ | $1.06\cdot10^{-6}$ |
| $a_3$ | ≥0.0340<br>≤0.0364 | 0.0397 | 0.0266 | 0.0418 | 0.0377 |
| $a_4$ | ≥0.290<br>≤0.296 | 0.290 | 0.2851 | 0.302 | 0.304 |
| $a_5$ | ≥0.000514<br>≤0.000995 | 0 | 0.00100 | 0 | 0 |
| $a_6$ | ≥0<br>≤12.3 | 0 | 0 | 0 | 0 |
| $a_7$ | ≥0<br>≤3.47 | 4.01 | 0 | 0 | 1.30 |

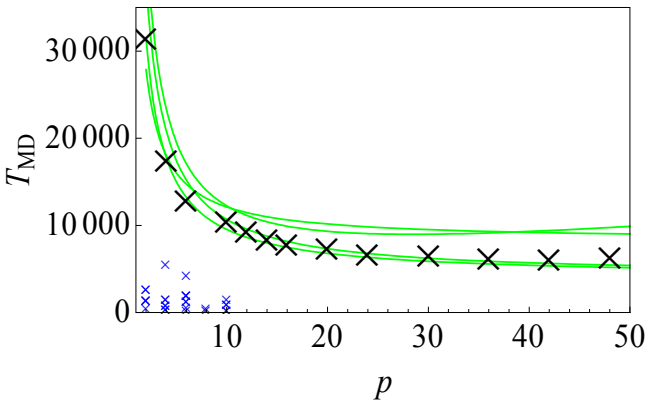


図 3　QE 法モデル（式(12)，表 3-1）の予測の確認

図は問題の大きさを $n$=96800 とした場合のプロセッサ $p$ に対する経過時間 $T_{MD}$ である．黒の×印は予測確認用の測定値である．緑線がモデルの予測である．緑色のモデルによる予測は，データの組み合わせを変えても黒の×印の値と傾向を予測できることを示す．

データとして用いた測定値を青色の×印で示す．モデル構築に使用したデータの時間域と予測される時間を比較するため，No.2 から 5 までの全てのデータをプロットした．

各パラメータの平均値と標準偏差を表 3-2 に示す．$c_3$, $c_4$ 以外の標準偏差はモデルパラメータのばらつきが大きいことを示す．

表 3-2　モデルパラメータの平均値 av と標準偏差 σ

| | $a_0$ | $a_1$ | $a_2(10^{-6})$ | $a_3(10^{-2})$ | $a_4$ | $a_5(10^{-4})$ | $a_6$ | $a_7$ |
|---|---|---|---|---|---|---|---|---|
| av | 32 | 3.42 | 3.36 | 3.65 | 0.295 | 2.5 | 0 | 1.35 |
| σ | 36 | 4.29 | 1.88 | 0.68 | 0.009 | 5.0 | 0 | 1.89 |

ところで表 3-1 を見ると $c_6$ は全て零，$c_0$, $c_1$, $c_5$, $c_7$ は零になる場合がある．そこでこれらの基底関数を過剰適合の要因と考え，モデルから除外した式(14)を用い，表 3-1 と同じデータを用いてモデルパラメータを決定した．これを表 4-1 に示す．表 4-2 は前のモデルの標準偏差に比べ，$c_2$ のばらつきが小さくなり，予測のばらつきが改善されたことを示す．予測の確認結果を図 4 に示す．図 3 に比べ予測の範囲が狭まり，予測力が向上していることがわかる．

$$T_{MD}(p,n) = a_2 \cdot \frac{n^2}{p} + a_3 \cdot n + a_4 \cdot \frac{n}{p} \qquad (14)$$

表 4-1　QE により決定した式(14)のモデルパラメータ

| *No* | *1* | *2* | *3* | *4* | *5* |
|---|---|---|---|---|---|
| $a_2$ | 2.73･$10^{-6}$ | 2.39･$10^{-6}$ | 3.08･$10^{-6}$ | 2.74･$10^{-6}$ | 2.48･$10^{-6}$ |
| $a_3$ | 0.0416 | 0.0481 | 0.0448 | 0.0418 | 0.0476 |
| $a_4$ | 0.303 | 0.283 | 0.292 | 0.302 | 0.285 |

表 4-2　モデルパラメータの平均値 av と標準偏差 σ

| | $a_2(10^{-6})$ | $a_3$ | $a_4$ |
|---|---|---|---|
| av | 2.68 | 0.0448 | 0.293 |
| σ | 0.27 | 0.0031 | 0.009 |

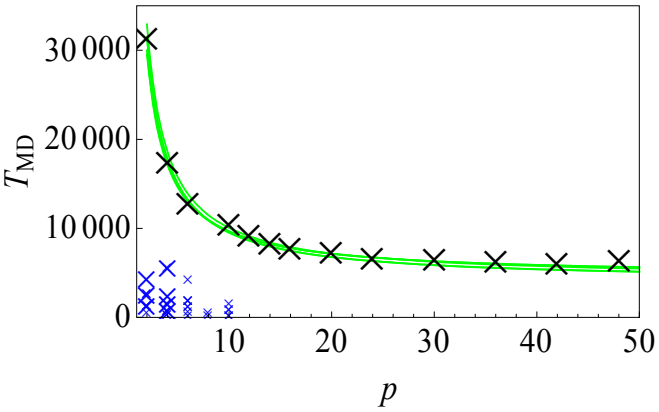


図 4　QE 法モデル（式(14)，表 4-1）の予測の確認

以上により，QE を用いると実際のモデル化においても不要な基底関数のモデルパラメータが零となり，過剰適合が緩和される場合があることが確かめられた．また零になったりならなかったりするモデルパラメータの基底関数を削除することにより，モデルの予測力が向上する場合があることがわかった．

QE により得られた予測力向上を客観的に確かめるため，式(12)に LSM を用い No.2 から No.5 のデータを用い，全てのモデルパラメータの初期値を零としてモデルパラメータを決定した．これを図 5 に示す．図 5 と図 3，4 を比較すると QE によるモデルは $p$ が増加した場合と $T_{MD}$ が大きい場合についての予測に優れていることがわかる．またプロセッサの増加に対する $T_{MD}$ の減少傾向を良く捉えていることがわかる．

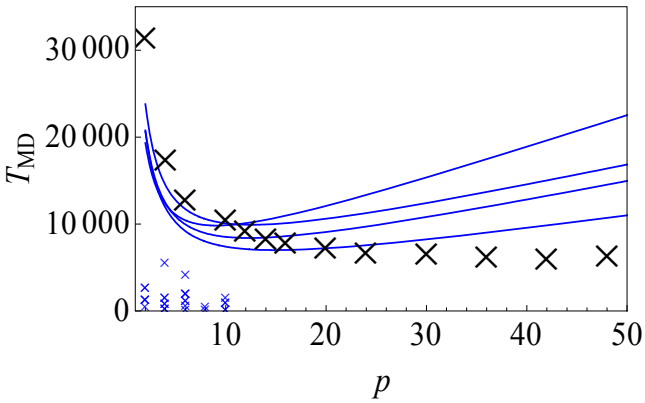


図 5　LSM 法で構築したモデルの予測の確認

## 6. おわりに

並列処理時間の回帰モデルの予測力向上のため，モデルパラメータに正負の条件を付けた連立残差不等式を，数式処理のアルゴリズムの一つ限量記号消去法(QE)を用いて解いてモデルパラメータを決定する方法を提案した．この方法は，モデルパラメータが零となることによりモデル構築の段階で不要な基底関数を自動的に排除して過剰適合を緩和する故，画一的な手順で予測力向上を図ることを期待できる．5.3 節ではこれを並列処理時間のモデル化に対して確かめたが，種々の現象を記述する回帰モデルにも適用できることが期待される．

## 参考文献

1) 折居茂夫：高並列処理における並列性能評価方法（Ⅱ），情報処理学会研究報告, 2010-HPC-126, No.48 (2010).
2) 折居茂夫 : 時間モデルを用いた並列処理の性能評価 - 並列化部に隠れた並列オーバーヘッド -, 2011-HPC-131, No.1 (2011).
3) Orii, O. and Anai, H.: Application of Quantifier Eliminator to Symbolic-Numeric Optimization in Biochemical Model, Research Communications in Biochemistry and Cell & Molecular Biology, Vol.12 (1&2), pp.73-89 (2008).
4) 折居茂夫 : 限量記号消去法による時間モデルパラメータの決定, 情報処理学会研究報告, 2012-HPC-134, No.11 (2012).
5) Mathematica, http://reference.wolfram.com/mathematica/guide/Mathematica.en.html
6) Watanabe, T. and Kaburaki, H. : Increase in chaotic motions of atoms in a large-scale self-organized motion, Phys, Rev. E, Vol.54, pp. 1504-1509 (1996).
7) 折居茂夫：レベル 1･2 並列ベンチマーク使用及びそれに基づくスカラ並列計算機 SP2 のベンチマークテスト，JAERI-Data/Code 98-020 (1998).