A priori inverse operator estimation for guaranteed error estimate

Akitoshi Takayasu (Waseda University, Japan)

Shin'ichi Oishi (Waseda University, Japan)
Takayuki Kubo (University of Tsukuba, Japan)

REC 2010 Mar.5

Purpose of this talk

Let $\Omega \subset \mathbf{R}^N \ \ (N = 1, 2)$ is convex polygonal domain.

We consider Drichlet boundary value problems of the form

$$\begin{cases} -\Delta u = f(u), & x \in \Omega, \\ u = 0, & x \in \partial\Omega. \end{cases}$$

Here, u : Exact solution, $\hat{u}$: Computational result,

'Guaranteed' Error Estimate:

$$\|u - \hat{u}\|_X \leqq Const.$$

X : suitable functional space,
$Const$: computable.

Purpose of this talk

Let $\Omega \subset \mathbf{R}^N \;\; (N = 1, 2)$ is convex polygonal domain.

We consider Drichlet boundary value problems of the form

$$\begin{cases} -\Delta u = f(u), & x \in \Omega, \\ u = 0, & x \in \partial\Omega. \end{cases}$$

Here, u : Exact solution, $\hat{u}$: Computational result,

'Guaranteed' Error Estimate:

$$\|u - \hat{u}\|_X \leq Const.$$

X : suitable functional space,
$Const$: computable.

Purpose of this talk

Let $\Omega \subset \mathbf{R}^N$ $(N = 1, 2)$ is convex polygonal domain.

We consider Drichlet boundary value problems of the form

$$\begin{cases} -\Delta u = f(u), & x \in \Omega, \\ u = 0, & x \in \partial\Omega. \end{cases}$$

Here, u : Exact solution, $\hat{u}$: Computational result,

'Guaranteed' Error Estimate:

$$\|u - \hat{u}\|_X \leqq Const.$$

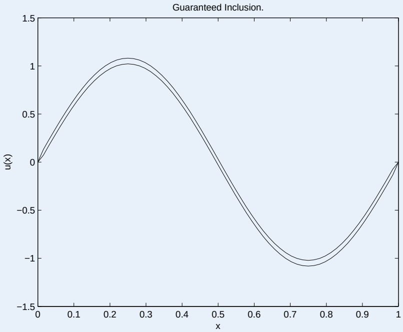
X : suitable functional space,
$Const$: computable.

'Guaranteed' Error Estimate

'Guaranteed' error proves...

- Existence of the exact solution.
- (Local) Uniqueness.
- Every error is included.
    - Discretization error.
    - Rounding error.

$\Rightarrow$ Mathematically guaranteed error by reliable computing.



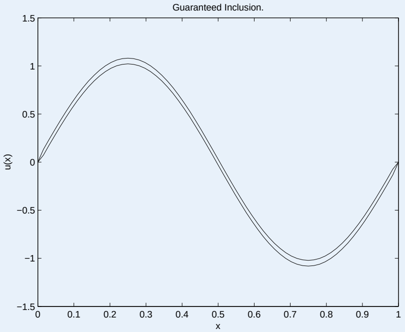
We can solve the problem with mathematically rigorous.

'Guaranteed' Error Estimate

'Guaranteed' error proves...

- Existence of the exact solution.
- (Local) Uniqueness.
- Every error is included.
  - Discretization error.
  - Rounding error.

$\Rightarrow$ Mathematically guaranteed error by reliable computing.



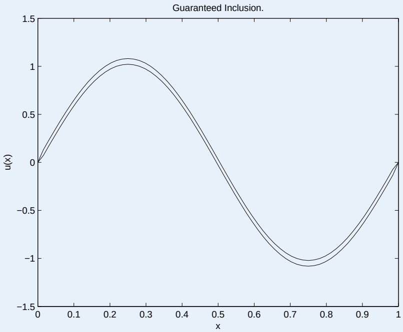
We can solve the problem with mathematically rigorous.

'Guaranteed' Error Estimate

'Guaranteed' error proves...

- Existence of the exact solution.
- (Local) Uniqueness.
- Every error is included.
  - Discretization error.
  - Rounding error.

$\Rightarrow$ Mathematically guaranteed error by reliable computing.



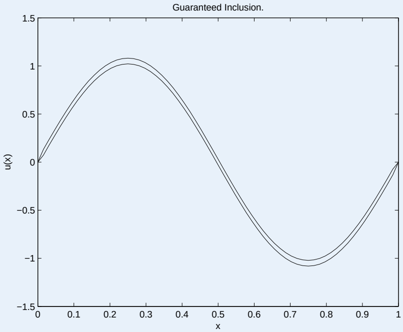
We can solve the problem with mathematically rigorous.

‘Guaranteed’ Error Estimate

‘Guaranteed’ error proves...

- Existence of the exact solution.
- (Local) Uniqueness.
- Every error is included.
  - Discretization error.
  - Rounding error.

$\Rightarrow$ Mathematically guaranteed error by reliable computing.



We can solve the problem with mathematically rigorous.

Preliminary

- $L^2(\Omega)$: Lebesgue-measurable square-integrable functions with

$$(f,g) = \int_0^1 f(x)g(x)dx \quad \text{and} \quad \|f\|_{L^2} = \sqrt{(f,f)}, \quad (f,g \in L^2).$$

- $H^m(\Omega)$: L^2-Sobolev space of order m
- With $(\nabla u, \nabla v)$ and $\|u\|_{H_0^1} = \|\nabla u\|_{L^2}$,

$$H_0^1(\Omega) = \{u \in H^1 : \mathrm{tr}(u) = 0 \ (\text{on } \partial\Omega)\}.$$

- $L^\infty(\Omega)$: Space of functions that are essentially bounded on Ω with the norm

$$\|u\|_\infty = \operatorname*{ess\,sup}_{a \leqq x \leqq b} |u(x)|.$$

Preliminary

- Let X and Y be Banach spaces. The set of linear bounded operators denoted by $\mathcal{L}(X,Y)$ with

$$\|L\|_{\mathcal{L}(X,Y)} = \sup_{u \in X \setminus \{0\}} \frac{\|Lu\|_Y}{\|u\|_X}, \quad (L \in \mathcal{L}(X,Y)).$$

  Here, $\|\cdot\|_X$ is the norm in X and $\|\cdot\|_Y$ is the norm in Y.
- Sobolev embedding theorem $(N = 2)$:
  (1) for $(k > l)$ the embedding $H^k(\Omega) \hookrightarrow H^l(\Omega)$ is compact and continuous,
  (2) let $q \in [1, \infty)$, the embedding $H_0^1(\Omega) \hookrightarrow L^q(\Omega)$ is compact and continuous,
  (3) and $H_0^1(\Omega) \subset L^p(\Omega)$ for $p \geq 2$ with

$$\|v\|_{L^p} \leq C_{e,p} \|v\|_{H_0^1}, \quad v \in H_0^1(\Omega), \quad \left(\text{ex. } N = 2, \quad C_{e,p} = \frac{p\sqrt{2}}{8}\right).$$

Preliminary

Let $\Omega = \mathbf{R}^N, \quad (N = 1, 2)$ be the convex bounded polygonal domain.

Consider Dirichlet boundary value problems

$$\begin{cases} -\Delta u = f(u), & x \in \Omega, \\ u = 0, & x \in \partial\Omega. \end{cases}$$

Let $f : H_0^1(\Omega) \to L^2(\Omega)$ be

$$f(u) = c_1 u + c_2 u^2 + c_3 u^3 + g.$$

$c_i \in L^\infty(\Omega), \ (i = 1, 2, 3), \ g \in L^2(\Omega).$

Preliminary

As an auxiliary problem, we consider Poisson equation.

For a given $g \in L^2$, the solution operator $\mathcal{K}$ corresponding to Poisson equation

$$\begin{cases} -\Delta u = g, & x \in \Omega, \\ u(x) = 0, & x \in \partial\Omega. \end{cases}$$

leads the weak solution $u = \mathcal{K}g \in H^2 \cap H_0^1$.

- Define the solution operator $\mathcal{K} : L^2 \to H^2 \cap H_0^1$.
- Since the embedding $H^2 \hookrightarrow H^1$ is compact, $\mathcal{K} \in \mathcal{L}(L^2, H_0^1)$ and $f : H_0^1 \to L^2$, the composite operator

$$\mathcal{K}f : H_0^1 \to H_0^1$$

  becomes compact operator.

Established Algorithm

Outline of established method

Dirichlet boundary problems

$$(P)\begin{cases} -\Delta u = f(u), & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases}$$

are transformed into

$$(P) \iff u = \mathcal{K}f(u),\ (u \in H_0^1(\Omega)).$$

by using solution operator $\mathcal{K}$.
Define the operator $\mathcal{F} : H_0^1 \to H_0^1$ Newton-Kantorovich theorem is applied to a nonlinear operator equation:

$$\iff \mathcal{F}(u) = u - \mathcal{K}f(u) = 0.$$

Newton-Kantorovich Theorem

Let $\hat{u} \in H_0^1$ *and* $\mathcal{F} : H_0^1 \to H_0^1$ *be Fréchet differentiable. Assume that*

(1) *the Fréchet derivative* $\mathcal{F}'(u)$ *is nonsingular and satisfies*

$$\|\mathcal{F}'(\hat{u})^{-1}\mathcal{F}(\hat{u})\|_{H_0^1} \leqq \alpha$$

for a certain positive α.

(2) $\forall v, w \in B(\hat{u}, 2\alpha) = \{v \in H_0^1 : \|v - \hat{u}\|_{H_0^1} \leqq 2\alpha\} \subset H_0^1$,

$$\|\mathcal{F}'(\hat{u})^{-1}(\mathcal{F}'(v) - \mathcal{F}'(w))\|_{\mathcal{L}(H_0^1, H_0^1)} \leqq \omega \|v - w\|_{H_0^1}$$

is obtained for a certain positive ω.

If $\alpha\omega \leqq \frac{1}{2}$, *then there exists a solution* $u^* \in H_0^1$ *of* $\mathcal{F}(u) = 0$ *satisfying*

$$\|u^* - \hat{u}\|_{H_0^1} \leqq \rho := \frac{1 - \sqrt{1 - 2\alpha\omega}}{\omega}.$$

Furthermore, the solution u^* *is unique in* $B(\hat{u}, \rho)$.

Newton-Kantorovich Theorem

For $u \in H_0^1$ we define an another linear operator $\mathcal{T} = \mathcal{T}(u) : H_0^1 \to H_0^1$ as

$$\mathcal{T}(u)v = \mathcal{K}(c_1 I + 2c_2 u + 3c_3 u^2)v.$$

Then the Fréchet derivative $\mathcal{F}'(u) : H_0^1 \to H_0^1$ is given as

$$\mathcal{F}'(u)v = (I - \mathcal{T}(u))v.$$

In order to valid the theorem, we compute three constants

$$\|\mathcal{F}'(\hat{u})^{-1}\|_{\mathcal{L}(H_0^1,H_0^1)} = \|(I - \mathcal{T}(\hat{u}))^{-1}\|_{\mathcal{L}(H_0^1,H_0^1)} \leqq C_1,$$

$$\|\mathcal{F}(\hat{u})\|_{H_0^1} = \|\hat{u} - \mathcal{K}f(\hat{u})\|_{H_0^1} \leqq C_2$$

and

$$\|\mathcal{F}'(v) - \mathcal{F}'(w)\|_{\mathcal{L}(H_0^1,H_0^1)} \leqq C_3\|v - w\|_{H_0^1}.$$

We set $\alpha = C_1 * C_2,\ \omega = C_1 * C_3$.

C_1 is tough to calculate

Theorem (Oishi 2000)

Let $\mathcal{T} : H_0^1 \to H_0^1$ be the compact operator, and $\mathcal{P}_n : H_0^1 \to X_n$ be the projection operator. Assume that there exist three constants K, L and M satisfying assumptions:

$$\|\mathcal{P}_n \mathcal{T}\|_{\mathcal{L}(H_0^1, H_0^1)} \leq K,$$

$$\|\mathcal{T} - \mathcal{P}_n \mathcal{T}\|_{\mathcal{L}(H_0^1, H_0^1)} \leq L$$

and

$$\|(I - \mathcal{P}_n \mathcal{T})^{-1}\|_{\mathcal{L}(H_0^1, H_0^1)} \leq M.$$

If $(1 + MK)L < 1$, then the operator $(I - \mathcal{K})$ has the unique inverse operator which is estimated as

$$\|(I - \mathcal{T})^{-1}\|_{\mathcal{L}(H_0^1, H_0^1)} \leq \frac{1 + MK}{1 - (1 + MK)L} =: C_1.$$

Inverse operator norm estimation algorithm

Compute rigorous upper bound of $\|(I-\mathcal{T})^{-1}\|_{\mathcal{L}(H_0^1,H_0^1)}$ by the following steps:

1 Compute $\|\hat{u}\|_{H_0^1}$ and calculate K and L by

$$K = C_{e,2}C_d(\|\hat{u}\|_{H_0^1}),$$

$$L = C_0 h C_d(\|\hat{u}\|_{H_0^1})$$

respectively.

2 Let D and G be $n \times n$ matrices whose i-j elements are given by

$$(\nabla\phi_j, \nabla\phi_i) \text{ and } (\nabla\phi_j, \nabla\phi_i) - (f'(\hat{u})\phi_j, \phi_i),$$

respectively. Let a lower triangular matrix L be the Cholesky decomposition of D, $D = LL^t$. If G is invertible, then set

$$M = \|L^t G^{-1} L\|_2.$$

When G is not invertible, stop with failure.

Inverse operator norm estimation algorithm

3 Check whether $(1+MK)L<1$ holds or not. If this holds, then the Fréchet derivative $\mathcal{F}'(u)$ is nonsingular. We have

$$\|\mathcal{F}'(\hat{u})^{-1}\|_{\mathcal{L}(H_0^1,H_0^1)} = \|(I-\mathcal{T})^{-1}\|_{\mathcal{L}(H_0^1,H_0^1)} \leqq \frac{1+MK}{1-(1+MK)L} =: C_1.$$

Otherwise, stop with failure.

Constant M takes much time to calculate.

Inverse operator norm estimation algorithm

3 Check whether $(1+MK)L<1$ holds or not. If this holds, then the Fréchet derivative $\mathcal{F}'(u)$ is nonsingular. We have

$$\|\mathcal{F}'(\hat{u})^{-1}\|_{\mathcal{L}(H_0^1,H_0^1)} = \|(I-\mathcal{T})^{-1}\|_{\mathcal{L}(H_0^1,H_0^1)} \leqq \frac{1+MK}{1-(1+MK)L} =: C_1.$$

Otherwise, stop with failure.

Constant M takes much time to calculate.

Omitting calculation of M

Theorem (Oishi 2000)

Let $\mathcal{T} : H_0^1 \to H_0^1$ *be the compact operator, and* $\mathcal{P}_n : H_0^1 \to X_n$ *be the projection operator. Assume that there exist three constants* K, L *and* M *satisfying assumptions:*

$$\|\mathcal{P}_n \mathcal{T}\|_{\mathcal{L}(H_0^1, H_0^1)} \le K,$$

$$\|\mathcal{T} - \mathcal{P}_n \mathcal{T}\|_{\mathcal{L}(H_0^1, H_0^1)} \le L$$

and

$$\|(I - \mathcal{P}_n \mathcal{T})^{-1}\|_{\mathcal{L}(H_0^1, H_0^1)} \le M.$$

If $(1 + MK)L < 1$, *then the operator* $(I - \mathcal{K})$ *has the unique inverse operator which is estimated as*

$$\|(I - \mathcal{T})^{-1}\|_{\mathcal{L}(H_0^1, H_0^1)} \le \frac{1 + MK}{1 - (1 + MK)L} =: C_1.$$

Omitting calculation of M

Proposition (Takayasu)

Let $\mathcal{T} : H_0^1 \to H_0^1$ be the compact operator, and $\mathcal{P}_n : H_0^1 \to X_n$ be the projection operator. Assume that there exist two constants K and L satisfying assumptions:

$$\|\mathcal{P}_n \mathcal{T}\|_{\mathcal{L}(H_0^1, H_0^1)} \leq K,$$

$$\|\mathcal{T} - \mathcal{P}_n \mathcal{T}\|_{\mathcal{L}(H_0^1, H_0^1)} \leq L$$

If $K + L < 1$, then the operator $(I - \mathcal{K})$ has the unique inverse operator which is estimated as

$$\|(I - \mathcal{T})^{-1}\|_{\mathcal{L}(H_0^1, H_0^1)} \leq \frac{1}{1 - K - L} =: C_1.$$

Omitting calculation of M

Proof:

Since

$$u = (I - \mathcal{T})u + (\mathcal{T} - \mathcal{P}_n\mathcal{T})u + \mathcal{P}_n\mathcal{T}u,$$

then we have

$$\begin{aligned}\|u\|_{H_0^1} &\leqq \|(I-\mathcal{T})u\|_{H_0^1} + \|\mathcal{T} - \mathcal{P}_n\mathcal{T}\|_{\mathcal{L}(H_0^1,H_0^1)}\|u\|_{H_0^1} + \|\mathcal{P}_n\mathcal{T}u\|_{H_0^1} \\ &\leqq \|(I-\mathcal{T})u\|_{H_0^1} + L\|u\|_{H_0^1} + K\|u\|_{H_0^1}.\end{aligned}$$

Therefore, if $K + L < 1$,

$$\|u\|_{H_0^1} \leqq \frac{1}{1-K-L}\|(I-\mathcal{T})u\|_{H_0^1}.$$

From this inequality, if $(I - \mathcal{T})u = 0$, $u = 0$ follows. This implies the operator $(I - \mathcal{T}) : H_0^1(\Omega) \to H_0^1(\Omega)$ is injective.

Omitting calculation of M

Since the operator $(I - \mathcal{T}) : H_0^1(\Omega) \to H_0^1(\Omega)$ is of Fredholm type with the index 0, it is also surjective. Thus, $I - \mathcal{T} : H_0^1(\Omega) \to H_0^1(\Omega)$ is invertible $u = (I - \mathcal{T})^{-1} v$ for $v \in H_0^1(\Omega)$ and enjoys

$$\|(I - \mathcal{T})^{-1} v\|_{H_0^1} \leqq \frac{1}{1 - K - L} \|v\|_{H_0^1}.$$

It shows

$$\|(I - \mathcal{T})^{-1}\|_{\mathcal{L}(H_0^1, H_0^1)} \leqq \frac{1}{1 - K - L} = C_1.$$

□

A priori estimation for Linear case $(c_2 = c_3 = 0)$

Linear Problem

Let us consider linear Drichlet boundary value problems of the form

$$(P)\begin{cases} -\Delta u = c_1 u + f, & x \in \Omega, \\ u = 0, & x \in \partial\Omega. \end{cases}$$

The same as the nonlinear case, by using the solutioin operator $\mathcal{K}$,

$$(P) \iff (I - \mathcal{K}c_1)u = \mathcal{K}f \iff (I - \mathcal{T})u = g.$$

where $(\mathcal{T} = \mathcal{K}c_1,\ g = \mathcal{K}f,\ f \in L^2(\Omega))$. If the operator $I - \mathcal{T}$ is invertible, there exits the unique solution with the error estimate

$$\begin{aligned} \|u - \hat{u}\|_{H_0^1} &= \|(I - \mathcal{T})^{-1}(I - \mathcal{T})(u - \hat{u})\|_{H_0^1} \\ &\le \|(I - \mathcal{T})^{-1}\|_{\mathcal{L}(H_0^1, H_0^1)} \|g - (I - \mathcal{T})\hat{u}\|_{H_0^1} \\ &\le C_1 C_2. \end{aligned}$$

Numerical Examples

Numerical example 1

Example (Nearly singular problem)

Let $\Omega = (0,1)$, we consider

$$\begin{cases} -u'' = cu + 1 & 0 < x < 1, \\ u(0) = u(1) = 0, \end{cases} \tag{ex1}$$

where $c = \pm 5, \pm 10$. Note that if $c = \pi^2 = 9.86...$ then this equation has no unique solution.

Numerical example 1

Verification Results of L^2-norm

Grid Points: 2^x	$c=-10$	$c=-5$	$c=5$	$c=10$
5	5.29×10^{-4}	1.06×10^{-3}	1.56×10^{-3}	$1.38\times10^{+0}$
6	1.35×10^{-4}	2.63×10^{-4}	3.89×10^{-4}	2.76×10^{-1}
7	3.40×10^{-5}	6.58×10^{-5}	9.72×10^{-5}	6.55×10^{-2}
8	8.54×10^{-6}	1.64×10^{-5}	2.43×10^{-5}	1.62×10^{-2}
9	2.14×10^{-6}	4.11×10^{-6}	6.07×10^{-6}	4.03×10^{-3}
10	5.36×10^{-7}	1.03×10^{-6}	1.52×10^{-6}	1.01×10^{-3}
11	1.34×10^{-7}	2.57×10^{-7}	3.80×10^{-7}	2.52×10^{-4}
12	3.35×10^{-8}	6.43×10^{-8}	9.52×10^{-8}	6.32×10^{-5}
13	-	1.64×10^{-8}	2.47×10^{-8}	-
14	-	4.85×10^{-9}	9.00×10^{-9}	-
Value of K	1.76	0.88	0.88	1.76

Numerical example 1

Execution Time ($c = 5$).

Grid Points: 2^x	Previous (s)	New (s)
5	0.061	0.037
6	0.092	0.043
7	0.151	0.048
8	0.279	0.051
9	0.704	0.055
10	2.279	0.063
11	10.42	0.091
12	44.62	0.136

CPU: 1.86 GHz Intel Core 2 Duo,
MATLAB 2009a with a toolbox for verified computations, INTLAB by Prof. Rump.

Numerical example 2

Example

$$\begin{cases} -u'' = u^3 - \cos 2\pi x & 0 < x < 1, \\ u(0) = u(1) = 0, \end{cases}$$

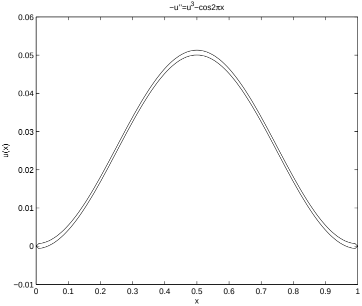


Figure: Guaranteed Inclusion of the Exact solution, $(n = 128)$.

Numerical example 2

Verification Results of H_0^1-norm

Grid Points: 2^x	Guaranteed Error: E_x	Ratio: $\frac{E_{x-1}}{E_x}$
5	4.99×10^{-3}	2.00
6	2.49×10^{-3}	2.00
7	1.25×10^{-3}	2.00
8	6.23×10^{-4}	2.00
9	3.12×10^{-4}	2.00
10	1.56×10^{-4}	2.00
11	7.79×10^{-5}	2.00
12	3.90×10^{-5}	2.00
13	1.95×10^{-5}	2.00
14	9.74×10^{-6}	2.00

Numerical example 2

The ratio of computational costs.

Grid Points: 2^x	Approximate (t_1)	Previous $(/t_1)$	New $(/t_1)$
5	8.86×10^{-4}	36.5	17.9
6	1.12×10^{-3}	43.8	15.1
7	1.08×10^{-3}	77.7	16.3
8	1.39×10^{-3}	118	12.7
9	2.03×10^{-3}	182	9.87
10	4.52×10^{-3}	216	7.11
11	1.86×10^{-2}	172	4.07
12	2.63×10^{-2}	423	3.26

t_1 is an approximate time.

CPU: 2.7 GHz, Quad-Core AMD Opteron(tm) Processor 8384,
MATLAB 2009a with a toolbox for verified computations, INTLAB.

Numerical example 3

Example

Let $\Omega = (0,1) \times (0,1)$,

$$\begin{cases} -\Delta u = u^2 + 10, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases}$$

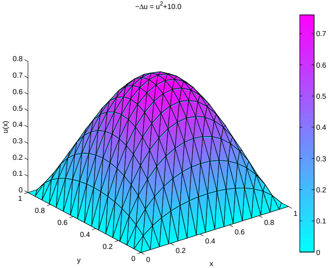


Figure: Shape of an approximate solution, (mesh size $1/16$).

Numerical example 3

Verification Results of H_0^1-norm

Grid Points: 2^x	Guaranteed Error: E_x	Ratio: $\frac{E_{x-1}}{E_x}$
5	2.08×10^{-1}	2.14
6	1.01×10^{-1}	2.06
7	4.98×10^{-2}	2.03
8	2.47×10^{-2}	2.01
9	1.23×10^{-2}	2.01
10	6.15×10^{-3}	2.00

Numerical example 3

The ratio of computational costs.

Mesh size: $1/2^x$	Approximate (t_1)	Previous ($/t_1$)	New ($/t_1$)
5	0.09	2.47	1.06
6	0.45	2.55	1.03
7	3.01	3.28	1.02
8	39.64	-	1.01
9	429.4	-	1.01

t_1 is an approximate time.

CPU: 2.7 GHz, Quad-Core AMD Opteron(tm) Processor 8384,
MATLAB 2009a with a toolbox for verified computations, INTLAB.

Summery

Conclusion

- Established method requires much time to compute M.
- Omitting M in case of an condition.
- For easy problem, it is possible to calculate with high speed.
- Ratio is 2 - 10 times more than approximation.

Future Works

- How do I solve for difficult problem (In case of $K > 1$) ?
- We use a posteriori estimate?

Thank You for your kind Attention!