平成 11 年度

卒業論文

Affine Arithmeticにおける除算について

On the Division of the Affine Arithmetic

平成 12 年 2 月 4 日

指導教授： 柏木 雅英 助教授

早稲田大学理工学部情報学科

G96P0568 白井 健一

目 次

1 序論 **5**

1.1 背景 6

1.2 本論文の目的 7

1.3 本論文の構成 7

2 区間演算 **8**

2.1 はじめに 9

2.2 区間 9

2.3 区間演算 9

2.4 むすび 11

3 Affine Arithmetic **12**

3.1 はじめに 13

3.2 Affine 形式 13

3.3 初期化 14

3.4 Affine 演算 15

3.5 Affine 演算 15

3.6 非 Affine 演算 16

3.7 単項演算の定義 16

3.8 非線形二項演算 17

3.9 むすび 19

4 Affine Arithmetic における除算 **20**

4.1 はじめに 21

4.2 従来の方法 21

4.3 新手法 21

4.4 最適な近似 24

4.4.1 $[x]$ が正の区間の場合 24

4.4.2 $[x]$ が負の区間の場合 26

4.4.3 [x] が 0 を含む区間の場合 28

4.5 むすび 29

5 従来法との比較 **30**

5.1 数値実験 31

5.2 考察 32

6 まとめ **34**

7 謝辞 **36**

参考文献 **38**

図 目 次

3.1 x と y に相関がない場合 14
3.2 x と y に相関がない場合 14
3.3 非線形関数の線形近似 17

4.1 $e_x(b)$ のグラフ 23
4.2 $e_x(b)$ の区間 [x] が正の区間の時のグラフ 25
4.3 $e_x(b)$ の区間 [x] が負の区間の時ののグラフ 27
4.4 $e_x(b)$ の区間 [x] が 0 を含む時のグラフ 28

5.1 $\tilde{e}_x(b)$ と $e_x(b)$ との間が狭い場合 33
5.2 $\tilde{e}_x(b)$ と $e_x(b)$ との間が広い場合 33

表 目 次

2.1 区間 [x],[y] の乗算 10
2.2 区間 [x],[y] の除算 10

4.1 $x \geq 0$ の場合の最大最小 22
4.2 $x < 0$ の時の最大最小 23

5.1 入力区間に対する係数の絶対値の値 32

第 1 章

序論

1.1 背景

現在のコンピュータ技術で連続数学の問題を計算する場合, 実数を浮動小数点で近似して計算するが, その場合, 計算機の浮動小数点数は有限個であるためこれだけでは連続関数の問題を扱うことは難しい. そのため, 2 つの浮動小数点を両端とする区間の中に連続関数の解を包み込み, その区間同士の加減剰余等の演算を演算結果として有り得る集合を包含するように定義する, 区間演算と言う方法が考えられた.

区間演算を使うために区間を数の拡張として考える方法が始められた. これは幅の充分狭い区間に求める解が含まれることがわかれば実用上充分だからである. 区間を数として考え, 四則演算を定義したものが区間演算である.

区間演算の大きな欠点として計算された区間は真の解を確かに含むものの, 区間幅が極端に大きくなることが挙げられる. 例として $f(x) = x^2 - 2x$ を例にとって考える. 通常の区間演算の手法で区間 $[0.9, 1.1]$ の演算を行なうと $[-1.39, -0.59]$ という結果になってしまう. 真の値は $[-1.0, -0.99]$ であるのでこれでは実に 80 倍の区間の開きになってしまう. この区間の増大の原因は二つの関数 x^2 と $2x$ が $[0.9, 1.1]$ の区間では同じ x の関数で互いに相関があり, しかも区間内で傾きが非常に近い値をとりそれの減算をするためであり, 区間演算における演算ではこの相関性を無視して独立な値として計算しているため例のように区間が爆発的に増大してしまうのである.

本卒業論文では, 区間演算で起こる爆発的な区間の増大を解決するために Mean Value Form 等の様々な手法が提案されているが, その方法の一つである Affine Arithmetic に注目する. この手法は区間演算の一種であり, 区間演算に似ているが, 区間演算と違うのは, 区間演算が一般の区間演算の区間が他の区間に全く影響しないのに対し,Affine Arithmetic では変量が互いに影響を及ぼすことである. この特徴により,Affine Arithmetic は計算式が膨大な場合や計算式が多くの変数を含んでいるような場合等に特にその効果を発揮する. 上の例を Affine Arithmetic で計算すると $[-1.0, -0.99]$ となる.

Affine Arithmetic は非線形関数をどのように計算するかによってその精度が大きく変化する.Affine Arithmetic では様々な単項計算を行なう非線形関数 ($\sin(x)$,$1/x$) は既にその最

良の計算方法が発見されているが,Affine 形式どうしの乗算を行なう方法は最良の方法がまだ見つかってない. しかし最良の近似を与えるわけではないが実用上の観点からは充分の精度をもつ. また,Affine 形式どうしの除算については今までは $x/y = x \times (1/y)$ というふうに, 単項演算 $1/x$ と Affine 形式どうしの乗算を合成して行なっていた. 最近では柏木により除算を直接計算する方法が発表された [2]. しかしその方法では計算時間を優先していて, 複雑な場合けが必要な部分を無視して近似して計算しているため最適ではない.

1.2 本論文の目的

本論文では柏木が提案した除算を直接計算する方法 [2] を改良し, 近似している部分の最適な解を求める方法を提案する. そして, 除算を 2 回にわけて行う方法と除算を直接計算する方法で近似した場合との差を考察する. 最後に実際の数値例を用いて本論文のアルゴリズムを実行する.

1.3 本論文の構成

本論文は次のような構成をとる.

- 第 2 章 精度保証付数値計算の基本的な手法である, 区間演算についての解説とその問題点を取り上げる.
- 第 3 章 Affine Arithmetic の定義とその有効性, また, 現状の解説をする.
- 第 4 章 従来の Affine Arithmetic の除算の手法を紹介し, 新たな除算のアルゴリズムを提案する.
- 第 5 章 実際にプログラムをくみ, 従来の方法との精度の比較を行い, 本手法の評価をする.

第 2 章

区間演算

2.1 はじめに

本章ではまず Affine Arithmetic の解説の前に区間演算の説明をする.

2.2 区間

区間演算は数値計算において最も一般的な手法であり,Affine Arithmetic の特徴は区間演算との比較で明確になる. 区間は実数値を [下限, 上限] という 2 つの浮動小数点で挟まれた区間で表現し

$$[\underline{x}, \overline{x}] = \{x \in R \mid \underline{x} \leq x \leq \overline{x}\}$$

とあらわす. ただし,$\underline{x} \leq \overline{x} \in R$ でそれぞれ区間の下端, 上端とする.

2.3 区間演算

区間演算は区間同士の加減乗除等の計算を「演算結果として有り得る集合を包含するように」定義することにより行なわれる.

2 つの区間 [x],[y] が与えられたときその区間の四則演算を次のように定義する.

$$[x] \circ [y] = \{x \circ y \mid x \in [x], y \in [y]\}$$

ただし $\circ \in \{+, -, \times, /\}$ とする. この定義では次が成立する.

$$\begin{aligned}
[x] + [y] &= [\underline{x} + \underline{y}, \overline{x} + \overline{y}] \\
[x] - [y] &= [\underline{x} - \overline{y}, \overline{x} - \underline{y}] \\
[x] \times [y] &= [min\{\underline{x}\underline{y}, \overline{x}\overline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}\}, max\{\underline{x}\underline{y}, \overline{x}\overline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}\}] \\
[x]/[y] &= [x] \times [1/\underline{y}, 1/\overline{y}], (0 \notin [y])
\end{aligned}$$

さらに乗算と除算については, 場合分けにより, より少ない手間で計算することができる (表 2.1, 表 2.2)

	$y \geq 0$	$y \ni 0$	$y \leq 0$
$x \geq 0$	$[\underline{x}\underline{y}, \overline{x}\overline{y}]$	$[\overline{x}\underline{y}, \overline{x}\overline{y}]$	$[\overline{x}\underline{y}, \underline{x}\overline{y}]$
$x \ni 0$	$[\underline{x}\overline{y}, \overline{x}\overline{y}]$	$[min\{\underline{x}\overline{y}, \overline{x}\underline{y}\}, max\{\underline{x}\underline{y}, \overline{x}\overline{y}\}]$	$[\overline{x}\underline{y}, \underline{x}\underline{y}]$
$x \leq 0$	$[\underline{x}\overline{y}, \overline{x}\underline{y}]$	$[\underline{x}\overline{y}, \underline{x}\overline{y}]$	$[\overline{x}\overline{y}, \underline{x}\underline{y}$

表 2.1: 区間 [x],[y] の乗算

	$y \geq 0$	$y \leq 0$
$x \geq 0$	$[\underline{x}/\overline{y}, \overline{x}/\underline{y}]$	$[\overline{x}/\overline{y}, \underline{x}/\underline{y}]$
$x \ni 0$	$[\underline{x}/\underline{y}, \overline{x}/\underline{y}]$	$[\overline{x}/\overline{y}, \underline{x}/\overline{y}]$
$x \leq 0$	$[\underline{x}/\underline{y}, \overline{x}/\overline{y}]$	$[\overline{x}/\underline{y}, \underline{x}/\overline{y}]$

表 2.2: 区間 [x],[y] の除算

区間演算について, 包含関係における単調性,

$$[x] \subseteq [x'], [y] \subseteq [y'] \longrightarrow [x] \circ [y] \subseteq [x'] \circ [y'], \circ \in \{+, -, \times, /\}$$

が成立する. また加法と乗法に関し交換則と結合則が成立する.

$$\begin{aligned} [x] \circ [y] &= [y] \circ [x], \circ \in \{+, \times\} \\ [x] \circ ([y] \circ [z]) &= ([x] \circ [y]) \circ [z], \circ \in \{+, \times\} \end{aligned}$$

しかし, 減法と除法の逆元は存在しない. すなわち,$-[x] = [-\overline{x}, -\underline{x}]$ であるが

$$0 = [0] \subseteq [x] - [x] = [\underline{x} - \overline{x}, \overline{x} - \underline{x}]$$

$$1 = [1] \subseteq [x]/[x]$$

となる. 上式で等号は $[x]$ が点区間のときのみ成立する.

また, 分配則も区間演算に対しては成立しない. その代わり次の劣分配則が成立する.

$$[x] \times ([y] + [z]) \subseteq [x] \times [y] + [x] \times [z]$$

上式で等号は区間 [y] と [z] が同じ符号を持つときに成立する.

2.4 むすび

本章では区間演算の概説を行った. 区間演算はその考え方が単純でわかりやすいが, 単純に区間演算を行っていくと, 区間の幅が増大し, その結果の値の使うのには非常に扱いにくくなる. それは区間が単純に演算可能になっているだけで, それぞれの変量の相関性などの内容についてはまったく考慮されないためである.Affine Arithmetic はこの区間演算の改良版であり, 区間演算よりも絞った区間で演算を行うことが可能である. 次章ではその仕組みと演算の実装を紹介する.

第 3 章

Affine Arithmetic

3.1 はじめに

本章では Affine Arithmetic の概念と定義の説明を行っていく.

3.2 Affine 形式

Affine Arithmetic は, 変数間の相関性を考慮することにより区間演算の区間の爆発的増大問題を解決する方法の一つである.

Affine Arithmetic では, 変量 x は affine 多項式

$$x = x_0 + x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n \tag{3.1}$$

で表される. ここで,x_i は実数であり,ε_i はそれが区間 [-1,1] に含まれることだけが分かっているような dummy 変数であり,x の変量の部分のそれぞれ独立した構成要素である.

x_0 をこの Affine form の central value, 係数 x_i を部分偏差 (partial deviations),ε_i を noise symbols という.noise symbol である,ε_i はそれぞれ独立した値をとって変量 x を構成し, 係数 x_i がそのふれ幅を決定する.

たとえば,Affine 形式 x, y があり, それが

$$x = 1 + 0.5\varepsilon_1 \tag{3.2}$$

$$y = 1 + 0.5\varepsilon_2 \tag{3.3}$$

とする. これは x と y に相関性がない状態で, このとき (x, y) がとり得る領域は図 3.1 のようになる.

これに対し,

$$x = 1 + 0.5\varepsilon_1 \tag{3.4}$$

$$y = 1 + 0.4\varepsilon_1 + 0.1\varepsilon_2 \tag{3.5}$$

では, それぞれがとり得る範囲は $[0.5, 1, 5]$ で変わらないが ε_1 の係数を見ると分かるように両者には強い相関があり (x, y) のとり得る範囲は図 3.2 のようになる.

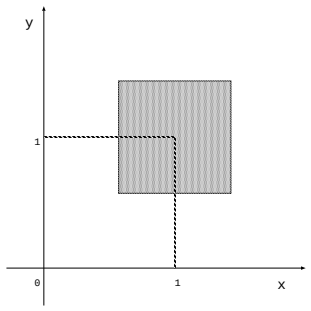

図 3.1: x と y に相関がない場合

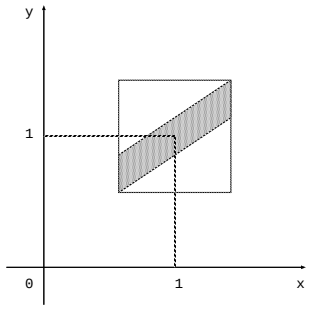

図 3.2: x と y に相関がない場合

3.3 初期化

n 変数の関数 $f(p_1, \cdots, p_n)$ の評価を行なう場合, 最初に与えられる n この変数 $p_1, \cdots, p_n$ は以下のような affine 多項式で初期化する.

$$\begin{aligned} p_1 &= \frac{\overline{p_1}+\underline{p_1}}{2} + \frac{\overline{p_1}-\underline{p_1}}{2}\varepsilon_1 \\ p_2 &= \frac{\overline{p_2}+\underline{p_2}}{2} + \frac{\overline{p_2}-\underline{p_2}}{2}\varepsilon_2 \\ &\vdots \\ p_n &= \frac{\overline{p_n}+\underline{p_n}}{2} + \frac{\overline{p_n}-\underline{p_n}}{2}\varepsilon_n \end{aligned} \tag{3.6}$$

但し, 入力変数 p_k の変域を $[\underline{p_k}, \overline{p_k}]$ とする.

Affine 多項式は, 次のようにして通常の区間に変換できる.

$$[x_0 - \delta, x_0 + \delta], \delta = \sum_{i=1}^{n} |x_i| \tag{3.7}$$

3.4 Affine 演算

2 つの Affine 多項式 x, y,

$$x = x_0 + x_1\varepsilon_1 + \cdots + x_n\varepsilon_n \tag{3.8}$$

$$y = y_0 + y_1\varepsilon_1 + \cdots + y_n\varepsilon_n \tag{3.9}$$

に対して, 以下に各種演算の定義を行う. 演算とは,

$$\begin{aligned} z &= f(x, y) \\ &= f(x_0 + x_1\varepsilon_1 + \cdots + x_n\varepsilon_n, y_0 + y_1\varepsilon_1 + \cdots + y_n\varepsilon_n) \end{aligned} \tag{3.10}$$

という $f(x, y)$ を,

$$z = z_0 + z_1\varepsilon_1 + \cdots + z_n\varepsilon_n \tag{3.11}$$

のように,Affine 多項式の形を満たすように表すことである.

各種 Affine 演算は次のように定義するが, それは線形なものと (Affine 演算), そうでないもの (非 Affine 演算) にわけられる.

3.5 Affine 演算

演算 f が線形な場合の演算は, 以下のように定義できる.

$$x \pm y = (x_0 \pm y_0) + (x_1 \pm y_1)\varepsilon_1 + \cdots + (x_n \pm y_n)\varepsilon_n \tag{3.12}$$

$$x \pm \alpha = (x_0 \pm \alpha) + x_1\varepsilon_1 + \cdots + x_n\varepsilon_n \tag{3.13}$$

定数倍も次のように定義できる.

$$\alpha x = (\alpha x_0) + (\alpha x_1)\varepsilon_1 + \cdots + (\alpha x_n)\varepsilon_n$$

3.6 非 Affine 演算

演算 f が非線形な場合の演算を考える. 次の式のように,z を ε の式で表すことを考える.

$$\begin{aligned} z &= f(x_0 + x_1\varepsilon_1 + \cdots + x_n\varepsilon_n, y_0 + y_1\varepsilon_1 + \cdots + y_n\varepsilon_n) \\ &= f^*(\varepsilon_1, \cdots, \varepsilon_n) \end{aligned} \tag{3.14}$$

この場合は一般的に z を正しく表すような $z_0, \cdots, z_n$ を決定することはできない.

よって次のような f^a を与え,

$$f^a(\varepsilon_1, \cdots, \varepsilon_n) = z_0 + z_1\varepsilon_1 + \cdots + z_n\varepsilon_n \tag{3.15}$$

f^* と f^a の誤差を $z_{n+1}\varepsilon_{n+1}$ として次のように表す.

$$\begin{aligned} z &= f^a(\varepsilon_1+, \cdots, \varepsilon_n) + z_{n+1}\varepsilon_{n+1} \\ &= z_0 + z_1\varepsilon_1 + \cdots + z_n\varepsilon_n + z_{n+1}\varepsilon_{n+1} \end{aligned} \tag{3.16}$$

ここで z_{n+1} とは f^* と f^a の誤差の最大値であるから,

$$|z_{n+1}| \geq max|f^*(\varepsilon_1, \cdots, \varepsilon_n) - f^a(\varepsilon_1, \cdots, \varepsilon_n| \tag{3.17}$$

である.

3.7 単項演算の定義

$f(x) = \sin(x)$ や $f(x) = e^2$ などの単項演算については, 一般に曲線 $y = f(x)$ を x_0 の付近で直線 $y = ax + b$ の形に近似し, そのときの最大誤差を δ として次のように計算できる.

$$f(x) = ax + b + \delta\varepsilon_{n+1} \tag{3.18}$$

ここで δ は関数 $f(x)$ と直線 $f(x) = ax + b$ の区間 $[\underline{x}, \overline{x}]$ における差の最大値であるから,

$$\delta = max|f(x) - (at + b)| \tag{3.19}$$

であり等号成立の時に最良の近似を与える. つまり直線 $y = ax + b$ をどのようにして設定し δ をいかに小さな値にするかが Affine の単項演算の焦点となる (図 3.3).

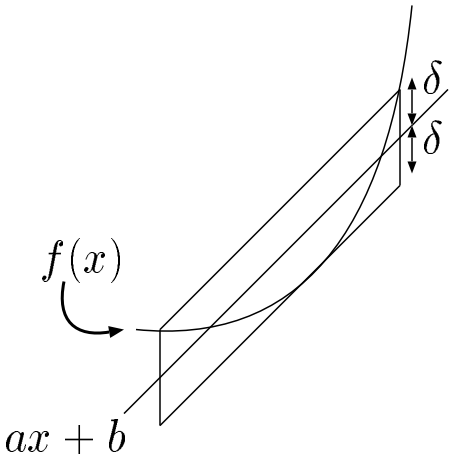

図 3.3: 非線形関数の線形近似

3.8 非線形二項演算

Affine 多項式 x, y:

$$x = x_0 + x_1\varepsilon + \cdots + x_n\varepsilon_n \tag{3.20}$$

$$y = y_0 + y_1\varepsilon + \cdots + y_n\varepsilon_n \tag{3.21}$$

に対する非線形な二項演算 $z = f(x, y)$ について考える. 二項演算の場合も,

$$\varepsilon_1, \cdots, \varepsilon_n \in [-1, 1] \times \cdots \times [-1, 1]$$

において,

$$\delta = \max_{-1 \leq \varepsilon_i \leq 1} |f(x, y) - (z_0 + z_1\varepsilon_1 + \cdots + z_n\varepsilon_n)| \tag{3.22}$$

を最小にするような Affine 多項式 z が求まれば最適であるが, これを直接解くことは極めて難しい. 単項演算のように $f(x,y)$ を最良近似するような一次方程式 $ax+by+c$ を求め,$ax+by+c+\delta\epsilon_{n+1}$ を結果とする方法を使うことはできるが, これが最良の近似を与えるとは限らない. このように二項演算の場合最良近似問題は未解決である. したがってそれに近い方法を使うことになる. 例えば乗算の場合,$x \times y$ は,

$$\begin{aligned} x \times y &= (x_0 + \sum_{i=1}^{n} x_i\varepsilon_i) \times (y_0 + \sum_{i=1}^{n} y_i\varepsilon_i) \\ &= x_0y_0 + \sum_{i=1}^{n}(x_0y_i + y_0x_i)\varepsilon + (\sum_{i=1}^{n} x_i\varepsilon_i) \times (\sum_{i=1}^{n} y_i\varepsilon_i) \\ &= f^a(\varepsilon_1, \cdots, \varepsilon_n) + (\sum_{i=1}^{n} x_i\varepsilon_i) \times (\sum_{i=1}^{n} y_i\varepsilon_i) \end{aligned} \tag{3.23}$$

ここで最後の項は ε どうしの積の形がでてしまい,Affine 形式にならない. そこでこの項の全体を新たに $\delta\varepsilon_{n+1}$ として計算する.

δ の値は単純に ε がその振れ幅の最大値をとるとして,

$$\begin{aligned} & max|(\sum_{i=1}^{n} x_i\varepsilon_i) \times (|sum_{i=1}^{n} y_i\varepsilon_i)| \\ &= (\sum_{i=1}^{n} |x_i|) \times (\sum_{i=1}^{n} |y_i|) \\ &= \delta \end{aligned}$$

として大きめに計算する.

よって乗算を計算するには,

$$x_0y_0 + \sum_{i=1}^{n}(x_0y_i + x_iy_0)\varepsilon_i + \delta\varepsilon_{n+1} \tag{3.24}$$

を行なえばよい.

この δ の値は, 先程大きめにとったので必ずしも常に最良の値をとるわけではないが, 必ず真の値を包含し, また計算しやすいので現時点では Affine 形式どうしの乗算の主流となっている.

また除算については今のところ, 第 3.7 節で紹介した単項演算 $f(x) = 1/x$ と本節で紹介

したAffine形式どうしの乗算を組み合わせて行なう方法

$$\frac{x}{y} = x \times (\frac{1}{y})$$

以外は知られていない.

3.9 むすび

第3.8節で説明したAffine Arithmeticの除算の方法では1回の計算で近似を2回行なうので計算の量が増えてしまうだけでなく,1回の計算でεが2つ増えてしまいロスが大きくなってしまう.次の章でこの除算を1回の近似で行なうアルゴリズムの解説をする.

第 4 章

Affine Arithmeticにおける除算

4.1 はじめに

本章では従来の Affine Arithmetic の除算方法と新手法の解説を行い, その後に新手法を改良するしたアルゴリズムの説明をする.

4.2 従来の方法

はじめに述べたように, 従来の除算の方法は $x/y = x \times (1/y)$ であるから, 逆関数と乗算の組み合わせにより, 計算している. この方法では,2 つの非線形演算を経由するので,ε は一回の除算で 2 つ増加してしまう.

4.3 新手法

まず柏木が提案した一回の非線形演算による除算の方法を説明する [2]. 乗算と同様に $\frac{x}{y}$ を $ax + by + c$ で近似する方法である. すなわち,

$$\max_{\substack{\underline{x} \le x \le \overline{x} \\ \underline{y} \le y \le \overline{y}}} |\frac{x}{y} - (ax + by + c)| \tag{4.1}$$

をなるべく小さくする a, b, c を決定する問題を考える.

ここで y を固定すると,$\frac{x}{y} - (ax + by + c)$ は x について線形, すなわち $|\frac{x}{y} - (ax + by + c)|$ を x の関数と見たとき最大となるのは,$x = \underline{x}$ または $x = \overline{x}$ である. すなわち式 4.1 を小さくする問題は,

$$max \left(\max_{\substack{\underline{x} \le x \le \overline{x} \\ \underline{y} \le y \le \overline{y}}} |\frac{\underline{x}}{y} - (ax + by + c)|, \max_{\substack{\underline{x} \le x \le \overline{x} \\ \underline{y} \le y \le \overline{y}}} |\frac{\overline{x}}{y} - (ax + by + c)| \right) \tag{4.2}$$

を小さくする問題と等価である. ここで,b を固定し,

$$p = \max_{\underline{y} \le y \le \overline{y}} (\frac{\underline{x}}{y} - by) \tag{4.3}$$

$$q = \min_{\underline{y} \le y \le \overline{y}} (\frac{\underline{x}}{y} - by) \tag{4.4}$$

$$r = \max_{\underline{y} \le y \le \overline{y}} (\frac{\overline{x}}{y} - by) \tag{4.5}$$

$$s = \min_{\underline{y} \le y \le \overline{y}} (\frac{\overline{x}}{y} - by) \tag{4.6}$$

とすると,

$$a\underline{x} + c = \frac{p+q}{2} \tag{4.7}$$

$$a\overline{x} + c = \frac{r+s}{2} \tag{4.8}$$

を満たすように a, c を決めれば式 4.2 は最小となり, その値は,

$$max\left(\frac{p-q}{2}, \frac{r-s}{2}\right) \tag{4.9}$$

となる. よって式 4.2 を小さくするような b を求める問題に帰着する. そこで,x をパラメータとする y の関数

$$f_x(y) = \frac{x}{y} - by \tag{4.10}$$

を考えその $\underline{y} \le y \le \overline{y}$ における最大最小を考えるとグラフの形から表 4.1, 表 4.2 のようになる.

b の値	最大値	最小値
$b \ge -\frac{x}{\overline{y}^2}$	$\frac{x}{\underline{y}} - b\underline{y}$	$\frac{x}{\overline{y}} - b\overline{y}$
$b \le -\frac{x}{\underline{y}^2}$	$\frac{x}{\overline{y}} - b\overline{y}$	$\frac{x}{\underline{y}} - b\underline{y}$
$\frac{x}{\underline{y}^2} < b < -\frac{x}{\overline{y}^2}$	$max\left(\frac{x}{\underline{y}} - b\underline{y}, \frac{x}{\overline{y}} - b\overline{y}\right)$	$2\sqrt{-bx}$

表 4.1: $x \ge 0$ の場合の最大最小

b の値	最大値	最小値
$b \leq -\frac{x}{\overline{y}^2}$	$\frac{x}{\overline{y}} - b\overline{y}$	$\frac{x}{\underline{y}} - b\underline{y}$
$b \geq -\frac{x}{\underline{y}^2}$	$\frac{x}{\underline{y}} - b\underline{y}$	$\frac{x}{\overline{y}} - b\overline{y}$
$\frac{x}{\overline{y}^2} < b < -\frac{x}{\underline{y}^2}$	$2\sqrt{-bx}$	$min\left(\frac{x}{\underline{y}} - b\underline{y}, \frac{x}{\overline{y}} - b\overline{y}\right)$

表 4.2: $x < 0$ の時の最大最小

$e_x(b)$ =(最大値 - 最小値) とすると,p,q は $x = \underline{x}$ のときの最大最小,r,s は $x = \overline{x}$ のときの最大最小であるから, 式 4.2 を小さくする問題は,

$$max(e_{\underline{x}}(b), e_{\overline{x}}(b)) \tag{4.11}$$

を小さくする問題になる.$e_{\underline{x}}(b)$ と $e_{\overline{x}}(b)$ のグラフは図 4.1 のようになるので, この 2 つのグラフの交点で最小値をとる.

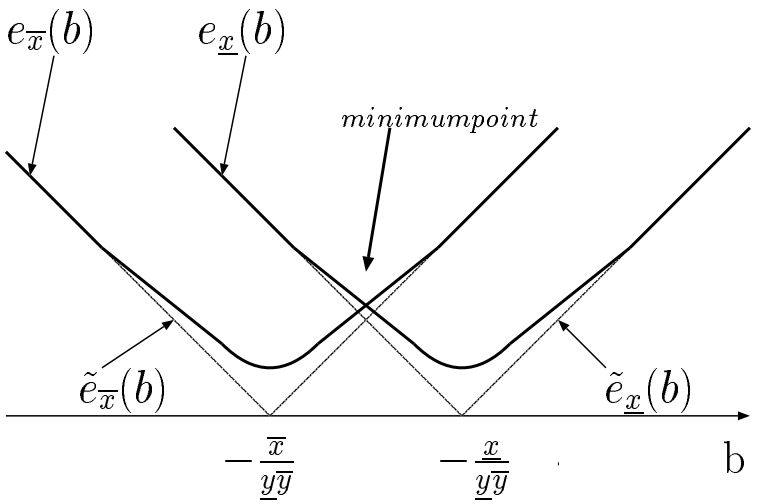

図 4.1: $e_x(b)$ のグラフ

グラフの式は分かっているので交点を求めることは可能であるが, 複雑な場合分けが必要となり, 計算量も増えることから今までは最適な b を求めるのは敬遠されてきた. そこで

$e_x(b)$ の近似

$$\tilde{e}_x(b) = (\overline{y} - \underline{y})|b + \frac{x}{\underline{y}\overline{y}}| \tag{4.12}$$

を使って b を近似的に算出している. この場合 b は,

$$b = -\frac{1}{\underline{y}\overline{y}} \frac{\underline{x} + \overline{x}}{2} \tag{4.13}$$

で求められる.

4.4 最適な近似

本節では第 4.3 節で説明した, 柏木の提案した方法を改良し最適な b を求める方法を説明する. 先程も述べたが複雑な計算が必要となる. まず,$[x] \geq 0$,$[x] \leq 0$,$[x] \in 0$ の三つの場合にわけて考える. 初期条件として、

$$bx \leq 0 \tag{4.14}$$

つまり,$x \geq 0$ のとき $b \leq 0$, $x \leq 0$ のとき $b \geq 0$ となる. ただし $b \neq 0$ である. これは $f_x(y)$ の増減を求める時の条件でもある.

4.4.1 $[x]$ が正の区間の場合

$\underline{x} \geq 0$ の場合, つまり区間 $[x]$ が正の場合から考える. 表 4.1 と図 4.1 から b の値が,$b \geq -\frac{x}{\overline{y}^2}$ のときと,$b \leq -\frac{x}{\underline{y}^2}$ のときに直線となり,$-\frac{x}{\underline{y}^2} \leq b \leq -\frac{x}{\overline{y}^2}$ の時曲線となる. この時 $e_{\underline{x}}$ と $e_{\overline{x}}$ ののグラフがどこで交わるかがポイントとなる. ここで重要なのは, 曲線と曲線が交わる時である. その時の条件は,

$$-\frac{\overline{x}}{\overline{y}^2} \geq -\frac{\underline{x}}{\underline{y}^2} \tag{4.15}$$

の時である (図 4.2).

この時の曲線の交点は次の 4 パターンある.

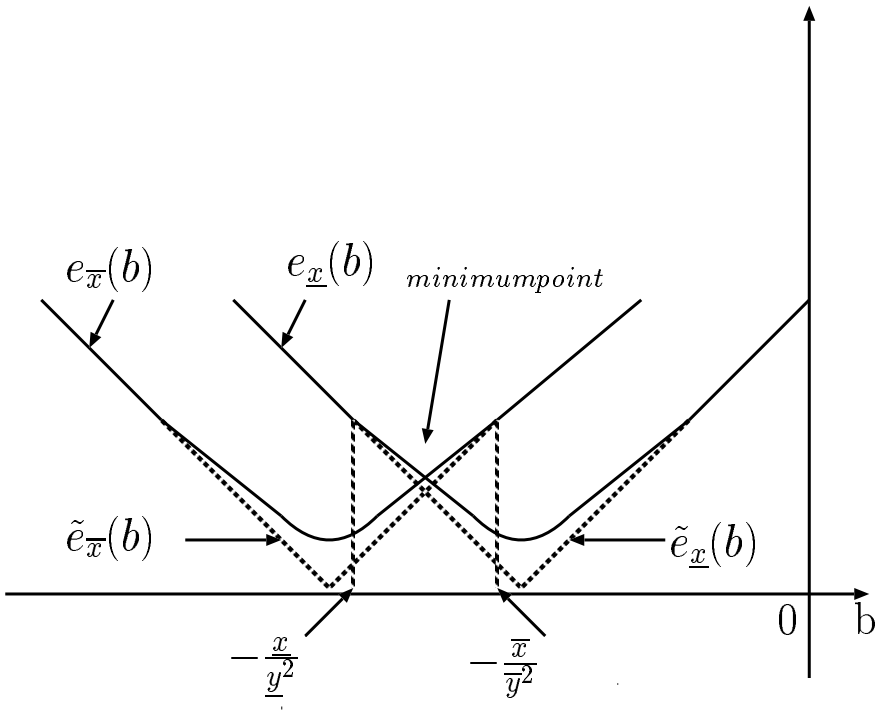

図 4.2: $e_x(b)$ の区間 [x] が正の区間の時のグラフ

$$\frac{\underline{x}}{\underline{y}} - b\underline{y} - 2\sqrt{-b\underline{x}} = \frac{\overline{x}}{\underline{y}} - b\underline{y} - 2\sqrt{-b\overline{x}} \tag{4.16}$$

$$\frac{\underline{x}}{\underline{y}} - b\underline{y} - 2\sqrt{-b\underline{x}} = \frac{\overline{x}}{\overline{y}} - b\overline{y} - 2\sqrt{-b\overline{x}} \tag{4.17}$$

$$\frac{\underline{x}}{\overline{y}} - b\overline{y} - 2\sqrt{-b\underline{x}} = \frac{\overline{x}}{\underline{y}} - b\underline{y} - 2\sqrt{-b\overline{x}} \tag{4.18}$$

$$\frac{\underline{x}}{\overline{y}} - b\overline{y} - 2\sqrt{-b\underline{x}} = \frac{\overline{x}}{\overline{y}} - b\overline{y} - 2\sqrt{-b\overline{x}} \tag{4.19}$$

これらの式を解くとそれぞれ,

$$b = \frac{-(\overline{x} - \underline{x})^2}{4\underline{y}^2(\overline{x} + \underline{x} + 2\sqrt{\overline{x}\underline{x}})} \tag{4.20}$$

$$b = \frac{B \pm \sqrt{B^2 - (\overline{y} - \underline{y})^2 A^2}}{(\overline{y} - \underline{y})^2} \tag{4.21}$$

$$
\begin{aligned}
(A &= \frac{\overline{x}\underline{y} - \underline{x}\overline{y}}{\overline{y}\underline{y}}, B = A(\overline{y} - \underline{y}) + 2(-\overline{x} - \underline{x} + 2\sqrt{\overline{x}\underline{x}})) \\
b &= \frac{-(\overline{x} - \underline{x})^2}{4\overline{y}^2(\overline{x} + \underline{x} + 2\sqrt{\overline{x}\underline{x}})} \qquad (4.22) \\
b &= \frac{B' \pm \sqrt{B'^2 - (\overline{y} - \underline{y})^2 A'^2}}{(\overline{y} - \underline{y})^2} \qquad (4.23) \\
(A' &= \frac{\underline{x}\underline{y} - \overline{x}\overline{y}}{\overline{y}\underline{y}}, B' = A'(\overline{y} - \underline{y}) + 2(-\overline{x} - \underline{x} + 2\sqrt{\overline{x}\underline{x}}))
\end{aligned}
$$

この 6 つの候補の中から, どの解を選ぶかが次の問題である.b の解の中で, 次の条件

$$
-\frac{\overline{x}}{\underline{y}\overline{y}} \leq b \leq -\frac{\underline{x}}{\underline{y}\overline{y}} \qquad (4.24)
$$

をみたし, その中で最小値をとるものを選べば良い. そうする事によって,$x \geq 0$ の場合, 式 4.11 は最小の値をとる. また, 式 4.15 を満たさない場合については直線と直線の場合なので,

$$
b = -\frac{1}{\underline{y}\overline{y}} \frac{\underline{x} + \overline{x}}{2} \qquad (4.25)
$$

で計算し求めることができる.

4.4.2 $[x]$ が負の区間の場合

負の区間の場合も, まず曲線の部分の条件は,

$$
-\frac{\underline{x}}{\underline{y}^2} \geq -\frac{\overline{x}}{\overline{y}^2} \qquad (4.26)
$$

である (図 4.3). そして, 正の区間の場合と同様にして, 曲線の交点を求めるには,

$$
\begin{aligned}
-2\sqrt{-b\underline{x}} - (\frac{\underline{x}}{\underline{y}} - b\underline{y}) &= -2\sqrt{-b\overline{x}} - (\frac{\overline{x}}{\underline{y}} - b\underline{y}) \qquad (4.27) \\
-2\sqrt{-b\underline{x}} - (\frac{\underline{x}}{\underline{y}} - b\underline{y}) &= -2\sqrt{-b\overline{x}} - (\frac{\overline{x}}{\overline{y}} - b\overline{y}) \qquad (4.28) \\
-2\sqrt{-b\underline{x}} - (\frac{\underline{x}}{\overline{y}} - b\overline{y}) &= -2\sqrt{-b\overline{x}} - (\frac{\overline{x}}{\underline{y}} - b\underline{y}) \qquad (4.29) \\
-2\sqrt{-b\underline{x}} - (\frac{\underline{x}}{\overline{y}} - b\overline{y}) &= -2\sqrt{-b\overline{x}} - (\frac{\overline{x}}{\overline{y}} - b\overline{y}) \qquad (4.30)
\end{aligned}
$$

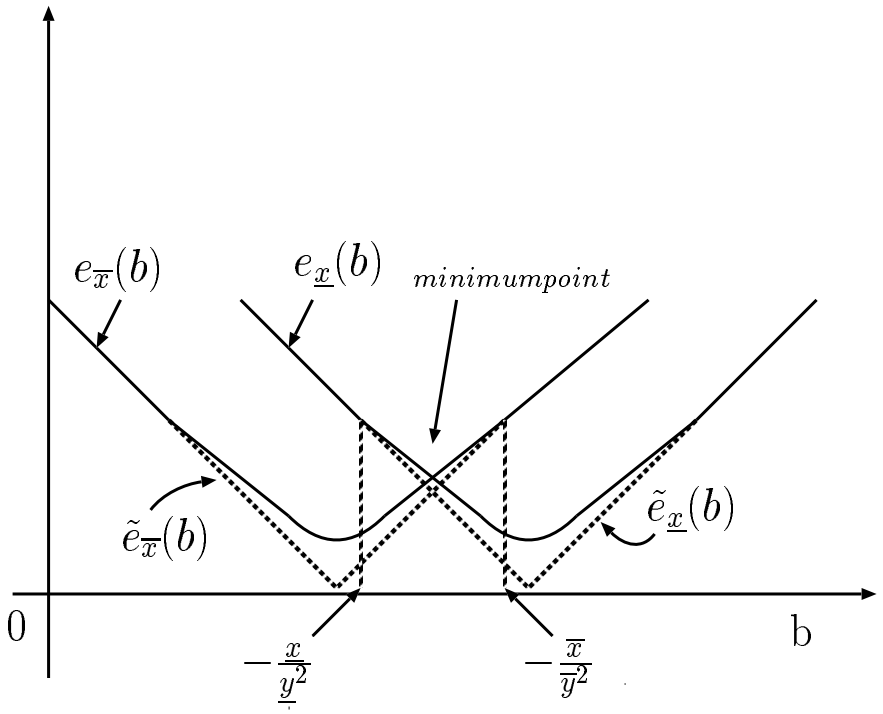

図 4.3: $e_x(b)$ の区間 [x] が負の区間の時ののグラフ

の 4 つのパターンがあり, これらを解いて

$$b = \frac{(\overline{x}-\underline{x})^2}{4\underline{y}^2(-\underline{x}-\overline{x}-2\sqrt{\underline{x}\overline{x}})} \tag{4.31}$$

$$b = \frac{-B \pm \sqrt{B^2 - A^2(\overline{y}-\underline{y})^2}}{(\overline{y}-\underline{y})^2} \tag{4.32}$$

$$(A = \frac{\underline{x}\overline{y}-\overline{x}\underline{y}}{\underline{y}\overline{y}}, B = A(\overline{y}-\underline{y}) - 2(-\underline{x}-\overline{x}-2\sqrt{\underline{x}\overline{x}}))$$

$$b = -\frac{(\overline{x}-\underline{x})^2}{4\overline{y}^2(-\underline{x}-\overline{x}-2\sqrt{\underline{x}\overline{x}})} \tag{4.33}$$

$$b = \frac{-B' \pm \sqrt{B'^2 - A'^2(\overline{y}-\underline{y})^2}}{(\overline{y}-\underline{y})^2} \tag{4.34}$$

$$(A' = \frac{\underline{x}\underline{y}-\overline{x}\overline{y}}{\underline{y}\overline{y}}, B' = A'(\underline{y}-\overline{y}) - 2(-\underline{x}-\overline{x}-2\sqrt{\underline{x}\overline{x}}))$$

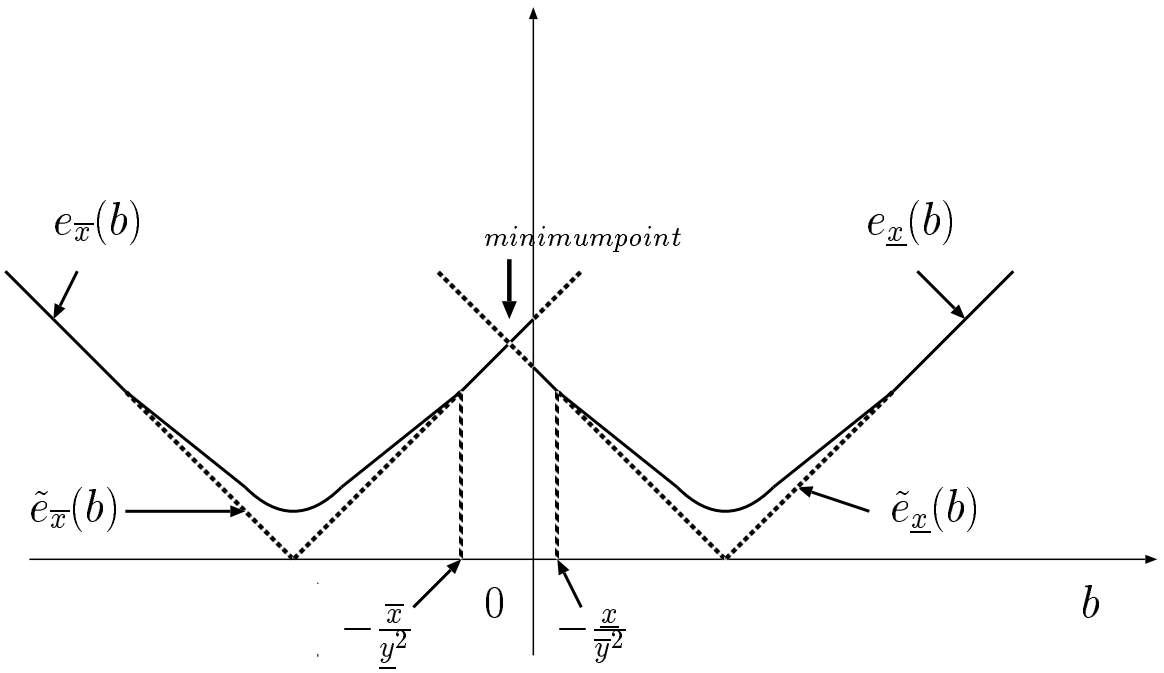

図 4.4: $e_{x}(b)$ の区間 [x] が 0 を含む時のグラフ

となる. この解の中で

$$-\frac{\underline{x}}{\underline{y}\overline{y}} \leq b \leq -\frac{\overline{x}}{\underline{y}\overline{y}} \tag{4.35}$$

を満たし, 最大のものを選ぶ. これで $x \leq 0$ の場合, 式 4.11 は最小値をとる. また, 式 4.26 を満たさないときは, 正の区間の時と同様に,

$$b = -\frac{1}{\underline{y}\overline{y}}\frac{\underline{x}+\overline{x}}{2} \tag{4.36}$$

で求める.

4.4.3 [x] が 0 を含む区間の場合

区間に 0 を含む場合を考える. その時のグラフは図 4.4 のようになる.b の値は,$x \geq 0$ の時 $b \leq 0$,$x \leq 0$ の時 $b \geq 0$ である. 図 4.4 のように 0 を含む区間では,$e_{\underline{x}}(b)$ のグラフと $e_{\overline{x}}(b)$ のグ

ラフは曲線部分同士で交わることは無い. それは, 直線と曲線を区切る点が必ず存在し,0 の付近では $b = 0$ の時以外は必ず直線でなければ解をもたないからである. よってこの場合は,

$$b = -\frac{1}{\underline{y}\overline{y}} \frac{\underline{x} + \overline{x}}{2} \tag{4.37}$$

で b を算出する.

4.5 むすび

本章では Affine Arithmetic の除算の新手法のと, それを改良したアルゴリズムの解説を行った. 次章で実際にプログラムで実行して従来法と最適な近似解を求めた場合の除算の比較をする.

第 5 章

従来法との比較

本章では従来法と近似の善し悪しをいくつかの例により比較して見る.

5.1 数値実験

本論文で考察した, 除算を 1 回の近似で行なう方法で,b を近似して求める方法と,b を近似せず最適に求める場合と,2 回の近似にわけて行なう方法の比較を行なう.

例えば区間 $[1,3]$ を区間 $[2,5]$ で除算する場合,2 回にわけて行う場合のプログラムでの実行結果は, 初期値を

$$x = 2 + \varepsilon_1 \tag{5.1}$$

$$y = 3.5 + 1.5\varepsilon_2 \tag{5.2}$$

として実行すると次の結果が得られる .

```
X=
1 3
Y=
2 5
X / Y = (0.632456 ,0.316228 ,-0.3 ,0.0675445 ,0.183772 )
```

ここで $X/Y = (x_0, x_1\varepsilon_1, \cdots, x_n\varepsilon_n, \delta\varepsilon_{n+1})$ とする. さらに 1 回の近似で行う方法で b を近似して求める方法では,

```
X=
1 3
Y=
2 5
b=-0.200000
X / Y = (0.685905 ,0.338692 ,-0.3 ,0.175403 )
```

となり, 最適な b を求める方法の実行結果は,

```
X=
1 3
Y=
2 5
b=-0.209
X / Y = (0.683544 ,0.335305 ,-0.314695 ,0.166456 )
```

となる.

次にいろいろな数値で試した結果を掲載する. なお除算を 1 回の近似で行う方法は, 増えた ε の係数の絶対値を加えるものとする. 括弧内が加えた数である.

入力区間	従来法 $(x \times (1/y)$	b を近似する方法	最適な b を求める方法
[1,3]/[2,5]	$0.184 + 0.0675(0.251)$	0.175	0.166
[3,4]/[0.1,0.2]	$1.464 + 1.501(2.965)$	2.292	2.206
[9,10]/[1,4]	$0.25 + 1.187\ (1.437)$	1.314	1.293
[9,10]/[3,4]	$0.0223 + 0.0284\ (0.051)$	0.041	0.039
[-2,-1]/[-2,-1]	$0.146 + 0.0643\ (0.210)$	0.125	0.125
[-3,-1]/[2,5]	$0.184 + 0.0675\ (0.251)$	0.175	0.166
[-10,10]/[2,3]	$0.918 + 0\ (0.918)$	0.833	0.833

表 5.1: 入力区間に対する係数の絶対値の値

5.2 考察

数値実験の結果, 近似しない場合と比べて, 曲線と曲線が交わる部分での値は確かに精度が上がっている. 上の例の中では,b を近似する場合と最適な b を求める場合との差は最大

で,0.086 の誤差がでている. 最小誤差は 0.002 である . この最大誤差と最小誤差の違いは, グラフの形から, 近似したときの直線 $\tilde{e}_x(b)$ のグラフの交点と本当の $e_x(b)$ のグラフとの交点では $e_x(b)$ の極小点と極小点が近ずくにつれ, 曲線部分と近似した直線 $\tilde{e}_x(b)$ と間が広がるために差が出るものと思われる (図 5.1, 図 5.2). 近似しない場合の結果と同じものは直線同士が交わる場合である.

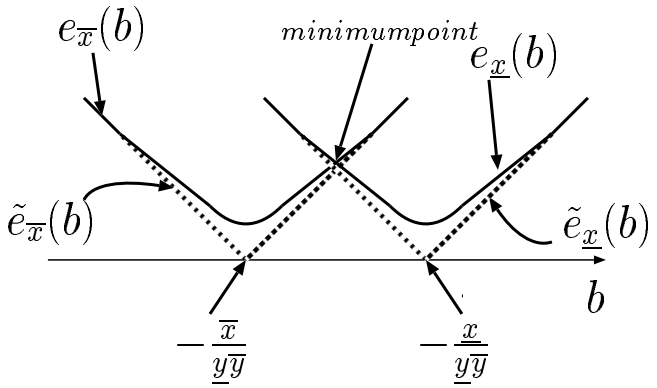

図 5.1: $\tilde{e}_x(b)$ と $e_x(b)$ との間が狭い場合

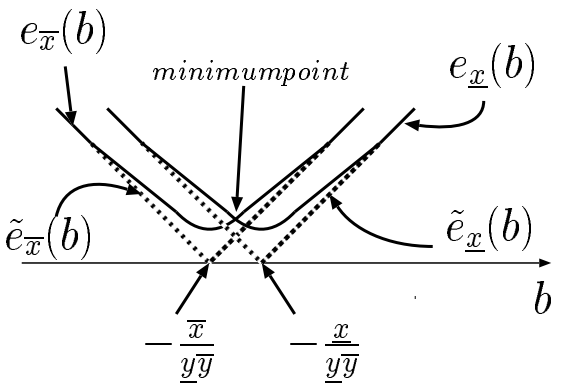

図 5.2: $\tilde{e}_x(b)$ と $e_x(b)$ との間が広い場合

第 6 章

まとめ

本論文では第 2 章で精度保証付数値計算の基礎である区間演算の解説を行ない, 第 3 章から Affine Arithmetic の定義の解説を行なった.

そして第 4 章で柏木の提案した Affine Arithmetic における除算の $x/y = x \times (1/y)$ の形に分解しない方法 [2] を解説し, さらに最適な b を求める方法を示した. 第 5 章では本論文で提案した方法を用いれば,ε の増加も一つで済み, 少しの場合であるが従来の方法よりも増加する ε の係数も小さくてすむことを示した. 本手法では近似をしない場合と比べて計算回数は多少増加する. 実際には気にするほどでもないがこの計算量をいかに減らすかが今後の課題である. これで Affine Arithmetic における最適な除算の方法を示すことができた.

第 7 章

謝辞

本研究を進めるに当たり常に正しい道を示し、終始丁寧な御指導及び御激励を賜り、その他多くの面でもいろいろと御面倒を見て下さり御助言を与えて下さいました柏木雅英 助教授に深く感謝致します。

また、卒業論文中間報告の際など機会のあるごとに御指導、御鞭撻を賜り、研究に方向正を与えて下さった大石進一 教授に深く感謝致します。

柏木研において柏木研助手 相馬隆郎 氏 , 博士課程 2 年 宮田孝富 氏には, 非線形の分野について熱心に御指導してくださり, また, 日頃の柏木研究室内でのあらゆる面において御教示頂いたことに, 心より感謝申し上げます.

柏木研修士課程 1 年高崎大輔 氏 , 岩折朱希嗣 氏には, 中間発表の準備の際に適切なアドバイスを頂き, また私のコンピュータに関する稚拙な疑問にも親身に御指導を賜り, 心より感謝申し上げます.

最後に, 意見の交換や協力などして下さいました柏木究室学部 4 年の皆様, 金谷 卓充 氏, 川上 修 氏, 小泉 健 氏, 櫻井幹夫 氏, 洲浜 陽一 氏, 長友 泰崇 氏, 波多野 伸哉 氏, 深谷 光統 氏, 村竹 範彦 氏, 山田 浩之 氏, 吉田 直史 氏, 渡部 啓 氏 に心より感謝致します.

参考文献

[1] 大石進一:“数値計算”,(応用解析セミナー), 裳華房,1999

[2] 柏木雅英: “Affine Arithmetic における除算について”, 信学技報 ,NLP98-114,PP23-30,(1999-03)

[3] 柏木雅英: “精度保証付シミュレーション [1]”, 日本シミュレーション学会,1999 年 12 月 第 18 巻 4 号

[4] 長嶋雄一郎:“Affine Arithmetic の除算の実現”, 平成 11 年度卒業論文 (柏木研究室所属)