Bibiliothèque efficace d'arithmétique d'intervalles en C

Samir SAYAH Guillaume NGUYEN
Vincent LIU Hala BEN ALI

Encadrant: Stef GRAILLAT

29 mai 2018

POLYTECH SORBONNE

Motivations

De nos jours, la fiabilité des calculs est primordiale

Inconvénients de la représentation des nombres réels avec la norme IEEE 754

Pourquoi l'arithmétique d'intervalles ?

- Prise en compte des incertitudes
- Calcul garanti

Motivations

De nos jours, la fiabilité des calculs est primordiale

Inconvénients de la représentation des nombres réels avec la norme IEEE 754

Pourquoi l'arithmétique d'intervalles ?

- Prise en compte des incertitudes
- Calcul garanti

Notre but va être de proposer un format compressé permettant de stocker un intervalle dans un flottant

Representation des nombres réels

Problème : Représenter l'ensemble $\mathbb{R}$ en machine

Representation des nombres réels

Problème : Représenter l'ensemble $\mathbb{R}$ en machine

Solution : Représenter l'ensemble **fini** $\mathbb{F}$ des nombres flottants

Representation des nombres réels

Problème : Représenter l'ensemble $\mathbb{R}$ en machine

Solution : Représenter l'ensemble **fini** $\mathbb{F}$ des nombres flottants

Norme IEEE 754

Soit $x \in \mathbb{F}$, alors :

$$x = (-1)^{Signe} \times 1, \underbrace{M_1 M_2 \ldots M_n}_{Mantisse} \times 2^{exposant}$$

où $M_i = 0$ ou 1, donc la mantisse est une suite de 0 et de 1.

Exemple

Comment représenter le nombre décimal 8,625 en norme IEEE 754 et en simple précision ?

Exemple

Comment représenter le nombre décimal 8,625 en norme IEEE 754 et en simple précision ?

- $8,625 = 2^3 + 2^{-1} + 2^{-3}$.
  Donc : 8,625 en binaire vaut : $8,625_{10}$=$1000,101_2$

Exemple

Comment représenter le nombre décimal 8,625 en norme IEEE 754 et en simple précision ?

- $8,625 = 2^3 + 2^{-1} + 2^{-3}$.
  Donc : 8,625 en binaire vaut : $8,625_{10}$=$1000,101_2$
- 1000,101 en écriture scientifique vaut : $1,\mathbf{000101}.2^{\mathbf{3}}$.

Exemple

Comment représenter le nombre décimal 8,625 en norme IEEE 754 et en simple précision ?

- $8,625 = 2^3 + 2^{-1} + 2^{-3}$.
  Donc : 8,625 en binaire vaut : $8,625_{10}$=$1000,101_2$
- 1000,101 en écriture scientifique vaut : $1,\mathbf{000101}.2^{\mathbf{3}}$.
- L'exposant décalé vaut :
  E = exposant + décalage = 3 + 127 = 130.
  (en simple précision, décalage = 127).
  E en binaire vaut : **10000010**.

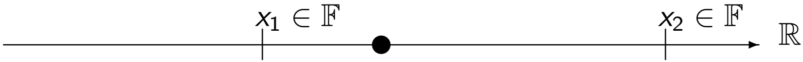
Exemple

Comment représenter le nombre décimal 8,625 en norme IEEE 754 et en simple précision ?

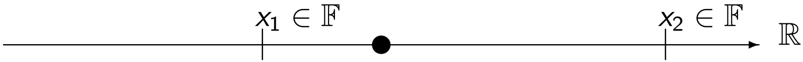
- $8,625 = 2^3 + 2^{-1} + 2^{-3}$.
  Donc : 8,625 en binaire vaut : $8,625_{10}$=$1000,101_2$
- 1000,101 en écriture scientifique vaut : $1,\mathbf{000101}.2^{\mathbf{3}}$.
- L'exposant décalé vaut :
  E = exposant + décalage = 3 + 127 = 130.
  (en simple précision, décalage = 127).
  E en binaire vaut : **10000010**.

Signe	**E**	**Mantisse**
0	**10000010**	**000101**00000000000000000

L'ensemble des nombres flottants



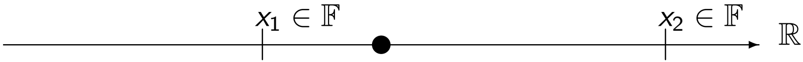
L'ensemble des nombres flottants



4 modes d'arrondis :

- Vers $-\infty$

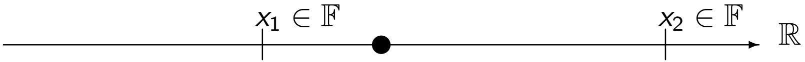
L'ensemble des nombres flottants



4 modes d'arrondis :

- Vers $-\infty$
- Vers $+\infty$

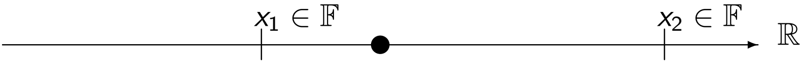
L'ensemble des nombres flottants



4 modes d'arrondis :

- Vers $-\infty$
- Vers $+\infty$
- Vers 0

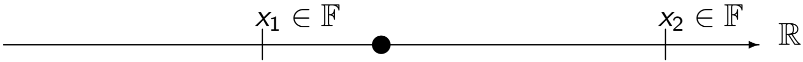
L'ensemble des nombres flottants



4 modes d'arrondis :

- Vers $-\infty$
- Vers $+\infty$
- Vers 0
- Au plus près

L'ensemble des nombres flottants



4 modes d'arrondis :

- Vers $-\infty$
- Vers $+\infty$
- Vers 0
- Au plus près

Problème : Erreur d'arrondis lors des calculs

L'arithmétique d'intervalles

Problème : Erreur d'arrondis lors des calculs impliquant des nombres flottants

L'arithmétique d'intervalles

Problème : Erreur d'arrondis lors des calculs impliquant des nombres flottants

Solution : Manipuler des intervalles contenant ces nombres

L'arithmétique d'intervalles

Problème : Erreur d'arrondis lors des calculs impliquant des nombres flottants

Solution : Manipuler des intervalles contenant ces nombres

Il existe deux manières de représenter un intervalle :

- Notation inf-sup : $I = [a, b] = [c - r, c + r]$

L'arithmétique d'intervalles

Problème : Erreur d'arrondis lors des calculs impliquant des nombres flottants
Solution : Manipuler des intervalles contenant ces nombres
Il existe deux manières de représenter un intervalle :

- Notation inf-sup : $I = [a, b] = [c - r, c + r]$
- Notation Centre-Rayon : $I =< c, r >$

L'arithmétique d'intervalles

Problème : Erreur d'arrondis lors des calculs impliquant des nombres flottants

Solution : Manipuler des intervalles contenant ces nombres

Il existe deux manières de représenter un intervalle :

- Notation inf-sup : $I = [a, b] = [c - r, c + r]$
- Notation Centre-Rayon : $I =< c, r >$

Inconvénient : Nécessite de devoir stocker deux nombres

La représentation FP-Intervalle

Avec le format FP_Interv, un flottant codé en binaire ne représentera plus un nombre mais un intervalle.

La représentation FP-Intervalle

Avec le format FP_Interv, un flottant codé en binaire ne représentera plus un nombre mais un intervalle.
Exemple :

Intervalle	Signe	E	Mantisse m'
I	0	10000010	00010110000000000000000

La représentation FP-Intervalle

Avec le format FP_Interv, un flottant codé en binaire ne représentera plus un nombre mais un intervalle.
Exemple :

Intervalle	Signe	E	Mantisse m'
I	0	10000010	00010110000000000000000

Ce nombre code l'intervalle $I =< 8.625, 2^{-4} >$, le dernier 1 indique la position du rayon qui vérifie la relation suivante :

$$r = 2^{exposant-position}.$$

La représentation FP-Intervalle

Avec le format FP_Interv, un flottant codé en binaire ne représentera plus un nombre mais un intervalle.
Exemple :

Intervalle	Signe	E	Mantisse m'
I	**0**	**10000010**	**00010110000000000000000**

Ce nombre code l'intervalle $I = < 8.625, 2^{-4} >$, le dernier 1 indique la position du rayon qui vérifie la relation suivante :

$$r = 2^{exposant - position}.$$

Attention, ce nombre peut être également interprété comme un simple flottant codé sous la norme IEEE 754 avec $I = 8,6875$.

Addition

Problème : Comment additionner deux FP-Intervalles ?

Addition

Problème : Comment additionner deux FP-Intervalles ?

Théorème Addition

$$
\begin{aligned}
&\textit{Soit } c_3 = \textit{arrondi}(c_1 + c_2)\\
&|z_1 + z_2 - c_3| = |z_1 + z_2 - (c_1 + c_2) + (c_1 + c_2) - c_3|\\
&\qquad\qquad\qquad \leq |z_1 - c_1| + |z_2 - c_2| + |(c_1 + c_2) - c_3|\\
&\qquad\qquad\qquad \leq r_1 + r_2 + u.2^{exposant(c_3)}\\
&\qquad\qquad\qquad \leq r_3
\end{aligned}
$$

Multiplication

Problème : Comment multiplier deux FP-Intervalles ?

Multiplication

Problème : Comment multiplier deux FP-Intervalles ?

Théorème Multiplication

Soit $c_3 = arrondi(c_1 \times c_2)$

$$\begin{aligned}
|z_1 \times z_2 - c_3| &= |z_1 \times z_2 - (c_1 \times c_2) + (c_1 \times c_2) - c_3| \\
&\leq |z_1 \times z_2 - (c_1 \times c_2)| + |(c_1 \times c_2) - c_3| \\
&\leq (|c_1| + r_1) \times r_2 + r_1 \times |c_2| + u \times 2^{exposant(c_3)} \\
&\leq r_3
\end{aligned}$$

Algorithme de Newton par Intervalle

Problème : Approcher $\sqrt{2}$

Algorithme de Newton par Intervalle

Problème : Approcher $\sqrt{2}$
Solution : Adapter la méthode de Newton pour les intervalles

Algorithme de Newton par Intervalle

Problème : Approcher $\sqrt{2}$
Solution : Adapter la méthode de Newton pour les intervalles

Algorithme de Newton

Soit $\tilde{x} \in X$ *et* $f(x) = x^2 - 2$
On définit :

$$N(\tilde{x}, X) := \tilde{x} - \frac{\tilde{x}^2 - 2}{2 \times X}$$

Si $N(\tilde{x}, X) \subset X$, *alors* X *contient une racine de* f.
Si $N(\tilde{x}, X) \cap X = \emptyset$ *alors* $f(x) \neq 0$ *pour tout* $x \in X$.

Algorithme de Newton par Intervalle

i	Itérations de Newton	Centre-Rayon
1	+1.011____ $\times 2^0$	$< 1.3750, 2^{-4} >$
2	+1.0110100011____ $\times 2^0$	$< 1.409, 2^{-11} >$
3	+1.0110101000001010001011111111101_____________________ $\times 2^0$	$< 1.41421, 2^{-31} >$
4	+1.01101010000010011110011001100111111110110011011_____ $\times 2^0$	$< 1.41421, 2^{-48} >$
5	+1.011010100000100111100110011001111111001110111100110_ $\times 2^0$	$< 1.4142135, 2^{-52} >$

Algorithme de Newton par Intervalle

i	Itérations de Newton	Centre-Rayon
1	+1.011____ $\times 2^{0}$	$< 1.3750, 2^{-4} >$
2	+1.0110100011____ $\times 2^{0}$	$< 1.409, 2^{-11} >$
3	+1.0110101000001010001011111111101_____________________ $\times 2^{0}$	$< 1.41421, 2^{-31} >$
4	+1.01101010000010011110011001100111111110110011011____ $\times 2^{0}$	$< 1.41421, 2^{-48} >$
5	+1.0110101000001001111001100110011111111001110111100110_ $\times 2^{0}$	$< 1.4142135, 2^{-52} >$

$\Longrightarrow$ **Intervalle de longueur** 2^{-52} **contenant** $\sqrt{2}$

Algorithme de Newton par Intervalle

i	Itérations de Newton	Centre-Rayon
1	+1.011____ $\times 2^{0}$	$< 1.3750, 2^{-4} >$
2	+1.0110100011____ $\times 2^{0}$	$< 1.409, 2^{-11} >$
3	+1.011010100000101000101111111101______________________ $\times 2^{0}$	$< 1.41421, 2^{-31} >$
4	+1.01101010000010011110011001100111111110110011011_____ $\times 2^{0}$	$< 1.41421, 2^{-48} >$
5	+1.011010100000100111100110011001111111001110111100110_ $\times 2^{0}$	$< 1.4142135, 2^{-52} >$

$\Longrightarrow$ **Intervalle de longueur** 2^{-52} **contenant** $\sqrt{2}$

$\Longrightarrow$ **Convergence Quadratique**

Exponentielle

Problème :Soit I de type FP-Intervalle. Comment obtenir $Exp(I)$?

Exponentielle

Problème :Soit I de type FP-Intervalle. Comment obtenir $Exp(I)$?

Solution :

$I =< C, r >= \{z : C - r \leq z \leq C + r\}$

$I = [a, b] = [C - r, C + r]$

$K = exp(I) =< C_k, r_k >= \{z : exp(a) \leq z \leq exp(b)\}$

$$\boxed{C_k = \frac{exp(a)+exp(b)}{2}}$$

$$\boxed{r_k = \frac{exp(a)-exp(b)}{2} + u \times 2^{exposant(C_k)}}$$

Conclusion

Matrice d'intervalles : Méthode similaire au produit matriciel classique en reprenant les résultats obtenus précédemment de l'addition et de la multiplication de deux FP-Intervalle. Complexité en $O(n^3)$.

Conclusion

Matrice d'intervalles : Méthode similaire au produit matriciel classique en reprenant les résultats obtenus précédemment de l'addition et de la multiplication de deux FP-Intervalle. Complexité en $O(n^3)$.

Perspective

Utilisation de la bibliothèque CBLAS qui permet d'améliorer notre produit matriciel et le rendre plus performant

Merci pour votre attention !

Merci pour votre attention !

Questions

Avez-vous des questions ?
Notre projet est disponible sur :
https ://gitlab.com/vinceliu/bibliIntervalleC