Práctica 1: Errores

Marco Praderio y Marta Cavero

1.

$$f(x) = \begin{cases} \frac{1-\cos^3(x)}{x^2} & \text{si } x \neq 0 \\ \frac{3}{2} & \text{si } x = 0 \end{cases}$$

Queremos evaluar $f(x_0)$ para el valor $x_0 = 1,1\times^{-5}$

1.1. Comprobar que $0 \leq f(x) < \frac{3}{2}$ para cada $x \neq 0$

Dado que $x^2 > 0$ para todo $x \neq 0$ y $1-\cos^3(x) \geq 0$ para todo x, entonces $f(x) \geq 0$ para todo x en cuanto producto de numeros positivos. Vamos a estudiar la función alrededor de 0. Calculando el límite cuando x tiende a 0 de $f(x)$ obtenemos una indeterminación del tipo zero partido zero que podemos resolvemos con el método de l'Hôpital.

$$\begin{aligned} \lim_{x\to 0} f(x) = \lim_{x\to 0} \frac{1-\cos^3(x)}{x^2} = \lim_{x\to 0} \frac{3\cos^2(x)\sin(x)}{2x} = \\ = \lim_{x\to 0} \frac{3\cos^3(x) + 6\sin^2(x)\cos(x)}{2} = \frac{3}{2} \end{aligned} \tag{1}$$

Si ahora demostramos que en el entorno $(0, \frac{\pi}{2}]$ la derivada de f és negativa mientras que en el entorno $[-\frac{\pi}{2}, 0)$ la derivada de f és positiva habremos de mostrado que $f < \frac{3}{2}$ en el conjunto $[-\frac{\pi}{2}, \frac{\pi}{2}]\backslash 0$.

La afirmación que acabamos de hacer, aunque intuitiva a partir del significado geometrico de derivada, merece un pequeño parentesis para ser demostrada. La demostraremos por reducción al absurdo. Supongamos que existe $x_0 \in [-\frac{\pi}{2}, 0)$ tal que $f(x_0) \geq \frac{3}{2}$[1] entonces, dado que $f'(x_0) > 0$ por hipòtesis, existe un entorno derecho de x_0 tal que para todo $x \neq x_0$ perteneciente a ese entorno (que llamaremos U) se cumple $f(x) > f(x_0) \geq \frac{3}{2}$. Cogemos entonces $x_1 \in U \cap [-\frac{\pi}{2}, 0)$ tal que $x_1 \neq x_0$ y notamos que, dado que la derivada de f en el entorno $[x_1, 0)$ és estrictamente positiva por hipòtesis, entonces. para todo $x \in [x_1, 0)$ se cumplirà $f(x) \geq f(x_1)$. Por lo tanto, si existe, el límite $\lim_{x\to 0^-} f(x)$ serà mayor o igual a $f(x_1)$ y, por lo tanto, estrictamente mayor a $\frac{3}{2}$. En particular serà diferente de $\frac{3}{2}$. Pero hemos demostrado en 1 que ese límite existe y és exactamente $\frac{3}{2}$. Por lo tanto llegamos a contradicción y queda así demostrado por reducción al absurdo que $f(x) < \frac{3}{2}$ para todo $x \in [-\frac{\pi}{2}, 0)$.

Demostremos ahora que $f'(x) > 0$ para todo $x \in [-\frac{\pi}{2}, 0)$ i que $f'(x) < 0$ para todo $x \in (0, \frac{\pi}{2}]$.
La derivada de la funcion $f(x)$ es:

$$f'(x) = \frac{2x(\cos^3(x) - 1) - 3\sin(x)x^2}{x^4}$$

Notemos que el denominador siempre es positivo por lo tanto el signo de $f'(x)$ vendrá dado por el numerador.
En $(0, \frac{\pi}{2}]$ se cumple que $-3\sin(x)x^2 < 0$ y que $2x(\cos^3(x) - 1) < 0$ por lo tanto el numerador cumplirá

$$2x(\cos^3(x) - 1) - 3\sin(x)x^2 < 0$$

por lo tanto $f'(x)$ será negativa en $(0, \frac{\pi}{2}]$. Analogamente vemos que en $[-\frac{\pi}{2}, 0)$ $f'(x)$ es positiva.
Podemos por tanto afirmar que, en $[-\frac{\pi}{2}, \frac{\pi}{2}]\backslash 0$, se cumplirá $f(x) < \frac{3}{2}$.

Además tenemos que, para $x^2 > 2$, se cumple

$$f(x) = \frac{1-\cos^3(x)}{x^2} \leq \frac{2}{x^2} < \frac{2}{2} < \frac{3}{2}$$

Y así vemos que $f(x) < \frac{3}{2}$ para $|x| > \sqrt{2}$. Como $\sqrt{2} < \frac{\pi}{2}$ entonces, como ya hemos demostrado que $f(x) < \frac{3}{2}$ para todo $x \neq 0$ tal que $|x| \leq \frac{\pi}{2}$, podemos concluir que $f(x) < \frac{3}{2}$ para todo $x \neq 0$.

[1] La demostración para el intervalo $(0, \frac{3}{2}]$ és analoga y nos abstendremos de hacerla

1.2. Escribir programas en C, uno en precisión simple y otro en precisión doble, que evalúen la función para valores de $x \neq 0$ y excluyan explícitamente $x = 0$. Calcular para cada uno de los programas el valor $f(x_0)$. Comparar los resultados con el que se obtiene utilizando una calculadora de bolsillo. Identifica la causa del error. Justifica las discrepancias y la magnitud de los errores cometidos.

Se puede encontrar el código de los programas en Programas práctica 1

Si evaluamos la función en el punto x_0 con una precisión simple, obtenemos como resultado

$$f(x_0) = 1{,}5000011921$$

mientas que, si la evaluamos con una precisión doble, obtenemos

$$f(x_0) = 1{,}500001225$$

Por otro lado, calculando con una calculadora de bolsillo, obtenemos como resultado

$$f(x_0) = 1{,}500024793$$

La causa principal de este error se debe a la operación de resta $1 - \cos^3(x)$ la cual cumple que $\lim_{x\to 0} 1 - \cos^3(x) = 0$, o, dicho de otra manera, produce una cancelación. Esta cancelación provoca que el error relativo se dispare llegando a asumir valores del orden de $10^{10} \cdot \varepsilon_r$ donde ε_r indica el epsilon de la máquina.
Evaluando la función en un punto tal que $1 - \cos^3(x)$ no produce una cancelación como por ejemplo π hemos obtenido 0.2026423514 con precisión simple y 0.2026423673 con precisión doble y 0.202642367 con una calculadora de bolsillo. Como vemos los resultados son más similares que cuando evaluábamos en x_0. En efecto el resultado del càlculo hecho con calculadora de bolsillo ahora coincide en los 9 decimales que muestra la calculadora con el resultado realizado con precisión doble mientras antes habia discrepancias a partir de quinto decimal. Esto confirma nuestra anterior afirmación de que la principal causa de error es debida a la cancelación.

1.3. Reescribir la función para reducir el error

Hagamos algunas modificaciones trigonométricas a la función para eliminar cancelaciones alrededor de 0.

$$\begin{aligned} f(x) = \frac{1-\cos^3(x)}{x^2} &= \frac{1+\cos(x)}{1+\cos(x)} \cdot \frac{(1-\cos^3(x))}{x^2} = \\ &= \frac{(1-\cos^2(x)) \cdot (1+\cos(x)+\cos^2(x))}{(1+\cos(x)) \cdot x^2} = \\ &= \frac{\sin^2(x)(1+\cos(x)+\cos^2(x))}{x^2(1+\cos(x))} \end{aligned}$$

[2] Al evaluar en la función modificada en el punto x_0 obtenemos 1.500000000 con la calculadora de bolsillo, con precisión simple 1.500000000000000 y con precisión doble 1.49999999989413. Podemos observar que la propagación del error inicial se ha reducido de manera considerable ya que los resultados con precisión simple y con precisión doble son mucho más cercanos (la diferencia pasa de ser del orden de 10^{-8} a ser del orden de 10^{-10}) y el resultado dado por la calculadora de bolsillo coincide en todas las cifras que la calculadora puede dar con el resultado obtenido mediante el calculo con precisión simple.

1.4. Discutid el error de calcelación

En la función inicial teníamos una cancelación que aumentaba el error relativo en un factor del orden de 10^{10}[3]. Mientras que con la función modificada se elimina el error de cancelación porque todos los términos son positivos y se suman.

[2] La función reescrita de esta manera solo presenta suma i productos entre valores que, en un entorno del 0, són positivos y, por lo tanto, no producirà errores debidos a cancelación.

[3] $fl(1) - fl(\cos^3(x_0)) = 1 - \cos^3(x_0)(1+15\varepsilon) = (1-\cos^3(x_0))(1 - \frac{\cos^3(x_0)}{1-\cos^3(x_0)} 15\varepsilon)$ y tenemos que

$$\frac{\cos^3(x_0)}{1-\cos^3(x_0)} \approx \frac{1}{(1-\cos(x_0))(1+\cos(x_0)+\cos^2(x_0))} \approx \frac{1}{3(1-\cos(x_0))} = \frac{1+\cos(x_0)}{3(1-\cos^2(x_0))} \approx \frac{2}{3}\frac{1}{\sin^2(x_0)} \approx \frac{2}{3}\frac{1}{x_0^2} = \frac{200}{363} \times 10^{10}$$

2. El coseno hiperbólico se define como

$$\cosh(t) = \frac{e^t + e^{-t}}{2}$$

y su función inversa

$$t = \operatorname{arccosh}(x)$$

es la solución de la equación $x = \frac{e^t+e^{-t}}{2}$ Resolviendo la equación obtenemos

$$\operatorname{arccosh}(x) = \log(x \pm \sqrt{x^2-1})$$

Antes de empezar vamos a demostrar la identidad $\operatorname{arccosh}(x) = \log(x \pm \sqrt{x^2-1})$

$$x = \frac{e^{-t}+e^t}{2} \Rightarrow x - e^t = \frac{e^{-t}-e^t}{2} \Rightarrow$$

$$\Rightarrow (x-e^t)^2 = \left(\frac{e^{-t}-e^t}{2}\right)^2 \Rightarrow$$

$$\Rightarrow (x-e^t)^2 = \frac{e^{-2t}-2+e^{2t}}{4} = \frac{e^{-2t}+2+e^{2t}}{4} - 1 = \left(\frac{e^{-t}+e^t}{2}\right)^2 - 1 = x^2 - 1 \Rightarrow$$

$$\Rightarrow x - e^t = \pm\sqrt{x^2-1} \Rightarrow e^t = x \pm \sqrt{x^2-1} \Rightarrow t = \log(x \pm \sqrt{x^2-1}) = \operatorname{arccosh}(x)$$

2.1. Demostrad que esta fórmula produce resultados contaminados de error de cancelación cuando se usa el signo menos y x és grande. Probar, por ejemplo, $x = \cosh(10)$ usando doble precisión.

Si x es grande y utilizamos el signo menos entonces, dado que $\lim_{x\to\infty} \frac{x}{\sqrt{x^2-1}} = \lim_{x\to\infty} \frac{x}{x} = 1$, entonces la operación $x - \sqrt{x^2-1}$ produce un error de cancelación que dispara el error relativo[4].

Para obtener un ejemplo que nos muestre este error de cancelación evaluaremos la función $\operatorname{arccosh}(x)$ en el punto $x = \cosh(10) = \cosh(-10)$[5]
Aplicamos la fórmula $\operatorname{arccosh}(y) = \log(x - \sqrt{x^2-1})$ y observamos que, aunque el resultado tendría que ser exactamente -10, obtenemos el valor -10.0000000135. Tal y como esperabamos el error de cancelación provoca que el resultado obtenido se distancie del resultado exacto. De hecho, si modificamos la fórmula del arccosh podemos evitar esta cancelación.

$$\log(x - \sqrt{x^2-1}) = \log\left((x-\sqrt{x^2-1})\frac{x+\sqrt{x^2-1}}{x+\sqrt{x^2-1}}\right) = \log\left(\frac{x^2-x^2+1}{x+\sqrt{x^2-1}}\right) = \log\left(\frac{1}{x+\sqrt{x^2-1}}\right) = -\log(x+\sqrt{x^2-1})$$

Y, por tanto, $\operatorname{arccosh}(y) = \pm\log(x+\sqrt{x^2-1})$ Aplicando esta nueva fórmula obtenemos:

$$\operatorname{arccosh}(\cosh(-10)) = -10$$

que es el resultado exacto.

---

[4]

$$fl(x) - fl(\sqrt{x^2-1}) = x(1+\varepsilon_1) - \sqrt{x^2-1}(1+\varepsilon_2) = (x-\sqrt{x^2-1})\left(1 + \frac{x}{x-\sqrt{x^2-1}}\varepsilon_1 + \frac{-\sqrt{x^2-1}}{x-\sqrt{x^2-1}}\varepsilon_2\right) =$$

$$= (x-\sqrt{x^2-1})\left(1 + \varepsilon_1 + \frac{-\sqrt{x^2-1}}{x-\sqrt{x^2-1}}(\varepsilon_2-\varepsilon_1)\right)$$

y dado que $\lim_{x\to\infty} \frac{\sqrt{x^2-1}-x}{\sqrt{x^2-1}} = 1 - \lim_{x\to\infty} \frac{x}{\sqrt{x^2-1}} = 0$ entonces $\frac{\sqrt{x^2-1}}{\sqrt{x^2-1}-x}$ tiende a infinito. En otras palabras el error se vuelve arbitrariamente grande cuando x tiende a infinito i utilizamos el signo menos debido a un error de cancelación.

[5] Esta igualdad es cierta ya que la función cosh és par, ademàs resulta estéticamente mejor escribir $\cosh(-10)$ ya que, tal y como hemos definido la función arccosh, cuando usamos el simbolo negativo nos da la solución negativa y queda feo poner $\operatorname{arccosh}(\cosh(10)) = -10$.

3. Calculo de la varianza muestral

En estadística la varianza muestral de n números se define como

$$s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

donde la media muestral $\overline{x}$ se calcula como

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

El calulo de la varianza muestral s_n^2 requiere dos bucles recorriendo los datos: uno para el cálculo de $\overline{x}$ y otro para s_n^2. Para un volumen de datos grande o para el caso en que los datos se generan a tiempo real, no es recomendable hacer dos bucles (uno tras otro). Una fórmula alternativa equivalente que aparece en numerosos textos de estadística que conlleva un número de operaciones similar pero con un solo bucle es la siugiente:

$$s_n^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2\right)$$

Esta fórmula puede sufrir error de cancelación!

3.1. Escribe programas en C que calculen la varianza muestral con dos bucles y con un bucle, respectivamente para simple precisión y para doble precisión donde el imput sea un vector de números reales y el output la varianza muestral.

Se puede encontrar el código del programa en Programas práctica 1

3.2. Considera el vector $x = \{10000, 10001, 10002\}^T$ y calcula la varianza con los programas generados y con la calculadora de bolsillo. Analiza discrepancias.

La varianza al cuadrado calculada con un solo bucle y precisión simple es:
0.00000000000000000000

La varianza al cuadrado calculada con dos bucles y precisión simple es:
1.00000000000000000000

La varianza al cuadrado calculada con un solo bucle y precisión doble es:
1.00000000000000000000

La varianza al cuadrado calculada con dos bucles y precisión doble es:
1.00000000000000000000

Las diferencias entre los cálculos con un solo bucle y con dos bucles se deben a errores de cancelación mientras que las diferencias entre precisión simple y doble tienen su origen en la finitud de la representación y son más evidentes cuantas más operaciones se hacen.

3.3. Construye ejemplos de vectores de dimensión grande (al menos 100 componentes) donde estas discrepancias sean más conspicuas.

Para conseguir mostrar esta cancelación construiremos una lista de números muy similares con una varianza pequeña. Esto provocará error de cancelación en el método de un solo bucle en cuanto se producirà una resta entre números muy similares. En nuestro caso hemos construïdo una lista de números que oscila alrededor del 10000 y cuya distancia máxima entre dos cualesquiera de ellos es siempre inferior a 0.02 (el i-ésimo elemento de la lista es $10000 + 0{,}001 \cdot \sin(i-1)$).

De esta manera obtenemos que la varianza al cuadrado calculada con un solo bucle y precisión simple es:
-62.06060791015625000000

La varianza al cuadrado calculada con dos bucles y precisión simple es:
6.5504900703672E-7

La varianza al cuadrado calculada con un solo bucle y precisión doble es:
5.2018599076705E-7

La varianza al cuadrado calculada con dos bucles y precisión doble es:
5.0518694653883E-7

Como podemos observar calculando la varianza al cuadrado con el método simple de un solo bucle obtenemos como resultado un valor negativo claramente erróneo. Este resultado absurdo se debe al error de cancelación provocado por la resta en el calculo de la varianza cuadrada mediante un solo bucle. Notemos además que este error parece no producirse en el caso de doble precisión. Pero esto esto se debe simplemente al hecho de que con precisión doble se oculta mejor la propagación de error dado que el error previo a la cancelación no es tan grande.

De hecho, si aumentamos el número de datos introducidos desde 100 hasta 1000 obtenemos que la varianza al cuadrado calculada con un solo bucle y precisión simple es:
-229.60560607910156250000

La varianza al cuadrado calculada con dos bucles y precisión simple es:
6.4437455193911E-7

La varianza al cuadrado calculada con un solo bucle y precisión doble es:
7.637031562813E-8

La varianza al cuadrado calculada con dos bucles y precisión doble es:
5.0003524117220E-7

Como podemos observar aumentando el número de datos introducidos la diferencia entre el calculo de la varianza con dos bucles y precisión simple y el calculo de la varianza con dos bucles y precisión doble se mantiene aproximadamente igual mientras que la diferencia entre el calculo de la varianza con un solo bucle y precisión doble y el calculo de la varianza con dos bucles y precisión doble va en aumento empezando a mostrar el error debido a cancelación (a medida que aumentamos el número de datos similares que se introducen el error debido a cancelación se nota cada vez más).

4. Suma de una serie

$$S = \sum_{i=1}^{\infty} \frac{(-1)^{k+1}}{k^2} = \frac{\pi^2}{12}$$

Deseamos calcular aproximadamente la suma S sumando términos (sumas parciales) de la serie y vamos a establecer dos estrategias para programarlas en C en simple y en doble precisión.

4.1. Realizar programas C que calculen la suma de los términos de la serie S en orden creciente hasta un término máximo tanto en simple como en doble precisión donde los datos serán el número de términos a sumar. Realizar el cálculo ejecutando los programas para 4096 términos.

Se puede encontrar el código de los programas en Programas práctica 1

Vamos a calcular una aproximación numérica de la serie $S = \sum_{k=1}^{\infty} (-1)^{k+1}/k^2$ que sabemos que da como resultado el valor $\pi^2/12 = 0{,}822467033424113\ldots$.
Sumamos los diferentes términos con precisión simple y doble.Y los resultados son:
0.8224670886993408203
En el caso de la precisión simple
0.8224670036290657738
En el caso de la precisión doble.

4.2. Realizar programas en C que sumen términos de la serie desde el más pequeño al más grande (orden decreciente). Realiza el cálculo para 4096 términos.

La segunda forma de hacerlo será sumando los términos en orden decreciente obtenemos los siguientes resultados:

0.8224669694900512695
En el caso de la precisión simple
0.8224670036290677722
En el caso de la precisión doble.

4.3. Comparar los resultados con el valor exacto y justifica los diferentes resultados obtenidos.

Como podemos observar los resultados obtenidos de las dos formas diferentes son bastante similares. No obstante, si comparamos los resultados obtenidos calculando con precisión simple podemos observar como el primer método se diferencia del resultado real en un factor del orden de 10^{-8} mientras que en el segundo es del orden de 10^{-6}. Lo cual nos indica que el primer método es mejor.

4.4. Proponer otra alternativa para el cálculo de la serie.

La tercera forma de hacerlo sera sumando los terminos de la serie de forma decreciente en modulo.
0.8224670886993408203
En el caso de la precisión simple
0.8224670036290622210
En el caso de la precisión doble.

Como podemos observar a partir del resultado, con este método obtenemos una solución del mismo orden de precisión que el primer método empleado. Y por tanto, dos ordenes de precisión mejor que el segundo método utilizado.

4.5. Realiza pruebas para obtener 8 dígitos correctos en doble precisión estimando el número de términos que se deberian sumar.

Haciendo 12100 iteraciones hemos obtenido 8 dígitos correctos utilizando los tres métodos presentados con precisión doble.

Con el primer método hemos obtenido
0.8224670300093295250

Con el segundo método
0.8224670300093306352

Con el tercer método
0.8224670300093201991

5. Fórmula de Heron para el área de un triangulo.

Sean $a \geq b \geq c$ las longitudes de los lados de un triángulo. Si $p = (a+b+c)/2$ (semi-perímetro) entonces el área del triángulo se puede calcular mediante la fórmula de Heron

$$A = \sqrt{p(p-a)(p-b)(p-c)}$$

5.1. Demuestra que dicha fórmula sufre error de cancelación cuando $a \approx b + c$.

Como $p = (a+b+c)/2$, entonces $p - a = (b+c-a)/2$ y dado que $b + c \approx a$. Entonces $\frac{b+c}{a} \approx 1$ y $p - a \approx 0$. Se producirá un error de cancelación.

5.2. Construye valores para los que el valor A sea muy inexacto. Computa el valor de A mediante programas en C con precisión simple y doble.

Para obtener un área inexacta necesitamos que la suma de dos lados del triángulo sea aproximadamente igual que el tercer lado. Para obtener un triángulo con estas características será suficiente imponer que uno de sus ángulos sea aproximadamente 0. Si esto ocurre tenemos dos casos posibles:

- Que otro de los ángulos también sea aproximadamente 0 y por tanto el tercer ángulo sería aproximadamente π. Y por el teorema del coseno, tendríamos un lado aproximadamente igual a la suma de los otros dos.

- Es el único ángulo pequeño y por el teorema del seno tendríamos un lado muy pequeño y , por lo tanto, la suma de ese lado y otro sería aproximadamente igual al lado mayor.

Un ejemplo de triángulo con estas características sería, por ejemplo, el triángulo de lados $10^{10}, 10^{10}, 10^{-7}$. Calculando el área de este triángulo empleando la fórmula de Heron obtenemos como resultado 0 tanto en precisión simple como en doble. Notemos que este resultado no puede ser correcto dado que, si cogemos como base el lado de 10^{-7} y aproximamos la altura por 10^{10}[6] obtenemos que el área del triángulo debería ser aproximadamente $A = \frac{b \cdot h}{2} \approx \frac{10^{-7} \cdot 10^{10}}{2} = 500$ muy diferente del resultado obtenido.

5.3. Investiga posibles reformulaciones de (6) que reduzcan el error de cancelación significativamente justificando analítica y numéricamente la mejora.

Desafortunadamente, en astronomía por ejemplo, ocurre muy frecuentemente que resulte necesario trabajar con triángulos con un ángulo muy pequeño. En este caso no podremos utilizar la fórmula de Herón escrita anteriormente si no que tendremos que reformularla para evitar cancelaciones. Las posibles reformulaciones que hemos encontrado són las siguientes.

$$\begin{aligned}
A = \sqrt{p(p-a)(p-b)(p-c)} &= \sqrt{\frac{((b+c)+a)((b+c)-a)(a+(c-b))(a-(c-b))}{16}} = \\
&= \frac{\sqrt{((b+c)^2-a^2)(a^2-(c-b))^2}}{4} = \\
&= \frac{\sqrt{(2cb)^2-(b^2+c^2-a^2)^2}}{4} = \\
&= \frac{\sqrt{2(a^2b^2+b^2c^2+c^2a^2)-(a^4+b^4+c^4)}}{4} = \\
&= \frac{\sqrt{(a^2+b^2+c^2)^2-2(a^4+b^4+c^4)}}{4}
\end{aligned}$$

Después de una larga meditación hemos notado que, cambiando los paréntesis de la segunda formulación podemos obtener

$$\begin{aligned}
A &= \sqrt{\frac{((b+c)+a)((b+c)-a)(a+(c-b))(a-(c-b))}{16}} = \\
&= \frac{\sqrt{(a+b+c)(c-(a-b))(c+(a-b))(a+(b-c))}}{4}
\end{aligned}$$

[6]Esta aproximación és plausible en cuanto, por el teorema de pitagoras i aprovechando el hecho de que el triangulo a estudiar és isoceles podemos calcular la hatura del triangulo como $h = \sqrt{(10^{10})^2 - \left(\frac{10^{-7}}{2}\right)^2} \approx 10^{10}$

Notemos que, si $a \geq b \geq c$ entonces $a - b \geq 0$ i las únicas operaciones que pueden producir errores de cancelación en la fórmula anterior són $x = a - b$, $y = c - x$ y $z = b - c$. Si utilizamos los datos del triangulo utilizado aneriormente tendremos que:

- $a = b$, esto causará que tanto a como b se guarden de la misma forma en la memoria del ordenador y, por lo tanto, la operación $x = a - b$ dará como resultado exactamente 0 y por lo tanto no provocará ningún error.
- por lo dicho anteriormente $x = 0$ y este resultado se guardará de manera exacta en la memoria del ordenador por lo tanto, al calcular $y = c - x$ el error dará como resultado $y = c$ manteniendo únicamente el error de representación de c.
- $b >> c$ por lo tanto la operación $z = b - c$ no conllevará ningún error de cancelación.

Para comprobar la estabilidad numérica de la reformulación propuesta la hemos aplicado para calcular el área del triángulo anterior obteniendo como resultado 500.0000305176 en el cálculo con precisión simple y 500.0000000000 en el cálculo con precisión doble. Notemos como el resultado obtenido en esta ocasión es coherente con nuestra predicción lo cual demuestra que hemos conseguido reducir los errores de cancelación con esta reformulación.