

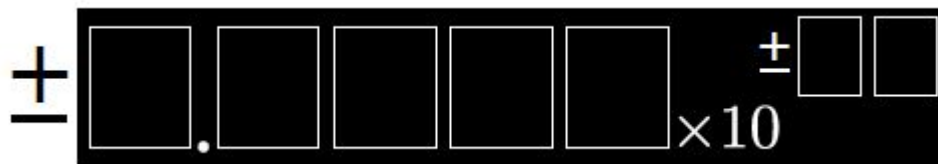
Aritmética de máquina

Modelado matemático II (2021-2)

Juan Andrés Olmos Rojas
Lizeth Daniela Jaimes González



Universidad
Industrial de
Santander



Representación
Punto-Flotante

Mantissa: 5 dígitos de precisión

Date	US population (thousands)
2011-02-01	311,189
2011-03-01	311,356

Table 1: Monthly U.S. population data for February and March 2011.

Cambio = 311,189 - 311,356

3.11136×10^5

- 3.11119×10^5

$= 1.7000 \times 10^2$ (realmente = 167)

Never compute a small number as the difference between two nearly equal large numbers.

Pérdida catastrófica de precisión numérica: Hay un gasto significativo en dígitos de precisión.

El gran kahuna de punto flotante en diferenciación en diferencias finitas

La aproximación directa en diferencias finitas a la primera derivada de $f(x)$ en un punto x es:

$$f'_{\text{FD}}(h, x) = \frac{f(x+h) - f(x)}{h}$$

Ejemplo: Diferenciaremos la función $f(x) = x$ en el punto $x = 1$ para dos pasos de tiempo diferentes:

$$h = \frac{2}{3}$$

$$h = \frac{2}{30}$$

$$f_{\text{float}}(h) = 0.66667$$

$$f_{\text{float}}(h) = 0.066667$$

$$f(x+h) = 1.6667$$

$$f(x+h) = 1.0667$$

$$f(x) = 1.0000$$

$$f(x) = 1.0000$$

$$f(x+h) - f(x) = 0.6667$$

$$f(x+h) - f(x) = 0.0667$$

$$f'_{\text{FD}}\left(h = \frac{2}{3}, x\right) = 1 + O(10^{-4})$$

$$f'_{\text{FD}}\left(h = \frac{2}{30}, x\right) = 1 + O(10^{-3})$$

$$f(x, \Delta) = \sqrt{x + \Delta} - \sqrt{x}. \quad \text{donde} \quad x = 900, \Delta = 4e-3,$$

$$= 30.00006667 - 30.00000000 = 0$$

Re escribiendo:

$$\left(\sqrt{x + \Delta} - \sqrt{x} \right) = \frac{\Delta}{\left(\sqrt{x + \Delta} + \sqrt{x} \right)}$$

$$= \frac{4e-3}{30.0000667 + 30.0000000} \approx 6.667e-5.$$

Acumulación error, caminos-aleatorios

```
function DirectSum(X, N)
    Sum=0.0;
    for n=1:N
        Sum += X;
    end
    Sum
end
```

```
function RecursiveSum(X, N)
    if N < BaseCaseThreshold
        Sum = DirectSum(X,N)
    else
        Sum = RecursiveSum(X,N/2) + RecursiveSum(X,N/2);
    end
    Sum
end
```

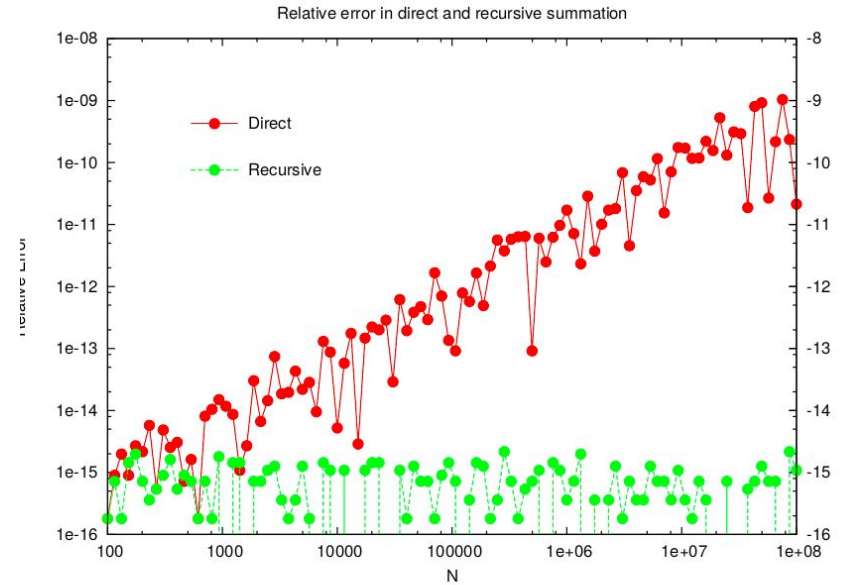


Figure 8: Relative error in the quantity $\text{RecursiveSum}\left(\frac{Y}{N}, N\right)$.