Numerical Analysis I   Question A   p3q10

~~p10q12~~

Model Answer   1999

Let x^* denote the floating point representation of x.

Absolute error ϵ is defined by $x^* = x + \epsilon$.

Relative error δ is defined by $x^* = x(1+\delta)$, so $\epsilon = \delta x$.

Machine epsilon is the smallest positive ϵ_m such that

$$(1+\epsilon_m)^* > 1.$$

Machine epsilon is a useful estimate of the maximum relative error in representing a number across most of the representable range of x. [4 marks]

In the round to even method numbers are rounded to the nearest representable value, except for half-way cases which are rounded so that the last digit is even, e.g. 7.3125 rounds to 7.312, but 7.3175 rounds to 7.318.

If $p=4$, $\beta=2$, 1 is represented as 1.000×2^0, so $1+\epsilon_m$ is represented as 1.001×2^0. ϵ_m is the value of the least significant bit of the mantissa, i.e. $\frac{1}{8}$.

$$1.0101 \to 1.010$$
$$1.1100 \to 1.110$$
$$1.0011 \to 1.010$$
$$1.1001 \to 1.100$$

[6 marks]

Writing $\cos 6 = 1 - c_1 + c_2 - \dots,$

$$c_1 = \frac{36}{1.2} = 18$$

$$c_2 = c_1 \cdot \frac{36}{3.4} = 54$$

$$c_3 = c_2 \cdot \frac{36}{5.6} = 64.8 \quad \text{largest}$$

$$c_4 = c_3 \cdot \frac{36}{7.8} < c_3$$

Absolute error in computing largest term $= 64.8 \times 10^{-7}$

Relative error in computing $\cos 6 \simeq \dfrac{64.8 \times 10^{-7}}{\cos 6} \simeq 10^{-5}$ [5 marks]

Guard digits are extra digits, in the arithmetic unit only, used to improve the accuracy of calculations.

$$\sqrt{x^2 - 2^{24}} = \sqrt{(x+2^{12})(x-2^{12})} = \sqrt{x+2^{12}}.\sqrt{x-2^{12}}$$

The latter formulation requires two square roots but does not require x^2 to be evaluated.

If guard digits are used, and x and 2^{12} are exactly represented, it is possible to compute $x \pm 2^{12}$ with relative error $< 2\epsilon_m$, so

relative error in $\sqrt{x \pm 2^{12}} < \epsilon_m$

relative error in $\sqrt{x+2^{12}}.\sqrt{x-2^{12}} < 2\epsilon_m$

$= 2 \times 10^{-7}$

[5 marks]