Q1

(a) $7.54 \cdot 10^2 + 4.26 \cdot 10^1 = 7.54 \cdot 10^2 + 0.426 \cdot 10^2$

$= 7.966 \cdot 10^2 \approx 7.97 \times 10^2$ ✓

Answer: 7.97×10^2

(b) $1.01 \cdot 10^5 \times 4.04 \cdot 10^{-2} = 4.0804 \times 10^3 \approx 4.08 \times 10^3$ ✓

Answer: 4.08×10^3

(c) $5.25 \cdot 10^{-7} \times (-2.02) \cdot 10^{-5} = -10.6050 \times 10^{-12}$

$= -0.106050 \times 10^{-10} \approx -0.11 \times 10^{-10}$ ✓

Answer: -0.11×10^{-10}

(d) $-6.06 \cdot 10^6 \times (2.02 \cdot 10^4) = -12.2412 \times 10^{10}$ too big!

Answer: $-$Infty ✓

(e) $((6.06 \cdot 10^6) \times (2.02 \cdot 10^4)) - ((5.05 \cdot 10^5) \times (3.03 \cdot 10^5))$

$= (+\text{Infty}) - (+\text{Infty}) = \text{NaN}$

Answer: NaN

✓

Q2

(a) Take $x = \frac{1}{4} \varepsilon_{machine}$ where $\varepsilon_{mach} = \beta^{1-p}$ in IEEE double-precision system. ✓

$fl(\sqrt{1+x} - 1) = fl(\sqrt{1} - 1) = fl(1-1) = 0 = A$

There is an rounding error in the computition $1+x$.

So the Relative Error:

$$RE = \frac{A-T}{T} = \frac{0-T}{T} = \frac{-T}{T} = -1$$ ✓

The magnitute of this R.E $\gg \varepsilon_{mach}$

Therefore, there is a very large relative error.

Q2(b) 5

$$\sqrt{1+x}-1 = \frac{(\sqrt{1+x}-1)(\sqrt{1+x}+1)}{\sqrt{1+x}+1} = \frac{1+x-1}{\sqrt{1+x}+1} = \frac{x}{\sqrt{1+x}+1}$$

New expression: $\frac{x}{\sqrt{1+x}+1}$ call it N.E

Proof: (δ_i's are rounding errors)

$$fl(N.E) = \frac{x}{((\sqrt{(1+x)(1+\delta_1)}\cdot(1+\delta_2)+1)(1+\delta_3)}\cdot(1+\delta_4)$$

$$\approx \frac{x}{\sqrt{1+x}+1}\cdot\frac{1+\delta_4}{(\sqrt{1+\delta_1}\cdot(1+\delta_2))\cdot(1+\delta_3)}$$

Explain: Since $(\sqrt{1+\delta_1}\cdot(1+\delta_2))$ is very small, we can just ignore the effect on 1 and pull it out.

$$R.E = \frac{A-T}{T} = \frac{fl(N.E)-T}{T}$$

$$= \frac{1+\delta_4}{(\sqrt{1+\delta_1}\cdot(1+\delta_2))\cdot(1+\delta_3)}$$

$$= (1+\hat{\delta}_1)(1+\hat{\delta}_2)(1+\hat{\delta}_3)(1+\delta_4)$$

Assume $\hat{\delta}_i \le \frac{1}{2}\cdot 1.01\,\varepsilon_{mach}$ and $\delta_i \le \frac{1}{2}\varepsilon_{mach}$.

Therefore, the R.E is very small under the condition that there is no overflow and underflow.

Q3 → Next page

Q3

(a) $\|x\|_1 = 3+5+1 = 9$ ✓

(b) $\|x\|_2 = \sqrt{9+25+1} = \sqrt{35}$ ✓

(c) $\|x\|_\infty = \max\{3,5,1\} = 5$ ✓

(d) $\|A\|_1 = \max\{6,10,7\} = 10$ ✓

(e) $\|A\|_\infty = \max\{6,9,8\} = 9$ ✓

Q4

Call $LHS = \|Ax\|_v$, $RHS = \|A\|_m \|x\|_v$

$$RHS = \left(\max_{x \neq 0} \frac{\|Ax\|_v}{\|x\|_v} \right) \cdot \|x\|_v$$

by definition (2).

Since $\|x\|_v$ is a number, we can put it inside the max function and it won't change the result. ✓

$$\Rightarrow RHS = \max_{x \neq 0} \left\{ \frac{\|Ax\|_v}{\|x\|_v} \|x\|_v \right\} = \max_{x \neq 0} \|Ax\|_v$$

By comparing LHS and RHS, it is obviously true that:

$$\|Ax\|_v \leq \max_{x \neq 0} \|Ax\|_v$$

Since a particular value of $\|Ax\|_v$ would less than taking max among all x's.

Therefore, $\|Ax\|_v \leq \|A\|_m \|x\|_v$. ✓

-1/ CASE $\vec{x} = \vec{0}$

ns or
nruled

hen-
first

other

TION ROOM

to

EXAMINER'S REPORT

1	5
2	10
3	5
4	5
5	
6	
7	
8	
9	
10	
11	
12	
Total	25.