

## 习题 01

班级\_\_\_\_\_学号\_\_\_\_\_姓名\_\_\_\_\_

### 1. 填空题

- 1) 如果  $y=1000000$ ，它的一个近似值为 999996，采用此近似值代表  $y$  时，其相对误差是0.0004%；具有5位有效数字。
- 2) 进行数值计算过程中产生的主要误差是：截断误差，舍入误差。
- 3) 双精度浮点数体系中，可以表示的最大实数为： $1.796 \times 10^{308}$ 。机器精度的数值约为： $2.22 \times 10^{-16}$ 。
- 4) MATLAB 的通用命令中，`clc` 的功能是清除命令窗口显示内容，`clear` 的功能是清除变量空间。

### 2. 判断题

- 1) 浮点数的使用是利用计算机进行数值计算时误差难以避免的原因。( ☒ )
- 2) 浮点运算时，加法和乘法交换律有可能不成立。( ☐ )
- 3) 浮点数加法中，需要将加数中较大数的指数进行调整，与另一个加数的指数相同，这一过程称为对阶。( ☐ )
- 4) 当除数较大时，进行一次除法运算后可能使结果的绝对误差减小。( ☒ )
- 5) 两个相邻浮点数之间的间距不一定相等，浮点数越大则间距越大。( ☐ )

3. 设  $P(x) = x^3 - 3x^2 + 3x - 1$  和  $Q(x) = ((x - 3)x + 3)x - 1$ ，试采用以上二式分别计算  $P(2.19)$  和  $Q(2.19)$  的值，计算过程中采用三位数字，超过位数的数字均被舍去，如 10.46 写为 10.4。

$$P(2.19) = 4.79 \times 2.19 - 3 \times 4.79 + 3 \times 2.19 - 1 = 10.4 - 14.3 + 6.57 - 1 = 1.67$$

$$Q(2.19) = ((2.19 - 3) \times 2.19 + 3) \times 2.19 - 1 = 1.69$$

$Q(2.19)$  的误差较小，嵌套乘积有助于多项式求值计算取得较为理想的结果。

4. 以下浮点数运算采用 IEEE 双精度格式，`eps` 表示从 1 到一个较大浮点数的距离；试写出其计算结果。

- |                                      |   |
|--------------------------------------|---|
| 1) $2 + \text{eps} - 2$              | 2) $2 - 2 + \text{eps}$                           |
| 3) $2.0\text{e}308 - 1.1\text{e}308$ | 4) $1.5 \times (1.8\text{e}308 + 0.5\text{e}308)$ |
| 5) $0 / ((1 + 1\text{e}-16) - 1)$    | 6) $1.8\text{e}308 - 1.8\text{e}308$              |
| 7) $(1 + 3\text{e}-16) - 1$          | 8) $((3 + 3 \times \text{eps}) - 3) / \text{eps}$ |

- 1) 0; `eps` 被舍掉
- 2) `eps`,
- 3) `Inf` 无穷减有限实数，因此结果为无穷

- 4) Inf 有限实数乘以无穷也为无穷
- 5) NaN 0/0 结果为非数
- 6) NaN 无穷减无穷结果为非数
- 7) eps 或  $2.22e-16$  距离 1 最近的实数是  $1+eps$ , 下个不小于  $1+2eps$ , 因此  $1+3e-16$  实际被认为是  $1+eps$
- 8) 4, 比 3 大的下一个浮点数是  $3+2eps$ , 然后是  $3+4eps, 3+6eps$ , 因此  $(3+3*eps)-3=4eps$