

Numerical Analysis Questions & Answers

Question by Student 201527145

Professor, I have a question about an assignment #1_Question #2_(d). Does the smallest possible number mean the "positive" smallest possible number? Or should I consider the "negative" smallest possible number?

Yes you are right: in Question A1Q2b and A1Q2d, we are seeking the smallest possible *positive* number. Thanks for pointing this out. I'll give you 2 points bonus boost.

Question by Student 201327107

Professor, I don't understand about float type denormal number. I learned that denormal number condition is

$$e = 00000000$$

then exponent should be $0 - 127$? But you explained exponent is -126 . How comes?

The exponent of the denormal number is the same as the smallest exponent (-126) but the difference is with the significant which is $0.f$ instead of $1.f$. Thus, the maximum denormal number is just below the minimum positive normal number. If the exponent would be -127 , then there would be a large gap between the smallest positive normal number and the largest denormal number. Not a good thing! I liked your question, I'll give you 2 points bonus boost.

Question by Student 201327102

Professor, I have a question. In double type, you taught us that

$$p = e - g$$

$$g = \frac{e_{max} - 1}{2}$$

And you substitute

$$2^{11} - 1 (e = 11111111111_{base2})$$

for

$$e_{max}$$

then in

$$p_{\max} = e_{\max} - g$$

we should use different value in

$$e_{\max}(\text{for } 1111111111_{\text{base2}} - 1)$$

Although $e=1111111111$ is not possible (because it indicates infinite # or NaN), why we use $e=1111111111$ in former

$$e_{\max}$$

'after all'? What's the matter that we use $e=1111111111-1$ in former

$$e_{\max}?$$

I understand what is confusing you. When determining g , $e_{\max}$ refers to the maximum possible positive exponent. But in other cases it refers to the maximum possible positive exponent minus one (because the maximum positive exponent is reserved). They should have been written with 2 symbols in class to avoid confusion. In your notes, rewrite $e_{\max}$ to $e'_{\max}$ when determining $p_{\max}$, with $e'_{\max} = e_{\max} - 1$. Good point, I'll give you 2 points bonus.

Question by Student 201029134

professor, I think you should give some interval like a to b in The Question#1 on your assignment but you just give initial point x_0 . and I think it is same in The Question#3 So, I wonder whether it is right or not.

For A2Q3, there is no initial interval. For A2Q1, I have made a change to the question formulation. I'll give you 1.5 points bonus boost for pointing this out — I would have given more if you had not made spelling mistakes.

Question by Student 201029134

professor, I am sorry but I actually ask the question above this about the Question#2 not the Question#3. I think secant method need two points like x_0 and x_1 for calculating.

The secant method only needs one x for the initial guess (it functions similarly to the Newton method). But you also need to specify a small Δx initially to calculate the derivatives. Just set it to a small value of your choice: this won't affect much the convergence history as long as it is not too small.

Question by Student 201527151

Professor, when we use function of $\sin(x)$, we include `<math.h>` library and also write a function code of

```
double sin(double x){
    double ret;
```

```
    ret = sin(x);  
    return ret;  
}
```

. However, the code still doesn't work, and when I compile, it says: undefined reference to 'sin'. Is there something else we should add to a code about sin then?

I don't understand. Are you trying to code a $\sin(x)$ function? If so, there is no need to code it because the function $\sin(x)$ is already defined in the math library.