

Errors

Scientific Notation

In **scientific notation**, a number can be expressed in the form

$$x = \pm r \times 10^m$$

where r is a coefficient in the range $1 \leq r < 10$ and m is the exponent.

$$1165.7 = 1.1657 \times 10^3$$

$$0.0004728 = 4.728 \times 10^{-4}$$

Binary numbers

- We will soon be learning about how numbers are represented in the computer.
- Before we introduce this new concept, make sure to revise binary number representation. You will need to know how to convert from binary numbers to decimal numbers and vice-versa.
- To help you with the review, I added a PrairieLearn homework assignment, where you can get extra credit points (counting towards the “short questions”).

Error in Numerical Methods

- Every result we compute in Numerical Methods contain errors!
- We always have them... so our job? Reduce the impact of the errors
- How can we model the error?

x : true value

\hat{x} : approximation

$$\left. \begin{aligned} e_a &= |x - \hat{x}| \\ e_r &= \frac{|x - \hat{x}|}{|x|} \end{aligned} \right\}$$

- Absolute errors can be misleading, depending on the magnitude of the true value x .
- Example:
- Relative error is independent of magnitude!

You are tasked with measuring the height of a tree which is known to be exactly 170 ft tall. You later realized that your measurement tools are somewhat faulty, up to a relative error of 10%. What is the maximum measurement for the tree height (numbers rounded to 3 sig figs)?

A) 153 ft

B) 155 ft

C) 187 ft

D) 189 ft

$$X = 170 \text{ ft}$$

$$e_r = 0.1$$

$$\begin{aligned} x - \hat{x} &= e_r x \\ \hat{x} - x &= e_r x \end{aligned}$$

$$e_r = \frac{|x - \hat{x}|}{|x|} \rightarrow |x - \hat{x}| = e_r |x|$$

$$\hat{x} = x(1 - e_r) \rightarrow 170(0.9)$$

$$\hat{x} = x(1 + e_r) \rightarrow 170(1.1) = 187 \text{ ft}$$

meet. ps / CS 357

You are tasked with measuring the height of a tree and you get the measurement as 170 ft tall. You later realized that your measurement tools are somewhat faulty, up to a relative error of 10%. What is the minimum height of the tree (numbers rounded to 3 sig figs) ?

A) 153 ft

$$\hat{x} = 170 \text{ ft}$$

B) 155 ft

$$e_r = 0.1$$

C) 187 ft

$$x = ?$$

D) 189 ft

$$\hat{x} = x(1 - e_r) \rightarrow x = \frac{\hat{x}}{(1 - e_r)} = \frac{170}{0.9} = 188.8$$

$$\hat{x} = x(1 + e_r) \rightarrow x = \frac{\hat{x}}{(1 + e_r)} = \frac{170}{1.1} = 154.5$$

Significant digits

Significant figures of a number are digits that carry meaningful information. They are digits beginning to the leftmost nonzero digit and ending with the rightmost “correct” digit, including final zeros that are exact.

The number 3.14159 has 6 significant digits.

The number 0.00035 has 2 significant digits.

The number 0.000350 has 3 significant digits.

Suppose x is the true value and \tilde{x} the approximation.

The number of significant figures tells us about how many positions of x and \hat{x} agree.

Suppose the true value is

$$x = 3.141592653$$

And the approximation is

$$\hat{x} = 3.14$$

We say that \hat{x} has 3 significant figures of x

Let's try the same for:

2) $\hat{x} = 3.14159 \rightarrow \hat{x}$ has 6 sig figs of x

3) $\hat{x} = 3.1415 \rightarrow \hat{x}$ has 4 sig figs of x Why

$\hat{x} = 3.1416 \rightarrow \hat{x}$ has 5 sig figs of x

\hat{x} has n significant figures of x if $|x - \hat{x}|$ has zeros in the first n decimal places counting from the leftmost nonzero (leading) digit of x , followed by a digit from 0 to 4.

$$x = 3.141592653 \times 10^0$$

1) $\hat{x} = 3.14 \rightarrow |x - \hat{x}| = \underbrace{0.001592653}_{(3)} = 1.59\dots \times 10^{-3}$

2) $\hat{x} = 3.14159 \rightarrow |x - \hat{x}| = \underbrace{0.000002653}_{(6)} = 2.653 \times 10^{-6}$

3) $\hat{x} = 3.1415 \rightarrow |x - \hat{x}| = \underbrace{0.000092653}_{(4)} = 0.92653 \times 10^{-4}$

4) $\hat{x} = 3.1416 \rightarrow |x - \hat{x}| = \underbrace{0.000007347}_{(5)} = 0.7347 \times 10^{-5}$

$$e_a = |x - \hat{x}| < 5 \times 10^{-n}$$

\hat{x} has **n significant figures** of x if $|x - \hat{x}|$ has zeros in the first n decimal places counting from the leftmost nonzero (leading) digit of x , followed by a digit from 0 to 4.

So far, we can observe that $|x - \hat{x}| \leq 5 \times 10^{-n}$. Note that the exact number in this example can be written in the scientific notation form $x = q \times 10^0$. *What happens when the exponent is not zero?*

$$x = q \times 10^p \quad \hat{x} = \hat{q} \times 10^p \quad |e_r| \leq 5 \times 10^{-n}$$

$$e_a = |x - \hat{x}| = |q - \hat{q}| \times 10^p$$

We use the relative error definition instead!

$$e_r = \frac{|x - \hat{x}|}{|x|} = \frac{|q - \hat{q}| \times 10^p}{q \times 10^p} < \frac{5 \times 10^{-n}}{q} = \frac{5}{q} \times 10^{-n}$$

$1 \leq q < 10$

Accurate to n significant digits means that you can trust a total of n digits. *Accurate digits* is a measure of relative error.

n is the number of accurate significant digits

x → Mariana
 \hat{x} → st.

Relative error:
$$error = \frac{|x_{exact} - x_{approx}|}{|x_{exact}|} \leq 5 \times 10^{-n}$$

In general, we will use the **rule-of-thumb:**

$$error = \frac{|x_{exact} - x_{approx}|}{|x_{exact}|} \leq 10^{-n+1}$$

For example, if relative error is 10^{-2} then \hat{x} has at most _____ significant figures of x

After rounding, the resulting number has 5 accurate digits. What is the tightest estimate of the upper bound on my relative error?

A) 10^5

B) 10^{-5}

C) 10^4

D) 10^{-4}

$$e_r < 10^{-n+1}$$

rule
of
thumb

$$e_r < 10^{-5+1}$$

Sources of Error

Main source of errors in numerical computation:

- **Rounding error:** occurs when digits in a decimal point ($1/3 = 0.3333\dots$) are lost (0.3333) due to a limit on the memory available for storing one numerical value.
- **Truncation error:** occurs when discrete values are used to approximate a mathematical expression (eg. the approximation $\sin(\theta) \approx \theta$ for small angles θ)

Let's first talk about plots...

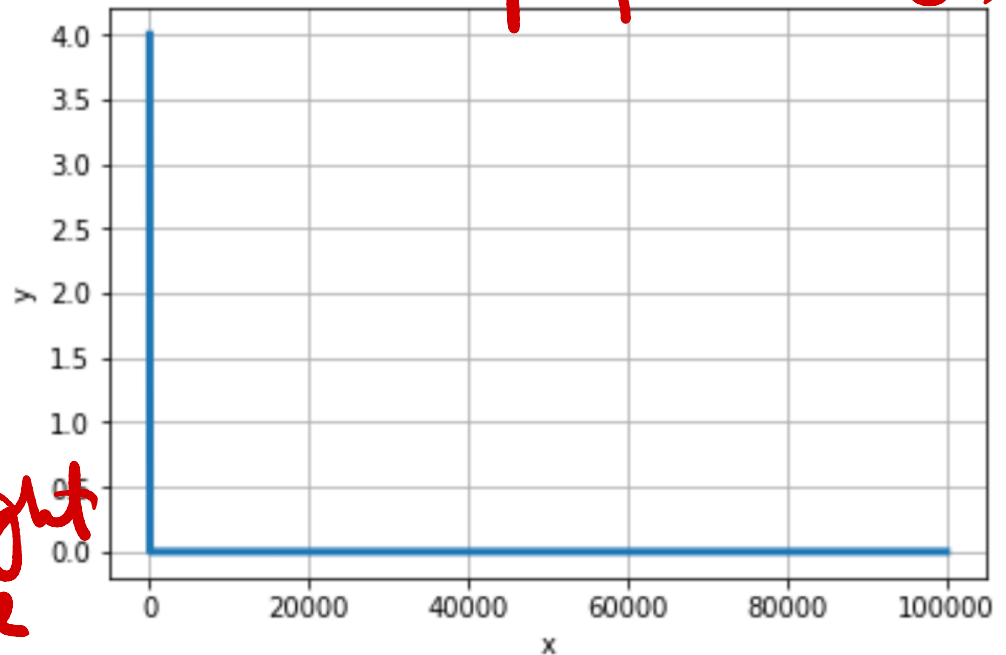
- Power functions:

$$y = a x^b$$

$$\log(y) = \log(ax^b) = \log(a) + \log(x^b)$$

$$\log(y) = \log(a) + b \frac{\log(x)}{x}$$

plt.plot(x,y)



$$\bar{y} = \bar{a} + b \bar{x}$$

plot log log → straight line

- Exponential functions:

$$y = e^x$$

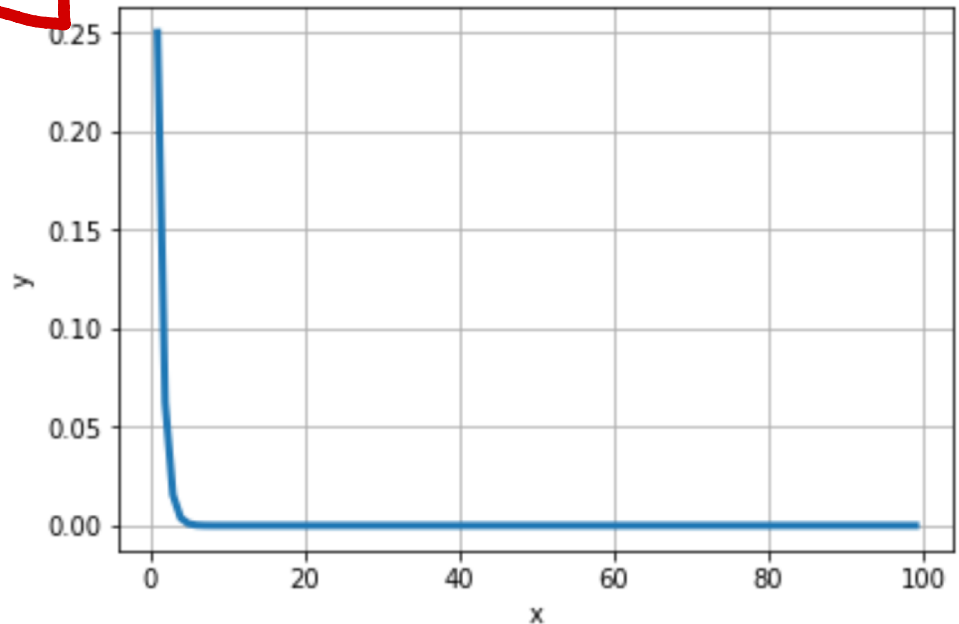
$$y = ab^x$$

$$\log(y) = \log(ab^x) = \log(a) + \log(b^x)$$

$$\underbrace{\log(y)}_y = \underbrace{\log(a)}_a + \underbrace{\left[\begin{array}{c} x \\ \sim \\ x \end{array} \right]}_x \underbrace{\log(b)}_b$$

$x \rightarrow$ linear

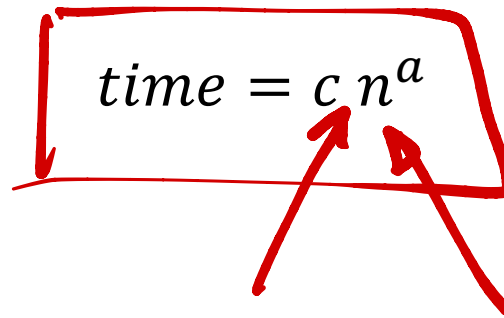
$y \rightarrow$ log



Big-O notation

Complexity: Matrix-matrix multiplication

For a matrix with dimensions $n \times n$, the computational complexity can be represented by a power function:

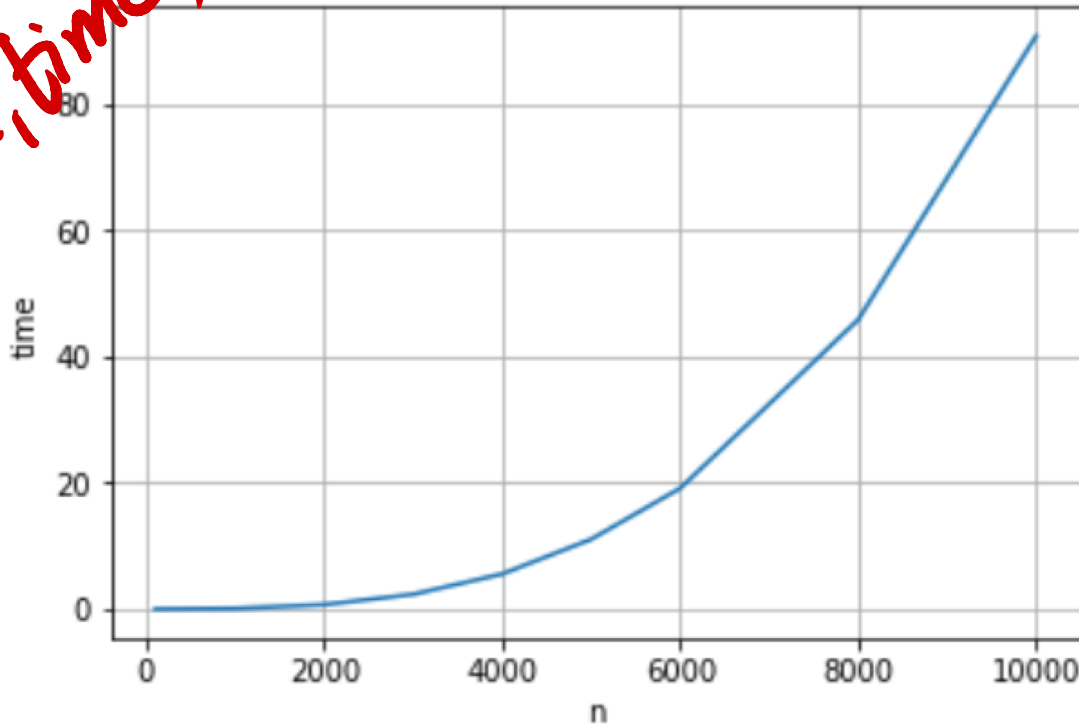

$$time = c n^a$$

We could count the total number of operations to determine the value of the constants above, but instead, we will get an estimate using a numerical experiment where we perform several matrix-matrix multiplications for vary matrix sizes, and store the time to take to perform the operation.

For a matrix with dimensions $n \times n$, the computational complexity can be represented by a power function:

$$time = c n^a$$

Plot. plot(n, time)

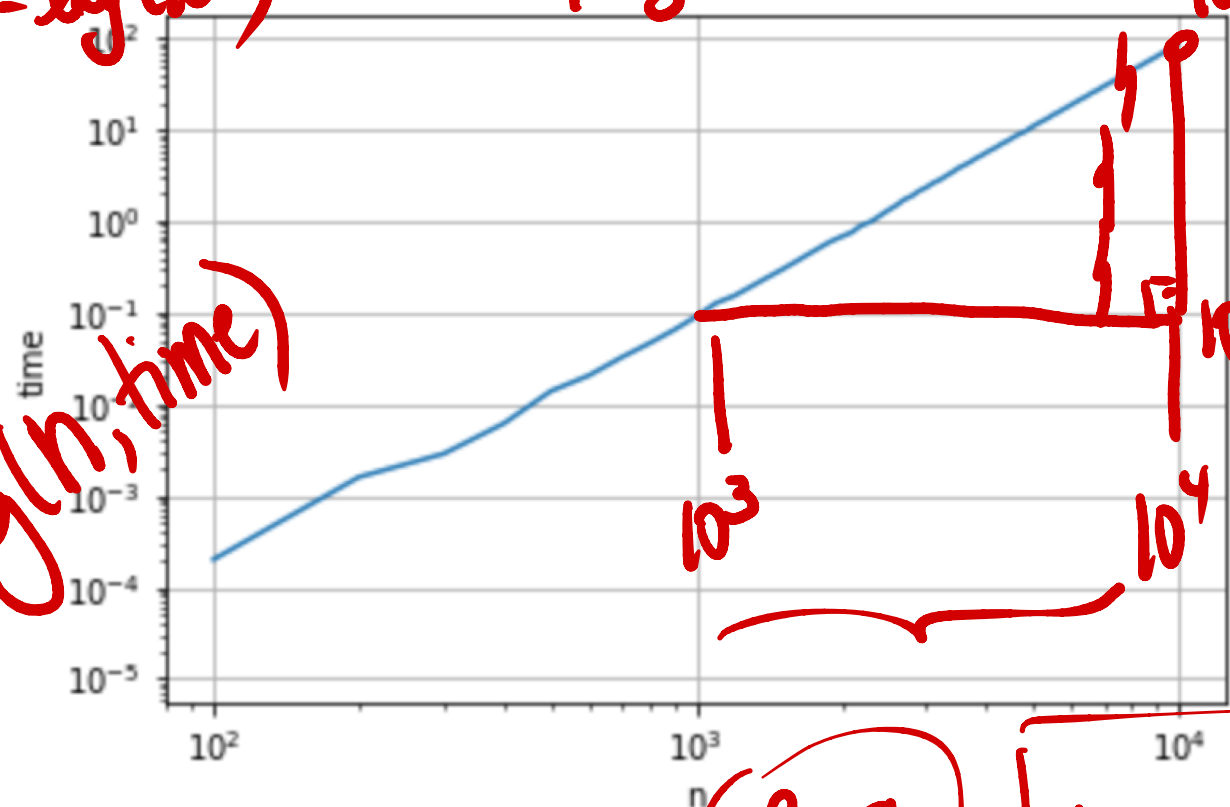


What type of plot will result in a straight line?

- A) semilog-x B) semilog-y C) log-log

Power functions are represented by straight lines in a log-log plot, where the coefficient a is determined by the slope of the line.

$$\frac{\log(10^2) - \log(10^{-1})}{\log(10^4) - \log(10^3)} = \frac{2 \log(10) - (-1) \log(10)}{4 \log(10) - 3 \log(10)} = \frac{2+1}{2-1} = \frac{3}{1}$$



Plt. log log(n, time)

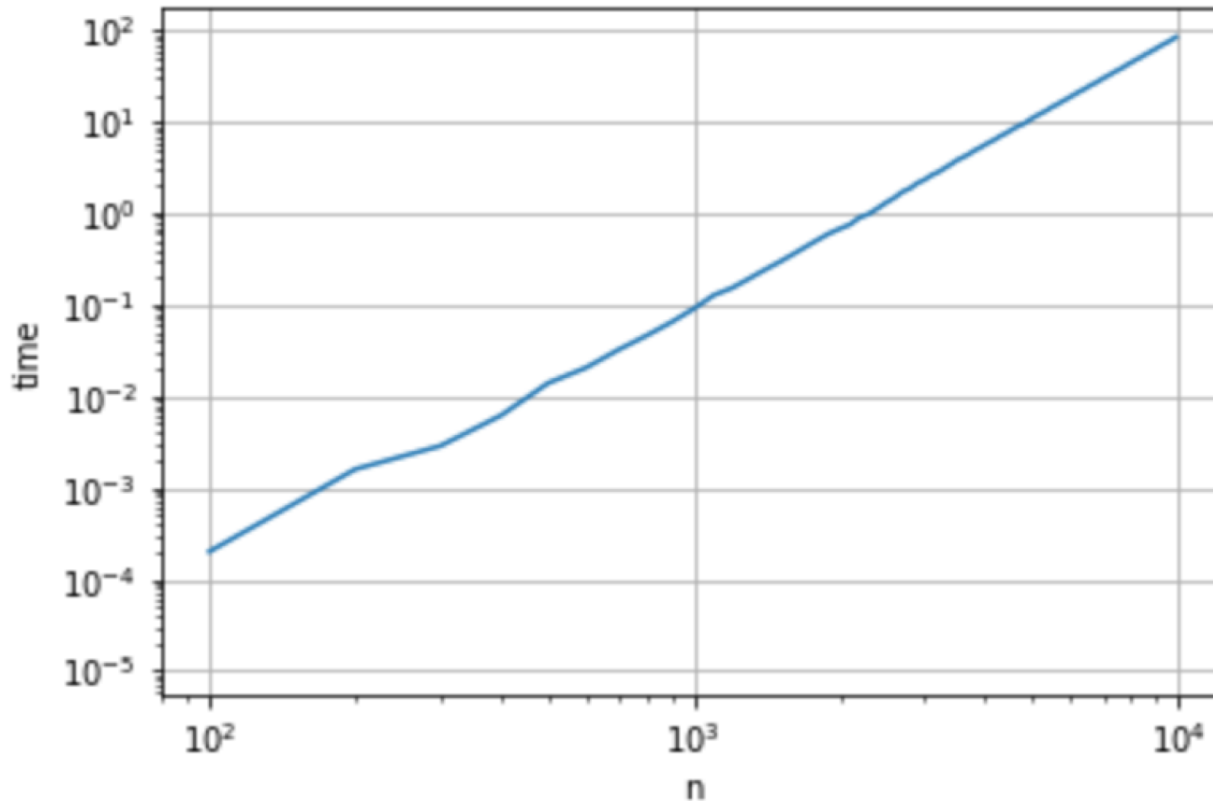
\log $\boxed{\text{time} = c n^3}$

Instead of predicting time using $time = c n^a$, we can use the big-O notation to write

$$time = O(n^a)$$

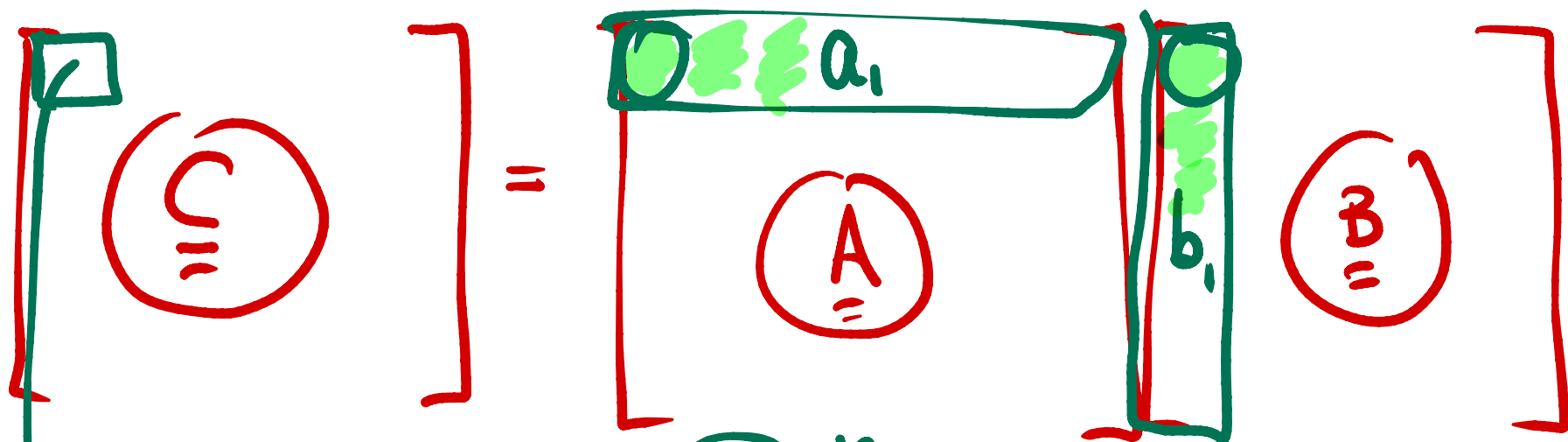
where a can be obtained from the slope of the straight line.
For a matrix-matrix multiplication, what is the value of a ?

$$time = O(n^3)$$



Demo: Cost of Matrix-Matrix Multiplication

As we mentioned previously, we can also get the complexity by counting the number of operations needed to perform the computation:



$x = a_i \cdot b_i$

→ (+) n

(x) n

$$\sum_{i=1}^n a_i b_i$$

→ 2n operations

2n flops

total
2n³

Big-Oh notation

$$f = c g$$
$$\text{time} = c n^a$$

Let f and g be two functions. Then

$$f(x) = O(g(x)) \text{ as } x \rightarrow \infty$$

If and only if there is a positive constant M such that for all sufficiently large values of x , the absolute value of $f(x)$ is at most multiplied by the absolute value of $g(x)$. In other words, there exists a value M and some x_0 such that:

$$|f(x)| \leq M |g(x)| \quad \forall x \geq x_0$$

Example:

Consider the function $f(x) = 2x^2 + 27x + 1000$

dominant

$x \rightarrow \infty$

$$2x^2 + 27x + 1000 \leq \frac{M}{x} \quad (???)$$

$x \rightarrow \infty$

$\rightarrow f(x) = O(x^2)$

Accuracy: approximating sine function

The sine function can be expressed as an infinite series:

$$f(x) = \sin(x) = x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040} + \dots$$

(we will discuss these approximations later)

Suppose we approximate $f(x)$ as $\tilde{f}(x) = x$

We can define the error as:

$$E = |f(x) - \tilde{f}(x)| = \left| -\frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040} + \dots \right| = O(x^3)$$

Or we can use the Big-O notation to say:

$$E = O(x^3)$$

Big-Oh notation (continue)

Let f and g be two functions. Then

$$f(x) = O(g(x)) \text{ as } x \rightarrow a$$

zero

If and only if there exists a value M and some δ such that:

$$\underbrace{|f(x)|} \leq \underbrace{M}_{\uparrow} \underbrace{|g(x)|} \quad \forall x \text{ where } 0 < |x - a| < \delta$$

Same example...

Consider the function $f(x) = 2x^2 + 27x + 1000$

$$x \rightarrow 0$$

$$2x^2 + 27x + 1000 \leq M (\quad)$$
$$x \rightarrow 0$$

$$f(x) = 0 (\quad 1 \quad)$$

Clicker question $5h^2 + 3h \leq M(h)$

Suppose that the truncation error of a numerical method is given by the following function:

$$E(h) = 5h^2 + 3h \implies O(h)$$

Which of the following functions are Oh-estimates of $E(h)$ as $h \rightarrow 0$

~~1) $O(5h^2)$~~

✓ 2) $O(h)$

✓ 3) $O(5h^2 + 3h)$

~~4) $O(h^2)$~~

Mark the correct

answer:

A) 1 and 2

B) 2 and 3

C) 2 and 4

D) 3 and 4

E) NOTA

Clicker question

$$5n^2 + 3n \leq M(5n^2 + 3n)$$

Suppose that the complexity of a numerical method is given by the following function:

$$c(n) = 5n^2 + 3n \quad n \rightarrow \infty$$

Which of the following functions are Θ -estimates of $c(n)$ as $n \rightarrow \infty$

- ✓ 1) $O(5n^2 + 3n)$
- ✓ 2) $O(n^2)$
- ✓ 3) $O(n^3)$
- ~~4) $O(n)$~~

Mark the correct

answer:

- A) 1,2,3
- B) 1,2,3,4
- C) 4
- D) 3
- E) NOTA

$$5n^2 + 3n \leq M n^3$$

~~$$5n^2 + 3n \leq M n$$~~