Taylor's theorem and big-oh notation

Course outcomes

Course outcome	What it is about
Scientific Literacy	I can explain and give reference to important results about numerical methods and theory in a precise manner with references.
Core Knowledge (Facts and Intuition)	I can explain important facts about numerical methods and the intuition behind them.
Analysis	I can give rigorous analysis of numerical methods.
Computation	I can write programs that implement numerical methods in an efficient and collaborative way, including analyzing and evaluating working programs.

Take-aways

After studying these notes, we will be able to

- use Taylor's theorem in an effective way, (*Analysis*)
  - choose a 'right' order of Taylor polynomial for use (trial and error is okay)
  - check the smoothnees (a.k.a. 'regularity') condition,
  - draw a meaningful conclusion by using the theorem correctly,
- translate big-oh notations to their precise definition, (*Scientific Literacy*)
  - state the definition with help of reference,
  - check whether specific examples satisfy the definition,
- explain what a big-oh notation tells us, (*Core Knowledge*)
  - describe the main tendency of a quantity in an intuitive language,
  - tell which part is negilible and which part matters,
  - use correct big-oh notation after operations.

Taylor's theorem

Taylor's theorem

> **Theorem** (Taylor's theorem with Lagrange remainder)

> Suppose and a real valued funtion defined on is times differentiable. Then, for any , we have
>
> where if or if .

proof of Taylor's theorem with Lagrange remainder

Big-oh notation

Motivating examples

Example: Big-oh notation with exponential in easy language

Example (Taylor series of exponential - easy language)

Recall from calculus, for any , we have

Note that this "equality" is "more true" near since the series is expanded around . For example, if , then the sum becomes

After the first few terms, the magnitude of the remaining terms are so small that it does not change the whole sum much.

	partial sum	error
0	1.000000	1.051709e-01
1	1.100000	5.170918e-03
2	1.105000	1.709181e-04
3	1.105167	4.251409e-06
4	1.105171	8.474231e-08
5	1.105171	1.408981e-09
6	1.105171	2.009237e-11
7	1.105171	2.511324e-13
8	1.105171	3.108624e-15
9	1.105171	4.440892e-16
10	1.105171	4.440892e-16

▶ code

For example, after eye-ball the table above, one may want to focus on the first four terms (the rows number 0 through 3) and thinks "just remember there are some other terms, but they are minor". We denote this as

The quotation marked part corresponds to the big-oh part.

Definition

There are slightly different variants of the definitions. We follow essentially Strichartz's definitions as defined in his excellent book (p. 147) [1].

Definition: Big-oh for finite quantities

> **Definition** (Big oh for finite quantities)
>
> Let and and are real valued functions defined near . If there exist (fixed) constants and such that
>
> or,

then we write

and say *f(x) is a big oh of near or f(x) is of order of near* .

Definition: Big-oh for growing quantities

Definition (Big oh for growing quantities)

Let and are real valued functions defined on near (i.e., for some). If there exist (fixed) constants and such that

then we write

and say *f(x) is a big oh of as grows* or *f(x) is of order of as grows*.

Notation (error form of big-oh)

Let be real valued functions defined near . If as , then we write

Interpretation

Interpretation (Controled/bounded by a simpler function)

Suppose as . To use big-oh notation in a useful way, we usually set:

function we want to study	a simpler function than

Then, the big-oh condition says that *the magnitude of can be controlled or bounded by* .

Interpretation (error form of big-oh)

The error form is the most common form of big-oh (e.g., as). Intuitively, we can interpret it as follows:

- *the error caused by in place of is no worse than (up to constant multiple)*, or
- *f(x) is made up of and something that behaves like (up to constant multiple), which is assumed to be small*.

Example: Big-oh notation with cosine

Example (Taylor series of cosine)

Recall from calculus, for any , we have

where the equality is in the sense of infinite sum: the sum on the right hand side converges to the left hand side. Note that this "equality" is "more true" near since the series is expanded around . For example, if , then the sum becomes

After the first few terms, the magnitude of the remaining terms are so small that it does not change the whole sum much. For example, one may want to focus on the first three terms and write

The last equality is not trivial because it involves infinitely many terms, but it is true (See Justification of lumping infinite sum).

Question

Give the interpretation of the big-oh notation in the previous example. (Type in a short answer)

(Reminder) This is **about atmosphere**, not getting it right.

1. Think for a short time.

2. Share your guess with your pair.
3. Type your answer in clicker.
4. Feel free to say out loud.

Example: Big-oh notation for complexity of matrix multiplication

Example (complexity of matrix multiplication)

$$\begin{bmatrix} \quad & \quad & \quad \end{bmatrix} \begin{bmatrix} \quad \end{bmatrix} \quad \begin{bmatrix} \quad \end{bmatrix}$$

Suppose we want to multiply matrix and column vector. Let us count how many operations are needed to do this. To get , we need 3 (real number) multiplications and 2 (real number) additions. We need the same amount of computation for and . If we increase the sizes to (matrix) and (vector), or more generally, to (matrix) and (vector), we need the following number of operations.

size	**( )**	**( )**	**total**
3			15
4			28
	?	?	?

Question

Fill the table and find the best big-oh notation.

Answer choice	**big-oh**
(A)	
(B)	
(C)	
(D)	I don't see a good answer.

(Reminder) This is **about atmosphere**, not getting it right.

1. Think for a short time.
2. Share your guess with your pair.

> 3. Type your answer in clicker.
> 4. Feel free to say out loud.

Properties

We state and prove some frequently used properties of big-oh relations for the finite quantities(see the boardwork for the proofs). Similar results hold also for the growing quantities.

> **Proposition** (sum)
>
> Suppose , , and as . Then,

Proof of sum of big-oh's

Example of sum of big-oh's

> **Corollary** (sum of the same big-oh terms)
>
> Suppose and as . Then,

Proof of sum of big-oh's: same order

Example of sum of big-oh's: same order

> **Proposition** (product)
>
> Suppose and as . Then,

Proof of product of big-oh's

> **Corollary** (product directly added to order)
>
> Suppose as . Then,

Proof of product of big-oh's: directly added to order

Example of product of big-oh's: directly added to order

Examples

Example: Application of Taylor's theorem

> **Example** (Accuracy of difference quotient)
>
> Suppose that is continuously differentiable near . Then, we have
>
> **Interpretation**

Accuracy of difference quotient

> **Question**
>
> Give a second look at the examples of and the complexity of matrix multiplication. Which part of each of those quantities is negligible and which part matters?
>
> (Reminder) This is **about atmosphere**, not getting it right.
>
> 1. Think for a short time.
> 2. Share your guess with your pair.
> 3. Type your answer in clicker.
> 4. Feel free to say out loud.

Remarks

Remark: Advantage of big-oh

> **Remark** (Advantage of big-oh)
>
> Big-oh notation allows us to forget about unimportant details (i.e., exact behaviors of a function) and to focus on what really matters (i.e., what is the magnitude of the function like) when it is helpful.

Remark: Big-oh term may or may not matter

> **Remark** (Big-oh term may or may not matter)

> Sometimes big-oh term is the one that matters. This is typically true for growing quanitities. Sometimes, big-oh term is the one that is negligible. This is usually the case for finite quantities.

Remark: Big-oh is about inequality, not equality

> **Warning** (Big-oh is about inequality, not equality)
>
> Consider and around . If is close enough to , say (hence), we have
>
> Using big-oh notation, it follows
>
> with the constant appearing in the definition.
>
> However, we do not have " as ."
>
> To see this, observe that it is impossible to choose a fixed constant that satisfies "for all , ": taking smaller and smaller 's, eventually becomes smaller than no matter how big you choose.
>
> Even if we try different , the situation is not changed. This is because the violation of big-oh condition (i.e., inequality flip) happens due to 's that are close to 0. Thus we conclude
>
> Observe, however, as you can see from the above reasoning, big-oh encapsulates inequalities.

Remark: The order of error term depends on the context

> **Remark** (The order of error term depends on the context)
>
> Consider the following example. If we want to focus on the terms up to 2nd order, we write
>
> However, while continuing working on, we may need more terms. If it turns out we need to

> include one more term, we write
>
> Both are true. But one thing may not be useful while the other is. It depends the situation.

Appendix

Justification of lumping infinite sum

In the example of Taylor seires of exponential, we claimed

Let us justify this.

(step 1)
We prove:

where is some constant and

for some . By the theorem on power series,

is continuous on the interval of convergence, which is in this case. Therefore,

where we use the continuity of functions defined by power series in the second equality (see the Theorem on power series). Thus, this proves the (step 1) with .

(step 2)

Since as , by the definition of the limit, for , there exists such that, for any , we have

Multiplying through by and using the inverse triangle inequality, i.e. for all , we have

Rearranging, we obtain

Q.E.D.

Error term via Taylor theorem

The most common situation where big-oh notation is used is when analyzing errors using Taylor theorem. Suppose that the context of the analysis somehow hints that knowing only the terms up to 2nd order is enough. Suppose you are studying the difference for some close to , say, . Then, by Taylor theorem, we have

for some or depending on or . If is continuous, then attains the maximum on a closed subinterval of (Extreme value theorem). Thus, for all , we have

where . Therefore, we conclude

Theorem on power seires

> **Theorem** (Power series)[2]

Suppose that has a Taylor expansion around with a radius of convergence or , that is,

Then, it is differentiable, hence continuous. Furthermore, its derivative and integral have the same radius of convergence as and they are given by

Extreme value theorem

Theorem (Extreme value theorem)

Suppose is continuous on a closed interval . Then, it attains the maximum and minimum. That is, there are such that , , and for all .

Question

What part would you make boldfaced in the Extreme value theorem if you were a professor? Why? If you feel like it, answer the same question to other theorems.

(Reminder) This is **about atmosphere**, not getting it right.

1. Think for a short time.
2. Share your guess with your pair.
3. Type your answer in clicker.
4. Feel free to say out loud.

Summary

- Big 'oh' notation allows us to focus only on the most important term out of (possibly) infinitely many terms. (See Advantage of big-oh)
  - In some cases, all terms except the big-oh term matter. (See Example: Big-oh notation with cosine)

    - In other cases, only the big-oh term matters. (See Example: Big-oh notation for complexity of matrix multiplication)
- Big 'oh' is essentially *inequality*. It is not an equality, though its notation uses ' ' sign. (See Remark: Big-oh is about inequality, not equality)
- Eventually, people use big-oh notation at an intuitive level. But it is good to see it works rigorously as well to convince yourself. (See Appendix for some rigorous approaches)

1. Strichartz (2000) *The Way of Analysis*, Jones & Bartlett Learning. ↩
2. James Stewart (2012) *Calculus Early Transcendentals*, 7th Edition, Brooks Cole ↩