Mathematical Writing and Typesetting in LaTeX

Hao Zhu

Department of Computer Science
University of Freiburg

July 11, 2025

About this talk

- guidelines for mathematical writing and typesetting in LaTeX

- useful in general for writing papers; can be very useful if math statements and proofs are included

- list some geneal rules that I am trying to follow, specific to optimization field

- accompanied with a note which is more formal than the slides

- covers both the LaTeX source as well as the output, *i.e.*, the PDF, which is intended to be read alongside its own source

- all material available at `https://github.com/nrgrp/math_latex_slides`

- the material was originally developed by Boyd *et al.* [BRP14] as guidelines for a course report

Outline

Some useful references

some famous guidelines for mathematical writing:

- [Hal70]: Halmos, *How to write mathematics*
- [KLR89]: Knuth $et\ al.$, *Mathematical Writing*

many respectable books follow similar rules, like

- [BV04]: Boyd and Vandenberghe, *Convex Optimization*
- [CT91]: Cover and Thomas, *Elements of Information Theory*
- [HTF01]: Hastie $et\ al.$, *The Elements of Statistical Learning*
- [Sip01]: Sipser, *Introduction to the Theory of Computation*
- [CSRL01]: Cormen $et\ al.$, *Introduction to Algorithms*
- [Rud76]: Rudin, *Principles of Mathematical Analysis*
- [Eva10]: Evans, *Partial Differential Equations*
- [Knu73]: Knuth, *The Art of Computer Programming, Volume I: Fundamental Algorithms*

Precision of mathmatical statements

- the sentence

  "Let $x^\star$ be the solution to the optimization problem."

  implicitly asserts that the solution is unique

- if the solution is not unique or need not be unique, write

  "Let $x^\star$ be a solution to the optimization problem."

- similarly, do not refer to "solving" an expression, as this is meaningless

- we can solve an equation or set of equations, evaluate an expression or function, or check that an equation or inequality holds

Punctuation in equations

- an equation is part of a sentence, so we may need to include a comma or a period at the end of an equation, whether or not inline or display math style is used
- an example for using a comma:

  We next discuss how to solve the problem

  $$\text{minimize} \quad (1/2)\|Ax - b\|_2^2,$$

  where $x \in \mathbf{R}^n$ is the optimization variable.

- an example for using a period:

  The objective function $f\colon \mathbf{R}^n \to \mathbf{R}$ is given by

  $$f(x) = (1/2)\|Ax - b\|_2^2, \quad x \in \mathbf{R}^n.$$

- an example where no punctuation is needed:

  The set

  $$E = \{q \in \mathbf{R} \mid q > 0,\ q^2 < 2\}$$

  has a supremum in $\mathbf{R}$.

Symbols in sentences

- don't start a sentence with a symbol since this hurts readability:

  Bad: f is smooth.
  Good: The function f is smooth.

  Bad: $x^n - a$ has n distinct zeros.
  Good: The polynomial $x^n - a$ has n distinct zeros.

- use words to separate symbols in different formulas if it might confuse the reader visually or in the actual meaning of the sentence:

  Bad: The sequences $x_1, x_2, \ldots, y_1, y_2, \ldots$ are Cauchy.
  OK: The sequences $x_1, x_2, \ldots,$ and $y_1, y_2, \ldots,$ are Cauchy.
  Good: The sequences (x_i) and (y_i) are Cauchy.

  OK: The image of S under f, $f(S) = \{x \mid x \in S\}$, is convex.
  Good: The image of S under f, given by $f(S) = \{x \mid x \in S\}$, is convex.

- do not insert superfluous words if the meaning is clear:

  Good: Consider the function $f + g + h$, where $f\colon \mathbf{R}^n \to \mathbf{R}$, $g\colon \mathbf{R}^m \to \mathbf{R}$, and $h\colon \mathbf{R}^p \to \mathbf{S}^n$ are closed proper convex.

English in math mode

- mathematical symbols should be typeset in math mode: write $Ax = b$, not Ax=b
- subscripts or superscripts that derive from English (or any human language) should not be italicized, for example, write f_{best}, not f_{best}
- the exception is subscripts based on a single letter: refer to a point that is the center of some set as x_c, not x_{c}
- similarly, use commands for special functions: use $\sin(x)$, $\log(x)$, and $\exp(x)$, not $sin(x)$, $log(x)$, or $exp(x)$
- a really heinous example would be the following:

  Consider the problem

  $$minimize \quad f(Ax - b)$$

  where x is the optimization variable and A and b are problem data.

Spacing

- a blank line ends a paragraph, so we shouldn't leave a blank line between an equation and the following text unless intending the equation to end the paragraph
- for example, in the LaTeX source, write:

  ```
  The image of $S$ under $f$,
  \[
  f(S) = \{ f(x) \mid x \in S \},
  \]
  is convex.
  ```

  inserting extra blank lines before `\[` or after `\]` will result in bad typesetting
- the following is fine, since a new paragraph is called for:

  ```
  The image of $S$ under $f$ is defined as
  \[
  f(S) = \{ f(x) \mid x \in S \}.
  \]

  We now turn to a different topic.
  ```

Use the right commands

there are certain special commands in LaTeX for notation that you otherwise might attempt to write in an ad-hoc manner, *e.g.*,

- norms:

  Bad: `$||x||$` ($\implies ||x||$)
  Good: `$\|x\|$` ($\implies \|x\|$)

- set-builder and conditional probability notation:

  Bad: `$|$` ($\implies \{x \in \mathbf{R}|x \geq 0\}$)
  Good: `$\mid$` ($\implies \{x \in \mathbf{R} \mid x \geq 0\}$)

- functions:

  Bad: `$:$` ($\implies f : \mathbf{R}^n \to \mathbf{R}$)
  Good: `$\colon$` ($\implies f\colon \mathbf{R}^n \to \mathbf{R}$)

- use `\ldots` (lower dots, $\ldots$) when the dots are surrounded by commas and `\cdots` (center dots, $\cdots$) when surrounded by other objects that have full height, as in

$$x_1, x_2, \ldots, x_n \quad \text{and} \quad x_1 + x_2 + \cdots + x_n$$

Outline

General guidelines (noncontroversial)

- don't use the same notation for two different things, *e.g.*, don't say "A_j for $1 \leq j \leq n$" in one place and "A_i for $i = 1, \ldots, n$" in another
- it can be useful to choose names for indices so, *e.g.*, i always varies from 1 to m and j always varies from 1 to n
- define all symbols before or near to where you use them
- a symbol like f refers to a function, while $f(x)$ refers to a function evaluated at a given point
  - avoid sloppy writing like "The function $f(x)$ is convex."
  - 'anonymous' functions defined inline are an exception to this rule, as in "The function $x^2 \cos x$ is a counterexample."
- try to use mnemonic notation, so x_c for a center point, c for a cost vector, S for a generic set, C for a convex set, B for a ball, and so on
- don't use symbols like $\forall$, $\exists$, and $\Longrightarrow$; use the corresponding words; these symbols are usually appropriate only in formal logic

- don't assign symbols to concepts that you never refer to, or can easily refer to without the symbol:

  Bad: Let X be a compact subset of a space Y. If f is a continuous real-valued function over X, it has a minimum over X.

  Good: A continuous real-valued function has a minimum over a compact set.

  similarly, do not say

  "The solution $x^\star$ is unique."

  if we never need to refer to $x^\star$ again

- do not write 'arg min' (and 'arg max') since 'argmin' is a single mathematical operator (which is different from $\liminf$ and $\limsup$)

  Very Bad: Let $x = \arg\min_u \left(f(u) + \frac{1}{2}\|u - z\|_2^2 \right)$.

  Bad: Let $x = \arg \underset{u}{\min} \left(f(u) + \frac{1}{2}\|u - z\|_2^2 \right)$.

  Good: Let $x = \operatorname*{argmin}_u \left(f(u) + \frac{1}{2}\|u - z\|_2^2 \right)$.

Symbols for some specific sets (controversial)

- it is common in analysis textbooks to use the bold face capital letters to represent some specific sets, $e.g.$,
  - $\mathbf{N}$: the set of natural numbers
  - $\mathbf{Q}$: the set of rational numbers
  - $\mathbf{Z}$: the set of integers
  - $\mathbf{R}$: the set of real numbers
  - $\mathbf{S}^n$: the set of $n \times n$ symmetric matrices

  the corresponding LaTeX macro is `\mathbf{}` (‘`bf`’ stands for *bold face*)

- recent years, people start to accept the blackboard bold face capital letters instead, $e.g.$, $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{S}^n$; the corresponding LaTeX macro is `\mathbb{}` (‘`bb`’ stands for *blackboard bold*)

- now we can choose to use either of them are as long as they are consistent in the same document; a bad example would be the following

  Bad: The set
  $$E = \{q \in \mathbb{Q} \mid q > 0,\ q^2 < 2\}$$
  has no supremum in $\mathbb{Q}$, but has a supremum in $\mathbf{R}$.

Writing optimization problems (controversial)

Rockafellar wrote optimization problems around 70s in his famous *Convex Analysis* book [Roc70] as follows:

Consider the problem

$$\begin{array}{ll} \text{minimize} & (1/2)\|Ax - b\|_2^2 + \lambda\|x\|_1 \\ \text{subject to} & 0 \preceq x \preceq \mathbf{1} \\ & \|x\|_2 \leq 1, \end{array} \tag{1}$$

where $x \in \mathbf{R}^n$ is the optimization variable, and $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$, and $\lambda > 0$ are problem data.

- the words ‘minimize’ and ‘subject to’ are considered as key *words* for instantiating an optimization problem
- it is always important to state which symbols refer to variables and which to problem data

sometimes for saving space, the problem (1) can be abbreviated as follows:

Consider the problem

$$\begin{array}{ll} \text{min.} & (1/2)\|Ax-b\|_2^2 + \lambda\|x\|_1 \\ \text{s.t.} & 0 \preceq x \preceq \mathbf{1} \\ & \|x\|_2 \leq 1, \end{array} \tag{2}$$

where $x \in \mathbf{R}^n$ is the optimization variable, and $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$, and $\lambda > 0$ are problem data.

- note the period after 'min', which says that 'min.' is the shorthand for 'minimize', instead of the $\min$ operator which is *only* defined for a finite set as

$$\min\{x_1, \ldots, x_n\} = x_k \text{ such that } x_k \leq x_i \text{ for all } i = 1, \ldots, n$$

more recently, people often integrate the sentence for specifying the variable and data into the definition of the problem, $e.g.$, for the problem (1) and (2):

Consider the problems

$$\begin{array}{ll} \underset{x\in\mathbf{R}^n}{\text{minimize}} & (1/2)\|Ax-b\|_2^2 + \lambda\|x\|_1 \\ \text{subject to} & 0 \preceq x \preceq \mathbf{1} \\ & \|x\|_2 \leq 1 \end{array} \quad \text{and} \quad \begin{array}{ll} \underset{x\in\mathbf{R}^n}{\text{min.}} & (1/2)\|Ax-b\|_2^2 + \lambda\|x\|_1 \\ \text{s.t.} & 0 \preceq x \preceq \mathbf{1} \\ & \|x\|_2 \leq 1. \end{array}$$

while avoiding the period after ‘min.’ for writing optimization problems is extremely widely accepted by people in various fields, I personally consider it to be sloppy math:

Let $f, g\colon \mathbf{R}^n \to \mathbf{R}$, and consider the problem

$$\begin{array}{ll} \displaystyle\min_{x \in \mathbf{R}^n} & \min\{f(x), g(x)\} \\ \text{s.t.} & 0 \preceq x \preceq \mathbf{1}. \end{array}$$

- in the above example, two different meanings are assigned to the three ASCII letters ‘min’:
  - the first ‘min’ is the key word for instantiating a minimization problem
  - the second ‘min’ is the operator of taking the smallest element of a finite set
- the LaTeX package `optidef` can be very useful in writing optimization problems in this style, especially when handling lots of constraints

Outline

General rules for mathematical typesetting

Mathematical notation and jargon

Miscellaneous comments

Sentence-ending periods

- LaTeX assumes all periods followed by a space are sentence-ending periods
- tell it otherwise when that is not the case
- for example:

  Bad:

  ```
  Let $x_1,x_2,\ldots ,x_n$ be i.i.d. normal random variables.
  ```

  $\Longrightarrow$ Let $x_1, x_2, \ldots, x_n$ be i.i.d. normal random variables.

  Good:

  ```
  Let $x_1,x_2,\ldots ,x_n$ be i.i.d.\ normal random variables.
  ```

  $\Longrightarrow$ Let $x_1, x_2, \ldots, x_n$ be i.i.d. normal random variables.

Commas

- know when commas should appear inside or outside math environments:

  Bad: `Note that $a,b,$ and $c$ are nonnegative.`

  $\Longrightarrow$ Note that $a, b,$ and c are nonnegative.

  Good: `Note that $a$, $b$, and $c$ are nonnegative.`

  $\Longrightarrow$ Note that a, b, and c are nonnegative.

  Bad:

  `We conclude that $x_1$, $x_2$, \dots, $x_n$ are decreasing.`

  $\Longrightarrow$ We conclude that x_1, x_2, $\dots$, x_n are decreasing.

  Good: `We conclude that $x_1, x_2, \ldots, x_n$ is decreasing.`

  $\Longrightarrow$ We conclude that $x_1, x_2, \ldots, x_n$ is decreasing.

Quotes

- use ‘ ’ and ‘ ‘ ’ ’, instead of ’ ’ and "" to type quotes, so that the left and right quote symbols are rendered correctly

  Bad: ”This is a bad example containing ’quotes’ in a sentence”.
  Good: “The ‘quotes’ in this sentence is good now.”

Dialects

- be aware when writing in mathematical dialect, $e.g.$, in statistics, machine learning, signal processing, control, vision, information theory, and so on
- unless the intended audience is only from this one field, try to avoid using dialect
- try to write in such a way that a general reader with a good understanding of basic mathematics can understand what we are saying

- use standard variable notation unless otherwise needed: x for variables, A for matrices, and so on
- a bad example would be to use
$$\Xi\beta = \chi$$
for a system of linear equations, unless it is really needed

No rule is absolute

- break any of these rules rather than write anything nasty

A bad example

- if you are interested in reading a really bad example document where almost all the rules mentioned before are violated, I can send you my master's thesis and the corresponding LaTeX source code

Reference I

[BRP14] S. Boyd, E. K. Ryu, and N. Parikh. LaTeX style guide for EE 364B. `https://web.stanford.edu/class/ee364b/latex_templates/template_notes.pdf`, 2014.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[CSRL01] T. H. Cormen, C. Stein, R. L. Rivest, and C. E. Leiserson. *Introduction to Algorithms*. McGraw-Hill Higher Education, 2nd edition, 2001.

[CT91] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 1991.

[Eva10] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 2010.

[Hal70] P. R. Halmos. How to write mathematics. *L'Enseignement Mathématique*, 16:123–152, 1970.

[HTF01] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer, 2001.

[KLR89] D. E. Knuth, T. Larrabee, and P. M. Roberts. *Mathematical Writing*, volume 14 of *MAA notes*. Mathematical Association of America, 1989.

Reference II

[Knu73] D. E. Knuth. *The Art of Computer Programming, Volume I: Fundamental Algorithms*. Addison-Wesley, 2nd edition, 1973.

[Roc70] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[Rud76] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill Book Co., third edition, 1976.

[Sip01] M. Sipser. *Introduction to the Theory of Computation*. Thomson Course Technology, 2001.