

论文格式排版你真的做对了吗?

常用格式及其 LaTeX 书写方法介绍

张皓

zhangh0214@gmail.com

引言

论文格式排版是你的文章留给审稿人的第一印象。一篇排版糟糕的文章很难会使审稿人相信这篇文章提出了卓有建树的思想。当论文提供模板时我们可以按照模板进行排版，而对于模板没有涵盖的地方甚至没有模板可用时，我们有必要了解大家约定俗成的排版格式。本文简要介绍了论文书写过程中常用的格式规范及其 LaTeX 书写方法。

基本符号

中文排版

不论中英文，科技文章通常均使用英文半角标点。中文文章的行内公式和两边正文之间要有空格。例如：

通过计算协方差矩阵~ Σ ~我们可以\ldots % 注意"~"

通过计算协方差矩阵 Σ 我们可以...

此外, 推荐使用 XeLaTeX 编译器编译中文文章. XeLaTeX 对中文支持较好, 比如它可以自动在行内公式和两边正文之间加空格, 无需人为指定.

括号/引用号

英文文章中括号, 以及引用符号(如[1], [2]等)要与前一个词留有空格, 而中文文章通常则不同. 例如:

Principal component analysis (PCA) is a statistical procedure that \ldots

Principal component analysis (PCA) is a statistical procedure that ...

引号

和 Word 这种所见即所得的排版软件不同, LaTeX 用不同的符号分别表示左引号和右引号. 左引号用```表示(键盘左上角, Tab 键上面), 右引号用`'`表示(平时输引号的那个键). 双引号用过输入两个引号来得到. 例如:

``machine', ``learning''.`

‘machine’, “learning”.

连字符

不同长度的"-"表示不同含义。一个"-"长度的连字符用于词中；两个"-"长度的连字符常用于制定范围；三个"-"长度的连字符是破折号；数学中的负数要用数学环境下的-得到。例如：

daughter-in-law, pages 13--67, yes --- or no?, \$-1\$.

daughter-in-law, pages 13–67, yes — or no?, −1.

着重强调

在正式文章中，通常不使用粗体进行强调，需要强调的词使用斜体标出。

例如：

PCA is the simplest of the true **eigenvector**-based multivariate analyses.

PCA is the simplest of the true *eigenvector*-based multivariate analyses.

数字分隔

文中出现的位数很长的数字要用逗号","进行分割,每 3 位数字进行分割,逗号两侧不加空格。例如:

```
The MNIST dataset contains 60,000 images \ldots
```

```
The MNIST dataset contains 60,000 images ...
```

图表格式

所有的图表都要有题目,图的题目要写在图的下方,而表的题目要写在表格的上方。图表的题目是一句话,所以要以句号"."结尾。在正文中对图/表/章节的引用可以用如下方式,注意 Figure/Table/Section 和数字之间的"~":

```
\ldots, as illustrated in Figure~1. \ldots, as shown in Figure~2. \
```

```
As Figure~3 illustrates, \ldots \
```

```
Figure~4 shows \ldots \
```

```
Table~1 summarizes \ldots \ldots will be discussed in Section~1. \ldots, as described in Section~2.
```

```
..., as illustrated in Figure 1.  
..., as shown in Figure 2.  
As Figure 3 illustrates, ...  
Figure 4 shows ...  
Table 1 summarizes ...  
... will be discussed in Section 1.  
..., as described in Section 2.
```

在实际写文章时,所有的引用(figure, table, equation, section)都应该用 ref 来管理,这样编号可以自动生成,免去手工调整编号的烦恼。

数学环境

LaTeX 自身只提供最基本的数学符号和环境，因此我们在需要书写数学公式时，通常在引言区导入 `amsmath` 包：

```
\usepackage{amsmath}
```

下面我们假定 `amsmath` 包已被导入。

数学符号

数域

实数域，自然数域等表示数域的符号需要用空心字母表示。空心字母在 LaTeX 中用 `mathbb` 输入，这需要在引言区导入 `amssymb` 包。例如：

```
\usepackage{amssymb}$\mathbb{R}, \mathbb{N}$.
```

$\mathbb{R}, \mathbb{N}.$

组合数

和国内通常用 C_n^k 表示组合数不同，国际上通常用括号表示组合数，读作 "n choose k"。在 LaTeX 中组合数可以用 `binom` 命令打出。例如：

```
\binom{n}{k}.
```

$$\binom{n}{k}.$$

集合

集合中的分割可以用竖线" $|$ ", 也可以用冒号" $:$ ". 分割符与左右要有空格. 例如

```
\{x \mid x \ge 0.\} % Note "\mid" instead of "/"
```

$$\{x \mid x \ge 0.\}.$$

省略号

省略号分两种. 列举一系列元素中的省略用 `ldots`; 数学运算表达式中的省略用 `cdots`. 即省略号高度要和左右符号的高度一致. 而在普通文本中的省略号一律用 `ldots`. 例如:

```
$x_1, x_2, \ldots, x_n$, $x_1 + x_2 + \cdots + x_n$.
```

$$x_1, x_2, \dots, x_n, x_1 + x_2 + \cdots + x_n.$$

优化问题

通常，优化问题解得的最优值用 star 表示，而不是"*"。优化问题的变量要写到"arg min"或"arg max"整体的下方。例如：

```
\begin{equation}
```

$$x^{\star} = \mathop{\arg\min}_x (x-1)^2 \text{ ,}$$

```
\end{equation}
```

$$x^{\star} = \arg \min_x (x - 1)^2 . \tag{1}$$

函数映射

函数映射通常有其固定的写法，例如：

```
$f\colon A \mapsto B$.
```

$$f: A \mapsto B.$$

矩阵与向量

矩阵和向量要用粗体表示。有人习惯于用直立粗体表示矩阵，用斜体粗体表示向量。使用 bm 需要在引言区导入 bm 包。例如：

```
\usepackage{bm}$\bf{A}, \bf{\Omega}, \bm{x}, \bm{\mu}$.
```

$$\mathbf{A}, \mathbf{\Omega}, \mathbf{x}, \mathbf{\mu}.$$

如果让大小写字母都变为斜体粗体，可以用 boldsymbol 命令。例如：

`\boldsymbol{A}, \boldsymbol{\Omega}, \boldsymbol{x}, \boldsymbol{\mu}`.

$$\boldsymbol{A}, \boldsymbol{\Omega}, \boldsymbol{x}, \boldsymbol{\mu}.$$

此外向量也可以用箭头标注. 例如

`\vec x, \vec \mu`.

$$\vec{x}, \vec{\mu}.$$

矩阵和向量可以用中括号"[]"或小括号"()"表示, 目前国际上用中括号表示更常见, 这是因为用小括号表示矩阵和向量容易和表示代数运算优先级的小括号相混淆. 在 LaTeX 中可以用 `bmatrix` 环境书写矩阵. 例如:

`\begin{equation}`

`\begin{bmatrix}`

`a_{11} & a_{12} & \cdots & a_{1n} \\`

`a_{21} & a_{22} & \cdots & a_{2n} \\`

`\vdots & \vdots & \ddots & \vdots \\`

`a_{m1} & a_{m2} & \cdots & a_{mn}`

`\end{bmatrix}`

`\end{equation}`

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (1)$$

数学公式中的直立体

有单位的量

表示单位量纲的符号不是变量，故应当以直立体出现，并且数字和单位之间要有空格，例如：

```
$m = 50\,\mathrm{kg}, V = 150\,\mathrm{m}^3$. % Note "\",
```

$$m = 50\,\mathrm{kg}, V = 150\,\mathrm{m}^3.$$

此外，siunitx 包含了标准单位，能自动调整字体和间距，还可以以非常自然的方式实现复杂的单位和科学记数法等表达。

微元符号

表示微元的 d 不是变量，应当以直立体出现。此外，在积分符号中，微元符号要与前面的被积函数保持一定空格距离。例如：

```
$\int_1^2 \log x \,\mathrm{d} x$. % Note "\",
```

$$\int_1^2 \log x \,\mathrm{d} x.$$

转置符号

表示转置的 T 不是变量，应当以直立体出现。例如

```
\boldsymbol{x}^{\mathrm{T}}$.
```

$$\boldsymbol{x}^{\mathrm{T}}.$$

此外，也有人用 `intercal`，`top`，或 `\mathsf{T}` 来表示转置。

函数名

函数名不是变量，故应当以直立体出现。LaTeX 内置支持一些常用的函数名，例如：

```
\sin(x), \exp(x)$.
```

$$\sin(x), \exp(x).$$

当我们需要自己定义函数名时，在引言区使用 `DeclareMathOperator` 命令声明需要定义的函数名。例如：

```
\DeclareMathOperator{\sign}{sign}
```

```
\sign(x)$.
```

$$\mathrm{sign}(x).$$

公式标点与标号

标号基础

重要的公式需要标号。包含在 `equation` 环境中的环境会被自动标号，如果你不需要标号，可以用 `equation*` 环境或将公式放于 `\[` 和 `\]` 之间。例如：

```
\begin{equation} % Equation with numbering
    e^{\mathrm{i}\pi} = -1
\end{equation}

\begin{equation*} % Equation without numbering
    a^2 + b^2 = c^2
\end{equation*}

\[ % Equation with numbering
    a^2 + b^2 = c^2
\]
```

$$e^{i\pi} = -1 \tag{1}$$
$$a^2 + b^2 = c^2$$
$$a^2 + b^2 = c^2$$

标点基础

公式后的标点是很多人忽视的地方。并不是所有公式后都要或不要加标点，公式(包括行内公式和行间公式)是句子的一部分，因此，我们需要根据公式在句子中所处的位置来判断在公式后加什么标点。行间公式中公式和标点之间要空格。例如：

Einstein introduced his formula

```
\begin{equation}
```

$$E = m \cdot c^2 \quad , \quad \% \text{ Note "\,,,"}$$

```
\end{equation}
```

which is at the same time the most widely known and the least well understood physical formula.

... from which follows Kirchhoff's current law:

```
\begin{equation}
```

$$\sum_{k=1}^n I_k = 0 \quad , \quad \% \text{ Note "\,,"}$$

```
\end{equation}
```

```
\begin{equation}
```

$$I_D = I_F - I_R \quad \% \text{ No break}$$

```
\end{equation}
```

is the core of a very different transistor model.

Einstein introduced his formula

$$E = m \cdot c^2 , \tag{1}$$

which is at the same time the most widely known and the least well understood physical formula.

... from which follows Kirchhoff's current law:

$$\sum_{k=1}^n I_k = 0 . \tag{2}$$

$$I_D = I_F - I_R \tag{3}$$

is the core of a very different transistor model.

多行公式的标点与标号

多行公式通常只在最后一行结尾时根据需要标点与标号。最新版本的 LaTeX 建议使用 `IEEEeqnarray` 环境替代 `eqnarray` 环境书写多行公式，这需要在引言区导入 `IEEEtranstools` 包。例如：

```
\usepackage[retainorgcmds]{IEEEtrantools}

\begin{IEEEeqnarray*}{rCl}

    a &=& b + c \\

    &=& d + e \\

    &=& f + g \quad \text{,. \IEEEyesnumber} \\

\end{IEEEeqnarray*}
```

$$\begin{aligned} a &= b + c \\ &= d + e \\ &= f + g. \end{aligned} \tag{1}$$

多分支公式的标点与标号

分支是公式的一部分，我们只对整体做一个标号，不对每个分支进行单独标号。每个分支需要单独进行标点：前面的分支结束后要用分号";"，最后一个分支结束后根据需要进行标点。此外，和写编程语言不同，与 `if` 分支对应的通常是 `otherwise`，而不是 `else`。LaTeX 中，分支可以用 `cases` 环境书写。例如：

```
\begin{equation}

    |x| =
```

```

\begin{cases}
-x & \text{if } x < 0 \\
x & \text{otherwise}
\end{cases}
\end{equation}

```

$$|x| = \begin{cases} -x & \text{if } x < 0; \\ x & \text{otherwise.} \end{cases} \quad (1)$$

优化问题的标点与标号

通常 max/min 和 s.t.(读作"subject to"而不是"such that")保持中心对齐，目标函数项和各约束项保持左对齐。只在目标函数项那一行标号但不标点，各约束项标点但不标号。例如：

```

\begin{alignat}{2}
\min_x & \quad f(x) \\
\mathrm{s.t.} & \quad g_i(x) \leq 0, \quad i = 1, 2, \dots, m, \\
& \quad h_j(x) = 0, \quad j = 1, 2, \dots, n.
\end{alignat}

```

$$\begin{aligned}
 \min_x \quad & f(x) \\
 \text{s.t.} \quad & g_i(x) \leq 0, \quad i = 1, 2, \dots, m, \\
 & h_j(x) = 0, \quad j = 1, 2, \dots, n.
 \end{aligned} \quad (1)$$

此外，也有人将优化问题看作一个整体，对整体作一居中标号。

数学证明

数学证明以"Proof."开始，以方框符号结尾. amsthm 包提供了证明环境.

例如:

```
\usepackage{amsthm}

\begin{proof}

  Trivial.

\end{proof}
```

Proof. Trivial.



致谢

感谢@蒋雨蒙对 siunitx 包的介绍，@胡不退对 ref 的建议，@风清扬对转置和优化问题标号的建议.

参考文献

1. Oetiker, Tobias, et al. "The not so short introduction to LATEX 2 ϵ ." (2001).