

LaTeX 优雅美观的文档编写


生活中总会有各种各样文字处理的需求，尤其在大学校园，很多时候需要提交纸质版电子文档，制作简洁的PPT等。

- 是否羡慕身边那些熟练掌握文字处理技巧和排版技能的同学呢！
- 是否被打印店Office版本不兼容导致内容排布错乱弄得直呼心累！！
- 是否被课程大作业、毕业论文中Word分节、各种格式难调搞得焦头烂额！！

不用羡慕，不必烦躁，因为我们有更加优雅高效的解决方案！—— LaTeX

实际上，LaTeX 离我们并不遥远。大学课本上长串的数学公式、化学方程式、电路控制图，Springer 等各大出版社的论文等等，都离不开 LaTeX 的排版支持。利用 LaTeX，想写出精致的文稿其实很简单！

Optimization Methods



Interim Report of Optimization Project

A Differential Equation for Modeling Nesterov's Accelerated Gradient Method: Theory and Insights

Kai Liu & Haiyan Wang

Shanghai Jiao Tong University

School of Mathematical Sciences

liu1995@sjtu.edu.cn / 1191942774@qq.com

Optimization Methods(Kiaoqun Zhang)

Contents

1 Introduction	3
2 Derivation	6
3 Connections and Interpretation	7
3.1 Analogous Convergence Rate	7
3.2 Quadratic f and Bessel Functions	8
3.3 Fluctuations of Strongly Convex f	9
3.4 Nesterov's Scheme Compared with Gradient Descent	10
4 The Magic Constant 3	11
4.1 High Friction	11
4.2 Low Friction	13
4.3 Strongly Convex f	14
4.4 Numerical Examples	15
5 Restarting	16
5.1 A new Restarting Scheme	16
5.2 Linear convergence	16
5.3 Numerical Examples	17
6 Summary	19

Derivation

We have figured out that Nesterov's scheme is not always monotone in the objective function value due to the momentum term, i.e., oscillations or overshoots along the trajectory of iterates approaching the minimizer are often observed when running the Nesterov's scheme. We can explain this by viewing the ODE as a damping system.

For **small** t , the damping rate $3/t$ is large, and the ODE tends to be an overdamping system, which will return to the equilibrium without oscillating.

For **large** t , the ODE with small damping rate $3/t$ behaves like an underdamping system, oscillating with the amplitude gradually decreasing to zero.

Here, we sketch an informal derivation of the ODE (3). Assume that $f \in \mathcal{F}_L$ for $L > 0$.³ Combining the two equations in (1) and applying a rescaling gives

$$\frac{x_{k+1} - x_k}{\sqrt{s}} = \frac{k-1}{k+2} \frac{x_k - x_{k-1}}{\sqrt{s}} - \sqrt{s} \nabla f(y_k). \quad (4)$$

Introduce the *Ansatz* $x_k \approx X(k\sqrt{s})$ for some smooth curve $X(t)$ defined for $t \geq 0$. Put $k = t/\sqrt{s}$. Then as the step size s goes to zero, $X(t) \approx x_{t/\sqrt{s}} = x_k$ and $X(t + \sqrt{s}) \approx x_{(t+\sqrt{s})/\sqrt{s}} = x_{k+1}$, and Taylor expansion gives

$$\begin{aligned} \frac{x_{k+1} - x_k}{\sqrt{s}} &= \dot{X}(t) + \frac{1}{2} \ddot{X}(t) \sqrt{s} + o(s) \\ \frac{x_k - x_{k-1}}{\sqrt{s}} &= \dot{X}(t) - \frac{1}{2} \ddot{X}(t) \sqrt{s} + o(s) \end{aligned}$$

and $\sqrt{s} \nabla f(y_k) = \sqrt{s} \nabla f(X(t)) + o(\sqrt{s})$. Thus (4) can be written as

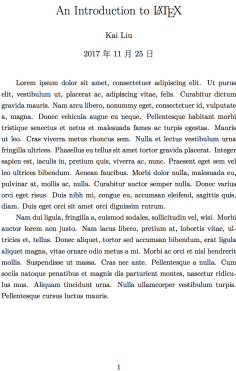
$$\begin{aligned} \dot{X}(t) + \frac{1}{2} \ddot{X}(t) \sqrt{s} + o(s) &= \left(1 - \frac{3\sqrt{s}}{t}\right) \left(\dot{X}(t) - \frac{1}{2} \ddot{X}(t) \sqrt{s} + o(s)\right) - \sqrt{s} \nabla f(X(t)) + o(\sqrt{s}) \end{aligned} \quad (5)$$

³we denote by \mathcal{F}_L the class of convex functions f with L -Lipschitz continuous gradients defined on \mathbb{R}^n

我是谁

一个Office菜鸟，自称二流Excel水平，三流Word水平，却又不愿向巨硬低头的研究生。利用 LaTeX 编写各类文档已有一年经验，包括但不限于：课程笔记/作业、个人简历、LaTeX幻灯片、LaTeX策划书、LaTeX海报、LaTeX申请信等。


```
1 \documentclass{ctexart}
2
3 \title{An Introduction to \LaTeX}
4 \author{Kai Liu}
5 \date{\today}
6 \usepackage{lipsum}
7
8 \begin{document}
9 \maketitle
10 \lipsum[1-2]
11 \end{document}
```



活动安排

此次活动主要介绍 LaTeX 从0到1的应用。时间定于本周三 **11月29日晚 7: 30-9: 00**，地点在**数学系大会议室**。

现场为各位同学准备了茶点零食，**前三十位**到场的同学还将获得由 LaTeX 制作的精美入门教程。

后续活动消息敬请关注[交大数学研究生会](#)