
Second Order Momentum

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 We study the effect of momentum on second order methods, both newton's method
2 as well as quasi-newton methods like LBFS. We compare the number of steps
3 required with and without momentum to reach close to the optimum value.

4 1 Project Objective

5 To achieve faster convergence second order methods were proposed over first order methods.
6 Although second order methods have better bounds than first order methods and they converge in
7 very few steps, second order methods are often not used practically due to the cost of each step.
8 To cope up with these problems, 2 approaches have been proposed and well studied: (i) Adding
9 momentum to first order methods to make convergence speed comparable to second order methods,
10 (ii) Quasi-newton methods which avoid the costly computations involved in second order, without
11 compromising the rate of convergence. In this project we aim to combine both these approaches
12 and attempt to verify if the classical second order methods as well as quasi newton methods can be
13 accelerated.

14 2 Novelty and Proposed Methodology

15 Momentum has been extensively studied on first order methods[Cite]*. But almost no work has been
16 done upon the effects of momentum on newton's method. Through this project we seek to address
17 this issue.

18 Moreover, to accelerate second order methods several quasi newton methods have been proposed
19 including BFGS[Cite], which was later improved for memory in LBFGS[Cite]. We would also check
20 the performance of momentum on BFGS/LBFGS.

21 2.1 Brief ideas

22 For first order methods the momentum equation is given as,

$$v^+ = \gamma v + (1 - \gamma) \nabla f(x) x^+ = x - t v^+ \quad (1)$$

23 Similarly, we propose the momentum equation for second order methods,

$$v^+ = \gamma v + (1 - \gamma) \nabla^2 f(x)^{-1} \nabla f(x) x^+ = x - t v^+ \quad (2)$$

24 Similarly we will explore momentum for BFGS. We intend to study both the classical form of
25 momentum[Cite] as well as the one used in deep learning proposed by Sutskever(?) [Cite]

26 3 Planned Experiments

27 We will compare the performance of first order methods, first order methods with momentum, sth..
28 We can use logistic regression func or $f(x) = (10x_1^2 + x_2^2)/2 + 5\log(1 + e^{-x_1 - x_2})$ etc. Basically
29 functions having non circular contours, on which second order methods perform better than first
30 order.

31 3.1 Citations within the text

32 The natbib package will be loaded for you by default. Citations may be author/year or numeric, as
33 long as you maintain internal consistency. As to the format of the references themselves, any style is
34 acceptable as long as it is used consistently.

35 The documentation for natbib may be found at

36 `http://mirrors.ctan.org/macros/latex/contrib/natbib/natnotes.pdf`

37 Of note is the command `\citet`, which produces citations appropriate for use in inline text. For
38 example,

39 `\citet{hasselmo}` investigated\dots

40 produces

41 Hasselmo, et al. (1995) investigated...

42 If you wish to load the natbib package with options, you may add the following before loading the
43 nips_2017 package:

44 `\PassOptionsToPackage{options}{natbib}`

45 If natbib clashes with another package you load, you can add the optional argument nonatbib
46 when loading the style file:

47 `\usepackage[nonatbib]{nips_2017}`

48 As submission is double blind, refer to your own published work in the third person. That is, use “In
49 the previous work of Jones et al. [4],” not “In our previous work [4].” If you cite your other papers
50 that are not widely available (e.g., a journal paper under review), use anonymous author names in the
51 citation, e.g., an author of the form “A. Anonymous.”

52 3.2 Footnotes

53 Footnotes should be used sparingly. If you do require a footnote, indicate footnotes with a number¹
54 in the text. Place the footnotes at the bottom of the page on which they appear. Precede the footnote
55 with a horizontal rule of 2 inches (12 picas).

56 Note that footnotes are properly typeset *after* punctuation marks.²

57 3.3 Figures

58 All artwork must be neat, clean, and legible. Lines should be dark enough for purposes of reproduction.
59 The figure number and caption always appear after the figure. Place one line space before the figure
60 caption and one line space after the figure. The figure caption should be lower case (except for first
61 word and proper nouns); figures are numbered consecutively.

62 You may use color figures. However, it is best for the figure captions and the paper body to be legible
63 if the paper is printed in either black/white or in color.

¹Sample of the first footnote.

²As in this example.

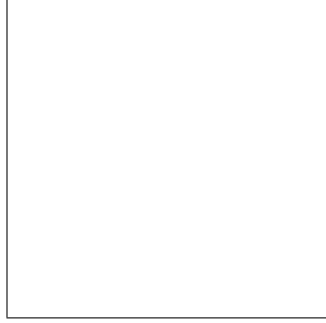


Figure 1: Sample figure caption.

Table 1: Sample table title

Part		
Name	Description	Size (μm)
Dendrite	Input terminal	~ 100
Axon	Output terminal	~ 10
Soma	Cell body	up to 10^6

64 3.4 Tables

65 All tables must be centered, neat, clean and legible. The table number and title always appear before
66 the table. See Table 1.

67 Place one line space before the table title, one line space after the table title, and one line space after
68 the table. The table title must be lower case (except for first word and proper nouns); tables are
69 numbered consecutively.

70 Note that publication-quality tables *do not contain vertical rules*. We strongly suggest the use of the
71 booktabs package, which allows for typesetting high-quality, professional tables:

72 `https://www.ctan.org/pkg/booktabs`

73 This package was used to typeset Table 1.

74 4 Final instructions

75 Do not change any aspects of the formatting parameters in the style files. In particular, do not modify
76 the width or length of the rectangle the text should fit into, and do not change font sizes (except
77 perhaps in the **References** section; see below). Please note that pages should be numbered.

78 5 Preparing PDF files

79 Please prepare submission files with paper size “US Letter,” and not, for example, “A4.”

80 Fonts were the main cause of problems in the past years. Your PDF file must only contain Type 1 or
81 Embedded TrueType fonts. Here are a few instructions to achieve this.

- 82 • You should directly generate PDF files using `pdflatex`.
- 83 • You can check which fonts a PDF files uses. In Acrobat Reader, select the menu
84 Files>Document Properties>Fonts and select Show All Fonts. You can also use the program
85 `pdffonts` which comes with `xpdf` and is available out-of-the-box on most Linux machines.
- 86 • The IEEE has recommendations for generating PDF files whose fonts are also ac-
87 ceptable for NIPS. Please see [http://www.emfield.org/icuwb2010/downloads/](http://www.emfield.org/icuwb2010/downloads/IEEE-PDF-SpecV32.pdf)
88 `IEEE-PDF-SpecV32.pdf`

- xfig "patterned" shapes are implemented with bitmap fonts. Use "solid" shapes instead.
- The `\bbold` package almost always uses bitmap fonts. You should use the equivalent AMS Fonts:

```
\usepackage{amsfonts}
```

followed by, e.g., `\mathbb{R}`, `\mathbb{N}`, or `\mathbb{C}` for \mathbb{R} , \mathbb{N} or \mathbb{C} . You can also use the following workaround for reals, natural and complex:

```
\newcommand{\RR}{I\!\!R} %real numbers
\newcommand{\Nat}{I\!\!N} %natural numbers
\newcommand{\CC}{I\!\!C} %complex numbers
```

Note that `amsfonts` is automatically loaded by the `amssymb` package.

If your file contains type 3 fonts or non embedded TrueType fonts, we will ask you to fix it.

5.1 Margins in L^AT_EX

Most of the margin problems come from figures positioned by hand using `\special` or other commands. We suggest using the command `\includegraphics` from the `graphicx` package. Always specify the figure width as a multiple of the line width as in the example below:

```
\usepackage[pdftex]{graphicx} ...
\includegraphics[width=0.8\linewidth]{myfile.pdf}
```

See Section 4.4 in the graphics bundle documentation (<http://mirrors.ctan.org/macros/latex/required/graphics/grfguide.pdf>)

A number of width problems arise when L^AT_EX cannot properly hyphenate a line. Please give LaTeX hyphenation hints using the `\-` command when necessary.

Acknowledgments

Use unnumbered third level headings for the acknowledgments. All acknowledgments go at the end of the paper. Do not include acknowledgments in the anonymized submission, only in the final paper.

References

References follow the acknowledgments. Use unnumbered first-level heading for the references. Any choice of citation style is acceptable as long as you are consistent. It is permissible to reduce the font size to `small` (9 point) when listing the references. **Remember that you can go over 8 pages as long as the subsequent ones contain *only* cited references.**

[1] Alexander, J.A. & Mozer, M.C. (1995) Template-based algorithms for connectionist rule extraction. In G. Tesauro, D.S. Touretzky and T.K. Leen (eds.), *Advances in Neural Information Processing Systems 7*, pp. 609–616. Cambridge, MA: MIT Press.

[2] Bower, J.M. & Beeman, D. (1995) *The Book of GENESIS: Exploring Realistic Neural Models with the GEneral NEural Simulation System*. New York: TELOS/Springer-Verlag.

[3] Hasselmo, M.E., Schnell, E. & Barkai, E. (1995) Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience* **15**(7):5249-5262.