GSoC Proposal for JuliaNLSolvers

Jiawei Li

Address: The Vale, Edgbaston, Birmingham, UK

Blog: http://lijiawei.cc

Email: li.jw@me.com

Github: iewaij

Hi, I'm Jiawei Li, currently an undergraduate Economics student in penultimate year. Being enthusiastic in programming and data science, I enjoy every minute solving problems, sharing thoughts, writing tutorials and making projects I'm proud of.

Why Julia and JuliaNLSolvers

I love Julia. It is fast, modern and easy to read. I used to write Python, the closed/open interval and weird self parameter is always a pain. As a language featuring OOP, it requires building an object first then build functions. But it is more intuitive to think of a function first rather than an object in my "functional" opinion. It also doesn't allow sharing same function across different objects unless inheriting. Julia deals with such issues elegantly and still keeps great performance, which proves that performance and high-level programming is not a trade-off.

I would be extremely excited to work for JuliaNLSolvers because it helps solve important quantitative problems in native Julia code. Optimization is key in nearly every scientific subjects, and nonlinear equations have major roles in Economics, such as solving general equilibrium. `LsqFit.jl` provides curve fitting function to estimate model using data. Building such packages in Julia code makes it possible to integrate their functionalities with other Julia packages. Meanwhile, contributing to these projects gives me an invaluable opportunity to move towards my goal to do further study in Computational Economics.

Project Overview

My plan mainly consists of three parts: documentation, benchmarks and improving functionality.

`LsqFit.jl`, and `NLsolve.jl` only have example codes in their READMEs. Documentation for these projects will be good references for users. Beginner's guide would dramatically reduce the learning curve for new users. Examples are also needed for `Optim.jl`, `LsqFit.jl` and `NLsolve.jl` to show people the Julia "pipeline" in areas such as Machine Learning, Statistics and Economics. Meanwhile, codes in documentation and examples can be used for testing.

Benchmarks are essential to show the advantage of Julia and therefore may persuade outside users to switch. By comparing with `SciPy`, it will also help guide development and find bugs.

`LsqFit.jl` is still on an early development stage and has large potential to improve. For example, allowing non-vectorized functions for `LsqFit.jl` will help it apply to more problems.

Deliverables

- Improved documentation with examples for `Optim.jl`, `LsqFit.jl` and `NLsolve.jl`
- Improved tests for `LsqFit.jl` and `NLsolve.jl`
- Speed and robustness benchmarks against `SciPy` for `LsqFit.jl` and `NLsolve.jl`
- Improved error handling for `NLsolve.jl`
- Convergence output after fitting data for `LsqFit.jl`
- Non-vectorized model functions for `LsqFit.jl`
- Trust Region Method for `LsqFit.jl`

Related Issues

- Documentation #154 `NLsolve.jl`
- Tests #2 `LsqFit.jl`
- Benchmarks/Experiments #150 `NLsolve.jl` #21 #34 `LsqFit.jl`
- Error handling #71 #133 `NLsolve.jl`
- Non-vectorized model functions #12 `LsqFit.jl`
- Trust Region Method #89 #116 `NLsolve.jl`

Contribution

I've already made following contributions:

- Decompose `estimate_errors()` and redefine alpha #64
- Improve JuliaCon 2017 Notebook
- Draft version of logistic regression example using `Optim.jl`

Project Plan

Before - 13 May

- ask for feature suggestions from community
- change development plan accordingly
- gain enough background knowledge and intuition
  - courses on convex optimization
  - courses on numeric methods
  - writing tutorials implementing what I've learnt using Julia
- fix minor issues of `LsqFit.jl` and `NLsolve.jl`

14 May - 31 May (Part-time)

- beginner's guide for `Optim.jl` based on Patrick's tutorial
- improve logistic regression example for `Optim.jl`
- population growth example for `Optim.jl`

1 Jun - 14 Jun

- build convergence output after fitting data for `LsqFit.jl`
- build documentation and beginner's guide for `LsqFit.jl` based on LsqFit.jl.ipynb
- improve tests for `LsqFit.jl` #2
- conditioning and valleys example for `Optim.jl` based on Patrick's tutorial

15 Jun - 29 Jun

- benchmarks/experiments for `LsqFit.jl` #21 #34 #59
- build documentation and beginner's guide for `NLsolve.jl` based on Patrick's tutorial
- improve tests for `NLsolve.jl`
- simple competitive equilibrium example based on p.187 Numerical Methods in Economics for `NLsolve.jl`
- Nash equilibrium example based on private provision of public good for `NLsolve.jl`

30 Jun - 13 Jul

- benchmarks for `NLsolve.jl` #150
- error handling for `NLsolve.jl` #71 #133
- exponential regression example for `Optim.jl`
- poisson regression example for `Optim.jl`

14 Jul - 27 Jul

- non-vectorized model functions for `LsqFit.jl` #12
- tests, documentation for above changes
- integrate `ForwardDiff.jl` with `Optim.jl` example

28 Jul - End

- Trust Region Method for `LsqFit.jl` #89 #116
- tests, benchmarks, documentation for above changes
- blogging for the project

Education

I'm now exchange student at University of Birmingham, UK. My home university is University of Nottingham, China. I've chosen lots of courses focusing on Mathematics and Statistics. Meanwhile, I took several online courses where I learned programming and data science by myself.

University Courses

- Optimization For Economists
- Econometrics
- Econometric Methods
- Mathematical Statistics for Economics

Online Courses

- Learning From Data (introductory Machine Learning)
- Introduction to Mathematical Thinking
- An Introduction to Interactive Programming in Python (Part 1)

Portfolio

Here're my code snippets covering basic algorithms, data processing and machine learning.
iewaij/code-examples

I'm also writing tutorials on data science which attract 681 followers in Zhihu and 36 stars in Github. Topics include exploratory data analysis, data collection, distributions, OLS, and Gauss-Markov Theorem.
iewaij/introDataScience

I like blogging on thoughts, mind hacks, programming, and Economic theories.
Pause Game

Work Experience

In 2016 and 2017, I led a team as project manager in Enactus Nottingham China, a university society focusing on social innovation. During the year, our team experimented social innovation ideas including culture database for fashion designers, campus vintage shop and food awareness codes.

In 2016 summer, I worked for China Next foundation, a charity foundation supporting NGO works in China, as project assistant. My job includes interviewing NGO leaders in China and producing reports on their project's impacts.

In 2015, I worked for the design team of Enactus Nottingham China. My job includes designing posters and Keynote slides.

Logistics

My university puts all academic year exams in one month, that is I'm having 10 exams during 3 May to 31 May. So I may have limited time developing in May. After that I'll be full-time working in Birmingham till early July, then I will return to Hangzhou, China.

- Work part-time till exam ends on 31 May
- Work full-time in Birmingham till early July
- Work full-time in Hangzhou till end