

Plan for Summer 2020

Contents

Introduction	3
Politics	3
MGGG GSoC project	3
Try and get my thesis published as a working paper	3
Try and get my work with Eubank and Rodden published as a working paper	3
Data Science and Programming	3
Build (parallel) path tracer (using BDRF) in Julia with Ross and Doyun	3
Self-study everything on the recommended reading list	4
Mathematics	4
Probability notes and exercises	4
Statistics notes and exercises	4
Simulation notes and exercises	5
Self-study: Introduction to Statistical Learning	6
Other fun stuff	6
Learn Classical Chinese	6
Learn German	6
(If got time) build that board game framework (during Game Jam July 10–12)?	6
Travel around	7
Get back to the gym	7
Admin things	7
Update blog and resume	7
Write blog post (year-in-review)	7
Update project page on blog	7
Send emails to the people I should email	7
Timeline	7
TTW8/Summer W0: 14–20 Jun	7
Summer W1 : 21–27 Jun	7

Summer W2 : 28 Jun-4 Jul	7
Summer W3 : 5-11 Jul	7
GMTK Game Jam: 10th - 12th July, 8pm-8pm	7
Summer W4 : 12-18 Jul	8
Summer W5 : 19-25 Jul	8
Summer W6 : 26 Jul-01 Aug	8
Summer W7 : 02 Aug-08 Aug	8
Summer W8 : 09 Aug-15 Aug	8
Summer W9 : 16 Aug-22 Aug	8
Summer W10 : 23 Aug-28 Aug	8

Introduction

This is what I plan to do this summer.

Politics

I aim to put up at least two working papers with a view towards publishing them. This closely relates to my upcoming Master's in data science. The summer grant would help defray the costs of running the many computationally-expensive simulations I will require for my working paper.

MGGG GSoC project

- Call Bhushan on Monday: see what I'll be doing.
 - Figuring out what communities of interest are (optimal transport? [arxiv link here](#))
 - Building geopandas in Julia? (have to learn Julia first)

Try and get my thesis published as a working paper

- Email Andy about next steps
- Maybe get Tak Huen to co-author with me
- Probably need to clean up the code and all the files

Try and get my work with Eubank and Rodden published as a working paper

- Wait for Nick's email back on Monday, see if he's still interested.
- Spend a few days looking through the codebase
- Double check the code to make sure it's giving correct results (how?)
 - Write tests? (Is it worth it??)
 - Move to Julia??? (no—absolutely not—horrible idea)
- See if there's a better way to memoise computations rather than saving the distance matrix as an enormous file (I already solved the “huge file in memory” problem with my thesis by reading line-by-line)

Data Science and Programming

Build (parallel) path tracer (using BDRF) in Julia with Ross and Doyun

- **How do I learn about path tracers? Theory?** Is there a textbook?
 - Ask Ross about the theory of path tracing
 - Write something really really simple myself before Ross comes in
 - Scope out the project

- Prime ray generation
- Learn the maths for Markov Chains to actually understand the BDRF function...
 - Stanford handout on Monte Carlo path tracing
 - Physically-based rendering textbook
 - Path tracing section

Self-study everything on the recommended reading list

Mathematics

- Matrix algebra – transpose, symmetric, rank, inverse, orthogonal, trace
- Matrix operations – multiplication, solution of linear system of equations
- Eigenvalues, eigenvectors
- Eigendecomposition and singular value decomposition of a matrix
- Differentiation – including partial differentiation, Hessian matrix
- Integration – including multiple integrals, and Jacobians and change of variables in multiple integrals
- Taylor series expansion
- Difference equations

The linear algebra notes 5 cover most of the required linear algebra. Jordan and Smith [12] covers the material at an appropriate level. The relevant material is commonly included in books on engineering mathematics: good examples are Kreyszig [13] and Stroud [18].

Probability notes and exercises

- Probability spaces
- Random variables – discrete and continuous
- Distributions, expectation, variance, covariance, independence
- Joint distributions, conditional distributions, method of change of variables using Jacobian
- Multivariate normal distribution – basic properties
- Moment generating functions
- Convergence of random variables, weak and strong law of large numbers, central limit theorem
- Basic properties of discrete-time Markov chains and Poisson processes

The Probability notes 1 cover the required material at an appropriate level, as do Chapters 1–6 of Grimmett and Stirzaker [11]. An alternative introductory probability book is Ross [17]. An alternative source for the material on Markov chains is Chapter 1 of Norris [14]. The books by Rice [15] and DeGroot and Schervish [9] cover most of this material.

Statistics notes and exercises

- Basic exploratory plots, e.g. histograms, boxplots, Q-Q plots

- Maximum likelihood estimation
- Properties of estimators – unbiasedness, consistency, mean squared error
- Delta method, asymptotic normality of maximum likelihood estimator
- Confidence intervals – exact intervals, approximate intervals using large sample theory
- Hypothesis testing, types of error, including t-tests (basic, paired, two sample)
- Likelihood ratio tests, asymptotic distribution of likelihood ratio statistic, applications to contingency tables, χ^2 goodness-of-fit tests
- Basic single and multiple linear regression
- Basic Bayesian statistics – conjugate prior and posterior, maximum a posteriori and expected posterior estimates, credible/highest posterior density intervals

The Statistics notes 2 cover the required material at an appropriate level. The books by Rice [15] and DeGroot and Schervish [9] provide good coverage of the required material. 2 Alternative references which are very suitable for MSc preparatory reading on statistics are Casella and Berger [7], Davison [8], Faraway [10], Wasserman [19]. The first five chapters of [19] also provide a summary of much of the prerequisite material on probability.

Simulation notes and exercises

Simulation methods: inversion, transformation, rejection. Importance sampling. Basic aspects of Markov Chain Monte Carlo. The basics of programming, in either R or Python.

Note: although simulation, and statistical programming in R, is taught in the MSc, it is covered at a fast pace and therefore some previous experience is a distinct advantage and expected. If you have not used R before, then you should gain some familiarity with it before the start of the MSc. If you have used Python but not R, we are assuming you would be able to make the transition to R quickly – you may still want to gain some experience with R before the start of the MSc.

The above material is covered in the Simulation notes [3] and the introduction to R and programming in the Statistical Programming notes [4]. Books that cover this material at an appropriate level are Braun and Murdoch [6], and Robert and Casella [16].

Statistical Programming notes and exercises: [here](#) Linear Algebra

Review and Reference: [here](#)

Self-study: Introduction to Statistical Learning

Other fun stuff

Learn Classical Chinese

Michael Fuller's *An Introduction to Literary Chinese*. It's much newer (1999, revised 2004), and takes into account recent scholarship into the grammar of Classical Chinese ("Classical" in its true sense, since most of the book's materials come from that period). It makes frequent references to Edwin Pulleyblank's *Outline of Classical Chinese Grammar*, so you may want to pick up a copy of that too for reference. There are other, more recent books than Fuller's, but from what I've seen none of them are of quite the same caliber.

Fuller starts out with 8 Lessons teaching the basic structure of the language. He also includes a little about reading classical commentary and resources to use when you have questions, such as large scholarly dictionaries. The second part, *Intermediate Texts*, consists of 16 lessons. They are all, like the *Beginning Texts*, from classical sources like Confucius, *Strategies of the Warring States*, Mencius, etc. The *Advanced Texts* section consists of longer readings from classical and Jin dynasty authors. The fourth part is selections of Tang and Song dynasty prose and poetry. He gradually gives you more work to do as the book goes on, including bibliographic exercises (verify what work X means by using dictionary Y; find country A on the map in atlas B, etc.) and further reading. The book is excellent, and really makes you think about why each phrase means what it does, from both grammatical and contextual points of view.

Goal: do one chapter a week, which will give us the basic grammar and two intermediate texts

Learn German

A Foundation Course in Learning German

16 chapters. Won't finish.

Goal: do one chapter a week, which will give me a lot of the grammar.

(If got time) build that board game framework (during Game Jam July 10–12)?

N.B. Before I came to Oxford I thought I would come out after three years being fluent in Latin. It turns out that's very far from the truth. (I have picked up several pretentious phrases in philosophy like *eo ipso*, *prima facie*, etc., but not much else)

Travel around

- After OberRamstadt, where?

Get back to the gym

- Deadlift, bench, (maybe) squat, pull-ups (rows), arms

Admin things

Update blog and resume

Write blog post (year-in-review)

Update project page on blog

Send emails to the people I should email

- Sergi
- Andy
- Prof Huang

Timeline

June 1 – Aug 31: GSOC internship

TTW8/Summer W0: 14–20 Jun

Do all the admin stuff

- Pack my stuff
- Send my stuff to Ober-Ramstadt
- Write emails
- Write blog posts
- Update website design
- Update projects page

Summer W1 : 21–27 Jun

Summer W2 : 28 Jun–4 Jul

Summer W3 : 5–11 Jul

GMTK Game Jam: 10th – 12th July, 8pm–8pm

Here would be a good time to either use : Julia path tracer, OR board game framework

Summer W4 : 12–18 Jul

Summer W5 : 19–25 Jul

Summer W6 : 26 Jul–01 Aug

Summer W7 : 02 Aug–08 Aug

Summer W8 : 09 Aug–15 Aug

Summer W9 : 16 Aug–22 Aug

Summer W10 : 23 Aug–28 Aug