Introduction and Logistics

Ryan P. Adams

COS 302 / SML 305
Mathematics for Numerical Computing and Machine Learning
Why does this course exist? A brief dialog

Setting: Bob and Alisha arrive back on campus after the summer.
Why does this course exist? A brief dialog

Setting: Bob and Alisha arrive back on campus after the summer.

Bob: Hi Alisha! What did you get up to this summer?

Alisha: Hey Bob! I spent it at Tesla. It was super cool.

Bob: Awesome. Did you get to meet Elon Musk?

Alisha: No, but I built a machine learning system for self-driving cars!

Bob: Whoa that's crazy. What did it do?

Alisha: Well, it used ML to predict how far a car is behind you. And it's deploying in the next update!

Bob: I guess taking COS 324 really paid off! How did the system work?
Why does this course exist? A brief dialog

Setting: Bob and Alisha arrive back on campus after the summer.

Bob: Hi Alisha! What did you get up to this summer?
Alisha: Hey Bob! I spent it at Tesla. It was super cool.
Why does this course exist? A brief dialog

Setting: Bob and Alisha arrive back on campus after the summer.

Bob: Hi Alisha! What did you get up to this summer?
Alisha: Hey Bob! I spent it at Tesla. It was super cool.
Bob: Awesome. Did you get to meet Elon Musk?
Why does this course exist? A brief dialog

Setting: Bob and Alisha arrive back on campus after the summer.

Bob: Hi Alisha! What did you get up to this summer?

Alisha: Hey Bob! I spent it at Tesla. It was super cool.

Bob: Awesome. Did you get to meet Elon Musk?

Alisha: No, but I built a machine learning system for self-driving cars!
Setting: Bob and Alisha arrive back on campus after the summer.

Bob: Hi Alisha! What did you get up to this summer?
Alisha: Hey Bob! I spent it at Tesla. It was super cool.
Bob: Awesome. Did you get to meet Elon Musk?
Alisha: No, but I built a machine learning system for self-driving cars!
Bob: Whoa that’s crazy. What did it do?
Why does this course exist? A brief dialog

Setting: Bob and Alisha arrive back on campus after the summer.

Bob: Hi Alisha! What did you get up to this summer?
Alisha: Hey Bob! I spent it at Tesla. It was super cool.
Bob: Awesome. Did you get to meet Elon Musk?
Alisha: No, but I built a machine learning system for self-driving cars!
Bob: Whoa that’s crazy. What did it do?
Alisha: Well, it used ML to predict how far a car is behind you.
    And it’s deploying in the next update!
Why does this course exist? A brief dialog

Setting: Bob and Alisha arrive back on campus after the summer.

Bob: Hi Alisha! What did you get up to this summer?
Alisha: Hey Bob! I spent it at Tesla. It was super cool.
Bob: Awesome. Did you get to meet Elon Musk?
Alisha: No, but I built a machine learning system for self-driving cars!
Bob: Whoa that’s crazy. What did it do?
Alisha: Well, it used ML to predict how far a car is behind you. And it’s deploying in the next update!
Bob: I guess taking COS 324 really paid off!
    How did the system work?
Alisha: Well, it was pretty straightforward.
Alisha: Well, it was pretty straightforward.

My team collected data with known ground-truth distances.
Alisha: Well, it was pretty straightforward. My team collected data with known ground-truth distances. The ML team has really good features for the car’s sensors.
Why does this course exist? A brief dialog

Alisha: Well, it was pretty straightforward. My team collected data with known ground-truth distances. The ML team has really good features for the car’s sensors. I used those feature vectors as basis functions for my model.
Alisha: Well, it was pretty straightforward. My team collected data with known ground-truth distances. The ML team has really good features for the car’s sensors. I used those feature vectors as basis functions for my model. Then I did a linear regression assuming Gaussian noise.
Why does this course exist? A brief dialog

Alisha: Well, it was pretty straightforward. My team collected data with known ground-truth distances. The ML team has really good features for the car’s sensors. I used those feature vectors as basis functions for my model. Then I did a linear regression assuming Gaussian noise. I set up an optimization problem to maximize the log likelihood.
Alisha: Well, it was pretty straightforward. My team collected data with known ground-truth distances. The ML team has really good features for the car’s sensors. I used those feature vectors as basis functions for my model. Then I did a linear regression assuming Gaussian noise. I set up an optimization problem to maximize the log likelihood. So I took the gradient of the resulting loss to find the zero.
Alisha: Well, it was pretty straightforward. My team collected data with known ground-truth distances. The ML team has really good features for the car’s sensors. I used those feature vectors as basis functions for my model. Then I did a linear regression assuming Gaussian noise. I set up an optimization problem to maximize the log likelihood. So I took the gradient of the resulting loss to find the zero. I found the weights by solving the overdetermined linear system.
Why does this course exist? A brief dialog

Bob (shocked and angry): What?!?!?!
Bob (shocked and angry): What?!?!?!?

It sounds like you just did math all summer!
Bob (shocked and angry): What?!?!?!? It sounds like you just did math all summer!

Machine learning is supposed to be COOL and FUN!
Why does this course exist? A brief dialog

Bob (shocked and angry): What?!?!?!?
   It sounds like you just did math all summer!
   Machine learning is supposed to be COOL and FUN!
   Gaussian basis log linear likelihood system optimizations?

Credit: Eleanor Adams
Why does this course exist? A brief dialog

Bob (shocked and angry): What?!?!?!?
It sounds like you just did math all summer!
Machine learning is supposed to be COOL and FUN!
Gaussian basis log linear likelihood system optimizations?
WHAT DOES ALL THAT MATH HAVE TO DO WITH SELF-DRIVING CARS?

Credit: Eleanor Adams
Bob (shocked and angry): What?!?!?!?
   It sounds like you just did math all summer!
   Machine learning is supposed to be COOL and FUN!
   Gaussian basis log linear likelihood system optimizations?
   WHAT DOES ALL THAT MATH HAVE TO DO WITH SELF-DRIVING CARS?

Alisha: Well, machine learning IS math.
   It’s probability, linear algebra, and optimization...
Bob (shocked and angry): What?!!?!??!
   It sounds like you just did math all summer!
   Machine learning is supposed to be COOL and FUN!
   Gaussian basis log linear likelihood system optimizations?
   WHAT DOES ALL THAT MATH
   HAVE TO DO WITH SELF-DRIVING CARS?
Alisha: Well, machine learning IS math.
   It’s probability, linear algebra, and optimization...
Bob: I’m a CS major.
   I took COS 340, but I don’t know what any of that means!
Why does this course exist?

- Historically, computer science has been centered on *discrete* math.
Why does this course exist?

- Historically, computer science has been centered on *discrete* math.
Why does this course exist?

- Historically, computer science has been centered on discrete math.
- Example: COS 340 uses a book called Mathematics for Computer Science, which covers logic, number theory, graphs and lots of great discrete mathematics.
- But now, a BIG chunk of what students want to learn is continuous math, like what you usually would learn in, say, electrical engineering.
Why does this course exist?

- Historically, computer science has been centered on discrete math.
- But now, a BIG chunk of what students want to learn is continuous math, like what you usually would learn in, say, electrical engineering.

- machine learning
Why does this course exist?

▶ Historically, computer science has been centered on *discrete* math.
▶ But now, a BIG chunk of what students want to learn is *continuous* math, like what you usually would learn in, say, electrical engineering.

▶ machine learning
▶ computer vision
Why does this course exist?

- Historically, computer science has been centered on *discrete* math.
- But now, a BIG chunk of what students want to learn is *continuous* math, like what you usually would learn in, say, electrical engineering.

- machine learning
- computer vision
- natural language processing
Why does this course exist?

- Historically, computer science has been centered on *discrete* math.
- But now, a BIG chunk of what students want to learn is *continuous* math, like what you usually would learn in, say, electrical engineering.

- machine learning
- computer vision
- natural language processing
- computational biology
Why does this course exist?

- Historically, computer science has been centered on \textit{discrete} math.
- Example: COS 340 uses a book called \textit{Mathematics for Computer Science}, which covers logic, number theory, graphs and lots of great \textit{discrete} mathematics.
- But now, a BIG chunk of what students want to learn is \textit{continuous} math, like what you usually would learn in, say, electrical engineering.
  
  - machine learning
  - computer vision
  - natural language processing
  - computational biology
  - theoretical neuroscience
Why does this course exist?

- Historically, computer science has been centered on discrete math.
- But now, a BIG chunk of what students want to learn is continuous math, like what you usually would learn in, say, electrical engineering.

- machine learning
- computer vision
- natural language processing
- computational biology
- theoretical neuroscience
- graphics
What is this course about?

- Linear algebra
  abstractions for reasoning about vector spaces
  Alisha: “...by solving the overdetermined linear system.”
What is this course about?

- **Linear algebra**
  - **abstractions for reasoning about vector spaces**
  - Alisha: “...by solving the overdetermined linear system.”

- **Vector calculus**
  - **understanding the geometry of multi-dimensional functions**
  - Alisha: “...took the gradient of the resulting loss.”
What is this course about?

- Linear algebra
  abstractions for reasoning about vector spaces
  Alisha: “...by solving the overdetermined linear system.”

- Vector calculus
  understanding the geometry of multi-dimensional functions
  Alisha: “...took the gradient of the resulting loss.”

- Probability
  reasoning about noise and uncertainty
  Alisha: “...assuming Gaussian noise for the labels.”
What is this course about?

- Linear algebra
  abstractions for reasoning about vector spaces
  Alisha: “...by solving the overdetermined linear system.”

- Vector calculus
  understanding the geometry of multi-dimensional functions
  Alisha: “...took the gradient of the resulting loss.”

- Probability
  reasoning about noise and uncertainty
  Alisha: “...assuming Gaussian noise for the labels.”

- Basic optimization
  finding the best fit to data
  Alisha: ‘...maximize the log likelihood.”
Linear algebra: weeks 1 through 5

- Linear systems
- Vector spaces
- Basis vectors
- Norms and inner products
- Orthogonality
- Projections
- Eigen-stuff
- Cholesky factorization
- Singular value decomposition
Differentiating functions of vectors
Probability: weeks 8-10

- Random variables
- Sampling
- Independence and dependence
- Monte Carlo
- Gaussian distributions
- Information theory
Optimization: weeks 11-12

- Constrained optimization
- Convex optimization
- Conjugate gradient method
Book: *Mathematics for Machine Learning*

Marc Peter Deisenroth  
A. Aldo Faisal  
Cheng Soon Ong  

(Cambridge University Press)

Freely available online at  
Is this class about machine learning?

No.

This course is about the math you need for machine learning and other computer science that depends on continuous mathematics.
Staff: TAs

Alexander
OH: Tue 5-7pm COS 201

Fangyin
OH: Fri 4-6pm COS 201

Haochen
OH: Mon 7-9pm FC Fishbowl

Ari
OH: Thu 5-7pm COS 201

Geoffrey
OH: Wed 4:30-6:30pm COS 201

Sulin
OH: Thu 7-9pm COS 201
Precepts

- Six weekly precepts, taught by TAs
- New material will be presented, so make sure you attend.
- Stick to your assigned precept.

- P01 (Haochen): Thu 9:00-9:50am, Friend Center 009
- P01A (Fangyin): Thu 9:00-9:50am, Friend Center 110
- P02 (Ari): Thu 10:00-10:50am, Friend Center 009
- P03 (Sulin): Fri 12:30-1:20pm, Friend Center 009
- P04 (Alexander): Fri 12:30-1:20pm, Friend Center 110
- P05 (Geoffrey): Fri 1:30-2:20pm, Friend Center 007
**Staff: Lab TAs**

Kenny  
OH: Mon 5-7pm COS 201

Alan  
OH: Sat 11am-1pm Fine 314

Michael  
OH: Sun 2-4pm COS 201
Typical weekly calendar

<table>
<thead>
<tr>
<th>Mon 2/10</th>
<th>Tue 2/11</th>
<th>Wed 2/12</th>
<th>Thu 2/13</th>
<th>Fri 2/14</th>
<th>Sat 2/15</th>
<th>Sun 2/16</th>
</tr>
</thead>
<tbody>
<tr>
<td>9am</td>
<td></td>
<td></td>
<td></td>
<td>9 - 9:30</td>
<td>Precept P0</td>
<td></td>
</tr>
<tr>
<td>10am</td>
<td></td>
<td></td>
<td>10 - 10:30</td>
<td>Precept P02 (Art)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11am</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12pm</td>
<td></td>
<td></td>
<td>12:30p - 1:20p</td>
<td>Lecture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1pm</td>
<td></td>
<td>1:30p - 3p</td>
<td>Staff Meeting</td>
<td>COS 401</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2pm</td>
<td>1p - 2p</td>
<td>1:30p - 3p</td>
<td>Ryan Office Hours</td>
<td>COS 411</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3pm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4pm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5pm</td>
<td>5p - 7p</td>
<td>5p - 7p</td>
<td>5p - 7p</td>
<td>5p - 7p</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6pm</td>
<td>Kenny Office Hours</td>
<td>COS 201</td>
<td>Alexander Office Hours</td>
<td>COS 201</td>
<td>Geoffrey Office Hours</td>
<td>COS 201</td>
</tr>
<tr>
<td>7pm</td>
<td>Sulin Office Hours</td>
<td>COS 201</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8pm</td>
<td>Friend Center Fish Bowl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9pm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Office hours available every day of the week!
Course website:
https://www.cs.princeton.edu/courses/archive/spring20/cos302/

Piazza:
http://piazza.com/princeton/spring2020/cos302

Staff email list:
cos302-s20@lists.cs.princeton.edu
Grading

- Homeworks: 60%
- In-class Midterm: 20%
- Final: 20%

- We’ll provide aggregate statistics about assignment/exam grades.
- You won’t get a letter grade until the final curve.
- Undergrads and grads will have different final curves.
- I generally will curve up but not down.
Homeworks

- 11 assignments, roughly weekly
- You’ll them individually.
- Will drop lowest grade.
- Some math, some code in Python (Colab) notebook
- Writeup in LaTeX
- Submit via Gradescope
- Regrade? Request via Gradescope up to two weeks after due date.

- Simple late policy: up to a week for 50% off. (No late days.)
Collaboration Policy

- We want you to be able to discuss the class material with each other, but we want the homework you submit to be your own work.
- You may never:
  - Share code.
  - Share writeups.
  - Search the internet for assignment solutions.
- You may always:
  - Discuss the related concepts and the high-level approach.
  - Discuss the results of your code at a high level, e.g., “I got 90% test accuracy.”
  - Search the internet for documentation on Python, Numpy, SciPy, etc.
- You should be wary of discussing details of proofs/derivations, your code, or results at an implementation level, rather than at the “big idea” level.
- In your assignment writeup, state who you discussed the problems with.
- You should assume that we will download your notebooks and run software to detect similarity between current and past code.
Python

- Python is the most widely used language for machine learning.
- It is gaining prevalence more broadly in scientific computation.
- Scientific stack: NumPy, SciPy, Matplotlib

- Isn’t it slow, since it is interpreted?
- For most things, no: BLAS and LAPACK libraries written in FORTRAN/C/C++
- Plus TensorFlow, PyTorch, JAX, etc. are in Python.

- We’ll use Jupyter notebooks for our workflow.
- In-browser execution with interleaved rich text and figures.
- Google Colaboratory: free Jupyter notebooks in the cloud. Avoids the need for you to get an identical Python installation on your laptop.
8. Detecting Outliers

An outlier is a point or set of points that are different from other points. Sometimes they can be very high or very low. It's often a good idea to detect and remove the outliers. Because outliers are one of the primary reasons for resulting in a less accurate model. Hence it's a good idea to remove them. The outlier detection and removing that I am going to perform is called IQR score technique. Often outliers can be seen with visualizations using a box plot. Shown below are the box plots of MSRP, Cylinders, Horsepower, and EngineSize. Herein all the plots, you can find some points are outside the box they are none other than outliers. The technique of finding and removing outlier that I am performing in this assignment is taken help of a tutorial from towards data science.

```python
[ ] sns.boxplot(x=df['Price'])
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f0d36a3b2e0>
```

```python
[ ] sns.boxplot(x=df['MP'])
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f0d369b3ba8>
```
\LaTeX\ (Usually pronounced LAH-tekh or LAY-tekh) is a typesetting system.\n
- Uses a markup language to produce nice-looking PDF files.
- Is the best system for mathematical documents.
- Very broadly used in computer science and other math-oriented fields.
- Learning \LaTeX\ is a good life skill.

- Assignments must be turned in using the COS 302 \LaTeX\ template.
- I strongly suggest you use the browser-based system Overleaf.
- We’ve made a screencast to show you how to get started.
Assignment #1
Due: 23:55pm February 12, 2020

Upload: [https://www.gradescope.com/courses/75501/assignments/316145](https://www.gradescope.com/courses/75501/assignments/316145)

Problem 1 (77pts)
In COS 302 you'll type your assignments in \LaTeX, which is the way that almost all documents are composed in mathematical fields. It's a good thing to learn and \LaTeX\ can produce beautiful documents. The first part of this assignment is going to get you started with it. There are a variety of ways to write and compile \LaTeX, from web-based tools like Overleaf, to various GUI tools for Mac, Linux, and Windows, to plaintext editors like Emacs.

If you aren't an experienced user of \LaTeX, with tools already installed on your computer, we strongly recommend that you use Overleaf for this course. Overleaf runs in your browser so you don't have to worry about Mac vs Linux vs Windows or Chromebook. To make this as easy as possible, we'll post links to templates of each homework to get you going. Just use the "copy project" option in the menu to the left. We have posted a screencast to walk you through it.

There are two files in the project. One of them, cos302.cls is a formatting (.cls) file that you should not have to worry about or edit. The other, hw1.tex is the file you'll actually edit. When you type it in \LaTeX, you're basically writing simple markup code. You then compile it and get a nice looking PDF file. Note that this compilation process just produces a PDF; it does not submit it for grading automatically or anything like that. To submit the PDF, you'll need to upload via Gradescope. A screencast is available to walk you through the process of submitting the homework.

To start out, look at the very top of the file where you see:

```latex
\begin{document}
\begin{problem} (77pts)
```

In COS 302 you'll type your assignments in \LaTeX, which is the way that almost all documents are composed in mathematical fields.

It's a good thing to learn and \LaTeX\ can produce beautiful documents. The first part of this assignment is going to get you started with it.

Here are a list of your assignments and template:

- [Assignment #1](https://www.gradescope.com/courses/75501/assignments/316145)
Additional resources

- Lots of great videos at http://videolectures.net and YouTube.
- Every lecture will have some optional resources for you to explore.
- Everybody learns in different ways.
- Sometimes hearing things from a different perspective is helpful.
Coaching versus refereeing: a thing to think about

The structure of our educational system conflates two contradictory roles.

instructor (coach): help you learn – most important

assessor (referee): assign you a grade – less important

The tragedy of this is viewing your professor only as antagonistic assessor.

Strategies I employ to try to reduce this effect:

▶ Dedicated graders so that your TAs can have a limited/no assessment role.
▶ Gradescope so that rubrics are clear.
▶ Simple late policy to reduce “assessment surface area”.
▶ Lots of staff office hours availability.
▶ Be as transparent as possible about assessment process.
A message to Ph.D. students

- This is an undergraduate course.
- The material will target COS undergraduates.
- You must participate constructively or I will ask you to leave.
- If you feel it’s going too slow, take a different class.
Anonymous Questionnaire

Help us understand your background and why you’re taking the course.

Fill out the anonymous survey at:

http://tiny.cc/cos302s20
Questions?
Looking forward to a great term!