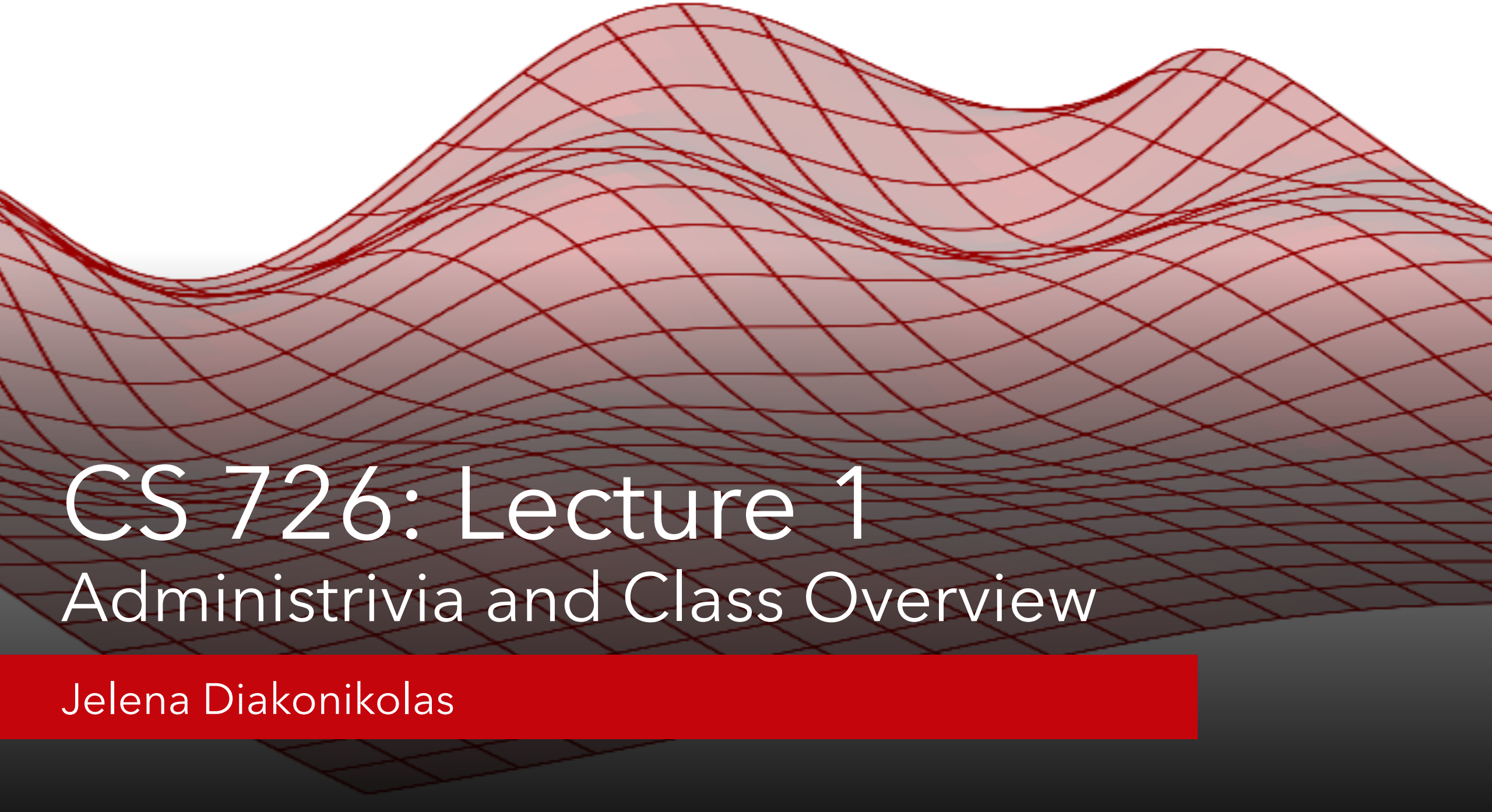

CS 726: Lecture 1

Administrivia and Class Overview

Jelena Diakonikolas

Nice to Meet You!

- About me:
  - My name: Jelena (pronounced as Yelena) Diakonikolas
  - Email: jelena@cs.wisc.edu
  - Office hours: Mon 9-10am and 4-5pm (email me if planning to show up)
  - Assistant professor in the computer science department since Jan 2020
  - Postdoctoral fellow at UC Berkeley 2018-2020; Simons Research Fellow in 2018
  - Postdoctoral associate in the CS department at Boston University 2016-2018
  - Graduated from Columbia University in 2016, in electrical engineering
  - My research is closely related to what you'll learn in this class!
- TA: Eric Lin clin353@wisc.edu; OH: Wed 9-10am and 4-5pm
- Communication Policy: use primarily piazza + office hours

Lecture Format

- Meet on **Mon and Wed** 2.30-3.45pm on BBC Ultra
- Fri slot is supplementary: we'll only use it as needed
- Attendance is not required, but is appreciated
- All lectures are recorded and available on Canvas
- Lectures are whiteboard-style (on ipad); notes posted after the class
- Focus is on theory (mathematical proofs)

Workload and Assessment

- Core graduate class in optimization: challenging (by design) and with a high workload
- Homework (5-6):
  - 30% of the grade
  - a combination of math problems and coding assignments
  - no collaboration allowed; any discussion must be verbal-only
- Midterm:
  - 30% of the grade
  - to be scheduled in mid-October
  - typically 4 multi-part questions of a similar format/difficulty as homework questions
- Final:
  - 40% of the grade
  - scheduled for 12/16/2020
  - similar format to midterm

Topics Covered in Class

- Introduction: optimization background; convex sets; convex functions; convergence rates.
- Background on smooth unconstrained optimization: Taylor theorem and optimality conditions.
- First-order methods:
  - gradient descent for convex and nonconvex optimization, line search methods;
  - projections, gradient mapping, and their use in projected gradient descent;
  - Bregman divergence and mirror descent;
  - Nesterov acceleration for convex optimization;
  - conjugate gradients; lower bound for smooth convex minimization;
  - conditional gradients (Frank-Wolfe methods);
  - stochastic gradient descent.
- Second-order methods:
  - Newton method;
  - trust-region Newton;
  - inexact Newton methods and Newton-CG;
  - cubic regularization;
  - quasi-Newton methods (DFP, BFGS, SR-1, general Broyden class);
  - limited-memory quasi-Newton (L-BFGS).

Planning for COVID-19

- Due to privacy concerns, I will **not** be notified if someone tests positive
- Campus is yet to specify whether any accommodations would be handled through the McBurney center
- For now: it is up to you to notify me if you test positive and require any accommodation (making up for missed coursework; getting an incomplete grade). I will provide accommodations where possible
- Homework will always be posted at least 10-12 days before the deadline
- HW extension policy: blanket approval for up to 6 days. The late days are counted in full days increments: if you are 1min late or 23h 59m late, both would count as a full day

Some Additional Policies

- Disability accommodation:
  - if you have a disability and require accommodation, please request it through the McBurney center – they determine the appropriate accommodation and notify me
  - this is the only way I am allowed to accommodate you
- Diversity & Inclusion:
  - this class is a welcoming community
  - please let me know if any issues arise that you make you feel like you do not belong
- Academic Honesty:
  - strictly enforced: no collaboration on assignments is allowed, unless specified otherwise; communication can be *verbal only*
  - any signs of copying/plagiarism will be harshly penalized – not worth it!

Other Remarks

- Waitlist: currently, there are 50 waitlisted students; the class capacity is 40 and I may open a few additional seats, but nothing beyond that
- Class materials:
  - "official" textbook is by Nocedal & Wright (Numerical Optimization, 2nd ed, 2006)
  - we will use the textbook only for some topics; for others, additional materials will be provided on Canvas
  - all class materials are for your use only – do not share them outside the class!
- Piazza:
  - if a question is not time-sensitive, we will give your peers the chance to respond first
  - if the answer is not complete (or there is no answer), we will post the answer
- Study rooms (proposal):
  - online rooms (probably on BBC Ultra) where you can meet your peers and study together; not supervised by me or the TA

How to Succeed in this Class?

- Do not miss lectures (either live or recorded)!
- Make sure you understand the material from lectures: ask questions!
- What does "understand" mean:
  - level 0: able to follow the steps I do in class
  - level 1: able to solve problems and reproduce proofs you have seen before
  - level 2: able to solve problems and produce proofs you have **not** seen before, but are similar to the ones from class/assignments (enough for doing well in class)
  - level 3: able to solve problems and produce proofs you have **not** seen before, and are **not** similar to the ones from class/assignments (this is research-level)
- "Practice your scales"!
- Start early on each homework assignment!