GEORGE MASON UNIVERSITY
Department of Mathematical Sciences
4400 University Drive, MS 3F2, Fairfax, Virginia 22030

Math 689, Spring 2021
Optimization, Nonlinear PDEs and Deep Learning

Instructor: Harbir Antil, Mathematical Sciences
Office: Exploratory Hall, room 4201
Phone: (703) 993-5777
Email: hantil@gmu.edu
Webpage: `http://math.gmu.edu/~hantil`

Goal: The goal of this course is to create a mathematical understanding of Deep Learning (DL). We shall cast the DL problems as constrained optimization problems. We will learn how to use a large class of algorithms and DL architectures to efficiently solve vast number of problems such as classification problems, problems in imaging science, parameterized partial differential equations (PDEs) and inverse problems constrained by PDEs.

The course will have both theoretical as well as numerical components. Every theoretical aspect will be tested numerically. This course is designed to provide a comprehensive foundation for the aforementioned topics which will be used to create a deeper mathematical understanding of real life problems.

Course description:[1] We will cover the following topics (suggested reading: [6, 2, 4, 3, 5, 7, 8, 1])

- **Week 1:** Quadratic optimization problem, Equality constrained quadratic program
- **Week 2:** Introduction to machine learning, Least square regression, Logistic regression and cross entropy, and Regularization.
- **Week 3:** Linear classification using cross entropy and Support vector machines
- **Week 4:** Universal approximation theorem, Activation functions and nonlinear classification models
- **Week 5:** Linear elliptic PDEs, Optimal control problems, Finite difference and finite element methods to solve PDEs
- **Week 6:** Stationary iterative methods: Jacobi, Gauss-Seidel and SOR, Applications to linear Elliptic PDEs and Least squares problems.
- **Week 7:** Krylov subspace methods: Conjugate gradient method, GMRES with applications to PDEs and Least square regression problems.
- **Week 8:** Iterative methods for unconstrained optimization: Gradient descent, Stochastic gradient descent, Line search, Trust-Region methods.
- **Week 9:** Single layer neural networks
- **Week 10:** Deep neural networks
- **Week 11:** Stability of DNN and relation to ODEs
- **Week 12:** Parameterized PDEs and application of DNNs
- **Week 13:** Paper reading
- **Week 14:** Paper reading

[1] Hand-written notes will be provided

Homework (70%): There will be several HOMEWORKS which will amount to 70% of the final grade. There will be a penalty of 10% per day late; homeworks will not be accepted after one week.

Students are encouraged to work in groups of up to three students but must hand in an individual self written proofs and answers.

Exams (30%): There will be a FINAL project and a presentation which will constitute 30% of the final grade.

Academic Integrity: GMU is an Honor Code university; please see the Office for Academic Integrity for a full description of the code and the honor committee process. The principle of academic integrity is taken very seriously and violations are treated gravely. What does academic integrity mean in this course? Essentially this: when you are responsible for a task, you will perform that task. When you rely on someone else's work in an aspect of the performance of that task, you will give full credit in the proper, accepted form. Another aspect of academic integrity is the free play of ideas. Vigorous discussion and debate are encouraged in this course, with the firm expectation that all aspects of the class will be conducted with civility and respect for differing ideas, perspectives, and traditions. When in doubt (of any kind) please ask for guidance and clarification.

Mason email accounts. Students must use their MasonLIVE email account for any correspondence during this course. For more information see: `http://masonlive.gmu.edu`.

Office of Disability Services. If you are a student with a disability and you need academic accommodations, please see me and contact the Office of Disability Services (ODS) at 993-2474, `http://ods.gmu.edu`. All academic accommodations must be arranged through the ODS.

University policies The University Catalog, `http://catalog.gmu.edu`, is the central resource for university policies affecting student, faculty, and staff conduct in university academic affairs. Other policies are available at `http://universitypolicy.gmu.edu`. All members of the university community are responsible for knowing and following established policies.

References

[1] Harbir Antil, Drew P. Kouri, Martin-D. Lacasse, and Denis Ridzal, editors. *Frontiers in PDE-constrained optimization*, volume 163 of *The IMA Volumes in Mathematics and its Applications*. Springer, New York, 2018. Papers based on the workshop held at the Institute for Mathematics and its Applications, Minneapolis, MN, June 6–10, 2016.

[2] Yoshua Bengio, Ian Goodfellow, and Aaron Courville. *Deep learning*, volume 1. MIT press Massachusetts, USA:, 2017.

[3] S. C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[4] H. C. Elman, D. J. Silvester, and A. J. Wathen. *Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2005.

[5] L.C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[6] Guanghui Lan. *First-order and Stochastic Optimization Methods for Machine Learning*. Springer, 2020.

[7] Jorge Nocedal and Stephen J. Wright. *Numerical optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2006.

[8] Fredi Tröltzsch. *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.