CMSC838B/CMSC498Z: Differentiable Programming Reading List

Differentiation, Jacobian and higher-order derivatives

- The Matrix Calculus You Need For Deep Learning (https://arxiv.org/pdf/1802.01528)
- Differentiation of Functions of Several Variables (https://math.libretexts.org/Bookshelves/Calculus/Calculus_(OpenStax)/14:_Differentiation_of_Functions_of_Several_Variables)
- The Matrix Cookbook (https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf)
- On the calculation of Jacobian matrices by the Markowitz rule (https://www.osti.gov/servlets/purl/5933713)
- Optimal accumulation of Jacobian matrices by elimination methods on the dual computational graph (https://link.springer.com/article/10.1007/s10107-003-0456-9)
- Edge Pushing is Equivalent to Vertex Elimination for Computing Hessians (https://par.nsf.gov/servlets/purl/10039361)
- Efficient Symbolic Differentiation for Graphics Applications (https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/main-65.pdf)
- Lazy Multivariate Higher-Order Forward-Mode AD (https://engineering.purdue.edu/~qobi/papers/popl2007a.pdf)
- Greedy Algorithms for Optimizing Multivariate Horner Schemes (https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.330.7430&rep=rep1&type=pdf)

Automatic Differentiation

- Automatic Differentiation in Machine Learning: A survey (https://arxiv.org/pdf/1502.05767)
- Differentiable Programming from Scratch (https://thenumb.at/Autodiff/)
- Automatic Differentiation: The most criminally underused tool in the potential machine learning toolbox? (https://justindomke.wordpress.com/2009/02/17/automatic-differentiation-the-most-criminally-underused-tool-in-the-potential-machine-learning-toolbox/)
- Introduction to Automatic Differentiation (https://alexey.radul.name/ideas/2013/introduction-to-automatic-differentiation/)
- Wengert 1964 “A Simple Automatic Derivative Evaluation Program” (https://www.cs.princeton.edu/courses/archive/fall19/cos597C/files/wengert1964.pdf)
- Fast reverse-mode automatic differentiation using expression templates in C++ (http://www.met.reading.ac.uk/~swrhgnrj/publications/adept.pdf)
- ADrien: an implementation of automatic differentiation in Maple (https://dl.acm.org/doi/10.1145/309831.309941)
- Hierarchical Approaches to Automatic Differentiation (https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.49.2929&rep=rep1&type=pdf)
- Simon Peyton Jones, “Automatic Differentiation for Dummies” (https://2019.ecoop.org/details/ecoop-2019-papers/11/Automatic-Differentiation-for-Dummies)
    - Not really for dummies but it covers how functional programming people think about this problem.
- Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges (https://arxiv.org/abs/2104.13478, https://www.youtube.com/watch?v=w6Pw4MOzMuo, https://towardsdatascience.com/geometric-foundations-of-deep-learning-94cdd45b451d)

- Randomized Automatic Differentiation (https://arxiv.org/abs/2007.10412)

- Perturbation confusion confusion (http://blog.sigfpe.com/2011/04/perturbation-confusion-confusion.html)
- Automatic Differentiation: Inverse Accumulation Mode (https://openreview.net/forum?id=Bygj2Ys6IS)
- Issues in Parallel Automatic Differentiation (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.57.2468)
- Who Invented the Reverse Mode of Differentiation? (https://ftp.gwdg.de/pub/misc/EMIS/journals/DMJDMV/vol-ismp/52_griewank-andreas-b.pdf)

Systems & Libraries

- PyTorch: An Imperative Style, High-Performance Deep Learning Library (https://arxiv.org/abs/1912.01703)
- How Usability Improves Performance in Pytorch (https://slideslive.com/38955641/how-usability-improves-performance-in-pytorch?ref=account-84635-latest)
- Tangent: Automatic differentiation using source-code transformation for dynamically typed array programming (https://arxiv.org/abs/1809.09569)
- Differentiable Programming for Image Processing and Deep Learning in Halide (https://people.csail.mit.edu/tzumao/gradient_halide/)
- Enoki (https://rgl.s3.eu-central-1.amazonaws.com/media/papers/NimierDavidVicini2019Mitsuba2_7.pdf)
- Opt (http://optlang.org/)
- Enzyme (https://enzyme.mit.edu/)
- The ADIFOR 2.0 System for the Automatic Differentiation of Fortran 77 Programs (https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.31.399&rep=rep1&type=pdf)
- Recipes for Adjoint Code Construction (TAMC) (http://twister.caps.ou.edu/OBAN2019/Giering_recipe4adjoint.pdf)
- Kotlin∇: A shape-safe DSL for differentiable programming (https://openreview.net/forum?id=SkluMSZ08H)
- CasADi: a software framework for nonlinear optimization and optimal control (https://link.springer.com/article/10.1007/s12532-018-0139-4)
- The Tapenade automatic differentiation tool (https://hal.inria.fr/hal-00913983/PDF/tapenadeRefV2.pdf)
- Relay: A New IR for Machine Learning Frameworks (https://arxiv.org/pdf/1810.00952.pdf)