

CMSC838B/CMSC498Z: Differentiable Programming Reading List

Jacobian and Higher-order Derivatives (Tues, Sept. 9, 2025 and Thur, Sept. 11, 2025)

- 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 (Tues, Sept. 16, 2025)

- 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=Byqj2Ys6IS>)

- 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-ismv/52_griewank-andreas-b.pdf)

Backpropagation (Thur, Sept. 18, 2025)

- Backpropagation as Lagrange Multipliers: Lecun 1988 “A Theoretical Framework for Back-Propagation” (<http://yann.lecun.com/exdb/publis/pdf/lecun-88.pdf>)
 - Backprop is not just the chain rule <https://timvieira.github.io/blog/post/2017/08/18/backprop-is-not-just-the-chain-rule/>
- Backpropagation-Friendly Eigendecomposition (<https://arxiv.org/abs/1906.09023>)

Differentiable Sorting & Optimization (Tues, Sept. 23, 2025)

- Fast Differentiable Sorting and Ranking (<https://arxiv.org/abs/2002.08871>)
- Differentiable Ranks and Sorting using Optimal Transport (<https://arxiv.org/abs/1905.11885>)
- SATNet: Bridging deep learning and logical reasoning using a differentiable satisfiability solver (<https://arxiv.org/abs/1905.12149>)
- Differentiable Convex Optimization Layers (<https://arxiv.org/abs/1910.12430>)
- Differentiable Learning of Submodular Models (differentiable graph cut) (<https://las.inf.ethz.ch/files/djolonga17learning.pdf>)
- OptNet: Differentiable Optimization as a Layer in Neural Network (<https://arxiv.org/abs/1703.00443>)
- Reverse accumulation and attractive fixed points (<https://www.tandfonline.com/doi/abs/10.1080/10556789408805572>)
- Derivative convergence for iterative equation solvers (<https://www.tandfonline.com/doi/abs/10.1080/10556789308805549#:~:text=When%20nonlinear%20equation%20solvers%20are,functions%20of%20these%20variable%20parameters.&text=We%20also%20formulate%20a%20constructive,on%20the%20implicit%20function%20theorem>)
- Gradient-based hyperparameter optimization through reversible learning (<http://proceedings.mlr.press/v37/maclaurin15.html>)
- Learning to Optimize Non-Rigid Tracking (<https://niessnerlab.org/projects/li2020learning.html>)
- Differentiable Dynamic Programming for Structured Prediction and Attention (<https://arxiv.org/abs/1802.03676>)
- Efficient and Modular Implicit Differentiation (<https://arxiv.org/abs/2105.15183>)
- Slice Sampling Reparameterization Gradients (https://openreview.net/forum?id=cT_RMSqVf4)

Checkpoint Methods

- Automatic Computation of Derivatives with the use of the Multilevel Differentiating Technique--I. Algorithmic Basis (<https://core.ac.uk/download/pdf/82685242.pdf>)
- Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation (<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.455.4143&rep=rep1&type=pdf>)
- Dynamic Tensor Rematerialization (https://openreview.net/forum?id=Vfs_2RnOD0H)
- Divide-and-conquer checkpointing for arbitrary programs with no user annotation (<https://arxiv.org/pdf/2006.09616.pdf>)

Differentiable Programming: Semantics, Correctness, Modularity and Functions

- QuanTaichi: A Compiler for Quantized Simulations <https://yuanming.taichi.graphics/publication/2021-quantaichi/>
- Taichi: A Language for High-Performance Computation on Spatially Sparse Data Structures <https://yuanming.taichi.graphics/publication/2019-taichi/taichi-lang.pdf>
- Reverse-Mode AD in a Functional Framework: Lambda the Ultimate Backpropagator <http://www.bcl.cs.may.ie/~barak/papers/toplas-reverse.pdf>
- Systematically Differentiating Parametric Discontinuities (<https://people.csail.mit.edu/sbanganu/projects/teg-2021/>)
- Beautiful differentiation (<http://conal.net/papers/beautiful-differentiation/>)
- The simple essence of automatic differentiation (<http://conal.net/papers/essence-of-ad/>)
- Decomposing reverse-mode automatic differentiation (<https://popl21.sigplan.org/details/lafi-2021-papers/9/Decomposing-reverse-mode-automatic-differentiation>)
- A Simple Differentiable Programming Language (<https://arxiv.org/abs/1911.04523>)
warning: not simple at all
- Automatic Differentiation in PCF (<https://arxiv.org/abs/2011.03335>)
- λ S: Computable semantics for differentiable programming with higher-order functions and datatypes (<https://dl.acm.org/doi/10.1145/3434284>)
- On Correctness of Automatic Differentiation for Non-Differentiable Functions (<https://cs.stanford.edu/people/wonyeol/papers/2020-neurips.pdf>)
- Backpropagation with Continuation Callbacks: Foundations for Efficient and Expressive Differentiable Programming (<https://proceedings.neurips.cc/paper/2018/hash/34e157766f31db3d2099831d348a7933-Abstract.html>)
- The if-problem in automatic differentiation (<https://www.sciencedirect.com/science/article/pii/S0377042794902941>)
- The Differentiable Curry (<https://openreview.net/forum?id=ryxuz9SzDB>)
- Getting to the Point. Index Sets and Parallelism-Preserving Autodiff for Pointful Array Programming (Dex) (<https://arxiv.org/abs/2104.05372>)
- Backpropagation in the Simply Typed Lambda-Calculus with Linear Negation (<https://arxiv.org/abs/1909.13768>)
- Finally, a Polymorphic Linear Algebra Language (<https://2019.ecoop.org/details/ecoop-2019-papers/5/Finally-a-Polymorphic-Linear-Algebra-Language>)

- Parallelism-preserving automatic differentiation for second-order array languages (<https://dl.acm.org/doi/abs/10.1145/3471873.3472975>)
- Automatic Differentiation for Message-Passing Parallel Programs (<https://www.osti.gov/servlets/purl/10572>)

Differentiable Geometry Processing (Thur, Sept. 25, 2025)

- HodgeNet: Learning Spectral Geometry on Triangle Meshes (<https://people.csail.mit.edu/smirnov/hodgenet/>)
- Learning Geometric Operators on Meshes (<https://rlgm.github.io/papers/28.pdf>)
- MeshCNN: A Network with an Edge (<https://ranahanocka.github.io/MeshCNN/>)
- Geodesic convolutional neural networks on Riemannian manifolds (https://www.cv-foundation.org/openaccess/content_iccv_2015_workshops/w22/papers/Masci_Geodesic_Convolutional_Neural_ICCV_2015_paper.pdf)
- TextureNet: Consistent Local Parametrizations for Learning from High-Resolution Signals on Meshes (<https://arxiv.org/abs/1812.00020>)
- DualConvMesh-Net: Joint geodesic and Euclidean convolutions on 3D meshes (<https://arxiv.org/abs/2004.01002>)

Differentiable Physics (Tues, Sept. 30, 2025 and Tues, Oct. 7, 2025 and Thur, Oct. 9, 2025)

- DiffTaichi: Differentiable Programming for Physical Simulation (<https://yuanming.taichi.graphics/publication/2020-difftaichi/>)
- ChainQueen: A Real-Time Differentiable Physical Simulator for Soft Robotics (<https://yuanming.taichi.graphics/publication/2019-chainqueen/>)
- PlasticineLab: A Soft-Body Manipulation Benchmark with Differentiable Physics (<https://openreview.net/pdf?id=xCcdBRQEDW>)
- Course: Underactuated Robotics (<http://underactuated.mit.edu/dp.html>, <http://underactuated.mit.edu/trajopt.html>, http://underactuated.mit.edu/policy_search.html)
- Fluid Control Using the Adjoint Method (<https://grail.cs.washington.edu/projects/control/fluidAdjoint.pdf>)
- End-to-End Differentiable Physics for Learning and Control (<https://papers.nips.cc/paper/2018/file/842424a1d0595b76ec4fa03c46e8d755-Paper.pdf>)
- Differentiable Cloth Simulation for Inverse Problems (<https://papers.nips.cc/paper/2019/hash/28f0b864598a1291557bed248a998d4e-Abstract.html>)
- Scalable Differentiable Physics for Learning and Control (<https://arxiv.org/abs/2007.02168>)
- Scalable Gradients for Stochastic Differential Equations (<https://arxiv.org/abs/2001.01328>)
- Functional Optimization of Fluidic Devices with Differentiable Stokes Flow (<https://people.csail.mit.edu/taodu/stokes/index.html>)
- Solver-in-the-Loop: Learning from Differentiable Physics to Interact with Iterative PDE-Solvers (<https://ge.in.tum.de/publications/2020-um-solver-in-the-loop/>)
- Learning to Control PDEs with Differentiable Physics (<https://ge.in.tum.de/publications/2020-iclr-holl/>)

- Lagrangian Fluid Simulation with Continuous Convolutions (<https://ge.in.tum.de/publications/2020-ummenhofer-iclr/>)
- ADD: Analytically Differentiable Dynamics for Multi-Body Systems with Frictional Contact (<https://arxiv.org/abs/2007.00987>)
- Machine Learning for Fluid Mechanics (<https://www.annualreviews.org/doi/abs/10.1146/annurev-fluid-010719-060214>)
- Underwater Soft Robot Modeling and Control with Differentiable Simulation (<https://people.csail.mit.edu/taodu/starfish/index.html>)

Differentiable Rendering

- Differentiable Monte Carlo Ray Tracing through Edge Sampling (<https://people.csail.mit.edu/tzuma/difftr/>)
- Flanders 1973, "Differentiation Under the Integral Sign" (http://sgpwe.izt.uam.mx/files/users/uami/jdf/proyectos/Derivar_inetegral.pdf)
 - Wikipedia "Leibniz integral rule" https://en.wikipedia.org/wiki/Leibniz_integral_rule
- Unbiased Warped-Area Sampling for Differentiable Rendering (<https://people.csail.mit.edu/sbanganu/projects/was-2020/index.html>)
- Monte Carlo Estimators for Differential Light Transport (<https://rgl.epfl.ch/publications/Zeltner2021MonteCarlo>)
- Path Replay Backpropagation: Differentiating Light Paths using Constant Memory and Linear Time (<https://rgl.epfl.ch/publications/Vicini2021PathReplay>)
- Antithetic Sampling for Monte Carlo Differentiable Rendering (<https://shuangz.com/projects/antithetic-sg21/>)
- Modular Primitives for High-Performance Differentiable Rendering (<https://arxiv.org/abs/2011.03277>)
- Vectorization for Fast, Analytic, and Differentiable Visibility (https://sites.cs.ucsb.edu/~lingqi/publications/paper_vectorization.pdf)
- Differentiable scattering matrix for optimization of photonic structures (https://www.osapublishing.org/DirectPDFAccess/0EA8DD40-C275-4D1E-9557DE98B0979B39_444132/oe-28-25-37773.pdf?da=1&id=444132&seq=0&mobile=no)

Physics-Informed Neural Networks

- Neural Ordinary Differential Equations (<https://arxiv.org/abs/1806.07366>)
- Neural Jump Stochastic Differential Equations (<https://arxiv.org/abs/1905.10403>)
- Hamiltonian Neural Networks (<https://arxiv.org/abs/1906.01563>)
- Lagrangian Neural Networks (<https://arxiv.org/abs/2003.04630>)
- Fourier Neural Operator for Parametric Partial Differential Equations (<https://arxiv.org/abs/2010.08895>)

Performance Optimization

- Efficient Differentiable Programming in a Functional Array-Processing Language (<https://dl.acm.org/doi/10.1145/3341701>)
- Automatic differentiation for adjoint stencil loops (<https://arxiv.org/abs/1907.02818>)

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/tzuma/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>)

Rotational-Equivariant Networks

- Group Equivariant Convolutional Networks (<https://arxiv.org/abs/1602.07576>)
- Spherical CNNs (<https://openreview.net/pdf?id=Hkbd5xZRb>)
- Gauge Equivariant Convolutional Networks and the Icosahedral CNN (<https://arxiv.org/abs/1902.04615>)
- Vector Neurons: A General Framework for SO(3)-Equivariant Networks (<https://arxiv.org/abs/2104.12229>)
- Euclidean Symmetry and Equivariance in Machine Learning (<https://www.sciencedirect.com/science/article/abs/pii/S2589597420302641>)
- SE(3)-Equivariant Graph Neural Networks for Data-Efficient and Accurate Interatomic Potentials (<https://arxiv.org/abs/2101.03164>)

Graph networks

- Graph Neural Networks: A Review of Methods and Applications (<https://arxiv.org/abs/1812.08434>)
- Graph Signal Processing for Machine Learning: A Review and New Perspectives (<https://ieeexplore.ieee.org/document/9244180>)
- Dynamic Graph CNN for Learning on Point Clouds (<https://arxiv.org/abs/1801.07829>)

Permutation invariance

- Deep Sets (<https://arxiv.org/abs/1703.06114>)
- PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation (<https://arxiv.org/abs/1612.00593>)

Others

- Categorical Reparameterization with Gumbel-Softmax (<https://arxiv.org/pdf/1611.01144.pdf>)
- Differential Manipulation <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.30.190&rep=rep1&type=pdf>
 - Through-the-lens camera control (<https://dl.acm.org/doi/10.1145/142920.134088>)
- Machine Learning Systems are Stuck in a Rut (<https://dl.acm.org/doi/10.1145/3317550.3321441>)