Applied Math 210: Algebraic Fundamentals of Representing Data

Anna Seigal

Fall 2023

Outline

Algebra gives mathematical abstractions that allow us to process information. Many optimization problems in data and learning are built on algebraic ideas. For example, principal component analysis finds a low-rank approximation of a matrix, a problem central to linear algebra. This course builds out from this example to study the algebraic fundamentals of optimization problems to find representations of data. The course is in four parts.

1. The course begins with *dimensionality reduction*, which seeks to simplify high dimensional data, projecting it to a lower dimensional space while preserving salient structure. We study algorithms to reduce dimensionality of data and the theorems that guide them. This takes us from the linear algebra of principal component analysis to the multilinear algebra of multidimensional data. We examine the algebra and geometry of low-rank tensors and connect to examples concerning biological data.

2. We then move to *statistics*, where we see the algebraic structure present in optimization problem to learn parameters in statistical models. We extract multilinear structure from the problem at hand to study the existence and uniqueness of solutions and the success or failure of algorithms to find an optimizer.

3. Next we visit *causality*, and the question of determining how a collection of entities relate causally to one another. We see the algebraic decompositions appearing in graphical models and structural causal models, which contain as special cases the matrix and tensor decompositions encountered up to this point. We study the combinatorial optimization problems of learning causal structures.

4. Finally, we consider how the multilinear and combinatorial problems studied so far can shed light on the nonlinear setting of *deep learning*. We look at expressivity of architectures, with connections to the study of piecewise linear functions, and identifiability of nonlinear algorithms.

This is a graduate course in applied algebra. The course combines mathematical theory, computational and numerical experiments, and exploration of real world data. The focus is on current research developments and connections to open problems.

Prerequisites. Recommended preparation for the course is familiarity with proofs in linear algebra at the level of two semesters of Math 22A/B or Math 25A or Math 121 as well as programming experience at the level of AM 120.

Assessment. Class attendance (10%), four problem sets (40%), a research-oriented final project (50%).

Learning outcomes. Students will have a unified algebraic toolbox to understand existing methods, to design new models, and to prove results on their theoretical underpinnings across the areas of dimensionality reduction, statistics, causality, and deep learning.

Syllabus

1. Dimensionality reduction:
   (a) low rank approximation of matrices; principal component analysis [GCKG22],
   (b) finding signals in multidimensional data [KB09].
2. Statistics:
   (a) the method of moments and connections to tensor decomposition [AGH+14],
   (b) blind sources separation, independent component analyis [Com94, CJ10],
   (c) maximum likelihood estimation; the EM algorithm [Sul18],
   (d) exponential families: examples and unifying theory [Sul18].
3. Causality:
   (a) graphical models, d-separation, and conditional independence [Sul18],
   (b) structural causal models,
   (c) causal structure learning, Markov equivalence [HDMM18, SU22].
4. Deep learning:
   (a) causal representation learning [SLB+21],
   (b) deep latent variable models, identifiability of nonlinear ICA [KKMH20],
   (c) expressivity in neural networks [MPCB14, KMMT22, ZNL18, GLM17].

References

[AGH+14] Animashree Anandkumar, Rong Ge, Daniel Hsu, Sham M Kakade, and Matus Telgarsky. Tensor decompositions for learning latent variable models. *Journal of machine learning research*, 15:2773–2832, 2014.

[CJ10] Pierre Comon and Christian Jutten. *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.

[Com94] Pierre Comon. Independent component analysis, a new concept? *Signal processing*, 36(3):287–314, 1994.

[GCKG22] Benyamin Ghojogh, Mark Crowley, Fakhri Karray, and Ali Ghodsi. *Elements of Dimensionality Reduction and Manifold Learning*. Springer International Publishing, 2022.

[GLM17] Rong Ge, Jason D Lee, and Tengyu Ma. Learning one-hidden-layer neural networks with landscape design. *arXiv preprint arXiv:1711.00501*, 2017.

[HDMM18] Christina Heinze-Deml, Marloes H Maathuis, and Nicolai Meinshausen. Causal structure learning. *Annual Review of Statistics and Its Application*, 5:371–391, 2018.

[KB09] Tamara G Kolda and Brett W Bader. Tensor decompositions and applications. *SIAM review*, 51(3):455–500, 2009.

[KKMH20] Ilyes Khemakhem, Diederik Kingma, Ricardo Monti, and Aapo Hyvarinen. Variational autoencoders and nonlinear ICA: A unifying framework. In *International Conference on Artificial Intelligence and Statistics*, pages 2207–2217. PMLR, 2020.

[KMMT22] Kathlén Kohn, Thomas Merkh, Guido Montúfar, and Matthew Trager. Geometry of linear convolutional networks. *SIAM Journal on Applied Algebra and Geometry*, 6(3):368–406, 2022.

[MPCB14] Guido F Montufar, Razvan Pascanu, Kyunghyun Cho, and Yoshua Bengio. On the number of linear regions of deep neural networks. *Advances in neural information processing systems*, 27, 2014.

[SLB+21] Bernhard Schölkopf, Francesco Locatello, Stefan Bauer, Nan Rosemary Ke, Nal Kalchbrenner, Anirudh Goyal, and Yoshua Bengio. Toward causal representation learning. *Proceedings of the IEEE*, 109(5):612–634, 2021.

[SU22] Chandler Squires and Caroline Uhler. Causal structure learning: a combinatorial perspective. *arXiv preprint arXiv:2206.01152*, 2022.

[Sul18] Seth Sullivant. *Algebraic statistics*, volume 194. American Mathematical Soc., 2018.

[ZNL18] Liwen Zhang, Gregory Naitzat, and Lek-Heng Lim. Tropical geometry of deep neural networks. In *International Conference on Machine Learning*, pages 5824–5832. PMLR, 2018.