Matrix Algebra and Optimization for Statistics and Machine Learning

Yiyuan She

Department of Statistics, Florida State University

- Yiyuan She, OSB 209F
- Tuesday and Thursday 12:30pm–1:45pm, OSB 215
- Office hour: Thursday 2:00pm–3:00pm and by appointment

Course Webpage

- http://campus.fsu.edu
- Syllabus
- Announcements
- Slides

Books

None. Some reference books though (optional)

- Matrix Algebra From A Statistician's Perspective, Harville, 1997
- Matrix Algebra, Gentle, 2007
- Convex Optimization, Boyd and Vandenberghe (`https://stanford.edu/~boyd/cvxbook/`)
- Numerical Optimization, Nocedal and Wright, 1999

Prerequisite

- Regression
- Linear algebra
- Optimization (basic)

Software

- R + necessary packages

Grades

- Project

Check the syllabus online.

Project

- Deadline: Dec 5 by 5pm
- Paper: typed, 10-15 pages
- It can either computational or theoretical
- (Potential arrangement: oral presentation)

Possible Topics

- Eigenvalues and eigenvectors, singular value decomposition, Cholesky decomposition, QR decomposition, partitioned matrices, Schur complement, generalized inverses, Hadamard products, Kronecker products, vectorization, projections, subspace intersection and canonical angles
- Matrix differentiation: Jacobian, the chain rule
- [Matrix inequalities]

Possible Topics (cont'd)

- Convex functions and convex optimization
- Lagrange duality
- Gradient descent and Newton's method
- Line search methods
- Inexact Newton, quasi-Newton, sketched Newton
- Proximal methods and linearization
- Bregman divergence and mirror descent
- MM algorithms
- Block coordinate descent and randomization
- Augmented Lagrangian,operator splitting
- J-L lemma, random sampling and random projection
- Stochastic approximation and optimization

Possible Topics (cont'd)

- Co-integration, support vector machines, SDP relaxation, canonical correlation analysis, low-rank matrix approximation, back-propagation in neural networks, multi-task learning, generalized linear models, graph learning, bandit algorithms, network clustering and ranking, approximate message passing, manifold learning, structural parsimony pursuit, boosted trees
- [Optional: gradient boosting, manifold optimization]

Examples

- Gaussian graph learning
- Multivariate meta-analysis
- Matrix completion
- Sufficient dimension reduction
- Neural network
- K-means for clustering
- Multi-task learning