BIOS:171:290/STAT:7190
Topics in High-Dimensional Data Analysis

Time and Location  Wednesdays 1:00-3:50pm, S030 CPHB

Instructor  Jian Huang
Biostatistics
Statistics and Actuarial Science
University of Iowa
Email: jian-huang@uiowa.edu

Departmental executive officer
Biostatistics: Kathryn Chaloner, N332A CPHB, Phone: 384-1581

Prerequisites  Familiarity with multivariate calculus and matrix algebra, mathematical statistics
(at the level of STAT:4100 and 4101 (22S:153 and 154) or higher), working knowledge of R and a
basic course in regression and ANOVA.

Grading  U/S based on class participation and projects.

Course description
This course covers methods, algorithms and theoretical developments in the area of high-
dimensional models and their applications in genetic and genomic data analysis.

Topics to be covered (tentative)

1. Introduction
   High-dimensional data problems arising from genetics and genomics will be used as
   motivations and examples throughout the course

2. Large scale hypothesis testing
   Large scale hypothesis testing problems
   FWER control

3. False discovery rate control
   False discovery proportions
   Benjamini-Hochberg’s FDR control procedure
   Direct FDR control

4. Variable selection: convex penalized selection
   Lasso in linear regression
   Weighted $\ell_1$ penalties
   Computational algorithms and R packages
   Theoretical properties

5. Clustering and classification
   Lasso in logistic regression
ROC analysis
Hierarchical and K-means clustering
Computational algorithms and R packages
Theoretical properties

6. Concave selection

Minimax concave penalties (MCP)
Smoothly clipped absolute deviation (SCAD)
Computational algorithms and R packages
Theoretical properties

7. Group selection

Group Lasso
Concave group selection
Computational algorithms and R packages
Theoretical properties

8. PCA and dimension reduction

Singular value decomposition and spectral decomposition
Principal components analysis (PCA)
Regularized PCA
Latent factor models
PCA regression
Regularized PCA regression

9. Large covariance matrix estimation (time permitting)

Graphical Lasso
Covariance matrix estimation based on thresholding
Other regularization methods

References

Books and research papers related to the topics listed above will be suggested in class. Also, lecture
notes will be made available on ICON.

Additional UI and CLAS Policy and Procedures

Administrative Home The College of Liberal Arts and Sciences is the administrative home of
this course and governs matters such as the add/drop deadlines, the second-grade-only option, and
other related issues. Different colleges may have different policies. Questions may be addressed to
120 Schaeffer Hall or see the CLAS Student Academic Handbook [www.clas.uiowa.edu/students/
academic_handbook/index.shtml].

Electronic Communication University policy specifies that students are responsible for all official
correspondences sent to their standard University of Iowa e-mail address (@uiowa.edu). Students
should check this account frequently. (Operations Manual, III.15.2. Scroll down to k.11.)
Accommodations for Disabilities A student seeking academic accommodations should first register with Student Disability Services and then meet privately with the course instructor to make particular arrangements. See www.uiowa.edu/~sds/ for more information.

Academic Fraud All CLAS students have, in essence, agreed to the College’s Code of Academic Honesty: “I pledge to do my own academic work and to excel to the best of my abilities, upholding the IOWA Challenge. I promise not to lie about my academic work, to cheat, or to steal the words or ideas of others; nor will I help fellow students to violate the Code of Academic Honesty.” Any student committing academic misconduct is reported to the College and placed on disciplinary probation or may be suspended or expelled (CLAS Academic Policies Handbook).

CLAS Final Examination Policies Final exams may be offered only during finals week. No exams of any kind are allowed during the last week of classes. Students should not ask their instructor to reschedule a final exam since the College does not permit rescheduling of a final exam once the semester has begun. Questions should be addressed to the Associate Dean for Undergraduate Programs and Curriculum.

Making a Suggestion or a Complaint Students with a suggestion or complaint should first visit the instructor, then the course supervisor, and then the departmental DEO. Complaints must be made within six months of the incident. See the CLAS Student Academic Handbook.

Understanding Sexual Harassment Sexual harassment subverts the mission of the University and threatens the well-being of students, faculty, and staff. All members of the UI community have a responsibility to uphold this mission and to contribute to a safe environment that enhances learning. Incidents of sexual harassment should be reported immediately. See the UI Comprehensive Guide on Sexual Harassment at www.uiowa.edu/~eod/policies/sexual-harassment-guide/index.html for assistance, definitions, and the full University policy.

Reacting Safely to Severe Weather In severe weather, class members should seek appropriate shelter immediately, leaving the classroom if necessary. The class will continue if possible when the event is over. For more information on Hawk Alert and the siren warning system, visit the Public Safety website http://www.uiowa.edu/~pubsfty/intlinks.htm.

Student Classroom Behavior The ability to learn is lessened when students engage in inappropriate classroom behavior, distracting others; such behaviors are a violation of the Code of Student Life. When disruptive activity occurs, a University instructor has the authority to determine classroom seating patterns and to request that a student exit the classroom, laboratory, or other area used for instruction immediately for the remainder of the period. One-day suspensions are reported to appropriate departmental, collegiate, and Student Services personnel (Office of the Vice President for Student Services and Dean of Students).