Errors using inadequate data are much less than those using no data at all.
-Charles Babbage


Additional Potential Resources:
Press et al., *Numerical Recipes*, various editions
Moler, *Numerical Computing with MatLab*, Mathworks, 2004
Menke, Geophysical data Analysis: Discrete Inverse Theory, various editions

Motivation: Like all science, the Earth Sciences require both observations and hypotheses to develop understanding of natural systems. It is the comparison of the two where science is advanced. By analyzing data, a scientist is simply looking for information about a process. This means that the data must contain predictive information in some sense about the process in question. To extract the information, various techniques may be applied. Fundamental methods are forward modeling of data, inverse modeling of data, and statistical analysis. This class will touch on each of these.

Problem Sets: Problem sets will be assigned throughout the semester. Show all work. Working with others is encouraged, but the final product must reflect your own work and understanding.

Exams: No exams.

Project: This is to be a project-oriented class. You are to choose a dataset (preferably pertinent to your research area) and apply techniques of analysis and visualization the data. The final project will consist of both a paper and a presentation to the class. The paper (~10-12 pages) is due Mon of finals week. Presentations will take place during the last week of class. Topic proposals are due 6th week of class (3–4 paragraphs describing topic, dataset, motivation, and goal).

Grading: The term project is worth 50% of your grade (evenly split between presentation and paper). Problem sets will comprise 40% and class participation 10%. Assignments turned in after due date will be penalized 10% per day late.

Academic integrity: Depending on the magnitude of the violation will result in a penalty ranging from a zero on the particular assignment to referral to the university for formal disciplinary action.
Violations of academic integrity include (but are not limited to) cheating, plagiarism, and passing off unoriginal work as own.

Topics

Basics of Matlab & scientific computing
  Basic programming (scripts, functions, etc)
  Visualization

Inverse modeling, data fitting
  Line & curve fitting
  Inverse theory
  2D surface fitting (isotropic, anisotropic)

Statistics for the scientist
  Univariate, bivariate, multivariate
  Distributions
  Error analysis
  Hypothesis testing
  Bayesian statistics
  R vs $R^2$
  Principal component analysis

Model selection
  F-test
  Akaike Information Criterion

Emergency Procedures: SIUC is committed to providing a safe and healthy environment for study and work. Because some health and safety concerns are beyond our control, we ask that you become familiar with the SIUC Emergency Response Plan and Building Emergency Response Plan (BERT) program. Emergency response information is available at HYPERLINK "http://www.bert.siu.edu" www.bert.siu.edu and on posters in buildings on campus.

Instructors will provide guidance and direction to students in the classroom in the event of an emergency affecting your location. The building Emergency Response Team will provide assistance to your instructor in evacuating the building or sheltering within the facility.
Geol 513/413: Data Analysis in Earth Sciences, Fall, 2015

Schedule

Week 1
Intro, motivation, types of data & data sets
sample data set

Week 2
Students’ data sets, more sample set
Introduce group project topic
Fitting models to data: Forward models

Week 3
RMAR
multiple regression, polyfit
Matlab basics

Week 4
Multiple parameters – grid search
Grid search, cont (weighted least squares)
Grid search – multigrids; regularization (under vs even vs overdetermined)

Week 5
Least squares inverse modeling – system of equations (even determined, N = M)
Matrices, linear algebra

Week 6
Inverse modeling – system of equations (over-determined, N > M)
Weighted least squares

Week 7
Univariate distributions, statistical descriptions
Example: Bad stats
Geol 413/535: Data Analysis in Earth Sciences, Spring, 2012

Week 8
Hypothesis testing – t-test
Hypothesis testing – $\chi^2$, K-S, type I vs type II errors

Week 9
Principle component analysis

Week 10
Student presentations of their datasets; Group project discussion

Week 11
Bivariate stats – covariance, correlation
Bivariate analysis – regression, bootstrapping, jackknifing

Week 12
F-test, AICc
Spatial methods: Interpolation, extrapolation, Kriging; clustering

Week 13
Inverse modeling – non-linear
Inverse modeling - minimum length solution (underdetermined, $N < M$)

Week 14
Inverse modeling – system of equations (mixed determined)
Inverse modeling – system of equations

Week 15
Discussion of Group project

Week 16 Presentations
2014: Students liked the Taiwan project. It was really on the fly. If were to lay out good tackling of problem, could likely solve it in class (& publish). What it needs is a more full integration with various methods.

Add principal component analysis. Ishman’s students use it regularly.