Research Class 3

December 23, 2017

The materials covered in this class are:

1. Detailed review: regressions.
2. Detailed review: Markov chain.

1 Regression Review

1.1 Linear Regression

From the last class: in what situation should we use linear regression? How do we fit a linear regression model? How do we interpret a linear regression model?

- Questions?
- Useful tool: matplotlib.pyplot (`https://matplotlib.org/xkcd/users/pyplot_tutorial.html`)
- Exercise: 11-3

Variance in linear regression model:

- What is random? The definition of a random variable: a variable whose possible values are numerical outcomes of a random phenomenon. A good start here: `https://www.khanacademy.org/math/statistics-probability/random-variables-stats-library/random-variables-discrete/v/random-variables`
- Definition of a confidence interval.

Quality of a linear regression model:

- R^2
- Significance hypothesis test.

1.2 Logistic Regression

From the last class: in what situation should we use logistic regression? How do we fit a logistic regression model? How do we interpret a logistic regression model?

- Questions?
- Derivation by hand? A general form of derivation follows the rule of "Maximum likelihood estimation".

For a general linear model, this course provides a comprehensive view: `https://onlinecourses.science.psu.edu/stat504/node/216`. We will use its focus of a model as follows:

- Objective
- Model structure (e.g. variables, formula, equation)

- Model assumptions
- Parameter estimates and interpretation
- Model fit (e.g. goodness-of-fit tests and statistics)
- Model selection

For logistic regression, the estimation process is complicated and there are no closed-form solutions to the maximization of the likelihood function, so we do not cover it here. However, it is very useful to get familiar with the fitting process using scikit-learn. Try 11-83 on Page 445 of the stats book.

2 Markov Chain Review

Concepts:

- States
- Transition probability, transition matrix
- Initial distribution
- Steady state distribution

Calculation: matrix manipulation in NumPy

- How can we input a matrix in NumPy?
- Matrix multiplication, inverse, transpose