LINEAR REGRESSION MODELS W4315

INTRO SURVEY QUESTIONS

September 5, 2011

Instructor: Frank Wood (Pupin 329, Tu-Th, 9:10am-10:25am)

1. (0 points)

(a) What is your field of study?

(b) What is your intended profession?

(c) Is this course required for you? If not, why are you taking it?

(d) Name some problems you would formulate and solve using regression?

(e) What, in addition to regression basics, would you like to learn in this course?

2. (0 points) Find the value of x that maximizes the function $x^2 + ln(x)$, i.e. find

$$\underset{x}{\operatorname{argmax}} \quad -x^2 + ln(x)$$

Simplify your answer.

3. (0 points) Pretend that you can't derive the answer to the previous question because the objective function is too complicated. Write pseudo-code for a program that finds the value of x that maximizes the function using nothing other than standard programming constructs (for, while, if, +,-,/,*, etc.) and evaluations of the function $y(x) = x^2 + ln(x)$.

4. (0 points)

(a) Given an algebraic solution to the following matrix equation

$$\mathbf{Ax} - \mathbf{b} = \mathbf{0}$$

(b) What conditions does your answer put on the matrix $\mathbf{A}$?

5. (0 points)

(a) Maximize the following quadratic form w.r.t. $\mathbf{x}$ ($\boldsymbol{\mu}$ is also a column vector, $\mathbf{A}$ and $\mathbf{x}$ are as in the previous problem)

$$(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{A}(\mathbf{x}-\boldsymbol{\mu})$$

6. (0 points) If $X|\lambda \sim \text{Poisson}(\lambda)$ and $\lambda|\alpha,\beta \sim \text{Gamma}(\alpha,\beta)$ how is $X|\alpha,\beta$ distributed? Hint, remember:

$$\begin{aligned} P(a|c) &= \int P(a|b)P(b|c)db \\ P(X|\lambda) &= \frac{1}{X!}\lambda^X e^{-\lambda} \\ P(\lambda|\alpha,\beta) &= \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda} \end{aligned}$$

Don't simplify the $\Gamma()$'s in the final result.

7. (0 points)

Let $\theta_i \sim \text{Exp}(\beta)$, $1 \leq i \leq n$ be samples from an exponential distribution. Remember that for an exponential distribution $P(\theta) = \frac{1}{\beta}e^{-\frac{\theta}{\beta}}$ for $\theta > 0$. Also remember that for the exponential distribution $E(\Theta) = \beta$ and $V(\theta) = \beta^2$ where $E()$ stands for expectation and $V()$ for variance.

(a) Derive the maximum likelihood estimator $\hat{\beta}$ for β given observations $\{\theta_i\}_{i=1}^n$.

(b) Is this estimator biased or unbiased? Show work. Reminder: the definition of bias is $B(\hat{\beta}) = \beta - E(\hat{\beta})$.

(c) Derive the sample variance of the estimator, $V(\hat{\beta})$. Remember $V(aX) = a^2V(X)$.

(d) Given a set of samples $\{\theta_i\}_{i=1}^n$ as above, illustrate the large sample (for instance, $N > 30$) symmetric confidence interval for $\hat{\beta}$. Remember that by the central limit theorem $\frac{\hat{\beta}-\beta}{\sqrt{V(\hat{\beta})}} \sim N(0,1)$ where $N(0,1)$ stands for a standard normal distribution with mean 0 and variance 1. Also, for a sufficiently large n remember that $\frac{\hat{\beta}-\beta}{\sqrt{V(\hat{\beta})}} \approx \frac{\hat{\beta}-\beta}{S/\sqrt{n}}$ where S is the sample population standard deviation. The sample variance is given by $S^2 = \frac{1}{n-1}\sum_i(\theta_i - \hat{\beta})^2$ but this expansion is unnecessary for the purposes of this question and is provided only for familiarity. What you should derive is c for a confidence interval

$$\hat{\beta} - c \leq \beta \leq \hat{\beta} + c$$

in terms of S, n, and k where k value of the standard normal inverse cumulative distribution for some $1-\alpha/2$ level of confidence. Also show a representative plot of the sampling distribution of the estimator.