MATH 408 Quiz 6 Solution Sketch

Qilin Ye

December 2, 2021

Questions can be found here.

Q1 (Least-Squares)

Proof. *We consider an alternate approach where we view $f(m,b)$ as a norm function. Let $(x_1,y_1),\dots,(x_n,y_n)$ be given accordingly. We define a $n\times 2$ matrix A, a vector $y\in\mathbb{R}^n$, and a function $u(\tilde{m},\tilde{b}):\mathbb{R}^2\to\mathbb{R}^2$ by*

$$A := \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \qquad u(\tilde{m},\tilde{b}) := \begin{bmatrix} \tilde{m} \\ \tilde{b} \end{bmatrix} \qquad \text{and} \qquad y := \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Then,

$$f(m,b) = \sum_{i=1}^n (y_i - (m_i x + b))^2 = \|y - Au(m,b)\|^2.$$

Therefore, geometrically, f is minimized when $y - Au(\tilde{m},\tilde{b})$ is the shortest, i.e., when $y - Au(\tilde{m},\tilde{b})$ is orthogonal to the column space of A, which contains $Au(\tilde{m},\tilde{b})$. Thus we want to find m,b satisfying

$$A^T(y - Au(m,b)) = 0 \implies A^Ty = A^TAu(m,b) \implies \begin{bmatrix} m \\ b \end{bmatrix} = (A^TA)^{-1}A^Ty.$$

Now let the computation begin:

$$A^TA = \begin{bmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} = \begin{bmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n \end{bmatrix} \tag{1}$$

$$(A^TA)^{-1} = \frac{1}{n\sum x_i^2 - (\sum x_i)^2} \begin{bmatrix} n & -\sum x_i \\ -\sum x_i & \sum x_i^2 \end{bmatrix} \tag{2}$$

$$(A^TA)^{-1}A^T = (2)\cdot \begin{bmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{bmatrix} = \frac{1}{n\sum x_i^2 - (\sum x_i)^2} \begin{bmatrix} nx_1 - \sum x_i & \cdots & nx_n - \sum x_i \\ \sum x_i^2 - x_1\sum x_i & \cdots & \sum x_i^2 - x_n \sum x_i \end{bmatrix} \tag{3}$$

and finally

$$(A^TA)^{-1}A^Ty = (3)\cdot \begin{bmatrix} y_1 \\ \cdots \\ y_n \end{bmatrix} = \frac{1}{n\sum x_i^2 - (\sum x_i)^2} \begin{bmatrix} n\sum x_iy_i - \sum x_i \sum y_i \\ \sum x_i^2 \sum y_i - \sum x_i \sum x_iy_i \end{bmatrix}. \tag{4}$$

It remains to verify that these correspond to the m and b given in the problem:

$$m = \frac{\sum x_i \sum y_i - n \sum x_i y_i}{(\sum x_i)^2 - n \sum x_i^2} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

and

$$\begin{aligned} b = \frac{1}{n}\left(\sum y_i - m \sum x_i\right) &= \frac{1}{n}\left(\sum y_i - \frac{n \sum x_i \sum x_i y_i - (\sum x_i)^2 \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}\right) \\ &= \frac{1}{n} \cdot \frac{n \sum x_i^2 \sum y_i - (\sum x_i)^2 \sum y_i - n \sum x_i \sum x_i y_i + (\sum x_i)^2 \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \\ &= \frac{1}{n} \frac{n(\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i)}{n \sum x_i^2 - (\sum x_i)^2} = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2}. \end{aligned}$$

□

Q3 (Turkeys? No Turkeys?)

Solution. We follow the notation used in lectures. For this sample, $\overline{Y}_1 = 6, \overline{Y}_2 = 5$, and $\overline{Y}_3 = 7$. Correspondingly, $\overline{Y} = 6$, $S_1^2 = S_3^2 = 1$, and $S_2^2 = 0$. Therefore, $S^2 = (2+0+2)/6 = 2/3$. Under our null hypothesis, all β's are identical. Using the proposition on F-test, we obtain the value

$$F = S^{-2} \sum_{i=1}^{3} 3(\overline{Y}_i - \overline{Y})^2 = 9.$$

Here we have a F-distribution with $3 - 1 = 2$ and $9 - 3 = 6$ degrees of freedom. Upon checking WolframAlpha we conclude that the corresponding p-value ≈ 0.0312, so we are confident in rejecting the null, i.e., claiming that the qualities of different groups of turkey differ.

Q4 (Concavity of Logistic Log-Likelihood)

Proof. From quiz 4 problem 5, it suffices to show that $t(\cdot, b)$ and $t(a, \cdot)$ are both strictly concave. We check the former. The log-likelihood is

$$\begin{aligned} t(a,b) &= \log\left([h(ax+b)]^y [1 - h(ax+b)]^{1-y}\right) \\ &= y \log h(ax+b) + (1-y)\log(1 - h(ax+b)) \\ &= y \log\left(\frac{1}{1+e^{-(ax+b)}}\right) + (1-y)\log\left(1 - \frac{1}{1+e^{-(ax+b)}}\right) \\ &= y \log\left(\frac{1}{1+e^{-(ax+b)}}\right) + (1-y)\log\left(\frac{e^{-(ax+b)}}{1+e^{-(ax+b)}}\right) \\ &= -y\log(1+e^{-(ax+b)}) + (1-y)\log(e^{-(ax+b)}) - (1-y)\log(1+e^{-(ax+b)}) \\ &= -\log(1+e^{-(ax+b)}) + (y-1)(ax+b). \end{aligned}$$

notice that $(y-1)(ax+b)$ is a decreasing affine function of b and $-\log(1+e^{-(ax+b)})$ is strictly concave. Therefore their sum, $t(\cdot, b)$, is strictly concave. A similar argument (omitted) shows that $t(a, \cdot)$ is strictly concave as well. Thus $t(a,b)$ is strictly concave. By quiz 4 problem 2, there exists at most one global maximum. □