UC Berkeley
Department of Electrical Engineering and Computer Sciences

EE126: Probability and Random Process

Solution 8
Spring 2017

Issued: Thursday, March 23, 2017

Self-graded Scores Due: 5pm, Monday, April 10, 2017
Submit your self-graded scores via the google form:
https://goo.gl/forms/VbulsZU1R9LCTwr53. Make sure that you use your **SORTABLE NAME** on bCourses.

Problem 1. See midterm solutions.

Problem 2. (a) We have

$$MLE[X|Y=y] = \arg\max_x f_{Y|X}[y|x].$$

Now,

$$f_{Y|X}[y|x] = xe^{-xy}, \forall x, y > 0.$$

Hence, $MLE[X|Y=y]$ is the value of x such that

$$0 = \frac{\partial}{\partial x}[xe^{-xy}] = e^{-xy} - xye^{-xy},$$

so that

$$MLE[X|Y] = \frac{1}{Y}.$$

(b) We have

$$MAP[X|Y=y] = \arg\max_x f_{X|Y}[x|y] = \arg\max_x f_{Y|X}[y|x]f_X(x).$$

Now,

$$f_{Y|X}[y|x]f_X(x) = xe^{-xy}e^{-x} = xe^{-x(1+y)}.$$

Hence, $MAP[X|Y=y]$ is the value of x such that

$$0 = \frac{\partial}{\partial x}[xe^{-x(1+y)}] = e^{-x(1+y)} - x(1+y)e^{-x(1+y)}.$$

Hence,

$$MAP[X|Y] = \frac{1}{1+Y}.$$

Problem 3. Let G be a random variable whose values are realizations of the graph and let A be a random variable representing the labeling of the two communities. We are interested in $\text{MAP}[A|G] = \text{argmax}_A P(G|A)P(A)$. Note that since each assignment of labels is equally likely, the MAP rule is equivalent to the MLE, and we are simply interested in $\text{argmax}_A P(G|A)$. Let $|E|$ be the number of edges across the partition in assignment A and let $|T|$ be the total number of edges in assignment A. We see that:

$$\begin{aligned} P(G|A) &= q^{|E|}(1-q)^{(\frac{n}{2})^2-|E|}p^{|T|-|E|}(1-p)^{2\binom{n/2}{2}-(|T|-|E|)} \\ &= \left(\frac{q}{1-q}\cdot\frac{1-p}{p}\right)^{|E|}\cdot\left(\frac{p}{1-p}\right)^{|T|}\cdot(1-p)^{2\binom{n/2}{2}}(1-q)^{n^2/4} \end{aligned}$$

Now, note that the last three terms are constant for any assignment of labels and do not affect the likelihood. Also note that since $p > q$, $\left(\frac{q}{1-q}\cdot\frac{1-p}{p}\right) < 1$, so we can see that increasing $|E|$ corresponds to decreasing the likelihood, and the MAP rule is to select the partition with the smallest number of edges across it, which is exactly the min-bisection of the graph.

Problem 4. (a) We observe x and y. Thus,

$$\begin{aligned} \epsilon_{ML} &= \arg\max_\epsilon \Pr(x, y|\epsilon) \\ &= \arg\max_\epsilon \Pr(x|\epsilon)\Pr(y|x,\epsilon) \\ &= \arg\max_\epsilon \Pr(y|x,\epsilon) \\ &= \arg\max_\epsilon \epsilon^{1\{y\neq x\}}(1-\epsilon)^{1\{y=x\}} \end{aligned}$$

Note that ϵ and X are independent so $\Pr(x|\epsilon) = \Pr(x)$ and does not change the maximizer over ϵ. Now if $x \neq y$, the expression is clearly maximized on the largest possible value of ϵ which is $\epsilon = 0.5$. If $x = y$, the expression is maximized for smallest value of ϵ which is 0.

(b) We observe $y_1, \ldots, y_n$. Thus, $\epsilon_{ML} = \arg\max_\epsilon \Pr(y_1, \ldots, y_n|\epsilon)$. Since every use of the channel is independent we have,

$$\begin{aligned} \Pr(y_1, \ldots, y_n|\epsilon) &= \prod_{i=1}^{n} \Pr(y_i|\epsilon) \\ &= \prod_i \big((0.6(1-\epsilon)+0.4\epsilon)1\{y_i = 1\} + (0.4(1-\epsilon)+0.6\epsilon)1\{y_i = 0\}\big) \\ &= \prod_i (0.6-0.2\epsilon)^{y_i}(0.4+0.2\epsilon)^{(1-y_i)} \\ &= (0.6-0.2\epsilon)^{\sum_i y_i}(0.4+0.2\epsilon)^{(n-\sum_i y_i)} \end{aligned}$$

Let $t = \sum_i y_i$. As we can see, what matters for estimating ϵ is t. To find the maximizer of the expression, we first take the log and then set the derivative to 0. Thus,

$$\frac{-0.2t}{0.6-0.2\epsilon} + \frac{0.2(n-t)}{0.4+0.2\epsilon} = 0.$$

Solving the equation, we get $\epsilon_{ML} = 3 - \frac{5t}{n}$. Of course, since we know that $0 \leq \epsilon < 0.5$, if ϵ_{ML} is not in the interval $[0, 0.5]$ we should pick the closest point to it which will be either 0 or 0.5.

(c) This time we want to maximize $\Pr(y_1, \ldots, y_n|\epsilon)f(\epsilon)$. Similar to the calculations of previous part, we want to maximize,

$$(4 - 8\epsilon)(0.6 - 0.2\epsilon)^t(0.4 + 0.2\epsilon)^{(n-t)}.$$

Taking the log and setting the derivative equal to 0 we have

$$\frac{-8}{4 - 8\epsilon} + \frac{-0.2t}{0.6 - 0.2\epsilon} + \frac{0.2(n-t)}{0.4 + 0.2\epsilon} = 0.$$

Then, we get the following quadratic equation.

$$-8(0.6-0.2\epsilon)(0.4+0.2\epsilon) - 0.2t(4-8\epsilon)(0.4+0.2\epsilon) + 0.2(n-t)(4-8\epsilon)(0.6-0.2\epsilon) = 0.$$

One can solve the long quadratic equation analytically, and find ϵ_{MAP}. We skip the painful algebra here. (You also get full credit, if you find the quadratic equation.)

Problem 5. (a) Note that since both outputs are equiprobable, the MAP rule is equivalent to the ML rule. Note that the likelihood ratio is:

$$L(y) = \frac{f(y|x = V_1)}{f(y|x = V_2)}$$

$$\hat{x} = \begin{cases} V_1 & \text{if } L(y) > 1 \\ V_2 & \text{if } L(y) \leq 1 \end{cases}:$$

Thus, we have:

$$f(y|x = V_1) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-V_1-V_3)^2/2\sigma^2}$$

$$f(y|x = V_2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-V_2-V_3)^2/2\sigma^2}$$

so:

$$L(y) = e^{(V_1-V_2)(y-(V_1+V_2)/2-V_3)/\sigma^2}$$

WLOG, we may assume $V_1 > V_2$, so the ML rule is:

$$\hat{x} = \begin{cases} V_1 & \text{if } y > \frac{(V_1+V_2)}{2} + V_3 \\ V_2 & \text{if } y \leq \frac{(V_1+V_2)}{2} + V_3 \end{cases}:$$

(b) We have:

$$\begin{aligned} P(\text{error}) &= P(\hat{x} = V_2|x = V_1)P(x = V_1) + P(\hat{x} = V_1|x = V_2)P(x = V_2) \\ &= P(y < \frac{V_1 + V_2}{2} + V_3|x = V_1)P(x = V_1) + P(y > \frac{V_1 + V_2}{2} + A_3|x = V_2)P(x = V_2) \\ &= \frac{1}{2}\left(1 - Q\left(\frac{V_2 - V_1}{2\sigma}\right)\right) + \frac{1}{2}Q\left(\frac{V_1 - V_2}{2\sigma}\right) \\ &= Q\left(\frac{V_1 - V_2}{2\sigma}\right) \end{aligned}$$

(c) We would like to minimize $Q\left(\frac{V_1-V_2}{2\sigma}\right)$ subject to the constraint: $\frac{V_1^2+V_2^2}{2} \leq E$. Note that this is equivalent to maximizing $V_1 - V_2$ which is equivalent to maximizing $(V_1 - V_2)^2$, subject to the same constraint. Now, we have:

$$(V_1 - V_2)^2 \leq (|V_1| + |V_2|)^2 \leq 4E$$

where we have equality iff $V_1 = -V_2$, so the optimal choice is $V_1 = \sqrt{E}$, $V_2 = -\sqrt{E}$.