MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.s077 Spring 2018

Problem Set 3 due Tuesday 3/6, in class

Problem 1.

$$\begin{aligned}\mathbb{P}(\Theta = 1|X = x) &= \frac{f_{X|\Theta}(x|1)\mathbb{P}(\Theta = 1)}{f_X(x)} \\ &= \frac{e^{-x}\cdot\frac{1}{2}}{e^{-x}\cdot\frac{1}{2} + 2e^{-2x}\cdot\frac{1}{2}} \\ &= \frac{1}{1+2e^{-x}} \qquad x \geq 0\end{aligned}$$

Since Θ can only take two values, to get the MAP estimate we note that $\mathbb{P}(\Theta = 1|X = x) \geq \frac{1}{2}$ when $2e^{-x} \leq 1 \Rightarrow x \geq \log 2$. Therefore

$$\hat{\Theta}_{MAP} = \begin{cases} 1, & \text{if X} \geq \log 2, \\ 2, & \text{if X} < \log 2. \end{cases}$$

Problem 2. Bayesian estimation with a prior that favors a zero value.

(a) We retain the terms which depend on θ.

$$\begin{aligned}f_{\Theta|Y}(\theta|y) &= \frac{f_{Y|\Theta}(y|\theta)\cdot f_\Theta(\theta)}{f_Y(y)} \\ &\propto \exp(-(y-\theta)^2/2 - \theta^2/2)\end{aligned}$$

We maximize the posterior to get the MAP estimate. Taking derivatives yields

$$(y-\theta) - \theta = 0 \Rightarrow \theta = \frac{y}{2} \Rightarrow \hat{\Theta}_{MAP} = \frac{Y}{2}$$

For an alternative derivation, note that the posterior is a normal distribution, and its peak is the same as its mean; i.e., the MAP and LMS estimatoes coincide. Furthermore, from the symmetry of $Y = \Theta + W$, we have $\mathbb{E}[\Theta|Y] = \mathbb{E}[W|Y]$. Therefore,

$$Y = \mathbb{E}[Y|Y] = \mathbb{E}[\Theta|Y] + \mathbb{E}[W|Y] = 2\mathbb{E}[\Theta|Y].$$

See also p. 448 of [BT]: Problem 13 and its solution (which is also given in the text) contains a more general version of this idea, extending it to the case of unequal variances.

(b)

$$\begin{aligned} f_{\Theta|Y}(\theta|y) &= \frac{f_{Y|\Theta}(y|\theta) \cdot f_\Theta(\theta)}{f_Y(y)} \\ &\propto \exp(-(y-\theta)^2/2 - |\theta|) \end{aligned}$$

We first consider the possibility that the optimum satisfies $\theta \geq 0$. Taking derivatives yields

$$(y-\theta) - 1 = 0 \Rightarrow \theta = y - 1$$

However for $\theta \geq 0$ to hold, this condition can be satisfied only if $y \geq 1$. By a symmetrical argument, a negative optimal θ can be found by setting the derivative to zero only if $y \leq -1$.

If $-1 < y < 1$, we notice that the function $(y-\theta)^2/2+|\theta|$ has a positive left derivative and a negative right derivative at $\theta = 0$, and so $\theta = 0$ is optimal.

Putting this all together yields

$$\hat{\Theta}_{MAP} = \begin{cases} Y+1, & \text{if } Y < -1, \\ 0, & \text{if } -1 \leq Y \leq 1, \\ Y-1, & \text{if } Y > 1. \end{cases}$$

Problem 3.

We have from repeated substitution

$$\begin{aligned} V_1 &= A_1 \\ V_2 &= \frac{2}{3}A_1 + A_2 \\ V_3 &= \frac{4}{9}A_1 + \frac{2}{3}A_2 + A_3 \end{aligned}$$

$$\begin{aligned} X_1 &= V_1 \\ &= A_1 \\ \Rightarrow Y_1 &= A_1 + W_1 \end{aligned}$$

$$\begin{aligned} X_2 &= X_1 + V_2 \\ &= \frac{5}{3}A_1 + A_2 \\ \Rightarrow Y_2 &= \frac{5}{3}A_1 + A_2 + W_2 \end{aligned}$$

$$
\begin{aligned}
X_3 &= X_2 + V_3 \\
&= \frac{19}{9}A_1 + \frac{5}{3}A_2 + A_3 \\
\Rightarrow Y_3 &= \frac{19}{9}A_1 + \frac{5}{3}A_2 + A_3 + W_3
\end{aligned}
$$

Letting $y = (y_1, y_2, y_3)$ and $\theta = (a_1, a_2, a_3)$ we get

$$
\begin{aligned}
&f_{\Theta|Y}(\theta|y) \propto f_{Y|\Theta}(y|\theta) f_\Theta(\theta) \\
&\qquad \propto \exp\left(-\frac{(y_1 - a_1)^2}{2} - \frac{(y_2 - \frac{5}{3}a_1 - a_2)^2}{2} - \frac{(y_3 - \frac{19}{9}a_1 - \frac{5}{3}a_2 - a_3)^2}{2}\right) \\
&\qquad\quad \times \exp\left(-\frac{a_1^2}{2} - \frac{a_2^2}{2} - \frac{a_3^2}{2}\right)
\end{aligned}
$$

Problem 4.

$$
\begin{aligned}
\mathbb{E}[(Z - \hat{Z})^2] &= \text{var}[Z - \hat{Z}] + \mathbb{E}\big[(Z - \hat{Z})\big]^2 \\
&= \text{var}[Z] + \text{var}[\hat{Z}] - 2\text{cov}[Z, \hat{Z}] + a_2^2 \\
&= 1 + \text{var}[a_2 Y_2 + a_3 Y_3] - 2a_2\text{cov}(Z, Y_2) - 2a_3\text{cov}(Z, Y_3) + a_2^2 \\
&= 1 + 10a_2^2 + 16a_3^2 + 2a_2a_3\text{cov}(Y_2, Y_3) - 2a_2\text{cov}(Z, Y_2) - 2a_3\text{cov}(Z, Y_3) \\
&= 1 + 10a_2^2 + 16a_3^2 + 2a_2a_3\text{cov}(Z^2, Z^3) - 2a_2\text{cov}(Z, Z^2) - 2a_3\text{cov}(Z, Z^3) \\
&= 1 + 10a_2^2 + 16a_3^2 - 6a_3
\end{aligned}
$$

So we should set $a_2 = 0$ and $a_3 = \frac{3}{16}$.

Problem 5.

(a) Let $g(Z) = \hat{Y} = \mathbb{E}[Y|Z] =$

$$
\begin{aligned}
\mathbb{E}\big[\tilde{Y}|Z = z\big] &= \mathbb{E}\Big[\Big(\mathbb{E}[Y|Z] - Y\Big)|Z = z\Big] \\
&= \mathbb{E}\big[g(Z)|Z = z\big] - \mathbb{E}[Y|Z = z] \\
&= g(z) - g(z) \\
&= 0
\end{aligned}
$$

(b)

$$
\begin{aligned}
\mathbb{E}[\tilde{Y}] &= \mathbb{E}\Big[\mathbb{E}[\tilde{Y}|Z]\Big] \\
&= \mathbb{E}[0] \\
&= 0
\end{aligned}
$$

(c) Since $\hat{Y}$ is a function of Z,

$$\mathbb{E}[\tilde{Y}\hat{Y}|Z = z] = \hat{Y}\mathbb{E}[\tilde{Y}|Z = z] = 0,$$

where the last equality follows from part (a).

(d) Same logic as in part (b).

(e) Substituting $Y = \hat{Y} - \tilde{Y}$ yields

$$\begin{aligned}\mathbb{E}[Y^2] &= \mathbb{E}[\hat{Y}^2] - 2\mathbb{E}[\hat{Y}\tilde{Y}] + \mathbb{E}[\tilde{Y}]^2 \\ &= \mathbb{E}[\hat{Y}^2] + \mathbb{E}[\tilde{Y}]^2\end{aligned}$$