ECE531 Screencast 5.3: General Bayesian Estimation

D. Richard Brown III

Worcester Polytechnic Institute
Bayesian Estimation Under Different Cost Assignments

Recall, under the squared error cost assignment $C_{\theta}(\hat{\theta}) = \epsilon^2$, we have the MMSE estimator and

$$\hat{\theta}_{\text{mmse}}(y) = \mathbb{E}[\Theta | Y = y].$$

What about other cost assignments?

- **Absolute error**: $C_{\theta}(\hat{\theta}) = |\epsilon|.$
- **Uniform error (“hit or miss”)**:

$$C_{\theta}(\hat{\theta}) = \begin{cases} 0 & |\epsilon| \leq \frac{\Delta}{2} \\ 1 & \text{otherwise} \end{cases}$$
Bayesian Estimation: Minimum Mean Absolute Error

Absolute error cost assignment: \( C_\Theta(g(y)) = \| g(y) - \Theta \|_1 \).

\[
\hat{\theta}_{\text{mmae}}(y) = \arg \min_{g(\cdot)} \mathbb{E}[\| g(y) - \Theta \|_1 \mid Y = y]
\]

\[
= \arg \min_u \mathbb{E} \left[ \sum_i |u_i - \Theta_i| \mid Y = y \right]
\]

\[
= \arg \min_u \sum_i \left\{ \int_{-\infty}^{u_i} (u_i - \theta_i) \pi_y(\theta_i) \, d\theta_i + \int_{u_i}^{\infty} (\theta_i - u_i) \pi_y(\theta_i) \, d\theta_i \right\}
\]

\[
:= q(u_i, y)
\]

where \( \pi_y(\theta_i) \) is the posterior density of parameter \( \theta_i \) given the observation \( Y = y \). The quantity \( q(u_i, y) \) is differentiable in \( u_i \), so we can solve this using Leibnitz's Rule...

\[
\frac{\partial}{\partial u_i} q(u_i, y) = \int_{-\infty}^{u_i} \pi_y(\theta_i) \, d\theta_i - \int_{u_i}^{\infty} \pi_y(\theta_i) \, d\theta_i
\]

What value of \( u_i \) makes this go to zero? The **median** of the random parameter \( \Theta_i \), conditioned on \( Y = y \). Hence, we can conclude that \( \hat{\theta}_{\text{mmae}}(y) = \text{Median}[\Theta \mid Y = y] \).
Hit-or-miss cost assignment:

\[ C_{\theta}(g(y)) = \begin{cases} 0 & \|g(y) - \theta\|_\infty \leq \frac{\Delta}{2} \\ 1 & \text{otherwise} \end{cases} \]

We seek to find an estimator that minimizes the Bayes risk

\[ \hat{\theta}_{\text{map}}(y) = \arg \min_{g(\cdot)} \mathbb{E}[C_{\theta}(g) \mid Y = y]. \]

To find \( \hat{\theta}_{\text{map}}(y) \), we fix \( y \) and \( g(y) = u = [u_1, \ldots, u_m]^\top \) and write

\[ \hat{\theta}_{\text{map}}(y) = \arg \min_u \int C_{\theta}(u)\pi_y(\theta) \, d\theta \]

\[ = \arg \min_u \left\{ 1 - \int_{u_1 - \Delta/2}^{u_1 + \Delta/2} \cdots \int_{u_m - \Delta/2}^{u_m + \Delta/2} \pi_y(\theta) \, d\theta \right\} \]
Bayesian Estimation: Maximum A Posteriori Probability

We can’t take this much further without two additional assumptions: (A1) the posterior density $\pi_y(\theta)$ is smooth and (A2) $\Delta$ is small. Under these assumptions, we can write

$$
\int_{u_1-\Delta/2}^{u_1+\Delta/2} \cdots \int_{u_m-\Delta/2}^{u_m+\Delta/2} \pi_y(\theta) d\theta \approx \Delta^m \pi_y(\theta)|_{\theta=u}.
$$

Hence we have

$$
\hat{\theta}_{\text{map}}(y) = \arg \min_u \{1 - \Delta^m \pi_y(u)\}.
$$

Since $\Delta > 0$, we can discard the $\Delta^m$ term and use our usual tricks to write

$$
\hat{\theta}_{\text{map}}(y) = \arg \max_u \pi_y(u).
$$

Remarks:

- $\hat{\theta}_{\text{map}}(y) = \text{Mode}[\Theta | Y = y]$.
- The solution is usually unique (but doesn’t have to be).
Bayesian Estimation: Summary of Common Approaches

- Bayesian MMSE, MMAE, and MAP estimators are distinguished only by the choice of cost assignment.
  - Squared error cost assignment: MMSE ⇒ conditional mean
    \[ \hat{\theta}_{\text{mmse}}(y) = \mathbb{E}[\Theta | Y = y]. \]
  - Absolute error cost assignment: MMAE ⇒ conditional median
    \[ \hat{\theta}_{\text{mmae}}(y) = \text{Median}[\Theta | Y = y]. \]
  - Uniform cost assignment with smooth posterior and small $\Delta$: MAP ⇒ conditional mode
    \[ \hat{\theta}_{\text{map}}(y) = \text{Mode}[\Theta | Y = y]. \]

- The observation $Y = y$ converts the given prior distribution $\pi(\theta)$ on the unknown parameter(s) to the posterior distribution $\pi_y(\theta)$.
- Each estimator is simply a feature of this posterior distribution.
Bayesian Estimation: Summary of Common Approaches

- Prior
- Posterior
- MMSE
- MMSE
- MAP
- MAE