Cost functions in Bayesian decision theory

Sensor fusion & nonlinear filtering

Lars Hammarstrand

BAYESIAN DECISION THEORY – SUMMARY

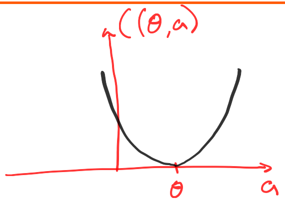
- Bayesian decision theory relies on
  1. Likelihood: $p(y|\theta)$
  2. Prior distribution: $p(\theta)$
  3. Loss function: $C(\theta, a)$
- Combining likelihood and prior gives posterior

$$p(\theta|y) \propto p(y|\theta)p(\theta).$$

Posterior and loss gives decisions

$$\hat{\theta} = \arg\min_a \int_\Theta C(\theta, a)p(\theta|y)\, d\theta.$$

THE QUADRATIC LOSS FUNCTION



Minimum mean squared error estimator, MMSE

- Parameter estimation, $\theta \in \Theta = \mathbb{R}^n$
- Most common loss function is the quadratic loss

$$C(\theta, a) = \|\theta - a\|_2^2 = (\theta - a)^T(\theta - a)$$

- Let: $\bar{\theta} = \mathbb{E}\{\theta | y\}$, $\mathbf{P} = \text{Cov}\{\theta | y\} = \mathbb{E}\left\{(\theta - \bar{\theta})(\theta - \bar{\theta})^T | y\right\}$

- Optimal estimator: $E\{C(\theta,a)|y\} = E\{(\theta-a)^T(\theta-a)|y\} = E\{(\overbrace{\theta-\bar\theta}^{\text{zero mean}}+\overbrace{\bar\theta-a}^{\text{Determ.}})^T(\theta-\bar\theta+\bar\theta-a)|y\}$

$$= \underbrace{E\{(\theta-\bar\theta)^T(\theta-\bar\theta)|y\}}_{Tr\{P\}} + \underbrace{E\{(\theta-\bar\theta)^T|y\}}_{=0}(\bar\theta-a) + 0 + (\bar\theta-a)^T(\bar\theta-a)$$

$$\hat\theta_{MMSE} = \min_a \arg E\{C(\theta,a)|y\} = \bar\theta$$

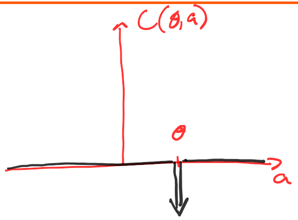
THE $0-1$ LOSS FUNCTION

Maximum a-posteriori estimator, MAP

- Parameter estimation, $\theta \in \Theta = \mathbb{R}^n$.
- Another common choice is the $0-1$ loss function

$$C(\theta, a) = -\delta(\theta - a),$$

$\delta(\cdot)$ is the Dirac's delta function.



- Optimal estimator: $\mathbb{E}\{C(\theta, a)\big|y\} = \mathbb{E}\{-\delta(\theta - a)\big|y\}$

$= -\int p(\theta\big|y)\delta(\theta - a)\,d\theta = -p(\theta\big|y)\Big|_{\theta=a}$

$$\Rightarrow \quad \hat{\theta}_{\text{MAP}} = \arg\min_{a} -p(\theta\big|y)\Big|_{\theta=a}$$

$$= \arg\max_{\theta} p(\theta\big|y)$$

SELF-ASSESSMENT

Suppose $p(\theta|y) = \mathcal{N}(\theta; \bar{\theta}, \mathbf{P})$. The MMSE and MAP estimators are, respectively,

- $\bar{\theta} + \mathrm{tr}\{\mathbf{P}\}$ and $\bar{\theta}$.
- $\bar{\theta}$ and $\mathcal{N}(\theta; \bar{\theta}, \mathbf{P})$.
- $\bar{\theta}$ and $\bar{\theta}$.
- $\bar{\theta} + \mathrm{tr}\{\mathbf{P}\}$ and $\mathcal{N}(\theta; \bar{\theta}, \mathbf{P})$.

Only one statement is correct.