

Egzamin: Statystyka Bayesowska

MIMUW, styczeń 2018.

1. Zakładamy, że obserwowane zmienne losowe X_1 i X_2 są, przy danej wartości parametru θ , warunkowo niezależne i $X_i|\theta \sim N(\theta, \sigma_i^2)$. Parametr θ ma rozkład *a priori* normalny, $\theta \sim N(\mu, v^2)$. Parametry σ_1 , σ_2 i v są znane.

- (a) Oblicz rozkład a posteriori $\theta|X_1, X_2$.
- (b) Oblicz $\text{Cov}(X_1, X_2)$.
- (c) Podaj rozkład brzegowy wektora (X_1, X_2) .

2. Obserwowane zmienne losowe $X_1, \dots, X_n$ są, przy danej wartości parametru θ , warunkowo niezależne i $X_i|\theta \sim U(0, \theta)$. Parametr θ ma rozkład *a priori* Pareto o gęstości

$$\pi(\theta) = \frac{1}{\theta^2} \mathbb{I}(\theta > 1).$$

- (a) Oblicz rozkład a posteriori $\theta|X_1, \dots, X_n$.
 - (b) Znajdź Bayesowski zbiór ufności (*credible region*) na poziomie ufności $1 - \alpha$ o najmniejszej mierze Lebesgue'a.
 - (c) Rozważamy zagadnienie testowania $H_0 : \theta = 1$ przeciw $H_1 : \theta \sim \pi(\cdot)$. Oblicz czynnik Bayesa (BF).
3. Rozważamy następujący model klasyfikacji: obserwujemy wektor losowy X w $\mathbb{R}^d$, który należy do jednej z dwóch klas, $c = 1$ lub $c = 2$. Prawdopodobieństwa *a priori* obu klas są równe, $\mathbb{P}(C = 1) = \mathbb{P}(C = 2) = 1/2$. Rozkłady prawdopodobieństwa w obu klasach są normalne, z jednakowymi macierzami kowariancji, $X|c \sim N(\mu_c, V)$. Macierz V jest znana i dodatnio określona. Poszukujemy „funkcji dyskryminacyjnej”, czyli $\delta : \mathbb{R}^d \rightarrow \mathbb{R}$, która minimalizuje ryzyko $r(\delta) = \mathbb{E}\ell(C, \delta(X))$ dla wykładniczej funkcji straty

$$\ell(c, \delta(x)) = \begin{cases} e^{-\delta(x)} & \text{dla } c = 2; \\ e^{\delta(x)} & \text{dla } c = 1. \end{cases}$$

- (a) Znajdź optymalną (Bayesowską) funkcję δ^* .
- (b) Podaj wzór na ryzyko $r(\delta^*)$ dla tej funkcji, w zależności od odległości Mahalanobisa $D(\mu_1, \mu_2) = ((\mu_2 - \mu_1)^\top V^{-1}(\mu_2 - \mu_1))^{1/2}$.
- (c) Rozważamy regułę klasyfikacyjną $c^* : \mathbb{R}^d \rightarrow \{1, 2\}$ zdefiniowaną wzorem

$$c^*(x) = \begin{cases} 2 & \text{jeśli } \delta^*(x) > 0; \\ 1 & \text{jeśli } \delta^*(x) < 0. \end{cases}$$

Dla tej reguły, podaj wzór na prawdopodobieństwo błędnej klasyfikacji $\mathbb{P}(c^*(X) \neq C)$ w zależności od odległości Mahalanobisa.

Wskazówki: Na ćwiczeniach obliczaliśmy iloraz $\mathbb{P}(C = 2|x)/\mathbb{P}(C = 1|x)$ w tym modelu. Pamiętaj, że jeśli $X \sim N(\mu, V)$, to $b^\top(X - a) \sim N(b^\top(\mu - a), b^\top V b)$, dla $a, b \in \mathbb{R}^d$.

4. W urnie znajduje się b białych kul i c kul czerwonych. Losujemy kolejno, jedną po drugiej, bez zwracania n kul.
 - (a) Oblicz prawdopodobieństwo, że pierwsza z wylosowanych kul była biała, jeśli wiadomo, że ostatnia, n -ta wylosowana kula jest biała.
 - (b) Oblicz $\text{Cov}(X_1, X_n)$, gdzie $X_i = \mathbb{I}(\text{biała kula w } i\text{-tym ciągnięciu})$.
 - (c) Odpowiedz na te same pytania, jeśli zamiast losowania bez zwracania rozpatrujemy schemat Polya: po wylosowaniu kolejnej kuli, zważamy ją i dokładamy do urny jeszcze jedną kulę tego samego koloru.
5. Niech $(X_0, X_1, \dots, X_k) = X_{0:k}$ będzie łańcuchem Markowa na przestrzeni stanów $\{1, 2\}$ o macierzy przejścia

$$P = \begin{pmatrix} 1 - \alpha & \alpha \\ \alpha & 1 - \alpha \end{pmatrix}.$$

Zakładamy, że rozkład początkowy jest rozkładem stacjonarnym, $\mathbb{P}(X_0 = 1) = \mathbb{P}(X_0 = 2) = 1/2$. Ciąg $X_0, X_1, \dots, X_k$ nie jest bezpośrednio obserwowany. Obserwujemy zmienne losowe $Y_0, Y_1, \dots, Y_k$ o wartościach w zbiorze $\{1, 2\}$ takie, że Y_i zależy tylko od X_i , przy czym

$$L(x, y) = \mathbb{P}(Y_i = y | X_i = x) = \begin{cases} 1 - \varepsilon & \text{dla } y = x; \\ \varepsilon & \text{dla } y \neq x, \end{cases} \quad x, y \in \{1, 2\}.$$

- (a) Interesuje nas rozkład a posteriori $\pi_{\text{post}}(x_{0:k}) = \mathbb{P}(X_{0:k} = x_{0:k} | y_{0:k})$ na przestrzeni $\{1, 2\}^{k+1}$. Konstruujemy próbnik Gibbsa zbieżny do rozkładu $\pi_{\text{post}}(\cdot)$. Podaj pełne rozkłady warunkowe $\pi_{\text{post}}(x_i | x_{-i})$, gdzie $x_{-i} = (x_0, \dots, x_{i-1}, x_{i+1}, \dots, x_k)$. (Możesz założyć, że $0 < i < k$, bo „skrajne przypadki” $i = 0$ lub $i = k$ wymagają tylko dość oczywistych modyfikacji.)
- (b) Oblicz $\mathbb{P}(X_i = 2 | Y_i = 2)$.
- (c) Oblicz $\mathbb{P}(X_{i+1} = 2 | Y_i = 2)$.

Uwaga: Łączny rozkład prawdopodobieństwa $(X_{0:k}, Y_{0:k})$ jest opisany wzorem

$$\mathbb{P}(X_{0:k} = x_{0:k}, Y_{0:k} = y_{0:k}) = \frac{1}{2} L(x_0, y_0) \prod_{i=1}^k P(x_{i-1}, x_i) L(x_i, y_i).$$

Mamy tu przykład „ukrytego modelu Markowa” (HMM). Tę rodzinę modeli rozważaliśmy na zajęciach, ale rozwiązanie zadania jest całkowicie elementarne i nie wymaga znajomości tej partii materiału.