Solution to Homework 4
Bayesian Methods

Anna Pavlova

Problem 1

1. We need determine $\lambda = (\pi, A, B)$ where

$$p(t_{i+1}|t_i) = A_{t_i t_{i+1}}$$

$$p(x_i|t_i) = B_{t_i x_i}$$

$$p(t_i) = \pi_{t_i}$$

Using these definitions:

$$\pi = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$$

$$B = \begin{bmatrix} 0.5 & 0.5 \\ 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix}$$

where first column black, second - white.

$$A = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

2. We need to calculate the probability of observing the marble color sequence $O = (1, 1, 2, 2, 1)$, if hidden sequence of boxes selected is $Q = (1, 1, 3, 3, 2)$.

$$\begin{array}{ccccccccc} 1 & \longrightarrow & 1 & \longrightarrow & 3 & \longrightarrow & 3 & \longrightarrow & 2 \\ {\scriptstyle 0.5}\downarrow & & {\scriptstyle 0.5}\downarrow & & {\scriptstyle 0.75}\downarrow & & {\scriptstyle 0.75}\downarrow & & {\scriptstyle 0.75}\downarrow \\ 1 & & 1 & & 2 & & 2 & & 1 \end{array}$$

$$p(Q, O|\lambda) = \pi_{q_1} \prod_{i=1}^{4} A_{q_i q_{i+1}} \prod_{i=1}^{5} B_{q_i o_i}$$

$$p(Q, O|\lambda) = \frac{1}{3}^5 \cdot (0.5)^2 \cdot (0.75)^3$$

3. $Q* = (2,2,3,3,2)$, $O = (1,1,2,2,1)$

$$\begin{array}{ccccccccc} 2 & \longrightarrow & 2 & \longrightarrow & 3 & \longrightarrow & 3 & \longrightarrow & 2 \\ {\scriptstyle 0.75}\big\downarrow & & {\scriptstyle 0.75}\big\downarrow & & {\scriptstyle 0.75}\big\downarrow & & {\scriptstyle 0.75}\big\downarrow & & {\scriptstyle 0.75}\big\downarrow \\ 1 & & 1 & & 2 & & 2 & & 1 \end{array}$$

$$p(Q*, O|\lambda) = \frac{1}{3}^5 \cdot 0.75^5$$

4.

$$\log \frac{P(Q_*|O,\lambda)}{P(Q|O,\lambda)} = \log \frac{P(Q_*|O,\lambda)P(O|\lambda)}{P(Q|O,\lambda)P(O|\lambda)} = \log \frac{P(Q_*,O,\lambda)}{P(Q,O,\lambda)} = \log \frac{0.75^5}{0.25^2 \cdot 0.75^3} = \log \frac{0.75^2}{0.25^2} = 2\log 3$$

Problem 2

Calculate $\sum\limits_{j=1}^{n} (\hat{y}_{-j} - y_j)^2$

Let us consider the quadratic form in the exponent of joint Normal distribution

$$\sum_{i,j} (y_i y_j (K^{-1})_{ij}$$

and select part that depends only on y_j

$$(K^{-1})_{jj} \Big[y_j^2 + 2\frac{(\sum_{i=1, i\neq j}^{n} (K^{-1})_{ij} y_i) y_j}{(K^{-1})_{ij}} \Big]$$

It is quadratic form of conditional distribution $p(y_j|\boldsymbol{y}_{-j})$, let us find mean:

$$\mu_j = -\frac{\sum\limits_{i=1, i\neq j}^{n} (K^{-1})_{ij} y_i}{(K^{-1})_{jj}} = \hat{y}_{-j}$$

$$\sum_{j=1}^{n} (\hat{y}_{-j} - y_j)^2 = \sum_{j=1}^{n} \Big[\frac{\sum\limits_{i=1, i\neq j}^{n} (K^{-1})_{ij} y_i}{(K^{-1})_{jj}} + y_j \Big]^2 =$$
$$= \sum_{j=1}^{n} \frac{(K^{-1} y)_j^2}{(K^{-1})_{jj}^2}$$

This formula demands calculation only of full covariance matrix K.

Problem 3

We need to Calculate a covariance function:$cov(y(x), \frac{\partial y(x)}{\partial x}|_{x=\tilde{x}}$

$$cov(y(x), y(\tilde{x}) = \mathbb{E} y(x) y(\tilde{x})$$

$$\frac{\partial}{\partial \tilde{x_1}} cov(y(x)y(\tilde{x})) = \frac{\partial}{\partial \tilde{x_1}} \mathbb{E} y(x)y(\tilde{x}) = \mathbb{E} y(x) \frac{\partial}{\partial \tilde{x_1}} y(\tilde{x}) = cov(y(x), \frac{\partial y}{\partial x_1}|_{x=\tilde{x}})$$

2-3. Difference of two normally distributed values is normal, because of it, $p(D, D_1)$ is normal distribution with zero mean and block covariance:

$$\mathcal{N}(0, \begin{bmatrix} K & K_1 \\ K_1^T & K_{11} \end{bmatrix})$$

Where K_1 covariance between D and D_1, K_{11} between D_1 and D_1 that can be calculated analogously (by twice differentiation). It means that $p(y(\boldsymbol{(}x)|D, D_1)$ is normal and we can find parameters of its distribution similarly to usual GP.

$$\mathbb{E} y(\boldsymbol{x}) = \hat{k}^T \hat{K}^{-1} \hat{y}$$

$$\mathbb{D} y(\boldsymbol{x}) = k(\boldsymbol{x}, \boldsymbol{x}) - \hat{k}^T \hat{K}^{-1} \hat{k}$$

"Hat"means concatenation with derivatives.