CK0146: Exercise 04 (2017.1)

Exercise Q.1 (Bayesian linear regression).

Forensic people say that the length of long bones in a human body is related to the person's height. These relationships appears to be so strong that you could predict a person's height, if you knew the length of one of his femur. The equation they use to estimate the height from the femur is of the linear form $y(x, \mathbf{w}) = w_0 + w_1 x$, in which y is used to denote the person's height, x is the length of the femur and $\mathbf{w} = (w_0, w_1)$ is the vector of coefficients relating x and y.

You are given a collection of data $\{(x_i, t_i)\}$ with $i = 1, \dots, N$ and your goal is to estimate the parameter vector $\mathbf{w}$. The data, in centimetres, can be downloaded `here`. In addition to assuming that the model equation is correct, we assume that the noise added when measuring the heights y is Gaussian with zero mean and known standard deviation, that is $t = y(x, \mathbf{w}) + \varepsilon$ with $\varepsilon \sim \mathcal{N}(\mu = 0, \sigma = 0.5)$. As always, we are implicitly also assuming that the measurements x of the femur are exact.

We are interested in three types of estimates of $\mathbf{w}$:

- The maximum likelihood estimate, $\mathbf{w}_{ML}$
- The batch Bayesian estimate, the posterior $p(\mathbf{w}|\mathtt{t})$
- The sequential Bayesian estimate, the posterior $p(\mathbf{w}|\mathtt{t})$

For the Bayesian estimates, assume some Gaussian prior $\mathbf{w} \sim \mathcal{N}(\mathbf{m}_0, \mathbf{S}_0 = \alpha^{-1}\mathbf{I})$. It is not exactly a lot of data that you have for learning ($N = 8$), so it might be a good idea to google for a good prior.

Comment on the results, show the posterior distributions $p(\mathbf{w}|\mathtt{t})$ and some draws from it. As for the sequential case, show also some intermediate solutions (0 observations, 1 observation, ...)

What is the estimated (posterior distribution of) height of a person whose femur has length 48 cm?