Homework 1 – Solution

1. If we define $R = diag(r_1, \ldots, r_N)$ to be diagonal matrix containing the weighting coefficients, then we can write the weighted sum-of-squares cost function in the form:

$$E_D(\mathrm{w}) = \frac{1}{2}(\mathrm{t} - \Phi\mathrm{w})^T\mathbf{R}(\mathrm{t} - \Phi\mathrm{w})$$

Setting the derivative with respect to w to zero, and rearranging, then gives:

$$\mathrm{w}^* = (\Phi^{\mathrm{T}}\mathrm{R}\Phi)^{-1}\Phi^{\mathrm{T}}\mathrm{Rt}$$

which reduces to the standard solution for the case $\mathrm{R} = \mathrm{I}$.

评分标准：$E_D(w)$表达式 1 分，最优解 1 分。如果不是按照参考答案的矩阵形式但是推导过程和结果正确也是满分。

2. (a) The likelihood of λ given $\mathcal{X}$ is

$$L(\lambda \mid \mathcal{X}) = \prod_{\ell=1}^{N}\left(\lambda\, e^{-\lambda x^{(\ell)}}\right) = \lambda^N \exp\left(-\lambda \sum_{\ell=1}^{N} x^{(\ell)}\right).$$

(b) The maximum likelihood estimate of λ can be found by setting the derivative of the log likelihood with respect to λ to zero:

$$\frac{d}{d\lambda}\left(N \log\lambda - \lambda \sum_{\ell=1}^{N} x^{(\ell)}\right) = \frac{N}{\lambda} - \sum_{\ell=1}^{N} x^{(\ell)} = 0,$$

which can be solved to give

$$\lambda_{\mathrm{ML}} = \frac{N}{\sum_{\ell=1}^{N} x^{(\ell)}} = \frac{1}{\frac{1}{N}\sum_{\ell=1}^{N} x^{(\ell)}} = \frac{1}{\bar{x}},$$

where $\bar{x}$ is the sample mean of $\mathcal{X}$.

(c) By Bayes' rule,

$$\begin{aligned}
p(\lambda \mid \mathcal{X}) &\propto p(\mathcal{X} \mid \lambda)\, p(\lambda) \\
&\propto \lambda^N \exp\left(-\lambda \sum_{\ell=1}^{N} x^{(\ell)}\right) \lambda^{\alpha-1} \exp(-\lambda\beta) \\
&= \lambda^{(\alpha+N)-1} \exp\left[-\lambda\Big(\beta + \sum_{\ell=1}^{N} x^{(\ell)}\Big)\right] \\
&= \lambda^{(\alpha+N)-1} \exp\Big[-\lambda(\beta + N\bar{x})\Big].
\end{aligned}$$

Thus $p(\lambda \mid \mathcal{X})$ has the same form as $p(\lambda)$ but with different parameter values:

$$\begin{aligned}
\alpha_{\mathrm{new}} &= \alpha + N \\
\beta_{\mathrm{new}} &= \beta + N\bar{x}.
\end{aligned}$$

评分标准：a, b, c 各 1 分