1. a) $L(\omega, c, \lambda) = -\sum_{k=1}^{m} c_k \ln \omega_k + \lambda\left(\sum_{k=1}^{m} \omega_k - 1\right)$

b) $\frac{\partial L}{\partial \omega} = \frac{\partial L}{\partial \lambda} = \frac{\partial L}{\partial c} = 0$

$$\frac{\partial L}{\partial \omega} = -\sum_{k=1}^{m} c_k \frac{1}{\omega_k} + m\lambda = 0$$

$$\frac{\partial L}{\partial c} = -\sum_{k=1}^{m} \ln \omega_k = 0$$

$$\frac{\partial L}{\partial \lambda} = \sum_{k=1}^{m} \omega_k - 1 = 0$$

$$\Rightarrow \frac{\partial L}{\partial \omega} = -c\,\omega^{-1} + m\lambda = 0$$

$$\Rightarrow \omega = \frac{1}{m\lambda} c$$

2. a) The Fisher kernel essentially is a function that demonstrates the similarity of two objects as criteria.

b) I expect lower training error since we are essentially increasing, and so increasing generalization error.