آیدین شکر - تکلیف شماره هفت درس شناسایی آماری الگو ۴۰۰۲۳۴۹۴

4.17 & 4.20. Refer to the MATLAB scripts.

4.22 $\phi(x) = \left[\frac{1}{\sqrt{2}}, \cos x, \cos 2x, \dots, \cos kx, \sin x, \sin 2x, \dots, \sin kx\right]^T$

$$K(x_i, x_j) = y_i^T y_j = \frac{1}{2} + \cos x_i \cos x_j + \cos 2x_i \cos 2x_j + \dots + \sin x_i \sin x_j + \sin 2x_i \sin 2x_j + \dots + \sin kx_i \sin kx_j$$

$$= \frac{1}{2} + \left[\cos x_i \cos x_j + \sin x_i \sin x_j\right] + \dots + \left[\cos kx_i \cos kx_j + \sin kx_i \sin kx_j\right]$$

$$= \frac{1}{2} + \cos(x_i - x_j) + \cos(2(x_i - x_j)) + \dots + \cos(k(x_i - x_j))$$

setting $\alpha = (x_i - x_j)$ we have:

$$S = \frac{1}{2} + \cos\alpha + \cos 2\alpha + \dots + \cos n\alpha$$

$$\Rightarrow 2S \times \sin\frac{\alpha}{2} = \sin\frac{\alpha}{2} + 2\cos\alpha \sin\frac{\alpha}{2} + 2\cos 2\alpha \sin\alpha/2 + \dots + 2\cos n\alpha \sin\frac{\alpha}{2}$$

$$= \sin\alpha/2 + \left[\sin(\cancel{\alpha + \alpha/2}) - \sin(\alpha - \alpha/2)\right]$$

using $2\cos A \sin B = \sin(A+B) - \sin(A-B)$

$$+ \left[\sin(\cancel{2\alpha + \alpha/2}) - \sin(\cancel{2\alpha - \alpha/2})\right]$$

$$+ \vdots$$

$$+ \left[\sin(n\alpha + \alpha/2) - \sin(\cancel{n\alpha - \alpha/2})\right]$$

$$\Rightarrow 2S \times \sin\alpha/2 = \cancel{\sin\alpha/2} - \cancel{\sin(\alpha - \alpha/2)} + \sin((n + \tfrac{1}{2})\alpha)$$

$$\Rightarrow S = \frac{\sin((n+\frac{1}{2})\alpha)}{2\sin\alpha/2} \longrightarrow \text{which is what we want}$$

4.23. We will first form the Lagrangian upon the constraint $\sum_{i=1}^{M} P(\omega_i|x)=1 \rightarrow P_{\text{combiner}}$

$$\Rightarrow L(D_{av}) = \frac{1}{L}\sum_{j=1}^{L}\sum_{i=1}^{M} P(\omega_i|x)\ln\frac{P(\omega_i|x)}{P_j(\omega_i|x)} - \lambda\left(\sum_{i=1}^{M}P(\omega_i|x)-1\right)$$

$$\rightarrow \frac{\partial L(D_{av})}{\partial p(\omega_i|x)} = \frac{1}{L}\sum_{j=1}^{L}\left[\ln\frac{P(\omega_i|x)}{P_j(\omega_i|x)}+1\right]-\lambda=0$$

(note that in the inner most summations of both terms only one summand has a derivative w.r.t. $P(\omega_i|x)$, where $i_{\text{summation}} = i_{\text{derivative}}$ )

$$\Rightarrow \frac{1}{L}\sum_{j=1}^{L}\ln P(\omega_i|x) - \frac{1}{L}\sum_{j=1}^{L}\ln P_j(\omega_i|x) + 1 - \lambda = 0$$

$$\Rightarrow \ln P(\omega_i|x) = (\lambda-1) + \frac{1}{L}\sum_{j=1}^{L}\ln P_j(\omega_i|x)$$

$$\Rightarrow P(\omega_i|x) = \exp\{(\lambda-1)\}\exp\left\{\frac{1}{L}\sum_{j=1}^{L}\ln P_j(\omega_i|x)\right\} \qquad (I)$$

Plugging (I) into the constraint equation we have:

$$\sum_{i=1}^{M}\exp\{(\lambda-1)\}\exp\left\{\left(\frac{1}{L}\sum_{j=1}^{L}\ln P_j(\omega_i|x)\right)\right\} = 1 \Rightarrow \exp\{(\lambda-1)\} = \frac{1}{\sum_{i=1}^{M}\exp\left\{\frac{1}{L}\sum_{j=1}^{L}\ln P_j(\omega_i|x)\right\}} \qquad (II)$$

plugging (II) back into (I):

$$P(\omega_i|x) = \frac{\exp\left\{\frac{1}{L}\sum_j \ln P_j(\omega_i|x)\right\}}{\sum_i \exp\left\{\frac{1}{L}\sum_j \ln P_j(\omega_i|x)\right\}} \longrightarrow \text{using } \sum_i \ln a_i = \ln\prod_i a_i \text{ we have}$$

$$P(\omega_i|x) = \frac{\exp\left\{\frac{1}{L}\ln\prod_j P_j(\omega_i|x)\right\}}{\sum_i \exp\left\{\frac{1}{L}\ln\prod_j P_j(\omega_i|x)\right\}} = \left[\frac{\prod_j P_j(\omega_i|x)}{\sum_i\prod_j P_j(\omega_i|x)}\right]^{1/L}$$

which is what we wanted.

4.24. we will follow the exact same logic as 4.23.

$$L(D_{av}^{*}) = \frac{1}{L}\sum_{j=1}^{L}\sum_{i=1}^{M} P_j(w_i|x) \ln \frac{P_j(w_i|x)}{p(w_i|x)} - \lambda\left(\sum_{i=1}^{M} P(w_i|x) - 1\right)$$

$$\rightarrow \frac{\partial L(D_{av}^{*})}{\partial p(w_i|x)} = \frac{1}{L}\sum_{j=1}^{L} \frac{-P_j(w_i|x)}{p(w_i|x)} - \lambda = 0$$

$$\rightarrow \frac{-1}{L\,p(w_i|x)}\sum_{j=1}^{L} P_j(w_i|x) = \lambda \rightarrow p(w_i|x) = \frac{1}{-L\lambda}\sum_{j=1}^{L} P_j(w_i|x) \quad (\mathrm{I})$$

Substituting (I) into the constraint:

$$\sum_{i=1}^{M} \frac{-1}{L\lambda}\sum_{j=1}^{L} P_j(w_i|x) = 1 \Rightarrow \lambda = \frac{-1}{L}\sum_{i=1}^{M}\sum_{j=1}^{L} P_j(w_i|x) \quad (\mathrm{II})$$

Plugging (II) back into (I):

$$P(w_i|x) = \frac{\sum_{j=1}^{L} P_j(w_i|x)}{\sum_{i=1}^{M}\sum_{j=1}^{L} P_j(w_i|x)}$$

, which is what we wanted to show because:

$$\sum_{i=1}^{M}\sum_{j=1}^{L} P_j(w_i|x) = \sum_{j=1}^{L} \underbrace{\boxed{\sum_{i=1}^{M} P_j(w_i|x)}}_{\text{Sums to } 1} = L$$

$$\Rightarrow P(w_i|x) = \frac{1}{L}\sum_{j=1}^{L} P_j(w_i|x)$$