Homework 1

Swaraj Khadanga

October 2, 2017

1. a. Estimate Maximum Likelihood for
$f(x) = \frac{x}{\theta^2} exp \frac{-x^2}{2\theta^2}$
so Maximum Likelihood function (ML) $= \prod_1^N \frac{x_i}{\theta^2} exp \frac{-x_i^2}{2\theta^2}$
$\implies$ Maximum Log Likelihood function (MLL) $=$
$\sum_1^N [\log x_i - 2\log\theta - \frac{x_i^2}{2\theta^2}]$

To maximize MLL we need to set $\frac{\partial MLL}{\partial \theta} = 0 \implies \sum_1^N [0 - \frac{2}{\theta} - \frac{x_i^2}{2} \frac{-2}{\theta^3}] = 0$
$\implies \sum_1^N [\frac{-2}{\theta} + \frac{x_i^2}{\theta^3}] = 0 \implies \frac{1}{\theta} \sum_1^N [2 + \frac{x_i^2}{\theta^2}] = 0$
$\implies \sum_1^N [-2 + \frac{x_i^2}{\theta^2}] = 0$
$\implies \sum_1^N -2 + \sum_1^N \frac{x_i^2}{\theta^2} = 0$
$\implies -2N + \frac{1}{\theta^2} \sum_1^N x_i^2 = 0$
$\implies \frac{1}{\theta^2} \sum_1^N x_i^2 = 2N$
$\implies \frac{\sum_1^N x_i^2}{2N} = \theta^2$
$\implies \hat{\theta} = \sqrt{\frac{\sum_1^N x_i^2}{2N}}$

b. $f(x|\alpha,\theta) = \alpha\theta^{-\alpha} x^{\alpha-1} \exp[-(\frac{x}{\theta})^\alpha]$
so Maximum Likelihood (ML) $= \prod_1^N \alpha\theta^{-\alpha} x_i^{\alpha-1} \exp[-(\frac{x_i}{\theta})^\alpha]$
Maximum Log Likelihood
(MLL) $= \sum_1^N [\log\alpha - \alpha\log\theta + (\alpha-1)\log x_i - (\frac{x_i}{\theta})^\alpha]$
To maximize MLL we need to set $\frac{\partial MLL}{\partial \theta} = 0$
$\implies \sum_1^N [0 - \frac{\alpha}{\theta} + 0 - \alpha(\frac{x_i}{\theta})^{\alpha-1} \frac{-x_i}{\theta^2}] = 0$
$\implies \sum_1^N [-\frac{\alpha}{\theta} + \frac{\alpha}{\theta}(\frac{x_i}{\theta})^\alpha] = 0$
$\implies \frac{\alpha}{\theta} \sum_1^N [-1 + (\frac{x_i}{\theta})^\alpha] = 0$
$\implies \sum_1^N [-1 + (\frac{x_i}{\theta})^\alpha] = 0$
$\implies -N + \sum_1^N (\frac{x_i}{\theta})^\alpha = 0$
$\implies \sum_1^N (\frac{x_i}{\theta})^\alpha = N$
$\implies \frac{1}{\theta^\alpha} \sum_1^N {x_i}^\alpha = N$
$\implies \theta^\alpha = \frac{1}{N} \sum_1^N {x_i}^\alpha$
$\implies \hat{\theta} = [\frac{1}{N} \sum_1^N {x_i}^\alpha]^{\frac{1}{\alpha}}$

c. $f(x) = \frac{1}{\theta} \forall 0 <= x <= \theta$
so Maximum Likelihood (ML) $= \prod_1^N \frac{1}{\theta}$
$\implies ML = \frac{1}{\theta^N}$
To maximize the above likelihood we need to minimize θ
But we also have a constraint that $0 <= x <= \theta$
so $\hat{\theta} = max[x_i]$

2. $p(x|C_1) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp^{\frac{-(x-\mu_1)^2}{2\sigma_1^2}}$
$p(x|C_2) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp^{\frac{-(x-\mu_2)^2}{2\sigma_2^2}}$

$\log p(x|C_1) = -0.5\log 2\pi - log\sigma 1 - \frac{(x-\mu_1)^2}{2\sigma_1^2}$
$\log p(x|C_2) = -0.5\log 2\pi - log\sigma 2 - \frac{(x-\mu_2)^2}{2\sigma_2^2}$
$p(C|x) = \frac{P(x|C)P(C)}{P(x)}$
$P(x)$ is constant for each class. So we can ignore this in discriminant function.
Discriminant function
$g_1 = \log p(x|C_1) + \log p(C)$
$g_2 = \log p(x|C_2) + \log p(C)$
for each test item, we can calculate g1 and g2 and classify them into classes by comparing g1 and g2.

3. Multivariate Gaussian Log Likelihood estimation
$p(x|C) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}}\exp[-\frac{1}{2}(x-\mu)^\intercal\Sigma^{-1}(x-\mu)]$

   - First read test and train file.
   - From training file learn μ_1, μ_2, S_1, S_2 by calling mean and cov function.
   - calculate discriminant function by fitting to function $g_i = log(p(x|C_i)) + log(C_i)$ where $log(C_i)$ is prior.
   - calculate g1 and g2 by passing $(\mu_1, S_1) and (\mu 2, S_2)$
   - during calculation we can ignore the constant terms as they are same for both.
   - compare g_1 and g_2 and chose class 1 if g1 is greater and 2 if g2 is greater.
   - calculate the error rate by taking the fraction of number of misclassified classes by total number of test data.
   - calculate shared S $= p(C_1)S_1 + p(C_2)S_2$
   - do the same steps for same covariance S by passing (μ_1,S) and (μ_2,S).