Problem Set

February 24, 2020

1. **Probabilities**

   ($\star\star$) Suppose that we have three coloured boxes r (red), b (blue), and g (green). Box r contains 3 apples, 4 oranges, and 3 limes, box b contains 1 apple, 1 orange, and 0 limes, and box g contains 3 apples, 3 oranges, and 4 limes. If a box is chosen at random with probabilities $p(r) = 0.2$, $p(b) = 0.2$, $p(g) = 0.6$, and a piece of fruit is removed from the box (with equal probability of selecting any of the items in the box), then what is the probability of selecting an apple? If we observe that the selected fruit is in fact an orange, what is the probability that it came from the green box?

2. **Beta** Use the realtionsip given by

$$\int_0^1 \mu^{a-1}(1-\mu)^{b-1}\,\mathrm{d}\mu = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

   to show that the variance is given by

$$\mathrm{var}[\mu] \quad = \quad \frac{ab}{(a+b)^2(a+b+1)}$$

3. **Gibbs sampling**

   In *over relaxation* in Gibbs sampling at each of the the conditional distribution for a a particular component z_i has some mean μ_i and some variance σ_i^2. In over-relaxation, the value z_i is replaced with :

   $$z_i' = \mu_i + \alpha(z_i - \mu_i) + \sigma_i(1 - \alpha_i^2)^{1/2}\nu$$

   where ν is a Gaussian random variable with zero nean and ubit variance and $\alpha \in [-1, 1]$. Verify that z_i' also has mean μ_i and variance σ_i^2.

4. **Support Vector Machines**

   (a) BRIEFLY explain the 'Kernel Trick.'

   (b) Show that the value ρ of the margin in the maximum-margin hyperplane is given by

   $$\frac{1}{\rho^2} = \sum_{n=1}^{N} \lambda_n$$

5. **Indepedent Component Analysis for audio signals** The standard Ng algorithm for recovering mixed audio signals uses the following sigmoid for the cdf of the signals of

   $$g(s) = \frac{1}{1 + e^{-s}}.$$

   However it may recover if the sigmoid had a hyper- parameter αthat allowed the slope of the sigmoid to vary to fit individual audio signals. so that

   $$g(s) = \frac{1}{1 + e^{-\alpha s}}.$$

   Show how the best α could be found by computing the gradient:

   $$\frac{\partial L}{\partial \alpha}$$