Machine Learning

Homework 1

Not collected, not graded.

1 Probabilities

1. **Fruit basket**: Suppose that we have three colored boxes r (red), b (blue), and g (green). The respective box contents and the probability p of choosing a certain box are:

	r	g	b
apples	3	1	3
oranges	4	1	3
bananas	3	0	4
p(box)	0.2	0.2	0.6

A box is selected at random (according to p(box)), and a piece of fruit is picked from the box (with equal probability of selecting any of the items in the box), then

(a) what is the probability of selecting an apple?

(b) if we observe that the selected fruit is an orange, what is the probability it came from the green box?

(c) is banana a fruit in the first place? (haha....)

2. **Variance**

(a) The variance of a function $f\colon \mathbb{R} \to \mathbb{R}$ is defined as $\mathrm{var}[f] := \mathbb{E}\left[(f(x) - \mathbb{E}[f(x)])^2\right]$.
Show that $\mathrm{var}[f] = \mathbb{E}\left[f(x)^2\right] - \mathbb{E}[f(x)]^2$.

(b) The **covariance between two random variables** is $\mathrm{cov}[x,y] := \mathbb{E}_{x,y}\left[(x - \mathbb{E}[x])(y - \mathbb{E}[y])\right]$.
Show that if two random variables x and y are *independent*, then their covariance is zero.

2 Gaussians

1. **Marginal and posterior Gaussians**: Read and study §2.3.2 and §2.3.3.

2. **Maximum likelihood estimation**: Given a set of observations $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}^T$, $\mathbf{x}_n \in \mathbb{R}^D$, assumed to be drawn independently from a multivariate Gaussian distribution, we can estimate the parameters of the distribution by maximum likelihood.

(a) Write the corresponding log-likelihood function $\ln p(\mathbf{X} \mid \mu, \Sigma)$

(b) Derive the MLE of the mean, μ_{ML}

(c) For the special case $D = 1$, derive the MLE of the covariance, σ^2_{ML}

3. **Bayesian inference**: Consider a D-dimensional Gaussian random variable x with distribution $\mathcal{N}(x \mid \mu, \Sigma)$, where $\Sigma \in \mathbb{R}^{D\times D}$ is known, and for which we want to infer $\mu \in \mathbb{R}^D$ from a set of observations $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}^T$, $\mathbf{x}_n \in \mathbb{R}^D$. Given a prior distribution $p(\mu) = \mathcal{N}(\mu \mid \mu_0, \Sigma_0)$, find the corresponding posterior distribution $p(\mu|\mathbf{X})$.