Exercise 1

More about multivariate Gaussians.

The general expression of a univariate Gaussian with mean $\mu$ and variance $\sigma^2$ is

$$N(x|\mu,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp \left\{ -\frac{1}{2\sigma^2} (x - \mu)^2 \right\} \quad (1)$$

The general expression of a multivariate Gaussian over a $D$ dimensional vector $x$ with $D$ dimensional mean vector $\mu$ and $D \times D$ covariance matrix $\Sigma$ is

$$N(x|\mu,\Sigma) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right\} \quad (2)$$

where $|\Sigma|$ is the determinant of $\Sigma$.

- Now consider a multivariate Gaussian $N(x|\mu,\Sigma)$ in which the covariance matrix $\Sigma$ is a diagonal matrix, i.e., its elements can be written as $\Sigma_{ij} = \sigma_i^2 I_{ij}$, where $I_{ij}$ are the matrix elements of the identity matrix (so $I_{ij} = 0$ if $i \neq j$ and $I_{ii} = 1$). Show, using (1) and (2) that a multivariate Gaussian with diagonal covariance matrix, $\Sigma_{ij} = \sigma_i^2 I_{ij}$, factorizes into a product of univariate Gaussians

$$N(x|\mu,\Sigma) = \prod_{i=1}^D N(x_i|\mu_i,\sigma_i^2)$$

- Show that for arbitrary positive definite covariance matrix $\Sigma$ the distribution Eq. 2 is properly normalized. Hint: transform onto the basis of eigenvectors of $\Sigma$ and use the result of Appendix C on eigenvectors.

Exercise 2

A factory produces products $X$. 75% is of quality $x = 1$ and the remainder of quality $x = 2$. There is a test $Z$, which can be a real number $z$ between 0 and 1. The conditional probability density of $z$, depending on the quality $x$ is

$$p(z|x = 1) = 2(1 - z)$$

$$p(z|x = 2) = 1$$

1. Interpret these equations and compute $p(x|z)$ using Bayes’ rule

2. Compute the Bayes optimal decision to minimize misclassification rate as function of $z$, i.e. for which $z$ should one classify $x = 1$ and for which $z$ should one classify $x = 2$. 

1
3. Suppose we have a loss matrix \( L_{kj} \), expressing the loss for classifying as \( x = j \) while the true class is \( k \). Suppose this matrix is given by

\[
L_{11} = L_{22} = 0, \quad L_{12} = 1, \quad L_{21} = 5
\]

Compute the optimal decision boundary to minimize expected loss.

**Exercise 3**

The Gaussian distribution in one dimension with mean \( \mu \) and variance \( \sigma^2 \) is

\[
\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}
\]  

(3)

The Kullback-Leibler divergence \( KL(p||q) \) is defined as

\[
KL(p(x)||q(x)) = -\int p(x) \ln q(x)dx + \int p(x) \ln p(x)dx
\]  

(4)

Compute the Kullback-Leibler divergence \( KL(p||q) \) between two Gaussians with the same variance \( \sigma^2 \), but different means \( \mu \) and \( m \). So \( p(x) = \mathcal{N}(x|\mu, \sigma^2) \) and \( q(x) = \mathcal{N}(x|m, \sigma^2) \). Verify that \( KL(p||q) \geq 0 \) and equal if and only if \( \mu = m \).

**Exercise 4**

Minimize \( f(x, y) = 3x^2 + xy + y^2 \) under constraint \( x + 2y = 3 \).

**Exercise 5**

Bishop 1.34

**Exercise 6**

Bishop 1.39. Except "Draw a diagram..."