1. Probability. Lois knows that on average radio station RANDOM-FM plays 1 out of 4 of her requests. If she makes 3 requests, what is the probability that at least one request is played?

2. Let \( A \) and \( v \) be defined as

\[
A = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 1 & -3 \\ 3 & -3 & -3 \end{pmatrix} \quad v = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}
\]

Calculate \( Av \). Is \( v \) an eigenvector of \( A \), and if so what is the corresponding eigenvalue?

3. A random vector \( x \) has zero mean a diagonal covariance

\[
E(xx^T) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\]

where \( E \) stands for expectation (or mean average) of a random variable. If \( y = A^T x \) (using \( A \) from Q1) what is the covariance of the resulting random vector \( y \): \( E(yy^T) \)? You may use the fact that expectation is linear: \( E(Rxx^TS) = RE(xx^T)S \). This shows how covariances change under linear transformations.

4. Find the partial derivatives of the function \( f(x, y, z) = (x + 2y)^2 \sin(xy) \).

5. Let \( x \) be a Gaussian random variable with mean \( \mu \) and variance \( \sigma^2 \). What is the expected value of \( 2x^2 \). Show what form the distribution of \( 2(x - \mu)^2 \) takes. Hint: the distribution of \( x^2 \) for a standard normal \( (N(0,1)) \) is chi-squared distributed.

The following ones should be discussed in the tutorial

6. You have data consisting of records of the form \((x_1, x_2, x_3, y)\) where \(x_1, x_2\) and \(x_3\) are features, and \(y\) is a class label. Each element takes the value 0 or 1.
Define clearly and precisely the Naive Bayes classification method for obtaining the probability of \( y \) given the values of the other attributes. Show that the maximum likelihood estimate of the parameter \( p(x_i = 0|y = 1) \) is proportional to the number of times attribute \( i \) is 0 for class 1 data. Convince yourself that you could use a similar method to show that \( p(y = 1) \) is proportional to the number of times class 1 occurs in the data.

The training data has the form

\[
(1, 0, 1, 1), (0, 1, 0, 1), (0, 1, 1, 0), (0, 0, 0, 0), (1, 1, 1, 1), (0, 0, 0, 0), (1, 0, 1, 0)
\]

- Using Naive Bayes on this data, what is the probability that \( y = 1 \) given \( x_1 = 1, x_2 = 1 \) and \( x_3 = 0 \)?
- Using Naive Bayes on this data, what is the probability that \( y = 0 \) given \( x_1 = 1, x_2 = 1 \) and \( x_3 = 0 \)?
- Using Naive Bayes on this data, what is the highest posterior classification for \( y \) given \( x_1 = 1 \)?

7. If \( a \) and \( b \) are \( D \times 1 \) column vectors and \( M \) is a \( D \times D \) symmetric matrix, show that \( a^T M b = b^T M a \).

8. Consider the Gaussian distribution

\[
p(x) \propto \exp \left\{ -\frac{1}{2} (x^T A x - 2 x^T b) \right\}
\]

where \( A \) is a symmetric matrix. Show that the mean and covariance are given by

\[
\text{mean}(x) = A^{-1} b, \quad \text{cov}(x) = A^{-1}.
\]

9. Consider a bivariate Gaussian \( p(x_1, x_2) = N(x|\mu, \Sigma) \) with

\[
\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}
\]

a. What is \( p(x_1) \)?

b. Suppose \( \mu = 0 \). \( p(x_1|x_2) \) is a Gaussian. Write down its mean and variance (using the formulae given in the Gaussian lecture).

c. Contrast your answers from parts a. and b. Sketch the conditional mean as a function of \( x_2 \).

d. **Bonus question.** Derive the form of \( p(x_1|x_2) \) from first principles. **HINT:** \( p(x_1|x_2) \propto p(x_1, x_2) \) when \( x_2 \) is viewed as fixed.

10. (If there is time.) Suppose that \( X \) and \( Y \) are independent and each is uniformly distributed on \((0, 1)\). Let \( U = X + Y \) and \( V = X - Y \).

- Sketch the range (or region of non zero probability) of \((X, Y)\) and the range of \((U, V)\).
- Find the density function of \((U, V)\).
- Find the density function of \( U \).
- Find the density function of \( V \).