MML - Review for Exam 1

We will have our first exam next Friday, Jan 30. The problems on that exam oughtta look something like the problems you see here.

Generally, I will expect *solutions* to the problems, as opposed to just answers. So, for example, if the answer to an optimization problem is $y = 5$, then the solution will consist of a clear explanation with correctly written supporting computations indicating *why* the answer is $y = 5$.

The problems

1. Figure 1 shows a scatter plot built from data for 343 penguins of three different species. Note that one point near the top is somewhat larger than the others and is of unknown species. Let's suppose we apply a K-Nearest Neighbors algorithm to classify the species of that lone unknown penguin.

    a. Describe this procedure. That is, how does the KNN algorithm work?
    b. What would the result be for that unknown penguin if
        i. $k = 1$?
        ii. $k = 2$?

2. Figure 2 shows a contour plot of $f(x, y) = x^2 + xy + 3x + y^2 + 2y$.

    a) Identify any max, min, or saddle points that you see on the graph.
    b) Note the grid of 9 points on figure 2. For each of those points, draw the corresponding gradient vector emanating from that point. Be sure to pay careful attention to the direction and relative magnitude of those vectors.

3. Suppose we wish to fit a least-squares regression line using a function of the form $f(x) = ax + b$ to the data

$$[[-1,-1],[0,2],[1,1]]$$

    a) Write down the function $E(a, b)$ that expresses the total squared error of the approximation.

b) Compute the two partial derivatives $\partial E/\partial a$ and $\partial E/\partial b$ of your error function.
c) Set those partial derivatives each to zero and solve the resulting system.
d) Interpret your solution to part (c) to obtain the least squares regression line.

4. Write down precise, mathematical definitions of each of the following:

   a. Critical point of a bivariate function
   b. Homogeneous system of equations
   c. Non-singular matrix
   d. Matrix transpose

5. Provide examples of linear systems of three equations in three unknowns satisfying each of the following situations:

   a) The system has no solutions,
   b) The system has infinitely many solutions, and
   c) The system has exactly one solution.

6. Let M denote the matrix
$$M = \begin{bmatrix} 2 & 1 & 5 & 10 \\ 1 & -3 & -1 & -2 \\ 4 & -2 & 6 & 12 \end{bmatrix}$$

   a) Place M into reduced row echelon form.
   b) Assuming that M is an augmented matrix representing a 3×3 system, use your reduced row echelon form to solve that system. If the solution set is infinite, you should parameterize that solution set.

7. The matrix A and its reduced row echelon form R are shown below.

   a) Explain why you can tell immediately from R that A is a singular matrix.
   b) Which columns of A are the pivot columns?
   c) Assuming that A is the coefficient matrix of a homogeneous system, find the solution set of that system.

$$A = \begin{bmatrix} -2 & 0 & 2 & 4 \\ 2 & -1 & -3 & -7 \\ -1 & -2 & -1 & -4 \\ 1 & 1 & 0 & 1 \end{bmatrix} \qquad R = \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
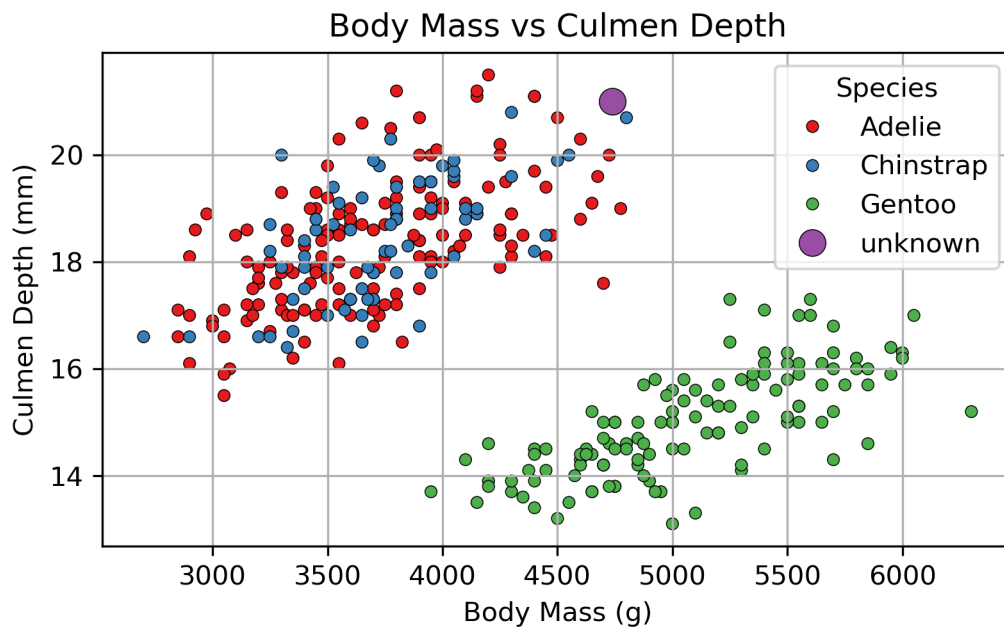
Figures

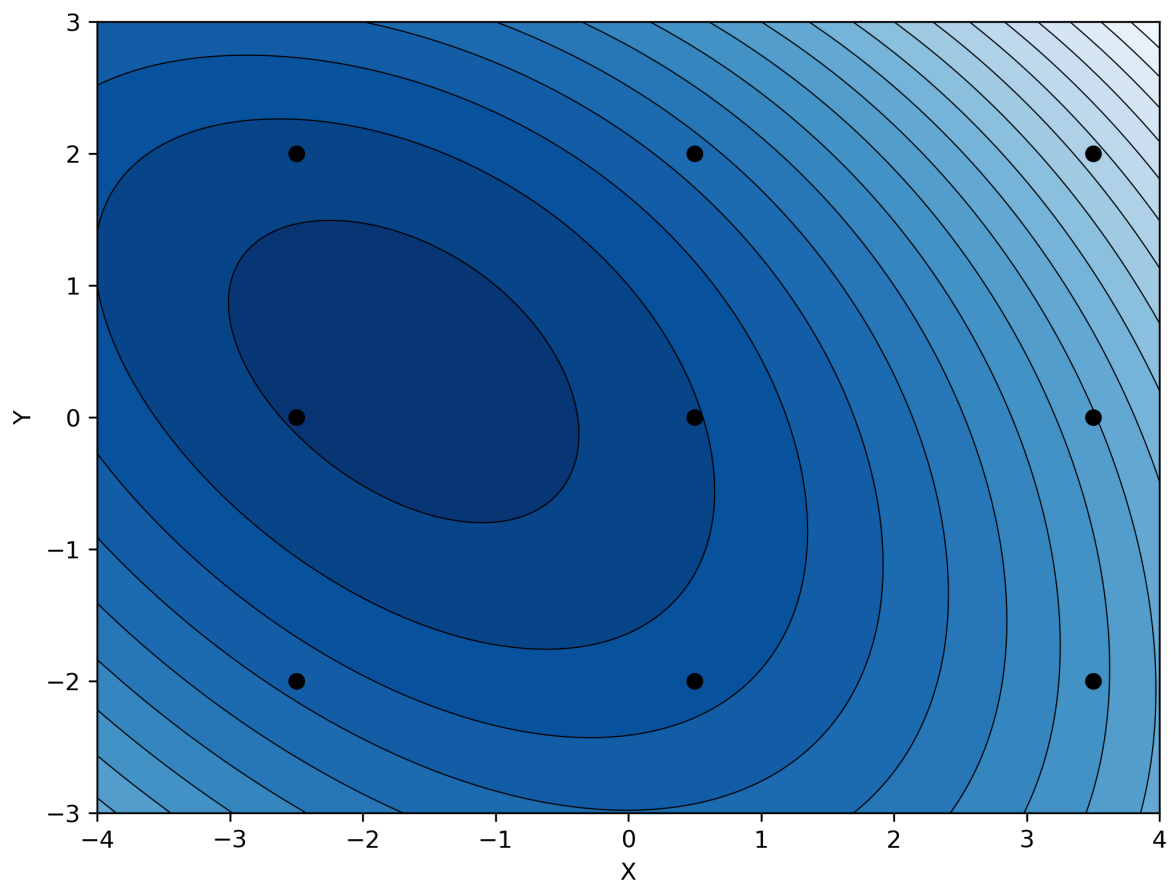Figure 1: A scatter plot off body mass vs culmen depth for penguins.

Figure 2: A contour plot of $f(x, y) = x^2 + xy + 3x + y^2 + 2y$.