CS/ECE/ME 532

Homework 3: least squares and quadratic forms

due: Sunday October 2, 2016

1. Products of PSDs. Suppose $P \succeq 0$ and $Q \succeq 0$ are (symmetric) positive semidefinite $n \times n$ matrices.

a) Prove that $PQP \succeq 0$.

b) Prove that $P^k \succeq 0$ for any $k = 1, 2, \ldots$

2. Simple least squares. Consider the following matrix and vector:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

a) Find the solution $\widehat{x}$ to $\min_x \|Ax - b\|^2$.

b) Make a sketch of the geometry of this particular problem in $\mathbb{R}^3$, showing the columns of A, the plane they span, the target vector b, the residual vector and the solution $\widehat{b} = A\widehat{x}$.

3. Tikhonov regularization. Sometimes we have competing objectives. For example, we want to find an x that minimizes $\|b - Ax\|^2$ (least-squares), but we also want the solution x to have a small norm. One way to achieve a compromise is to solve the following problem:

$$\underset{x}{\text{minimize}} \quad \|b - Ax\|^2 + \lambda \|x\|^2 \tag{1}$$

where $\lambda > 0$ is a parameter we choose that determines the relative weight we want to assign to each objective. This is called *Tikhonov regularization* (also known as L_2 regularization).

a) The optimization problem (1) has its own "normal equations" similar to those we derived for the standard least squares problem. Find them.

Hint: one approach is to reformulate (1) as a modified least squares problem with different "A" and "b" matrices. Another approach is to use the vector derivative method seen in class.

b) Suppose that $A \in \mathbb{R}^{m \times n}$ with $m < n$. Is there a unique least squares solution? Is there a unique solution to (1) ? Explain your answers.

4. Polynomial fitting. Suppose we observe pairs of points (a_i, b_i), $i = 1, \ldots, m$. Imagine these points are measurements from a scientific experiment. The variables a_i are the experimental conditions and the b_i correspond to the measured response in each condition. Suppose we wish to fit a degree $d < m$ polynomial to these data. In other words, we want to find the coefficients of a degree d polynomial p so that $p(a_i) \approx b_i$ for $i = 1, 2, \ldots, m$. We will set this up as a least-squares problem.

a) Suppose p is a degree d polynomial. Write the general expression for $p(a_i) = b$. Then, express the $i = 1, \ldots, m$ equations as a system in matrix form $Ax = b$. Specifically, what is the form/structure of b in terms of the given a_i.

b) Write a Matlab or Python script to find the least-squares fit to the $m = 30$ data points in `polydata.csv`. Plot the points and the polynomial fits for $d = 1, 2, 3$.

5. Calorie prediction for cereal, revisited. Recall the cereal calorie prediction problem discussed in class. The data matrix for this problem is

$$A = \begin{bmatrix} 25 & 0 & 1 \\ 20 & 1 & 2 \\ 40 & 1 & 6 \end{bmatrix}$$

Each row contains the grams/serving of carbohydrates, fat, and protein, and each row corresponds to a different cereal (*Frosted Flakes, Froot Loops, Grape-Nuts*). The total calories for each cereal are

$$b = \begin{bmatrix} 110 \\ 110 \\ 210 \end{bmatrix}$$

a) Write a short program (e.g., in Matlab or Python) that solves the system of equations $Ax = b$. Recall the solution b gives the calories/gram of carbohydrate, fat, or protein. Verify that the solution you find is the same as the solution we found in class.

b) The solution does not agree with the known calories/gram, which are 4 for carbs, 9 for fat and 4 for protein. We suspect this may be due to rounding the grams to integers, especially for the grams of fat. Assuming the true value for calories/gram is

$$x^\star = \begin{bmatrix} 4 \\ 9 \\ 4 \end{bmatrix}$$

and that the total calories, grams of carbs, and grams of protein *are correctly reported above*, determine the "correct" grams of fat in each cereal.

c) Now suppose that we predict total calories using a more refined breakdown of carbohydrates, into total carbohydrates, complex carbohydrates and sugars (simple carbs). So now we will have 5 features to predict calories (the three carb features + fat and protein). So let's suppose we measure the grams of these features in 5 different cereals to obtain this data matrix

$$A = \begin{bmatrix} 25 & 15 & 10 & 0 & 1 \\ 20 & 12 & 8 & 1 & 2 \\ 40 & 30 & 10 & 1 & 6 \\ 30 & 15 & 15 & 0 & 3 \\ 35 & 20 & 15 & 2 & 4 \end{bmatrix}$$

and the total calories in each cereal

$$b = \begin{bmatrix} 104 \\ 97 \\ 193 \\ 132 \\ 174 \end{bmatrix}$$

Can you solve $Ax = b$? Carefully examine the situation in this case. Is there a solution that agrees with the true calories/gram?