Further Mathematical Methods (Linear Algebra) 2002

Problem Sheet 8

(To be discussed in week 9 classes. Please submit answers to the asterisked questions only.)

In this sheet, we look at orthogonal projections and the analysis of data sets using least squares fits.

1. * Let X be the subspace of $\mathbb{R}^3$ spanned by the vectors $[1, 2, 3]^t$ and $[1, 1, -1]^t$. Find a matrix P such that $\mathsf{P}\mathbf{x}$ is the orthogonal projection of $\mathbf{x} \in \mathbb{R}^3$ onto X.

2. * Assume that we have a real $n \times n$ matrix A which is orthogonally diagonalisable and whose eigenvalues and eigenvectors are all real. Derive a spectral decomposition of A, i.e.

$$\mathsf{A} = \sum_{i=1}^{n} \lambda_i \mathsf{E}_i,$$

where $\mathsf{E}_i = \mathbf{x}_i \mathbf{x}_i^t$ and the set of vectors $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ form an orthonormal set of eigenvectors corresponding to the eigenvalues λ_i. Prove that

$$\mathsf{E}_i \mathsf{E}_j = \begin{cases} \mathsf{E}_i & \text{if } i = j \\ \mathsf{0} & \text{if } i \neq j \end{cases}$$

where $\mathsf{0}$ is the $n \times n$ zero matrix. Hence show that the $n \times n$ matrix E_i represents an orthogonal projection. Further, establish that for any vector $\mathbf{x} \in \mathbb{R}^n$ written in terms of the vectors in S, say $\mathbf{x} = \sum_{j=1}^{n} \alpha_j \mathbf{x}_j$, we have $\mathsf{E}_i \mathbf{x} = \alpha_i \mathbf{x}_i$, i.e. E_i represents an orthogonal projection onto $\text{Lin}\{\mathbf{x}_i\}$. Consequently, find $\mathsf{A}\mathbf{x}$ in terms of the vectors in S and describe the effect of the linear transformation represented by A on $\mathbf{x}$. (This result can be extended to the complex case, but we shall not do this here.)

3. Quantities x and y are known to be related by a rule of the form $y = ax + b$ for some constants a and b. An experiment is set up so that for various values of x, the value of y can be measured. At the end of the experiment, the following data has been collected:

x	1	2	3	4
y	5	3	2	1

Find the least squares estimate of a and b.

4. * Quantities x and y are known to be related by a rule of the form

$$y = \frac{m}{x} + c,$$

for some constants m and c. Quantity y is measured for various values of x, resulting in the following set of readings:

x	1/5	1/4	1/3	1/2	1
y	4	3	2	2	1

Find the least squares estimate of m and c. Further, explain why it would be wrong to suppose that this was equivalent to the problem of fitting a curve of the form

$$z = xy = cx + m,$$

through the data points (xy, x).

5. Suppose that we want to find the least squares fit of the form $y = m^* x + c^*$ through the data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. By using the matrix approach developed in the lectures, show that the parameters m^* and c^* are given by

$$m^* = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{n \sum x_i^2 - (\sum x_i)^2} \quad \text{and} \quad c^* = \frac{(\sum y_i)(\sum x_i^2) - (\sum x_i)(\sum x_i y_i)}{n \sum x_i^2 - (\sum x_i)^2},$$

where all summations in these formulae run from $i = 1$ to n.

Other Problems. (These are *not* compulsory, they are *not* to be handed in, and they will *not* be covered in classes.)

Here are some more questions on the analysis of data sets using least squares fits.

6. An input variable θ and a response variable y are related by a law of the form

$$y = a + b\cos^2\theta,$$

where a and b are constants. The observation of y is subject to error; use the following data to estimate a and b using the method of least squares.

θ	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$
y	4.1	3.4	2.7	2.1	1.6

7. Find the least squares solution to the system given by $x_1 = a_1, a_2, \ldots, a_n$. Indicate what you have found.

8. Show that:

> If $\mathsf{A}\mathbf{x} = \mathbf{b}$ is consistent, then every solution of $\mathsf{A}^t\mathsf{A}\mathbf{x} = \mathsf{A}^t\mathbf{b}$ also solves the original matrix equation.

Explain the importance of this result in the context of least squares analyses.

Harder Problems. (These are *not* compulsory, they are *not* to be handed in, and they will *not* be covered in classes.)

Here are some more results on orthogonal projections and direct sums for you to prove.

9. Let X be a subspace of the vector space V and let P denote the orthogonal projection onto X. Show that:

> If $\mathsf{Q} = \mathsf{I} - \mathsf{P}$, then for any subspace Y of V,
>
> $$\text{Lin}(X \cup Y) = \text{Lin}(X \cup Q(Y)),$$
>
> where $Q(Y)$ is given by
>
> $$Q(Y) = \{\mathsf{Q}\mathbf{y} \,|\, \mathbf{y} \in Y\},$$
>
> (i.e. it is the set of vectors which is found by multiplying each of the vectors in Y by Q.)

Interpret this result geometrically in the case where X and Y are one-dimensional subspaces of $\mathbb{R}^3$.

10. Let L and M be subspaces of the vector space V. Show that: $V = L \oplus M$ iff

$$L \cap M = \{\mathbf{0}\} \text{ and } \text{Lin}(L \cup M) = V.$$

(Recall that, by Theorem 2.4, if S is a set of vectors, then $\text{Lin}(S)$ is the smallest subspace that contains all of the vectors in S.) Assuming that L and M are such that $L \oplus M = V$, prove that

- $L^\perp \cap M^\perp = \{\mathbf{0}\}$
- $[\text{Lin}(L^\perp \cup M^\perp)]^\perp = \{\mathbf{0}\}$

Hence, deduce that $L^\perp \oplus M^\perp = V$.