THE UNIVERSITY OF SYDNEY

MATH2902 Vector Spaces

(http://www.maths.usyd.edu.au/u/UG/IM/MATH2902/)

Semester1, 2001 — Lecturer: R. Howlett

Assignment 2

1. Let $y = a + bx$ be the equation of the least squares line of best fit for the following points (x_i, y_i):

$$(0,1), \qquad (1,2), \qquad (2,2), \qquad (3,5), \qquad (4,5).$$

Calculate a and b.

Solution.

If all the points were on the line we would have $y_i = a + bx_i$ for each i, and so

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 5 \\ 5 \end{pmatrix}.$$

If this were true, the 5-component column vector on the right-hand side would be in the column space of the 5×2 matrix A on the left-hand side. In fact, it is not possible to find a and b to solve the equations exactly; instead a and b must be chosen so that

$$a \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

is as close as possible to ${}^t(1,2,2,5,5)$. Thus $a\,{}^t(1,1,1,1) + b\,{}^t(2,3,4,5)$ must be the projection of ${}^t(1,2,2,5,5)$ onto the column space of A. According to the theory described in the lectures, to find a and b we must solve the system of linear equations

$${}^tA\,A \begin{pmatrix} a \\ b \end{pmatrix} = {}^tA \begin{pmatrix} 1 \\ 2 \\ 2 \\ 5 \\ 5 \end{pmatrix}.$$

We find that

$${}^tA\,A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 10 \\ 10 & 30 \end{pmatrix}$$

and

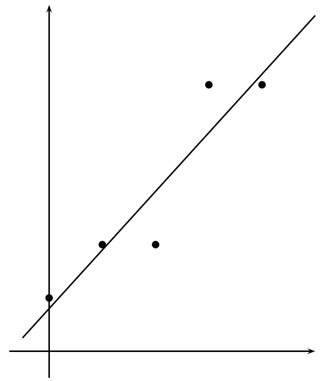
$${}^tA \begin{pmatrix} 1 \\ 2 \\ 2 \\ 5 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 2 \\ 5 \\ 5 \end{pmatrix} = \begin{pmatrix} 15 \\ 41 \end{pmatrix},$$

and so it follows that

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 5 & 10 \\ 10 & 30 \end{pmatrix}^{-1} \begin{pmatrix} 15 \\ 41 \end{pmatrix} = \frac{1}{50} \begin{pmatrix} 30 & -10 \\ -10 & 5 \end{pmatrix} \begin{pmatrix} 15 \\ 41 \end{pmatrix}$$
$$= \frac{1}{10} \begin{pmatrix} 8 \\ 11 \end{pmatrix} = \begin{pmatrix} 0.8 \\ 1.1 \end{pmatrix}$$

Thus the line of best fit is $y = 0.8 + 1.1x$.

The points and the line of best fit are shown in the diagram.



2. Let A be a square matrix which satisfies $A^2 - 3A + 2I = 0$. Prove that if λ is an eigenvalue of A then λ must be 1 or 2.

Solution.

Let v be an eigenvector corresponding to the eigenvalue λ. By definition v is nonzero, and $Av = \lambda v$. Multiplying this equation by A gives $A^2 v = \lambda Av$, and since $Av = \lambda v$ we deduce that $A^2 v = \lambda^2 v$. Now

$$(A^2 - 3A + 2I)v = A^2 v - 3Av + 2v = (\lambda^2 + 3\lambda + 2)v,$$

and since $A^2 - 3A + 2I = 0$ it follows that $(\lambda^2 + 3\lambda + 2)v = 0$. Since $v \neq 0$ this gives $\lambda^2 + 3\lambda + 2 = 0$, whence λ is 1 or 2.

3. If $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_dx^d$ is any polynomial, where the coefficients a_i are elements of a field F, and if A is any square matrix over F, we define $f(A)$ to be the matrix $a_0I + a_1A + a_2A^2 + \cdots + a_dA^d$.

(*i*) Suppose that D is an $n \times n$ diagonal matrix and $f(x) = \det(D - xI)$ its characteristic polynomial. Show that $f(D)$ is the zero matrix.

(*ii*) Use Part (*i*) to show that if A is any diagonalizable matrix and $f(x)$ its characteristic polynomial then $f(A) = 0$.

(*iii*) In fact it is true for all square matrices A that if $f(x)$ is the characteristic polynomial then $f(A) = 0$. Check this by direct calculation in the following cases:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Solution.

(*i*) Suppose that the diagonal entries of D are $d_1, d_2, \ldots, d_n$. Then

$$f(x) = \det(D - xI) = \det \begin{pmatrix} d_1 - x & 0 & 0 & \ldots & 0 \\ 0 & d_2 - x & 0 & \ldots & 0 \\ 0 & 0 & d_3 - x & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & d_n - x \end{pmatrix}$$
$$= (d_1 - x)(d_2 - x) \cdots (d_n - x)$$

and so we see that $f(d_1) = f(d_2) = \cdots = f(d_n) = 0$. Now for every integer $k \geq 0$ the matrix D^k is diagonal, and its diagonal entries are $d_1^k, d_2^k, \ldots, d_n^k$. So for any polynomial $p(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_r x^r$ we find that

$$p(D) = \alpha_0 I + \alpha_1 D + \cdots + \alpha_r D^r$$
$$= \begin{pmatrix} \alpha_0 & 0 & \ldots & 0 \\ 0 & \alpha_0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \alpha_0 \end{pmatrix} + \begin{pmatrix} \alpha_1 d_1 & 0 & \ldots & 0 \\ 0 & \alpha_1 d_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \alpha_1 d_n \end{pmatrix} + \cdots$$
$$\cdots + \begin{pmatrix} \alpha_r d_1^r & 0 & \ldots & 0 \\ 0 & \alpha_r d_2^r & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \alpha_r d_n^r \end{pmatrix}$$
$$= \begin{pmatrix} p(d_1) & 0 & \ldots & 0 \\ 0 & p(d_2) & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & p(d_n) \end{pmatrix}.$$

In particular, $f(D)$ is a diagonal matrix, and its diagonal entries are, respectively, $f(d_1), f(d_2), \ldots, f(d_n)$, which are all 0. So $f(D)$ is the zero matrix.

Alternatively, from the formula for $f(x)$ above we can see that

$$f(D) = (d_1 I - X)(d_2 I - D) \cdots (d_n I - D)$$
$$= \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 0 & d_1 - d_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & d_1 - d_n \end{pmatrix} \begin{pmatrix} d_2 - d_1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & d_2 - d_n \end{pmatrix}$$
$$\cdots \begin{pmatrix} d_n - d_1 & 0 & \ldots & 0 \\ 0 & d_n - d_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}.$$

In the i-th factor the i-th diagonal entry is zero, and so when we compute the product we find that all the diagonal entries are zero.

(*ii*) If A is diagonalizable then there exists an invertible matrix P such that $P^{-1}AP = D$, where the diagonal entries of D are the eigenvalues of A. Thus if $f(x)$ is the characteristic polynomial of A then $f(x) = (d_1 - x)(d_2 - x) \cdots (d_n - x)$, where $d_1, d_2, \ldots, d_n$ are the diagonal entries of D. In particular, $f(x)$ is also the characteristic polynomial of D. (It is true in general—this was proved in lectures—that similar matrices have the same characteristic polynomial.)

Write $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$. Since $A = PDP^{-1}$ it follows that for all positive integers r,

$$A^r = (PDP^{-1})^r = (PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1}) = PD^rP^{-1}$$

since the $P^{-1}P$'s in the middle of the expression cancel out. So

$$f(A) = a_0I + a_1A + a_2A^2 + \cdots + a_nA^n$$
$$= a_0I + a_1PDP^{-1} + a_2PD^2P^{-1} + \cdots + a_nPD^nP^{-1}$$
$$= P(a_0I + a_1D + a_2D^2 + \cdots + a_nPD^n)P^{-1} = Pf(D)P^{-1} = 0$$

since $f(D) = 0$ by Part prt (i).

(*iii*) In the second case the characteristic polynomial of A is $f(x) = (\lambda - x)^3$, and so the task is to show that $(\lambda I - A)^3$ is the zero matrix. Now $\lambda I - A = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$, and it is trivial to check that the cube of this is zero.

In the other case $f(x) = x^2 - (a + d)x + (ad - bc)$, and

$$f(A) = \begin{pmatrix} a^2 + bc & ab + bd \\ ca + dc & bc + d^2 \end{pmatrix} - (a + d) \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix}$$

which is indeed zero.