THE UNIVERSITY OF SYDNEY

MATH2008 Introduction to Modern Algebra

(http://www.maths.usyd.edu.au/u/UG/IM/MATH2008/)

Semester2, 2003 | Lecturer: R. Howlett

Tutorial 2

1. Calculate the projection of the vector $\underset{\sim}{v} = (1, 1, 0) \in (\mathbb{R}^3)'$ onto the one-dimensional space spanned by $\underset{\sim}{a} = (1, -2, 1)$. Check that if $\underset{\sim}{p}$ is this projection then $\underset{\sim}{p} - \underset{\sim}{v}$ is orthogonal to $\underset{\sim}{a}$.

Solution.

We have $\underset{\sim}{p} = \frac{(\underset{\sim}{a}\cdot\underset{\sim}{v})}{(\underset{\sim}{a}\cdot\underset{\sim}{a})}\underset{\sim}{a} = \frac{(-1)}{6}(1, -2, 1) = (-\frac{1}{6}, \frac{1}{3}, -\frac{1}{6})$; so $\underset{\sim}{p} - \underset{\sim}{v} = (-\frac{7}{6}, -\frac{2}{3}, -\frac{1}{6})$, and consequently $(\underset{\sim}{p} - \underset{\sim}{v}) \cdot \underset{\sim}{a} = -\frac{7}{6} + \frac{4}{3} - \frac{1}{6} = 0$.

2. Calculate the projection $\underset{\sim}{p}$ of $\underset{\sim}{v} = (1, 2, 3, 4)^T$ onto the subspace of $\mathbb{R}^4$ spanned by the following three vectors:

$$\underset{\sim}{a}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \qquad \underset{\sim}{a}_2 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \qquad \underset{\sim}{a}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Check that $\underset{\sim}{v} - \underset{\sim}{p}$ is orthogonal to each of $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$ and $\underset{\sim}{a}_3$.

Solution.

Let $W = \operatorname{Span}(\underset{\sim}{a}_1, \underset{\sim}{a}_2, \underset{\sim}{a}_3)$. To apply the formulas given in lectures, we need a basis for the space W. In fact the vectors $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$, $\underset{\sim}{a}_3$ are linearly independent, and hence form a basis. To check this, suppose that

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The first equation gives $\lambda_1 = \lambda_2$ and the third gives $\lambda_1 = \lambda_3$. So λ_1, λ_2 and λ_3 are all equal; let us call their common value λ. The fourth equation now gives $2\lambda = 0$; so all the coefficients λ_i must be zero, and this proves that the $\underset{\sim}{a}_i$ are linearly independent.

Let A be the 4×3 matrix whose columns are $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$, $\underset{\sim}{a}_3$. Then

$$A^T A = \begin{pmatrix} 1 & 0 & -1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}$$

and

$$A^T \underset{\sim}{v} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \\ 7 \end{pmatrix}.$$

The projection is given by $\underset{\sim}{p} = A\underset{\sim}{x}$, where $A^T A\underset{\sim}{x} = A^T\underset{\sim}{v}$; so we start by solving this system.

$$\left(\begin{array}{ccc|c} 2 & -1 & -1 & -2 \\ -1 & 2 & 1 & 3 \\ -1 & 1 & 2 & 7 \end{array}\right) \xrightarrow[\substack{R_2:=R_2+2R_1 \\ R_3:=R_3-R_1}]{R_1\leftrightarrow R_2} \left(\begin{array}{ccc|c} -1 & 2 & 1 & 3 \\ 0 & 3 & 1 & 4 \\ 0 & -1 & 1 & 4 \end{array}\right)$$

$$\xrightarrow[R_3:=R_3+3R_2]{R_2\leftrightarrow R_3} \left(\begin{array}{ccc|c} -1 & 2 & 1 & 3 \\ 0 & -1 & 1 & 4 \\ 0 & 0 & 4 & 16 \end{array}\right)$$

and now back substitution gives $x_3 = 4$, $x_2 = 0$ and $x_1 = 1$ (where x_1, x_2 and x_3 are the entries of $\underset{\sim}{x}$. So

$$\underset{\sim}{p} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 3 \\ 4 \end{pmatrix}.$$

So $\underset{\sim}{v} - \underset{\sim}{p} = (0, 2, 0, 0)^T$. If we compute the dot product of this with an arbitrary column vector $\underset{\sim}{a} \in \mathbb{R}^4$ then it is clear that the answer will be zero if the second component of $\underset{\sim}{a}$ is zero, since

$$\begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} = 0\alpha + 2\beta + 0\gamma + 0\delta = 2\beta = 0 \quad \text{if and only if } \beta = 0.$$

Since $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$ and $\underset{\sim}{a}_3$ all satisfy this condition, it is true that $\underset{\sim}{v} - \underset{\sim}{p}$ is orthogonal to $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$ and $\underset{\sim}{a}_3$.

3. Let $\{\underset{\sim}{a}_1, \underset{\sim}{a}_2, \ldots, \underset{\sim}{a}_k\}$ be a basis for a subspace W of $\mathbb{R}^n$, and let $\underset{\sim}{v}$ be any vector in $\mathbb{R}^n$. Show that $\underset{\sim}{v}$ is orthogonal to each of $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$, …, $\underset{\sim}{a}_k$ if and only if $\underset{\sim}{v}$ is orthogonal to every vector in W.

Solution.

Since $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$, …, $\underset{\sim}{a}_k$ are vectors in W, if $\underset{\sim}{v}$ is orthogonal to every vector in W then it is certainly orthogonal to $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$, …, $\underset{\sim}{a}_k$. Conversely, suppose that $\underset{\sim}{v}$ is orthogonal to each of $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$, …, $\underset{\sim}{a}_k$ and let $\underset{\sim}{w}$ be an arbitrary vector in W. Then $\underset{\sim}{w} = \lambda_1\underset{\sim}{a}_1 + \lambda_2\underset{\sim}{a}_2 + \cdots + \lambda_n\underset{\sim}{a}_k$ for some scalars λ_i (since the $\underset{\sim}{a}_i$ span W), and so $\underset{\sim}{w} \cdot \underset{\sim}{v} = \lambda_1\underset{\sim}{a}_1 \cdot \underset{\sim}{v} + \lambda_2\underset{\sim}{a}_2 \cdot \underset{\sim}{v} + \cdots + \lambda_n\underset{\sim}{a}_n \cdot \underset{\sim}{v} = 0$, because $\underset{\sim}{a}_i \cdot \underset{\sim}{v} = 0$ for all i. Since $\underset{\sim}{w}$ was chosen as an arbitrary element of W, this shows that $\underset{\sim}{v}$ is orthogonal to all elements of W, as required.

4. Find the least squares line of best fit for the four points $(0,1)$, $(2,0)$, $(3,1)$ and $(3,2)$.

Solution.

Let A be the 4×2 matrix whose 1st column consists of 1's and whose 2nd column consists of the x-coordinates of the data points. We must solve $A^T A \binom{a}{b} = A^T \underset{\sim}{y}$, where the entries of $\underset{\sim}{y}$ are the y-coordinates of the data points. We have

$$A^T A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 8 \\ 8 & 22 \end{pmatrix},$$

and

$$A^T \underset{\sim}{y} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}.$$

The equation to be solved is therefore

$$\begin{pmatrix} 4 & 8 \\ 8 & 22 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 4 \\ 9 \end{pmatrix},$$

and the solution is

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{24} \begin{pmatrix} 22 & -8 \\ -8 & 4 \end{pmatrix} \begin{pmatrix} 4 \\ 9 \end{pmatrix} = \begin{pmatrix} 2/3 \\ 1/6 \end{pmatrix}.$$

Thus the line of best fit has equation $y = \frac{2}{3} + \frac{1}{6}x$.

5. For each collection of data points below, find the parabola $y = a + bx + cx^2$ of best fit.

(i) $(-1,0)$, $(0,0)$, $(0,1)$, $(1,2)$.
(ii) $(-1,0)$, $(0,0)$, $(0,1)$, $(1,1)$.
(iii) $(-1,0)$, $(0,0)$, $(0,1)$, $(1,0)$.

Solution.

The 1st column of A should be all 1's, the 2nd should consist of the x-coordintates of the data points, the third should consist of the squares of these x-coordinates. We see that for all parts of this question, A is the same, namely

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

We must solve the equation $A^T A \underset{\sim}{a} = A^T \underset{\sim}{y}$ for $\underset{\sim}{a}$, where the entries of $\underset{\sim}{y}$ are the y-coordinates of the data points. Observe that

$$A^T A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

In Part (i), $\underset{\sim}{y} = (0,0,1,2)^T$ and so

$$A^T \underset{\sim}{y} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix}.$$

Solving

$$\begin{pmatrix} 4 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix}$$

gives $a = 1/2$, $b = 1$ and $c = 1/2$. So the parabola of best fit is $y = \frac{1}{2} + x + \frac{1}{2}x^2$.
For Part (ii) we have $\underset{\sim}{y} = (0,0,1,1)^T$, giving $A^T \underset{\sim}{y} = (2,1,1)^T$. Solving $A^T A \underset{\sim}{a} = A^T \underset{\sim}{y}$ gives $(a,b,c) = (1/2, 1/2, 0)$. So the "parabola" is actually the straight line $y = \frac{1}{2} + \frac{1}{2}x$.
For Part (iii), $\underset{\sim}{y} = (0,0,1,1)^T$ and $A^T \underset{\sim}{y} = (1,0,0)^T$. Solving $A^T A \underset{\sim}{a} = A^T \underset{\sim}{y}$ gives $(a,b,c) = (1/2, 0, -1/2)$. So the parabola of best fit is $y = \frac{1}{2} - \frac{1}{2}x^2$.
(For each part of the question it would be a good idea to plot the parabola and the four given points on graph paper to see if the parabola of best fit is reasonable.)

6. Find the cubic curve $y = a + bx + cx^2 + dx^3$ that best fits the following data points: $(-1,-14)$, $(0,-5)$, $(1,-4)$, $(2,1)$, $(3,22)$.

Solution.

Let A be the 5×4 matrix whose 1st column consists of 1's, 2nd column the x-coordinates of the data points, 3rd column the squares of these x-coordinates, 4th column the cubes of the x-coordinates, and solve $A^T A \underset{\sim}{a} = A^T \underset{\sim}{y}$ for $\underset{\sim}{a}$, the entries of $\underset{\sim}{y}$ being the y-coordinates of the data points. The answer is $\underset{\sim}{a} = (-5, 3, -4, 2)^T$, and so the cubic of best fit is $y = -5 + 3x - 4x^2 + 2x^3$. (It actually goes through all of the data points.)

7. If $A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_k \end{pmatrix}$, show that $A^T A = \begin{pmatrix} k & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix}$.

Solution.

The (i,j) entry of $A^T A$ is the dot product of the i-th and j-th columns of A. So the $(1,1)$ entry is $\sum_{i=1}^k 1^2 = k$, the (1,2) and (2,1) entries are both $\sum_{i=1}^k 1x_i$, and the (2,2) entry is $\sum_{i=1}^k x_i^2$.