NATIONAL UNIVERSITY OF SINGAPORE
Department of Mathematics

AY2021, Semester 1 MA1508E Linear Algebra for Engineering Tutorial 9

Solutions

1. Apply Gram-Schmidt Process to convert

(a) $\left\{\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\-1\\-1\end{pmatrix}, \begin{pmatrix}1\\2\\0\\1\end{pmatrix}\right\}$ into an orthonormal basis for $\mathbb{R}^4$.

Let $\mathbf{u}_1 = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$, $\mathbf{u}_2 = \begin{pmatrix}1\\-1\\1\\0\end{pmatrix}$, $\mathbf{u}_3 = \begin{pmatrix}1\\1\\-1\\-1\end{pmatrix}$, and $\mathbf{u}_4 = \begin{pmatrix}1\\2\\0\\1\end{pmatrix}$.

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1 \\
\mathbf{v}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_1\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 = \frac{1}{4}\begin{pmatrix}3\\-5\\3\\-1\end{pmatrix}, \\
\mathbf{v}_3 &= \mathbf{u}_3 - \frac{\mathbf{u}_3\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{u}_3\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2}\mathbf{v}_2 = \frac{2}{11}\begin{pmatrix}7\\3\\-4\\-6\end{pmatrix}, \\
\mathbf{v}_4 &= \mathbf{u}_4 - \frac{\mathbf{u}_4\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{u}_4\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2}\mathbf{v}_2 - \frac{\mathbf{u}_4\cdot\mathbf{v}_3}{\mathbf{v}_3\cdot\mathbf{v}_3}\mathbf{v}_3 = \frac{1}{10}\begin{pmatrix}1\\-1\\-2\\2\end{pmatrix}.
\end{aligned}$$

For easy computation, we can let each $\mathbf{v}_i$ to be the vector without the fraction part. Then by normalizing, we obtain an orthonormal basis

$$\left\{\frac{1}{2}\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \frac{1}{2\sqrt{11}}\begin{pmatrix}3\\-5\\3\\-1\end{pmatrix}, \frac{1}{\sqrt{110}}\begin{pmatrix}7\\3\\-4\\-6\end{pmatrix}, \frac{1}{\sqrt{10}}\begin{pmatrix}1\\-1\\-2\\2\end{pmatrix}\right\}$$

(b) $\left\{\begin{pmatrix}1\\2\\2\\1\end{pmatrix}, \begin{pmatrix}1\\2\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\2\\1\end{pmatrix}\right\}$ into an orthonormal set. Is the set obtained an orthonormal basis? Why?

Let $\mathbf{u}_1 = \begin{pmatrix}1\\2\\2\\1\end{pmatrix}$, $\mathbf{u}_2 = \begin{pmatrix}1\\2\\1\\0\end{pmatrix}$, $\mathbf{u}_3 = \begin{pmatrix}1\\0\\1\\0\end{pmatrix}$, and $\mathbf{u}_4 = \begin{pmatrix}1\\0\\2\\1\end{pmatrix}$.

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1 \\
\mathbf{v}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_1 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \frac{1}{10}\begin{pmatrix}3\\6\\-4\\-7\end{pmatrix}, \\
\mathbf{v}_3 &= \mathbf{u}_3 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \frac{2}{11}\begin{pmatrix}4\\-3\\2\\-2\end{pmatrix}, \\
\mathbf{v}_4 &= \mathbf{u}_4 - \frac{\mathbf{u}_4 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{u}_4 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \frac{\mathbf{u}_4 \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3}\mathbf{v}_3 = \mathbf{0}.
\end{aligned}$$

The orthonormal set obtained is

$$\left\{ \frac{1}{\sqrt{10}}\begin{pmatrix}1\\2\\2\\1\end{pmatrix}, \frac{1}{\sqrt{110}}\begin{pmatrix}3\\6\\-4\\-7\end{pmatrix}, \frac{1}{\sqrt{33}}\begin{pmatrix}4\\-3\\2\\-2\end{pmatrix} \right\}$$

It is not a basis since it only contains 3 vectors. Since $\mathbf{v}_4 = 0$, it means that $\mathbf{u}_4$ is contained in span$\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ =span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ since $\mathbf{u}_4$ minus its projection onto span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is the zero vector.

2. Let $\mathbf{A} = \begin{pmatrix}0 & 1 & 1 & 0\\1 & -1 & 1 & -1\\1 & 0 & 1 & 0\\1 & 1 & 1 & 1\end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix}6\\3\\-1\\1\end{pmatrix}$.

(a) Show that the linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ is inconsistent.

$$\left(\begin{array}{cccc|c}0 & 1 & 1 & 0 & 6\\1 & -1 & 1 & -1 & 3\\1 & 0 & 1 & 0 & -1\\1 & 1 & 1 & 1 & 1\end{array}\right) \longrightarrow \left(\begin{array}{cccc|c}1 & 0 & 0 & 1 & 0\\0 & 1 & 0 & 1 & 0\\0 & 0 & 1 & -1 & 0\\0 & 0 & 0 & 0 & 1\end{array}\right).$$

So the linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ is inconsistent.

(b) Find a least squares solution to the system. Is the solution unique? Why?

To find a least squares solution, we compute $\mathbf{A}^T\mathbf{A}$, $\mathbf{A}^T\mathbf{b}$ and solve the system $(\mathbf{A}^T\mathbf{A})\mathbf{x} = \mathbf{A}^T\mathbf{b}$.

$$\left(\mathbf{A}^T\mathbf{A} \mid \mathbf{A}^T\mathbf{b} \right) = \left(\begin{array}{cccc|c}3 & 0 & 3 & 0 & 3\\0 & 3 & 1 & 2 & 4\\3 & 1 & 4 & 0 & 9\\0 & 2 & 0 & 2 & -2\end{array}\right) \longrightarrow \left(\begin{array}{cccc|c}1 & 0 & 0 & 1 & -6\\0 & 1 & 0 & 1 & -1\\0 & 0 & 1 & -1 & 7\\0 & 0 & 0 & 0 & 0\end{array}\right).$$

A general solution to $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$ is

$$\begin{cases} x_1 &= -6 - s \\ x_2 &= -1 - s \\ x_3 &= 7 + s \\ x_4 &= s \end{cases}$$

So a least squares solution can be (when $s = 0$) $x_1 = -6, x_2 = -1, x_3 = 7, x_4 = 0$, that is, $\mathbf{v} = \begin{pmatrix} -6 \\ -1 \\ 7 \\ 0 \end{pmatrix}$. There are infinitely many least squares solutions. Since the matrix $\mathbf{A}$ is singular, the columns are linearly dependent, and thus $\mathbf{A}^T\mathbf{A}$ is not invertible. Hence, the system $\mathbf{A}^T\mathbf{A}\mathbf{v} = \mathbf{A}^T\mathbf{b}$ must have infinitely many solution.

(c) Use your answer in (b), compute the projection of $\mathbf{b}$ onto the column space of $\mathbf{A}$. Is the solution unique? Why?

The projection of $\mathbf{b}$ onto the column space of $\mathbf{A}$ is given by $\mathbf{A}\mathbf{v}$, which is

$$\mathbf{A}\mathbf{v} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -6 \\ -1 \\ 7 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 1 \\ 0 \end{pmatrix}.$$

This is unique since projection is unique. In fact, we can check that indeed

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -6 - s \\ -1 - s \\ 7 + s \\ s \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 1 \\ 0 \end{pmatrix}$$

for any choice of s.

We will breifly mention why the projection $\mathbf{A}\mathbf{v}$ is unique. We need a fact.

Fact. $Null(\mathbf{A}^T\mathbf{A}) = Null(\mathbf{A})$.

Recall also that every solution of $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$ is of the form $\mathbf{x}_p + \mathbf{w}$ for $\mathbf{w}$ in $Null(\mathbf{A}^T\mathbf{A})$ and $\mathbf{x}_p$ a particular solution. Using the fact above, the solution is of the form $\mathbf{x}_p + \mathbf{w}$ for $\mathbf{w}$ in $Null(\mathbf{A})$. Hence, the projection is

$$\mathbf{A}(\mathbf{x}_p + \mathbf{w}) = \mathbf{A}\mathbf{x}_p + \mathbf{A}\mathbf{w} = \mathbf{b}_{proj} + \mathbf{0} = \mathbf{b}_{proj},$$

where $\mathbf{b}_{proj}$ is the projection of $\mathbf{b}$ onto the column space of $\mathbf{A}$.

3. **(MATLAB)** Let W be the nullspace of $\mathbf{A} = \begin{pmatrix} 1 & 1 & 2 & 1 & 0 \\ 1 & 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{pmatrix}$.

(a) Find a basis S for W.

$\mathbf{A} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix}$. So, a basis of $\mathbf{A}$ is

$$S = \left\{ \begin{pmatrix} -1 \\ -2 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

(b) Find the projection of the i-th vector in the standard basis $\mathbf{e}_i$ of $\mathbb{R}^5$ onto W for $i = 1, ..., 5$. (**Hint:** Let $\mathbf{N}$ be a matrix whose columns are vectors in S. Consider the equation $\mathbf{N}^T\mathbf{N} = \mathbf{N}^T\mathbf{b}$ for some $\mathbf{b}$.)

One might consider applying Gram-Schmidt Process to the basis above and use that to compute the orthogonal projection. The other way is to let $\mathbf{N}$ be the matrix whose columns consist of the vectors in the basis, solve for $\mathbf{N}^T\mathbf{N}\mathbf{x} = \mathbf{N}^T\mathbf{e}_i$, then the projection will be $\mathbf{N}\mathbf{v}$ for any v a solution of $\mathbf{N}^T\mathbf{N}\mathbf{x} = \mathbf{N}^T\mathbf{e}_i$. In this case, the second method is significantly easier to compute.

Let $\mathbf{N} = \begin{pmatrix} -1 & 1 \\ -2 & -1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$.

```
>> N=[-1 -2 1 1 0;1 -1 0 0 1]';
```

`>> N'*N`, ans$=\begin{pmatrix} 7 & 1 \\ 1 & 3 \end{pmatrix}$.

This matrix is invertible. Now $\mathbf{N}^T\mathbf{e}_i$ is just the i-th column of $\mathbf{N}^T$, which is the transpose of the i-th row of $\mathbf{N}$. To retrieve the i-th row, we simply type `N(i,:)`. So the solution of $\mathbf{N}^T\mathbf{N}\mathbf{x} = \mathbf{N}^T\mathbf{e}_i$ will be

$$\mathbf{x} = (\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\mathbf{e}_i,$$

and the projection of $\mathbf{e}_i$ onto W will thus be

$$\mathbf{N}(\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\mathbf{e}_i.$$

Observe that $\mathbf{N}(\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\mathbf{e}_i$ is just the i-th column of $\mathbf{N}(\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T$. We shall now compute them in MATLAB.

```
>> N*inv(N'*N)*N'
```

$$\text{ans} = \begin{pmatrix} 3/5 & 0 & -1/5 & -1/5 & 2/5 \\ 0 & 3/4 & -1/4 & -1/4 & -1/4 \\ -1/5 & -1/4 & 3/20 & 3/20 & -1/20 \\ -1/5 & -1/4 & 3/20 & 3/20 & -1/20 \\ 2/5 & -1/4 & -1/20 & -1/20 & 7/20 \end{pmatrix}.$$

So

$$\mathbf{e}_{1\,proj} = \frac{1}{5}\begin{pmatrix}3\\0\\-1\\-1\\2\end{pmatrix}, \mathbf{e}_{2\,proj} = \frac{1}{4}\begin{pmatrix}0\\3\\-1\\-1\\-1\end{pmatrix}, \mathbf{e}_{3\,proj} = \frac{1}{20}\begin{pmatrix}-4\\-5\\3\\3\\-1\end{pmatrix},$$

$$\mathbf{e}_{4\,proj} = \frac{1}{20}\begin{pmatrix}-4\\-5\\3\\3\\-1\end{pmatrix}, \mathbf{e}_{5\,proj} = \frac{1}{20}\begin{pmatrix}8\\-5\\-1\\-1\\7\end{pmatrix}.$$

(c) Find the projection of $\mathbf{x} = \begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix}$ onto W.

The projection of $\mathbf{x}$ onto W will be

$$\mathbf{N}(\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\mathbf{x} = \begin{pmatrix}\frac{3}{5}x_1 - \frac{1}{5}x_3 - \frac{1}{5}x_4 + \frac{2}{5}x_5\\ \frac{3}{4}x_2 - \frac{1}{4}x_3 - \frac{1}{4}x_4 - \frac{1}{4}x_5\\ -\frac{1}{5}x_1 - \frac{1}{4}x_2 + \frac{3}{20}x_3 + \frac{3}{20}x_4 - \frac{1}{20}x_5\\ -\frac{1}{5}x_1 - \frac{1}{4}x_2 + \frac{3}{20}x_3 + \frac{3}{20}x_4 - \frac{1}{20}x_5\\ \frac{2}{5}x_1 - \frac{1}{4}x_2 - \frac{1}{20}x_3 - \frac{1}{20}x_4 + \frac{7}{20}x_5\end{pmatrix}.$$

In fact, in general, to find the projection of a vector $\mathbf{v}$ onto a subspace W, we can let $\mathbf{A}$ be a matrix whose columns form a basis for W. Then the columns of $\mathbf{A}$ must be linearly independent and thus $\mathbf{A}^T\mathbf{A}$ is invertible. Then the projection will be $\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{v}$.

4. **(Application)** A line

$$p(x) = a_1x + a_0$$

is said to be the *least squares approximating line* for a given a set of data points (x_1, y_1), (x_2, y_2), ..., (x_m, y_m) if the sum

$$S = [y_1 - p(x_1)]^2 + [y_2 - p(x_2)]^2 + \cdots + [y_m - p(x_m)]^2$$

is minimized. Writing

$$\mathbf{x} = \begin{pmatrix}x_1\\x_2\\\vdots\\x_m\end{pmatrix}, \mathbf{y} = \begin{pmatrix}y_1\\y_2\\\vdots\\y_m\end{pmatrix}, \text{ and } p(\mathbf{x}) = \begin{pmatrix}p(x_1)\\p(x_2)\\\vdots\\p(x_m)\end{pmatrix} = \begin{pmatrix}a_1x_1 + a_0\\a_1x_2 + a_0\\\vdots\\a_1x_m + a_0\end{pmatrix}$$

the problem is now rephrased as finding a_0, a_1 such that

$$S = ||\mathbf{y} - p(\mathbf{x})||^2$$

is minimized. Observe that if we let

$$\mathbf{N} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix} \text{ and } \mathbf{a} = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix},$$

then $\mathbf{Na} = p(\mathbf{x})$. And so our aim is to find $\mathbf{a}$ that minimizes $||\mathbf{y} - \mathbf{Na}||^2$.

It is known the equation representing the dependency of the resistance of a cylindrically shaped conductor (a wire) at 20^oC is given by

$$R = \rho \frac{L}{A},$$

where R is the resistance measured in Ohms Ω, L is the length of the material in meters m, A is the cross-sectional area of the material in meter squared m^2, and ρ is the resistivity of the material in Ohm meters Ωm. A student wants to measure the resistivity of a certain material. Keeping the cross-sectional area constant at $0.002m^2$, he connected the power sources along the material at varies length and measured the resistance and obtained the following data.

L	0.01	0.012	0.015	0.02
R	2.75×10^{-4}	3.31×10^{-4}	3.92×10^{-4}	4.95×10^{-4}

It is known that the Ohm meter might not be calibrated. Taking that into account, the student wants to find a linear graph $R = \frac{\rho}{0.002}L + R_0$ from the data obtained to compute the resistivity of the material.

(a) Relabeling, we let $R = y$, $\frac{\rho}{0.002} = a_1$ and $R_0 = a_0$. Is it possible to find a graph $y = a_1x + a_0$ satisfying the points?

Substituting in the data into the equation $y = a_1x + a_0$, we get the augmented matrix

$$\left(\begin{array}{cc|c} 1 & 0.01 & 2.75 \times 10^{-4} \\ 1 & 0.012 & 3.31 \times 10^{-4} \\ 1 & 0.015 & 3.92 \times 10^{-4} \\ 1 & 0.02 & 4.95 \times 10^{-4} \end{array}\right).$$

This linear system is inconsistent. Hence, no such graph exists.

(b) Find the least square approximating line for the data points and hence find the resistivity of the material. Would this material make a good wire?

Let $M = \begin{pmatrix} 1 & 0.01 \\ 1 & 0.012 \\ 1 & 0.015 \\ 1 & 0.02 \end{pmatrix}$ and $b = \begin{pmatrix} 2.75 \times 10^{-4} \\ 3.31 \times 10^{-4} \\ 3.92 \times 10^{-4} \\ 4.95 \times 10^{-4} \end{pmatrix}$. Solve for $\mathbf{M}^T\mathbf{Mx} = \mathbf{M}^T\mathbf{b}$.

Since the columns of $\mathbf{M}$ are linearly independent, the least square solution is

$$(\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{b} = \begin{pmatrix} 0.0001 \\ 0.0216 \end{pmatrix}.$$

So the least square approximating line is $y = 0.0216x + 0.0001$. So $\frac{\rho}{0.002} = 0.0216\Omega$, and hence $\rho = 4.32 \times 10^{-5}\Omega m$. It would not make a good wire, the resistivity of metals is in the $10^{-8}\Omega m$ range.

5. **(Application, MATLAB)** Suppose the equation governing the relation between data pairs is not known. We may want to then find a polynomial

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

of degree n, $n \leq m-1$, that best approximates the data pairs (x_1, y_1), (x_2, y_2), ..., (x_m, y_m). A *least square approximating polynomial* of degree n is such that

$$||\mathbf{y} - p(\mathbf{x})||^2$$

is minimized. If we write

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \ \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}, \ \mathbf{N} = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{pmatrix} \text{ and } \mathbf{a} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix},$$

then $p(\mathbf{x}) = \mathbf{Na}$, and the task is to find $\mathbf{a}$ such that $||\mathbf{y} - \mathbf{Na}||^2$ is minimized. Observe that $\mathbf{N}$ is a matrix minor of the Vandermonde matrix. If at least $n+1$ of the x-values $x_1, x_2, ..., x_m$ are distinct, the columns of $\mathbf{N}$ are linearly independent, and thus $\mathbf{a}$ is uniquely determined by

$$\mathbf{a} = (\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\mathbf{y}.$$

We shall now find a quartic polynomial

$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$$

that is a least square approximating polynomial for the following data points

x	4	4.5	5	5.5	6	6.5	7	8	8.5
y	0.8651	0.4828	2.590	-4.389	-7.858	3.103	7.456	0.0965	4.326

Enter the data points.

```
>> x=[4 4.5 5 5.5 6 6.5 7 8 8.5]';
```

```
>> y=[0.8651 0.4828 2.590 -4.389 -7.858 3.103 7.456 0.0965 4.326]';
```

Next, we will generate the 10×10 Vandermonde matrix.

```
>> N=fliplr(vander(x));
```

We only want the matrix minor up to the 4-th power, that is, up to the the 5-th column,

```
>> N=N(:,1:5);
```

Use this to find the least square approximating polynomial of degree 4.

`>> a=inv(N'*N)*N'*y`, ans= $\begin{pmatrix} -204.0716 \\ 169.2099 \\ -49.7013 \\ 6.1528 \\ -0.2720 \end{pmatrix}$. So the polynomial is

$$-0.2720x^4 + 6.1528x^3 - 49.7013x^2 + 169.2099x - 204.0716.$$

6. Compute the eigenvalues of the following matrices $\mathbf{A}$.

   (a) $\mathbf{A} = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}$.

   $\begin{vmatrix} x-1 & 3 & -3 \\ -3 & x+5 & -3 \\ -6 & 6 & x-4 \end{vmatrix} = x^3 - 12x - 16 = (x+2)^2(x-4)$. The eigenvalues are $\lambda = -2$ and $\lambda = 4$.

   (b) $\mathbf{A} = \begin{pmatrix} 9 & 8 & 6 & 3 \\ 0 & -1 & 3 & -4 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

   Since $\mathbf{A}$ is a triangular matrix, the eigenvalues are the diagonals $\lambda = -1$, $\lambda = 2$, $\lambda = 3$, $\lambda = 9$.

7. (a) Show that λ is an eigenvalue of $\mathbf{A}$ if and only if it is an eigenvalue of $\mathbf{A}^T$.

   $$\det(x\mathbf{I} - \mathbf{A}) = \det((x\mathbf{I} - \mathbf{A})^T) = \det((x\mathbf{I})^T - \mathbf{A}^T) = \det(x\mathbf{I} - \mathbf{A}^T).$$

   So the roots of $\det(x\mathbf{I} - \mathbf{A})$ are exactly the roots of $\det(x\mathbf{I} - \mathbf{A}^T)$.

   (b) Suppose $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ associated to eigenvalue λ. Show that $\mathbf{v}$ is an eigenvector of $\mathbf{A}^k$ associated to eigenvalue λ^k for any positive integer k.

   By definition, we have $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Then

   $$\mathbf{A}^k\mathbf{v} = \mathbf{A}^{k-1}\mathbf{A}\mathbf{v} = \lambda\mathbf{A}^{k-2}\mathbf{A}\mathbf{v} = \lambda^2\mathbf{A}^{k-3}\mathbf{A}\mathbf{v} = \cdots = \lambda^{k-1}\mathbf{A}\mathbf{v} = \lambda^k\mathbf{v}.$$

   Since $\mathbf{v} \neq \mathbf{0}$, $\mathbf{v}$ is a witness to λ^k being an eigenvalue of $\mathbf{A}^k$.

   (c) If $\mathbf{A}$ is invertible, show that $\mathbf{v}$ is an eigenvector of $\mathbf{A}^k$ associated to eigenvalue λ^k for any negative integer k.

   Suppose $k = -1$. First note that since $\mathbf{A}$ is invertible, $k \neq 0$. Then $\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \iff \lambda^{-1}\mathbf{v} = \mathbf{A}^{-1}\mathbf{v}$. Hence, λ^{-1} is an eigenvalue of $\mathbf{A}^{-1}$. The rest of the argument follows from (a).

   (d) Recall that a matrix is nilpotent if there is a positive integer k such that $\mathbf{A}^k = \mathbf{0}$. Show that if $\mathbf{A}$ is nilpotent, then 0 is the only eigenvalue.

   Let λ be an eigenvalue of $\mathbf{A}$ and $\mathbf{v}$ be an eigenvector associated to λ. By (a), $\mathbf{0} = \mathbf{A}^k\mathbf{v} = \lambda^k\mathbf{v}$. Since $\mathbf{v} \neq \mathbf{0}$, necessarily $\lambda^k = 0$, and hence $\lambda = 0$.

Supplementary Problems

8. (**QR-factorisation**) Let $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{pmatrix}$, where $\mathbf{u}_i$ is the i-th column of $\mathbf{A}$ for $i = 1, 2, 3$.

   (a) Use Gram-Schmidt Process to transform $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ into an orthonormal basis $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ for the column space of $\mathbf{A}$. (Do not change the order of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ when applying the Gram-Schmidt Process.)

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \\ \mathbf{v}_2 &= \mathbf{u}_2 - \left(\frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \right) \mathbf{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ \mathbf{v}_3 &= \mathbf{u}_3 - \left(\frac{\mathbf{u}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \right) \mathbf{v}_1 - \left(\frac{\mathbf{u}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \right) \mathbf{v}_2 = \frac{1}{3} \begin{pmatrix} -1 \\ -1 \\ 2 \\ 0 \end{pmatrix}. \end{aligned}$$

   Then

$$\begin{aligned} \mathbf{w}_1 &= \frac{\mathbf{v}_1}{||\mathbf{v}_1||} = \frac{1}{\sqrt{3}} \mathbf{u}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \\ \mathbf{w}_2 &= \frac{\mathbf{v}_2}{||\mathbf{v}_2||} = \mathbf{u}_2 - \mathbf{u}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ \mathbf{w}_3 &= \frac{\mathbf{v}_3}{||\mathbf{v}_3||} = \sqrt{\frac{3}{2}} (\mathbf{u}_3 - \frac{1}{3}\mathbf{v}_1 - \mathbf{v}_2) = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 2 \\ 0 \end{pmatrix}. \end{aligned}$$

   (b) Write each of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ as a linear combination of $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$.

   From (a), we have

$$\begin{aligned} \mathbf{u}_1 &= \sqrt{3}\mathbf{w}_1 \\ \mathbf{u}_2 &= \sqrt{3}\mathbf{w}_1 + \mathbf{w}_2 \\ \mathbf{u}_3 &= \frac{1}{\sqrt{3}}\mathbf{w}_1 + \mathbf{w}_2 + \sqrt{\frac{2}{3}}\mathbf{w}_3 \end{aligned}$$

(c) Hence or otherwise, write $\mathbf{A} = \mathbf{QR}$ where $\mathbf{Q}$ is a 4×3 matrix with orthonormal columns and $\mathbf{R}$ is a 3×3 upper triangular matrix with positive entries along its diagonal. (**Hint:** Recall that if $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ is a solution to $\mathbf{Mx} = \mathbf{b}$, then $\mathbf{b} = a\mathbf{m}_1 + b\mathbf{m}_2 + \mathbf{m}_3$, where $\mathbf{m}_i$ is the i-th column of $\mathbf{M}$.)

From (b), we have

$$\begin{pmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 \end{pmatrix} \begin{pmatrix} \sqrt{3} \\ 0 \\ 0 \end{pmatrix} = \mathbf{u}_1,$$

$$\begin{pmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 \end{pmatrix} \begin{pmatrix} \sqrt{3} \\ 1 \\ 0 \end{pmatrix} = \mathbf{u}_2,$$

$$\begin{pmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} \\ 1 \\ \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix} = \mathbf{u}_1$$

In other words,

$$\begin{pmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 \end{pmatrix} \begin{pmatrix} \sqrt{3} & \sqrt{3} & \frac{1}{\sqrt{3}} \\ 0 & 1 & 1 \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{pmatrix} = \mathbf{A}$$

Let $\mathbf{Q} = \begin{pmatrix} \mathbf{w_1} & \mathbf{w_2} & \mathbf{w_3} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ 0 & 1 & 0 \end{pmatrix}$ and $\mathbf{R} = \begin{pmatrix} \sqrt{3} & \sqrt{3} & \frac{1}{\sqrt{3}} \\ 0 & 1 & 1 \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}$.

Then the columns of $\mathbf{Q}$ are orthonormal, $\mathbf{R}$ is an upper triangular matrix and $\mathbf{A} = \mathbf{QR}$.

In general, we have

Theorem. *If* $\mathbf{A}$ *is an* $m \times n$ *matrix with linearly independent columns, then* $\mathbf{A}$ *can be factorised into* $\mathbf{A} = \mathbf{QR}$ *where* $\mathbf{Q}$ *is an* $m \times n$ *matrix whose columns form an orthonormal basis for the column space of* $\mathbf{A}$ *and* $\mathbf{R}$ *is an* $n \times n$ *invertible upper triangular matrix.*

Remark: $\mathbf{QR}$-factorisation is widely used in computer algorithms for various computations concerning matrices. We can show easily that in general, the matrix $\mathbf{R}$ is always invertible. Let $\mathbf{x}$ be a solution to the linear system $\mathbf{Rx} = \mathbf{0}$. Pre-multiplying $\mathbf{Q}$ on both sides, we have

$$\mathbf{Q}(\mathbf{Rx}) = \mathbf{Q0} \Rightarrow (\mathbf{QR})\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{Ax} = \mathbf{0}.$$

Since the columns are $\mathbf{A}$ are linearly independent, the rank of $\mathbf{A}$ is equal to the number of columns, and thus the nullity of $\mathbf{A}$ zero. Hence, the nullspace is trivial, and so necessarily $\mathbf{x} = \mathbf{0}$. This means the trivial solution is the only solution to $\mathbf{Rx} = \mathbf{0}$. Thus $\mathbf{R}$ must be invertible.

9. Let $\mathbf{v}_1$ be an eigenvector of $\mathbf{A}$ associated to the eigenvalue λ_1 and $\mathbf{v}_2$ an eigenvector of $\mathbf{A}^T$ associated to eigenvalue λ_2. Suppose $\lambda_1 \neq \lambda_2$. Show that v_1 and v_2 are orthogonal.

   Taking transpose of the equation $\mathbf{A}^T\mathbf{v}_2 = \lambda_2\mathbf{v}_2$ we obtain $\mathbf{v}_2^T\mathbf{A} = \lambda\mathbf{v}_2^T$. So,

$$\lambda_2\mathbf{v}_2 \cdot \mathbf{v}_1 = \lambda_2\mathbf{v}_2^T\mathbf{v}_1 = (\mathbf{v}_2^T\mathbf{A})\mathbf{v}_1 = \mathbf{v}_2^T(\mathbf{A}\mathbf{v}_1) = \lambda_1\mathbf{v}_2^T\mathbf{v}_1 = \lambda_1\mathbf{v}_2 \cdot \mathbf{v}_1.$$

   In other words, $(\lambda_2 - \lambda_1)\mathbf{v}_2 \cdot \mathbf{v}_1 = \mathbf{0}$. Since $(\lambda_2 - \lambda_1) \neq 0$, necessarily $\mathbf{v}_2 \cdot \mathbf{v}_1 = \mathbf{0}$.