Three solutions for Exercise 10-1

Isao Sasano

2015 June 16

Exercise Approximate a column vector $\boldsymbol{u} = \begin{pmatrix} 4 \\ 3 \\ 3 \end{pmatrix}$ by a linear combination of the column vectors $\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $\boldsymbol{e}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ (i.e., $\sum_{k=1}^{2} c_k \boldsymbol{e}_k = c_1 \boldsymbol{e}_1 + c_2 \boldsymbol{e}_2$ for some c_1 and c_2). That is, obtain c_1 and c_2 so that $c_1 \boldsymbol{e}_1 + c_2 \boldsymbol{e}_2$ is closest to $\boldsymbol{u}$. As for the measure of the distance, use (the half of) the square of the norm of the difference of $c_1 \boldsymbol{e}_1 + c_2 \boldsymbol{e}_2$ and $\boldsymbol{u}$.

$$J = \frac{1}{2} \left\| \sum_{k=1}^{2} c_k \boldsymbol{e}_k - \boldsymbol{u} \right\|^2$$

The norm of a column vector $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ is defined as follows.

$$\|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})} = \sqrt{\sum_{k=1}^{3} x_k^2}$$

We show three solutions. One is by solving $\dfrac{\partial J}{\partial c_i} = 0$ for $i = 1, 2$, as we did in Exercise 4. Another is by considering the special case. The other is by calculating the projection of $\boldsymbol{u}$ on the subspace spanned by $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$.

Solution 1 Firstly calculate J as follows.

$$
\begin{aligned}
J &= \frac{1}{2}\left\|\sum_{k=1}^{2} c_k \boldsymbol{e}_k - \boldsymbol{u}\right\|^2 \\
&= \frac{1}{2}\left(\sum_{k=1}^{2} c_k \boldsymbol{e}_k - \boldsymbol{u}, \sum_{k=1}^{2} c_k \boldsymbol{e}_k - \boldsymbol{u}\right) \\
&= \frac{1}{2}\left\{\left(\sum_{k=1}^{2} c_k \boldsymbol{e}_k, \sum_{k=1}^{2} c_k \boldsymbol{e}_k\right) - 2\left(\boldsymbol{u}, \sum_{k=1}^{2} c_k \boldsymbol{e}_k\right) + \|\boldsymbol{u}\|^2\right\} \\
&= \frac{1}{2}\left\{\sum_{k,l=1}^{2} c_k c_l (\boldsymbol{e}_k, \boldsymbol{e}_l) - 2\sum_{k=1}^{2} c_k (\boldsymbol{u}, \boldsymbol{e}_k) + \|\boldsymbol{u}\|^2\right\}
\end{aligned}
$$

Partially differenciate this with recpect to c_i $(i = 1, 2)$.

$$
\begin{aligned}
\frac{\partial J}{\partial c_i} &= \frac{\partial}{\partial c_i}\frac{1}{2}\left\{\sum_{k,l=1}^{2} c_k c_l (\boldsymbol{e}_k, \boldsymbol{e}_l) - 2\sum_{k=1}^{2} c_k (\boldsymbol{u}, \boldsymbol{e}_k) + \|\boldsymbol{u}\|^2\right\} \\
&= \frac{1}{2}\left\{\frac{\partial}{\partial c_i}\sum_{k,l=1}^{2} c_k c_l (\boldsymbol{e}_k, \boldsymbol{e}_l) - 2\frac{\partial}{\partial c_i}\sum_{k=1}^{2} c_k (\boldsymbol{u}, \boldsymbol{e}_k)\right\} \\
&= \frac{1}{2}\left\{2\sum_{k=1}^{2} c_k (\boldsymbol{e}_k, \boldsymbol{e}_i) - 2(\boldsymbol{u}, \boldsymbol{e}_i)\right\} \\
&= \sum_{k=1}^{2} c_k (\boldsymbol{e}_k, \boldsymbol{e}_i) - (\boldsymbol{u}, \boldsymbol{e}_i)
\end{aligned}
$$

By writing $\dfrac{\partial J}{\partial c_1} = 0$ and $\dfrac{\partial J}{\partial c_2} = 0$ in matrix form, we obtain

$$
\begin{pmatrix} (\boldsymbol{e}_1, \boldsymbol{e}_1) & (\boldsymbol{e}_2, \boldsymbol{e}_1) \\ (\boldsymbol{e}_1, \boldsymbol{e}_2) & (\boldsymbol{e}_2, \boldsymbol{e}_2) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} (\boldsymbol{u}, \boldsymbol{e}_1) \\ (\boldsymbol{u}, \boldsymbol{e}_2) \end{pmatrix}
$$

Since $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are orthogonal, we obtain

$$
\begin{pmatrix} \|\boldsymbol{e}_1\|^2 & 0 \\ 0 & \|\boldsymbol{e}_2\|^2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} (\boldsymbol{u}, \boldsymbol{e}_1) \\ (\boldsymbol{u}, \boldsymbol{e}_2) \end{pmatrix}.
$$

Hence we obtain $c_1 = \dfrac{(\boldsymbol{u}, \boldsymbol{e}_1)}{\|\boldsymbol{e}_1\|^2} = \dfrac{10}{3}$ and $c_2 = \dfrac{(\boldsymbol{u}, \boldsymbol{e}_2)}{\|\boldsymbol{e}_2\|^2} = \dfrac{1}{2}$. Thus the linear combination of $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ that is closest to the vector $\boldsymbol{u}$ is obtained as follows.

$$c_1\boldsymbol{e}_1 + c_1\boldsymbol{e}_2 = \frac{10}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 23/6 \\ 10/3 \\ 17/6 \end{pmatrix}$$

Solution 2 Assume the following equation holds.

$$\boldsymbol{u} = c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2$$

(There is actually no coefficients c_1 and c_2 that satisfy the equation, but it's ok.) Take the inner produce with $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ in the both side.

$$(\boldsymbol{u}, \boldsymbol{e}_1) = c_1(\boldsymbol{e}_1, \boldsymbol{e}_1) + c_2(\boldsymbol{e}_2, \boldsymbol{e}_1)$$
$$(\boldsymbol{u}, \boldsymbol{e}_2) = c_1(\boldsymbol{e}_1, \boldsymbol{e}_2) + c_2(\boldsymbol{e}_2, \boldsymbol{e}_2)$$

We write the above two equations in the matrix form.

$$\begin{pmatrix} (\boldsymbol{e}_1, \boldsymbol{e}_1) & (\boldsymbol{e}_2, \boldsymbol{e}_1) \\ (\boldsymbol{e}_1, \boldsymbol{e}_2) & (\boldsymbol{e}_2, \boldsymbol{e}_2) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} (\boldsymbol{u}, \boldsymbol{e}_1) \\ (\boldsymbol{u}, \boldsymbol{e}_2) \end{pmatrix}$$

Notice that this is the same as the one obtained in Solution 1. So we omit the remaining calculation.

Solution 3

A linear combination of $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ is written as $c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2$. It is closest to $\boldsymbol{u}$ when it is the projection of $\boldsymbol{u}$ on the subspace spanned by $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$. That is, the vector $\boldsymbol{u} - (c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2)$ is orthogonal to the subspace spanned by $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$. So the following two equations should hold.

$$(\boldsymbol{u} - (c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2), \boldsymbol{e}_1) = 0$$
$$(\boldsymbol{u} - (c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2), \boldsymbol{e}_2) = 0$$

We expand the inner product in each of the equations.

$$(\boldsymbol{u}, \boldsymbol{e}_1) - (c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2, \boldsymbol{e}_1) = 0$$
$$(\boldsymbol{u}, \boldsymbol{e}_2) - (c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2, \boldsymbol{e}_2) = 0$$

We move the second inner product into RHS in each of the equations.

$$(\boldsymbol{u}, \boldsymbol{e}_1) = (c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2, \boldsymbol{e}_1)$$
$$(\boldsymbol{u}, \boldsymbol{e}_2) = (c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2, \boldsymbol{e}_2)$$

We expand the inner product in the RHS in each of the equations.

$$\begin{aligned}(\boldsymbol{u}, \boldsymbol{e}_1) = c_1(\boldsymbol{e}_1, \boldsymbol{e}_1) + c_2(\boldsymbol{e}_2, \boldsymbol{e}_1) \\ (\boldsymbol{u}, \boldsymbol{e}_2) = c_1(\boldsymbol{e}_1, \boldsymbol{e}_2) + c_2(\boldsymbol{e}_2, \boldsymbol{e}_2)\end{aligned}$$

We write the above two equations in the matrix form.

$$\begin{pmatrix} (\boldsymbol{e}_1, \boldsymbol{e}_1) & (\boldsymbol{e}_2, \boldsymbol{e}_1) \\ (\boldsymbol{e}_1, \boldsymbol{e}_2) & (\boldsymbol{e}_2, \boldsymbol{e}_2) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} (\boldsymbol{u}, \boldsymbol{e}_1) \\ (\boldsymbol{u}, \boldsymbol{e}_2) \end{pmatrix}$$

Notice that this is the same as the one obtained in Solution 1 and 2. So we omit the remaining calculation.