

Three solutions for Exercise 10-1

Isao Sasano

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

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

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

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

We show three solutions. One is by solving $\frac{\partial J}{\partial c_i} = 0$ 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 \mathbf{u} on the subspace spanned by \mathbf{e}_1 and \mathbf{e}_2 .

Solution 1 Firstly calculate J as follows.

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

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

Partially differentiate this with respect to c_i ($i = 1, 2$).

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

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

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

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

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

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

$$c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 = \frac{10}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 23/6 \\ 10/3 \\ 17/6 \end{pmatrix}$$

Solution 2 Assume the following equation holds.

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

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

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

We write the above two equations in the matrix form.

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

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

$$\begin{aligned}(\mathbf{u} - (c_1\mathbf{e}_1 + c_2\mathbf{e}_2), \mathbf{e}_1) &= 0 \\(\mathbf{u} - (c_1\mathbf{e}_1 + c_2\mathbf{e}_2), \mathbf{e}_2) &= 0\end{aligned}$$

We expand the inner product in each of the equations.

$$\begin{aligned}(\mathbf{u}, \mathbf{e}_1) - (c_1\mathbf{e}_1 + c_2\mathbf{e}_2, \mathbf{e}_1) &= 0 \\(\mathbf{u}, \mathbf{e}_2) - (c_1\mathbf{e}_1 + c_2\mathbf{e}_2, \mathbf{e}_2) &= 0\end{aligned}$$

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

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

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

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

We write the above two equations in the matrix form.

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

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