Aalto University

Linear algebra

Exercise sheet 11 / Model solutions

1. Let $\boldsymbol{a}_1 = [1, 1, 0]^T$ and $\boldsymbol{a}_2 = [1, 0, 1]^T$.

   (a) Find two vectors $\boldsymbol{q}_1, \boldsymbol{q}_2 \in \mathbb{R}^3$ such that $\operatorname{span}(\boldsymbol{q}_1, \boldsymbol{q}_2) = \operatorname{span}(\boldsymbol{a}_1, \boldsymbol{a}_2)$ and such that $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are orthogonal with respect to the Euclidian inner product.

   (b) Find two vectors $\boldsymbol{q}_1, \boldsymbol{q}_2 \in \mathbb{R}^3$ such that $\operatorname{span}(\boldsymbol{q}_1, \boldsymbol{q}_2) = \operatorname{span}(\boldsymbol{a}_1, \boldsymbol{a}_2)$ and such that $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are orthogonal with respect to the inner product defined as

$$\left\langle \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \right\rangle = x_1 y_1 + 2 x_2 y_2 + 3 x_3 y_3.$$

Solution.

(a) Following the Gram–Schmidt process, we get

$$\begin{aligned}
\boldsymbol{q}_1 &= \frac{\boldsymbol{a}_1}{\|\boldsymbol{a}_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \\
\widetilde{\boldsymbol{q}}_2 &= \boldsymbol{a}_2 - \langle \boldsymbol{a}_2, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 = \boldsymbol{a}_2 - (\boldsymbol{a}_2^T \boldsymbol{q}_1) \boldsymbol{q}_1 \\
&= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{\sqrt{2}} (1 \cdot 1 + 0 \cdot 1 + 1 \cdot 0) \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}, \\
\boldsymbol{q}_2 &= \frac{\tilde{\boldsymbol{q}}_2}{\|\tilde{\boldsymbol{q}}_2\|} = \frac{1}{\sqrt{1/4 + 1/4 + 1}} \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} \\
&= \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}.
\end{aligned}$$

(b) With the same process, but applied to the other inner product, we get

$$
\begin{aligned}
\boldsymbol{q}_1 &= \frac{\boldsymbol{a}_1}{\|\boldsymbol{a}_1\|} = \frac{\boldsymbol{a}_1}{\langle \boldsymbol{a}_1, \boldsymbol{a}_1 \rangle^{1/2}} = \frac{1}{\sqrt{1^2 + 2 \cdot 1^2 + 3 \cdot 0^2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\
&= \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},
\end{aligned}
$$

$$
\begin{aligned}
\widetilde{\boldsymbol{q}}_2 &= \boldsymbol{a}_2 - \langle \boldsymbol{a}_2, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 \\
&= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \left(\frac{1}{\sqrt{3}} \cdot 1 + \frac{2}{\sqrt{3}} \cdot 1 + 3 \cdot 1 \cdot 0 \right) \cdot \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, \\
\boldsymbol{q}_2 &= \frac{\tilde{\boldsymbol{q}}_2}{\|\tilde{\boldsymbol{q}}_2\|} = \frac{1}{\frac{1}{3}\sqrt{4 + 2 \cdot 1 + 3 \cdot 3^2}} \frac{1}{3} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} \\
&= \frac{1}{\sqrt{33}} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}.
\end{aligned}
$$

2. Let

$$
A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 2 \\ -1 & 1 & 0 \end{bmatrix}
$$

and use the Euclidian inner product to measure orthogonality in this problem.

(a) Modify the Gram–Schmidt orthogonalisation process so that you can use it to find an orthonormal basis for a possibly linearly dependent set of vectors.

(b) Find an orthonormal basis for $R(A)$.

Solution.

(a) *Gram–Schmidt for linearly independent sets:* Denoting by $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$, the starting linearly independent vectors, the Gram–Schmidt process produces orthogonal vectors $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ from which in turn we get orthonormal vectors $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n$ in the following way:

$$\begin{aligned}
\boldsymbol{w}_1 &= \boldsymbol{a}_1 & \boldsymbol{q}_1 &= \frac{\boldsymbol{w}_1}{\|\boldsymbol{w}_1\|} \\
\boldsymbol{w}_2 &= \boldsymbol{a}_2 - \langle \boldsymbol{a}_2, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 & \boldsymbol{q}_2 &= \frac{\boldsymbol{w}_2}{\|\boldsymbol{w}_2\|} \\
\boldsymbol{w}_3 &= \boldsymbol{a}_3 - \langle \boldsymbol{a}_3, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 - \langle \boldsymbol{a}_3, \boldsymbol{q}_2 \rangle \boldsymbol{q}_2 & \boldsymbol{q}_3 &= \frac{\boldsymbol{w}_3}{\|\boldsymbol{w}_3\|} \\
&\vdots & &\vdots \\
\boldsymbol{w}_k &= \boldsymbol{a}_k - \sum_{j=1}^{k-1} \langle \boldsymbol{a}_k, \boldsymbol{q}_j \rangle \boldsymbol{q}_j & \boldsymbol{q}_k &= \frac{\boldsymbol{w}_k}{\|\boldsymbol{w}_k\|}.
\end{aligned}$$

The vectors $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n$ that we get with the algorithm are still linearly independent. They form an orthonormal basis for $\operatorname{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$.

Modified Gram–Schmidt process: If $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ are linearly *dependent*, for some value of k we get $\boldsymbol{w}_k = 0$ (meaning that $\boldsymbol{a}_k$ is a linear combination of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{k-1}$), hence in the computation of $\boldsymbol{q}_k$ we would divide by zero. To prevent this, in the algorithm we can add a step in which we check whether $\boldsymbol{w}_k = 0$, and if yes we simply forget this vector, replacing $\boldsymbol{a}_k$ with the next vector $\boldsymbol{a}_{k+1}$. This way we still get an orthonormal basis for $\operatorname{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$, which might have less than n elements in case we started from a linearly dependent set of vectors.

(b) The set $R(A)$ is generated by the columns of A. So the starting vectors for the algorithm are

$$\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \boldsymbol{a}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{a}_3 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

Observe that they are linearly dependent because $\boldsymbol{a}_1 + \boldsymbol{a}_2 = \boldsymbol{a}_3$. Then if we apply the modified algorithm described in part (a) we get

$$\begin{aligned}
\boldsymbol{w}_1 &= \boldsymbol{a}_1 = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T \neq \mathbf{0} \\
\boldsymbol{q}_1 &= \frac{\boldsymbol{w}_1}{\|\boldsymbol{w}_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T \\
\boldsymbol{w}_2 &= \boldsymbol{a}_2 - \langle \boldsymbol{a}_2, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 \\
&= \begin{bmatrix} \frac{1}{2} & 2 & \frac{1}{2} \end{bmatrix}^T \neq \mathbf{0} \\
\boldsymbol{q}_2 &= \frac{\boldsymbol{w}_2}{\|\boldsymbol{w}_2\|} = \frac{\sqrt{2}}{3} \begin{bmatrix} \frac{1}{2} & 2 & \frac{1}{2} \end{bmatrix}^T = \frac{1}{3\sqrt{2}} \begin{bmatrix} 1 & 4 & 1 \end{bmatrix}^T \\
\boldsymbol{w}_3 &= \boldsymbol{a}_3 - \langle \boldsymbol{a}_3, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 - \langle \boldsymbol{a}_3, \boldsymbol{q}_2 \rangle \boldsymbol{q}_2 \\
&= \boldsymbol{a}_3 - \left(\frac{1}{\sqrt{2}} \right) \boldsymbol{q}_1 - \left(\frac{3}{\sqrt{2}} \right) \boldsymbol{q}_2 \\
&= \begin{bmatrix} 1 & 2 & 0 \end{bmatrix}^T - \frac{1}{2} \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T - \frac{1}{2} \begin{bmatrix} 1 & 4 & 1 \end{bmatrix}^T = \mathbf{0}.
\end{aligned}$$

So we can discard the the last vector, and $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ alone form an orthonormal basis for $R(A)$.

3. Let
$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$
Compute by hand a QR decomposition of the matrix A and use it to solve the least-squares problem
$$\min_{\boldsymbol{x}\in\mathbb{R}^2} \|A\boldsymbol{x} - \boldsymbol{b}\|_2^2.$$
That is, solve the equation $R\boldsymbol{x} = Q^T\boldsymbol{b}$.

Hint: Problem 1(a).

Solution. From Problem 1(a) we get immediately
$$A = \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{6}} \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{3}{2}} \end{bmatrix}}_{R}.$$
Since the columns of A are linearly independent, $N(A) = N(A^TA) = \{\mathbf{0}\}$, that is, for the normal equation $A^TA\boldsymbol{x} = A^T\boldsymbol{b}$ there exists a unique solution, which is also the solution of the least-squares homework problem. (See Theorem 1.1 on page 43 of the lecture notes.)

On the other hand we can write $A^TA\boldsymbol{x} = A^T\boldsymbol{b}$ as $(QR)^TQR\boldsymbol{x} = (QR)^T\boldsymbol{b}$, and then multiply by the inverse of R^T, finding that an quivalent problem to $A^TA\boldsymbol{x} = A^T\boldsymbol{b}$ is $R\boldsymbol{x} = Q\boldsymbol{b}$, that is,
$$\begin{cases} \sqrt{2}x_1 + \frac{1}{\sqrt{2}}x_2 = \frac{1}{\sqrt{2}} \\ \sqrt{\frac{3}{2}}x_2 = \frac{1}{\sqrt{6}} \end{cases}$$
which in turn means that
$$\begin{cases} x_1 = \frac{1}{3} \\ x_2 = \frac{1}{3}. \end{cases}$$

4. Let
$$A = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 1 & 2 \\ 2 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}.$$
   (a) Find an orthonormal basis for $R(A)$ by using the modified Gram–Schmidt process.
   (b) Find the QR-decomposition of A by using the modified Gram–Schmidt process.

Use Matlab and the function `my_gsmith.m` given on page 51 of the lecture notes.

Solution. To answer is enough to use the Matlab code mentioned in the hint. The code is as follows:

```
function [Q,R] = my_gsmith(A)

Q = [ ];
for i=1:size(A,2)
  q = A(:,i);

  for k=1:size(Q,2)
      R(k,i) = q'*Q(:,k);
      q = q - R(k,i)*Q(:,k);
  end
  R(i,i) = norm(q);
  Q(:,i) = q/R(i,i);
end
```

The output given by Matlab is that

$$Q = \begin{bmatrix} 0.3162 & 0.7980 & 0.1147 \\ 0.3162 & 0.0725 & 0.8030 \\ 0.6325 & -0.5804 & 0.1147 \\ 0.6325 & 0.1451 & -0.5735 \end{bmatrix}, \qquad R = \begin{bmatrix} 3.1623 & 2.8460 & 2.5298 \\ 0 & 1.3784 & 1.3059 \\ 0 & 0 & 1.3765 \end{bmatrix}.$$

(a) An orthonormal basis for the space $R(A)$ consists of the columns of Q.

(b) The matrices just computet satisfy $A = QR$.