NATIONAL UNIVERSITY OF SINGAPORE
Department of Mathematics

MA1522 Linear Algebra for Computing **Tutorial 8**

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 $\mathbb{R}^4$.

Solution: 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 $\mathbf{u}_4 = \begin{pmatrix}1\\2\\0\\1\end{pmatrix}$.

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

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?

Solution: 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. The vector $\mathbf{v}_4 = 0$ means that $\mathbf{u}_4$ minus its projection onto $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is the zero vector. Hence $\mathbf{u}_4$ is contained in $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \text{span}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.

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) Is the linear system $\mathbf{Ax} = \mathbf{b}$ inconsistent?

   Solution:
$$\left(\begin{array}{cccc|c}0 & 1 & 1 & 0 & 6\\1 & -1 & 1 & -1 & 3\\1 & 0 & 1 & 0 & -1\\1 & 1 & 1 & 1 & 1\end{array}\right) \longrightarrow \left(\begin{array}{cccc|c}1 & 0 & 0 & 1 & 0\\0 & 1 & 0 & 1 & 0\\0 & 0 & 1 & -1 & 0\\0 & 0 & 0 & 0 & 1\end{array}\right).$$
   So the linear system $\mathbf{Ax} = \mathbf{b}$ is inconsistent.

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

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

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

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

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

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.

From the solution of (a), the matrix $\mathbf{A}$ is singular (Why?), 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?

Solution: 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.

3. **(Application)** A line

$$p(x) = a_1 x + a_0$$

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?

Solution: 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?

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

We will solve for $\mathbf{M}^T\mathbf{M}\mathbf{x} = \mathbf{M}^T\mathbf{b}$. Since the columns of $\mathbf{M}$ are linearly independent, the least square solution is

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

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.

4. **(Application)** 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 $p(\mathbf{x}) = \mathbf{N}\mathbf{a}$, and the task is to find $\mathbf{a}$ such that $||\mathbf{y} - \mathbf{N}\mathbf{a}||^2$ 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

$$p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$$

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.

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

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

5. Let $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.

   (a) Find a QR factorization of $\mathbf{A}$.

Solution:

$$
\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \\
\mathbf{v}_2 &= \mathbf{u}_2 - \left(\frac{\mathbf{u}_2 \cdot \mathbf{v_1}}{\mathbf{v}_1 \cdot \mathbf{v_1}} \right) \mathbf{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\
\mathbf{v}_3 &= \mathbf{u}_3 - \left(\frac{\mathbf{u}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \right) \mathbf{v}_1 - \left(\frac{\mathbf{u}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \right) \mathbf{v}_2 = \frac{1}{3} \begin{pmatrix} -1 \\ -1 \\ 2 \\ 0 \end{pmatrix}.
\end{aligned}
$$

So, let

$$
\mathbf{Q} = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ 0 & 1 & 0 \end{pmatrix}.
$$

Then

$$
\mathbf{R} = \mathbf{Q}^T \mathbf{A} = \begin{pmatrix} \sqrt{3} & \sqrt{3} & \frac{1}{\sqrt{3}} \\ 0 & 1 & 1 \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}.
$$

(b) Use your answer in (a) to find the least square solution to $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

Solution: Write $\mathbf{A} = \mathbf{QR}$. Then $\mathbf{A}^T\mathbf{A} = \mathbf{R}^T\mathbf{Q}^T\mathbf{QR} = \mathbf{R}^T\mathbf{R}$, and $\mathbf{A}^T\mathbf{b} = \mathbf{R}^T\mathbf{Q}^T\mathbf{b}$. Hence, solving for $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$ is equivalent to solving for $\mathbf{R}^T\mathbf{Rx} = \mathbf{R}^T\mathbf{Q}^T\mathbf{b}$, which is equivalent to solving for $\mathbf{Rx} = \mathbf{Q}^T\mathbf{b}$, since $\mathbf{R}$ is invertible (and hence, so is $\mathbf{R}^T$). Solving

$$
\begin{pmatrix} \sqrt{3} & \sqrt{3} & 1/\sqrt{3} \\ 0 & 1 & 1 \\ 0 & 0 & \sqrt{2}/\sqrt{3} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2/\sqrt{3} \\ 0 \\ -2/\sqrt{6} \end{pmatrix},
$$

we have $\mathbf{u} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ is the least square solution to $\mathbf{Ax} = \mathbf{b}$.

Extra problems

1. (a) Let S be the set of least square solutions to $\mathbf{Ax} = \mathbf{b}$. Show that there exists a $\mathbf{b}'$ such that $\mathbf{Av} = \mathbf{b}'$ for all $\mathbf{v} \in S$. This proves that the projection of $\mathbf{b}$ onto the column space of $\mathbf{A}$ is unique even though the least square solutions may not unique.

   Solution: Let $\mathbf{v}$ be any vector in S and let $\mathbf{Av} = \mathbf{b}'$. Now suppose $\mathbf{u}$ is another least square solution. This means that

   $$\mathbf{A}^T\mathbf{Au} = \mathbf{A}^T\mathbf{Av} = \mathbf{A}^T\mathbf{b}.$$

   This means that $\mathbf{A}^T\mathbf{A}(\mathbf{u} - \mathbf{v}) = \mathbf{0}$, or $\mathbf{u} - \mathbf{v} \in \mathrm{Null}(\mathbf{A})$.
   Recall that $\mathrm{Null}(\mathbf{A}^T\mathbf{A}) = \mathrm{Null}(\mathbf{A})$. Hence, $\mathbf{A}(\mathbf{u} - \mathbf{v}) = \mathbf{0}$. This means that

   $$\mathbf{Au} = \mathbf{Av} = \mathbf{b}'.$$

   (b) Suppose a linear system $\mathbf{Ax} = \mathbf{b}$ is consistent. Show that the solution set of $\mathbf{Ax} = \mathbf{b}$ is equal to the solution set of $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$.

   Solution: Suppose $\mathbf{u}$ is a solution to $\mathbf{Ax} = \mathbf{b}$, that is, $\mathbf{Au} = \mathbf{b}$. Premultiplying both sides by $\mathbf{A}^T$, we have $\mathbf{A}^T\mathbf{Au} = \mathbf{A}^T\mathbf{b}$, that is, $\mathbf{u}$ is a solution to $\mathbf{A}^T\mathbf{Au} = \mathbf{A}^T\mathbf{b}$ too.

   Now suppose $\mathbf{u}$ is a solution $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$, that is, $\mathbf{A}^T\mathbf{Au} = \mathbf{A}^T\mathbf{b}$. Let $\mathbf{v}$ be a solution to $\mathbf{Ax} = \mathbf{b}$. Then by the above argument, $\mathbf{v}$ is also a solution to $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$. Using the same argument as above, we have $\mathbf{u} - \mathbf{v} \in \mathrm{Null}(\mathbf{A})$, and hence,

   $$\mathbf{Au} = \mathbf{Av} = \mathbf{b}.$$

   Alternatively, recall that solving for $\mathbf{A}^T\mathbf{A} = \mathbf{A}^T\mathbf{b}$ is solving for $\mathbf{Ax} = \mathbf{b}'$, where $\mathbf{b}'$ is the projection of $\mathbf{b}$ onto the column space of $\mathbf{A}$. Since $\mathbf{Ax} = \mathbf{b}$ is consistent, $\mathbf{b} = \mathbf{b}'$ is the projection of itself onto the $\mathrm{Col}(\mathbf{A})$. So, the solution set to $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$ is equal to the solution set to $\mathbf{Ax} = \mathbf{b}' = \mathbf{b}$.

2. (Uniqueness of orthogonal projection)
   Let V be a subspace of $\mathbb{R}^n$ and $\mathbf{u}$ a vector in $\mathbb{R}^n$. Show that $\mathbf{u}$ can be written uniquely as

   $$\mathbf{u} = \mathbf{u}_p + \mathbf{u}_n,$$

   such that $\mathbf{u}_n$ is a vector orthogonal to V and $\mathbf{u}_p$ is a vector in V. Remark: By the Gram-Schmidt process, $\mathbf{u}$ can be written as $\mathbf{u} = \mathbf{u}_p + \mathbf{u}_n$. We need to prove show that $\mathbf{u}_p$ and $\mathbf{u}_n$ are unique. Hint: if $\mathbf{u} = \mathbf{u}_p + \mathbf{u}_n = \mathbf{u}_p' + \mathbf{u}_n'$, where $\mathbf{u}_n, \mathbf{u}_n'$ are orthogonal to V and $\mathbf{u}_p, \mathbf{u}_p' \in V$, then $\mathbf{u}_n = \mathbf{u}_n'$ and $\mathbf{u}_p = \mathbf{u}_p'$).

Solution: From the identity given in the hint, we have

$$\mathbf{u}_p - \mathbf{u}_p = \mathbf{u}'_n - \mathbf{u}_n. \tag{1}$$

The vector on the left hand side belongs to V and the vector on the right hand side is orthogonal to V. This means that

$$0 = (\mathbf{u}_p - \mathbf{u}_p) \cdot (\mathbf{u}'_n - \mathbf{u}_n) = (\mathbf{u}_p - \mathbf{u}_p) \cdot (\mathbf{u}_p - \mathbf{u}_p),$$

where the second equality follows from the identity in 1. But this means that

$$\mathbf{0} = \mathbf{u}_p - \mathbf{u}_p = \mathbf{u}'_n - \mathbf{u}_n,$$

which proves that $\mathbf{u}_n = \mathbf{u}'_n$ and $\mathbf{u}_p = \mathbf{u}'_p$.

3. Let $S = \{\mathbf{w}_1, ..., \mathbf{w}_k\}$ be an orthonormal basis for a subspace V in $\mathbb{R}^n$. Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in V.

   (a) Prove that $\mathbf{u} \cdot \mathbf{v} = [\mathbf{u}]_S \cdot [\mathbf{v}]_S$.

   Solution: Let $\mathbf{u} = c_1\mathbf{w}_1 + \cdots c_k\mathbf{w}_k$ and $\mathbf{v} = d_1\mathbf{w}_1 + \cdots d_k\mathbf{w}_k$, that is, $[\mathbf{u}]_S = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix}$ and $[\mathbf{v}]_S = \begin{pmatrix} d_1 \\ \vdots \\ d_k \end{pmatrix}$. Then

$$\begin{aligned}
\mathbf{u} \cdot \mathbf{v} &= (c_1\mathbf{w}_1 + \cdots c_k\mathbf{w}_k) \cdot (d_1\mathbf{w}_1 + \cdots d_k\mathbf{w}_k) = \sum_{i,j} c_i d_j \mathbf{w}_i \cdot \mathbf{w}_j \\
&= c_1 d_1 \mathbf{w}_1 \cdot \mathbf{w}_1 + \cdots + c_1 d_k \mathbf{w}_1 \cdot \mathbf{w}_k + c_2 d_1 \mathbf{w}_2 \cdot \mathbf{w}_1 + \cdots + c_k d_k \mathbf{w}_k \cdot \mathbf{w}_k \\
&= c_1 d_1 \mathbf{w}_1 \cdot \mathbf{w}_1 + c_2 d_2 \mathbf{w}_2 \cdot \mathbf{w}_2 + \cdots + c_k d_k \mathbf{w}_k \cdot \mathbf{w}_k \\
&= c_1 d_1 + c_2 d_2 + \cdots + c_k d_k = [\mathbf{u}]_S \cdot [\mathbf{v}]_S,
\end{aligned}$$

   where the fourth equality follows from the fact that $\mathbf{w}_j \cdot \mathbf{w}_j = 0$ for all $i \neq j$, and the fifth equality follows from the fact that $\mathbf{w}_i \cdot \mathbf{w}_i = 1$ for all $i = 1, ..., k$.

   (b) Prove that $\|[\mathbf{u}]_S\| = \|\mathbf{u}\|$.

   Solution: Replace $\mathbf{v}$ with $\mathbf{u}$ in (a), we have $\mathbf{u} \cdot \mathbf{u} = [\mathbf{u}]_S \cdot [\mathbf{u}]_S$. Hence, $\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{[\mathbf{u}]_S \cdot [\mathbf{u}]_S} = \|[\mathbf{u}]_S\|$.

   (c) Prove that the angle between $\mathbf{u}$ and $\mathbf{v}$ is equal to the angle between $[\mathbf{u}]_S$ and $[\mathbf{v}]_S$.

Solution: From (a) and (b), we have

$$\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = \frac{[\mathbf{u}]_S\cdot[\mathbf{v}]_S}{\|[\mathbf{u}]_S\|\|[\mathbf{v}]_S\|}.$$

Since cos is one-to-one on $[0, \pi]$,

$$\cos^{-1}\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = \cos^{-1}\frac{[\mathbf{u}]_S\cdot[\mathbf{v}]_S}{\|[\mathbf{u}]_S\|\|[\mathbf{v}]_S\|},$$

that is, the angle between $\mathbf{u}$ and $\mathbf{v}$ is equal to the angle between $[\mathbf{u}]_S$ and $[\mathbf{v}]_S$.