Example:

The following are the vectors of $\mathbb{R}^n$. Then which mark the each statement true or false.

a. If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis for W, then multiplying $\mathbf{v}_3$ by a scalar c gives a new orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, c\mathbf{v}_3\}$.

b. The Gram–Schmidt process produces from a linearly independent set $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ an orthogonal set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ with the property that for each k, the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ span the same subspace as that spanned by $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

c. If $A = QR$, where Q has orthonormal columns, then $R = Q^T A$.

Example:

The following are the vectors of $\mathbb{R}^n$. Then which mark the each statement true or false.

a. If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis for W, then multiplying $\mathbf{v}_3$ by a scalar c gives a new orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, c\mathbf{v}_3\}$.

b. The Gram–Schmidt process produces from a linearly independent set $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ an orthogonal set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ with the property that for each k, the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ span the same subspace as that spanned by $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

c. If $A = QR$, where Q has orthonormal columns, then $R = Q^T A$.

Solution:

(a) False
(b) True
(c) True

Example:

The following are the vectors of $\mathbb{R}^n$. Then which mark the each statement true or false.

a. If $W = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ with $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ linearly independent, and if $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set in W, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for W.

b. If $\mathbf{x}$ is not in a subspace W, then $\mathbf{x} - \text{proj}_W \mathbf{x}$ is not zero.

c. In a QR factorization, say $A = QR$ (when A has linearly independent columns), the columns of Q form an orthonormal basis for the column space of A.

Example:

The following are the vectors of $\mathbb{R}^n$. Then which mark the each statement true or false.

a. If $W = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ with $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ linearly independent, and if $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set in W, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for W.

b. If $\mathbf{x}$ is not in a subspace W, then $\mathbf{x} - \text{proj}_W\, \mathbf{x}$ is not zero.

c. In a QR factorization, say $A = QR$ (when A has linearly independent columns), the columns of Q form an orthonormal basis for the column space of A.

Solution:

(a) False
(b) True
(c) True

Matrix Factorization:

There are a couple of matrix factorizations, also called decomposition, that every Data Scientist should be very familiar with. These are important because they help find methods for actually computing and estimating results for the models and algorithms we use.

The ***QR*** factorization is one of these matrix factorizations that is very useful and has very important applications in Data Science, Statistics, and Data Analysis. One of these applications is the computation of the solution to the Least Squares (LS) Problem.

The QR Factorization

If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an orthonormal basis for Col A and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

Example:

Find the QR decomposition of the following matrix

$$A = \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix}$$

Solution:

Since A has independent column, then by QR decomposition, we can decompose A into product of two matrices,

$$A = QR,$$

where, Q is orthonormal matrix (orthonormal basis for the column space of A),

R is upper triangular matrix with positive entries of the diagonal.

Let us calculate the matrix Q.

Before, let us calculate orthogonal basis of the column space of A.

By using Gram-Schmidt orthogonalization process:

$\{u_1, u_2\}$ is the orthogonal basis of the column space of A.

$$u_1 = x_1, u_2 = x_2 - \left(\frac{x_2 \cdot u_1}{u_1 \cdot u_1}\right) u_1$$

$$x_1 = \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 9 \\ 7 \\ -5 \\ 5 \end{bmatrix}$$

$$u_1 = x_1 = \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix}$$

$$x_2 . u_1 = \begin{bmatrix} 9 \\ 7 \\ -5 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} = 45 + 7 + 15 + 5 = 72$$

$$u_1 \cdot u_1 = \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} = 25 + 1 + 9 + 1 = 36$$

$$u_2 = \begin{bmatrix} 9 \\ 7 \\ -5 \\ 5 \end{bmatrix} - \left(\frac{72}{36}\right) \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 9 \\ 7 \\ -5 \\ 5 \end{bmatrix} - 2 \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 5 \\ 1 \\ 3 \end{bmatrix}$$

Orthogonal basis= $\left\{ \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 1 \\ 3 \end{bmatrix} \right\}$

Orthonormal vectors are $v_1 = \frac{u_1}{\|u_1\|}, v_2 = \frac{u_2}{\|u_2\|}$

$\|u_1\| = \sqrt{25 + 1 + 9 + 1} = 6$

$\|u_2\| = \sqrt{1 + 25 + 1 + 9} = 6$

$$v_1 = \frac{1}{6}\begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix}, v_2 = \frac{1}{6}\begin{bmatrix} -1 \\ 5 \\ 1 \\ 3 \end{bmatrix}$$

$\{v_1, v_2\}$ forms an orthonormal basis.

Therefore, $Q = \begin{bmatrix} \frac{5}{6} & -\frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} \\ -\frac{3}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{3}{6} \end{bmatrix}$.

Since, we have $A = QR$,

Multiplying Q^T on both sides, we have

$Q^T A = Q^T Q R \Rightarrow R = Q^T A$

$$R = Q^T A$$

$$R = \begin{bmatrix} \frac{5}{6} & \frac{1}{6} & -\frac{3}{6} & \frac{1}{6} \\ -\frac{1}{6} & \frac{5}{6} & \frac{1}{6} & \frac{3}{6} \end{bmatrix} \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 6 & 12 \\ 0 & 6 \end{bmatrix}$$

$$\text{Therefore,} \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} \frac{5}{6} & -\frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} \\ -\frac{3}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{3}{6} \end{bmatrix} \begin{bmatrix} 6 & 12 \\ 0 & 6 \end{bmatrix}.$$

Example:

Find a QR factorization of $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

Solution:
Home work.

Hint:
$\{u_1, u_2, u_3\}$ is the orthogonal basis for the column space of the matrix A.

$$A = [x_1\ x_2\ x_3]$$

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

$$u_1 = x_1$$

$$u_2 = x_2 - \left(\frac{x_2 \cdot u_1}{u_1 \cdot u_1}\right) u_1$$

$$u_3 = x_3 - \left(\frac{x_3 \cdot u_1}{u_1 \cdot u_1}\right) u_1 - \left(\frac{x_3 \cdot u_2}{u_2 \cdot u_2}\right) u_2$$

do the rest ....

Note:

1) To have solution for the system $A\mathbf{x} = \mathbf{b}$, $\boldsymbol{b}$ must be in the column space of A.

Proof: For example consider $A = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix}$, $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is the solution of the system.$A\mathbf{x} = \mathbf{b}$

$$\Rightarrow \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

It can be written as follows

$$\begin{bmatrix} a_1 \\ a_4 \\ a_7 \end{bmatrix} x_1 + \begin{bmatrix} a_2 \\ a_5 \\ a_8 \end{bmatrix} x_2 + \begin{bmatrix} a_3 \\ a_6 \\ a_9 \end{bmatrix} x_3 = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$\Rightarrow c_1x_1 + c_2x_2 + c_3x_3 = b$, where c_1, c_2, c_3 are the columns of the matrix A.

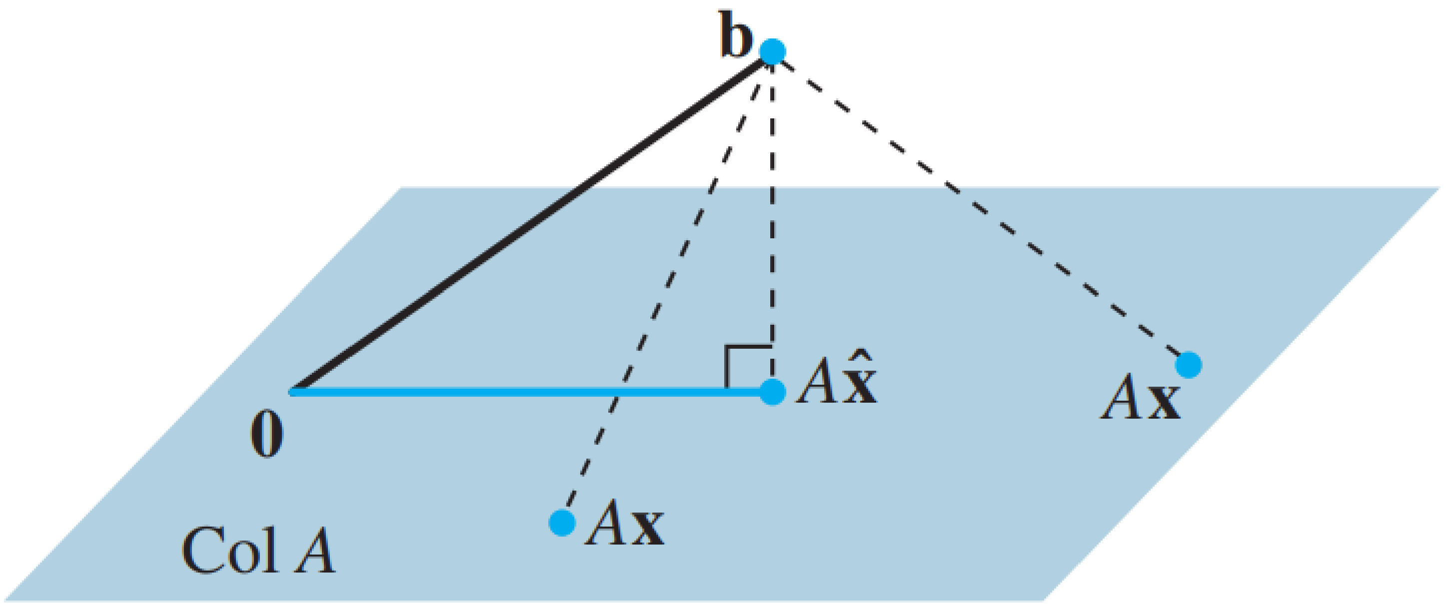
From the expression it is clear that the matrix $\boldsymbol{b}$ expressed as the linear combination of the columns of the matrix A. That is, we can say that $\boldsymbol{b}$ is in the column space of A.

LEAST-SQUARES PROBLEMS

If A is $m \times n$ and $\mathbf{b}$ is in $\mathbb{R}^m$, a **least-squares solution** of $A\mathbf{x} = \mathbf{b}$ is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

for all $\mathbf{x}$ in $\mathbb{R}^n$.



Example:

Find a least-squares solution of $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

Also, determine the least-squared error in the least square solution of the given system.

Solution:

Clearly is it can be verified that $\boldsymbol{b}$ is not in the column space of A.
Therefore, $A\mathbf{x} = \mathbf{b}$ is inconsistent.

We can find the least square solution for the given system.

First we will take the projection of $\boldsymbol{b}$ on to the column space of A. We denote that with $\widehat{\boldsymbol{b}}$.

Let c_1, c_2 are the columns of the matrix A.

$\hat{b}$ can be written as $\hat{b} = Proj_A b$.

$$\hat{b} = \left(\frac{b \cdot c_1}{c_1 \cdot c_1}\right) c_1 + \left(\frac{b \cdot c_2}{c_2 \cdot c_2}\right) c_2$$

$$b = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}, c_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, c_2 = \begin{bmatrix} -6 \\ -2 \\ 1 \\ 7 \end{bmatrix}$$

$$b \cdot c_1 = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = -1 + 2 + 1 + 6 = 8; \qquad c_1 \cdot c_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 1 + 1 + 1 + 1 = 4$$

$$b \cdot c_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} -6 \\ -2 \\ 1 \\ 7 \end{bmatrix} = 6 - 4 + 1 + 42 = 45; \qquad \mathrm{c}_2 \cdot c_2 = \begin{bmatrix} -6 \\ -2 \\ 1 \\ 7 \end{bmatrix} \cdot \begin{bmatrix} -6 \\ -2 \\ 1 \\ 7 \end{bmatrix} = 36 + 4 + 1 + 49 = 90$$

$$\hat{b} = \frac{8}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} + \frac{90}{45}\begin{bmatrix}-6\\-2\\1\\7\end{bmatrix} = 2\begin{bmatrix}1\\1\\1\\1\end{bmatrix} + 2\begin{bmatrix}-6\\-2\\1\\7\end{bmatrix} = \begin{bmatrix}-10\\-2\\4\\16\end{bmatrix}$$

Now, solve $A\mathbf{x} = \widehat{\boldsymbol{b}}$

$$\begin{bmatrix}1 & -6\\1 & -2\\1 & 1\\1 & 7\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}-10\\-2\\4\\16\end{bmatrix}$$

Augmented matrix $\begin{bmatrix}1 & -6 & -10\\1 & -2 & -2\\1 & 1 & 4\\1 & 7 & 16\end{bmatrix}$

By Gauss elimination method we need to make the elements below the principal diagonals to zeros.

$R_2 \to R_2 - R_1; R_3 \to R_3 - R_1; R_4 \to R_4 - R_1$

$$\begin{bmatrix} 1 & -6 & -10 \\ 0 & 4 & 8 \\ 0 & 7 & 14 \\ 0 & 13 & 26 \end{bmatrix} \Rightarrow It\ is\ clear\ that\ R_2, R_3, R_4 \text{ are dependent.}$$

Therefore, we have only two L.I. equations.

Which are, $x_1 - 6x_2 = -10$,

$$4x_2 = 8 \Rightarrow x_2 = 2$$

$$7x_2 = 14 \Rightarrow x_2 = 2$$

$$13x_2 = 26 \Rightarrow x_2 = 2$$

$$x_1 = -10 + 6(2) = 2$$

Therefore, the least square solution is $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

The Least square error is $\|b - A\hat{x}\| = \left\| \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix} - \begin{bmatrix} -10 \\ -2 \\ 4 \\ 16 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 9 \\ 4 \\ -3 \\ -10 \end{bmatrix} \right\| = \sqrt{49 + 16 + 9 + 100} = \sqrt{174}.$

Example:

Find a least-squares solution of the inconsistent system $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} \text{ using QR factorization.}$$

Also, determine the least-squared error in the least square solution of the given system.

Solution:

Given $A\mathbf{x} = \mathbf{b}$ is a inconsistent system.

Hint:

Since A has independent columns. Therefore, we can decompose A into

$$A = QR$$

Where, Q is orthonormal matrix (columns of Q forms an orthonormal basis of the column space of A) and R is the upper triangular matrix with positive diagonal elements.

We can find the columns of Q by Grams-Schmidt orthogonalization process.

After finding Q, we can find R by $R = Q^T A$

Given $A\mathbf{x} = \mathbf{b} \Rightarrow (\mathrm{QR})\mathbf{x} = \mathbf{b} \Rightarrow \mathrm{Rx} = \mathrm{Q}^{\mathrm{T}}\mathbf{b}$.

By back substitution, we can find the x.

Normal Equations:

Let $A\mathbf{x} = \mathbf{b}$ be the given system then $A^T A\mathrm{x} = \mathrm{A}^{\mathrm{T}}\mathbf{b}$ is called the normal equations.

Example:

Find a least-squares solution of the inconsistent system $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Solution:

From the normal equations $A^T A\mathrm{x} = \mathrm{A}^{\mathrm{T}}\boldsymbol{b}$

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Then the equation $A^T A\mathbf{x} = A^T\mathbf{b}$ becomes

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Then the equation $A^TA\mathbf{x} = A^T\mathbf{b}$ becomes

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Row operations can be used to solve this system, but since A^TA is invertible and 2×2, it is probably faster to compute

$$(A^TA)^{-1} = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$$

and then to solve $A^TA\mathbf{x} = A^T\mathbf{b}$ as

$$\begin{aligned} \hat{\mathbf{x}} &= (A^TA)^{-1}A^T\mathbf{b} \\ &= \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}\begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84}\begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{aligned} \qquad \blacksquare$$

Example:

Find a least-squares solution of $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}$$ using normal equations.

Also, determine the least-squared error in the least square solution of the given system.