$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}_{m\times n}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1}}_{x} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}_{m\times 1}}_{b}$$

CX1104: **Linear Algebra for Computing**

Chap. No : **6.3.1**

Lecture : **Orthogonality**

Topic : **Gram–Schmidt Process**

Concept : **Motivation and Review of Concepts**

Instructor: **A/P Chng Eng Siong**

TAs: **Zhang Su, Vishal Choudhari**

Rev: 29th June 2020

Motivation

Consider the problem of solving:

$$Ax = b$$

$$\begin{array}{c c}
 & \\
\text{M = 3 Rows} & \begin{bmatrix} 1 & 0 \\ 4 & 3 \\ 2 & 3 \end{bmatrix} \\
 & \text{N = 2 Columns}
\end{array}
\quad
\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
\quad
\begin{bmatrix} 0.4 \\ 2.5 \\ 1.7 \end{bmatrix}$$

Interpretation:

$$A\text{'s Rows} \Longrightarrow M \quad \textit{Examples or Equations}$$

$$A\text{'s Columns} \Longrightarrow N \quad \textit{Features or Unknowns}$$

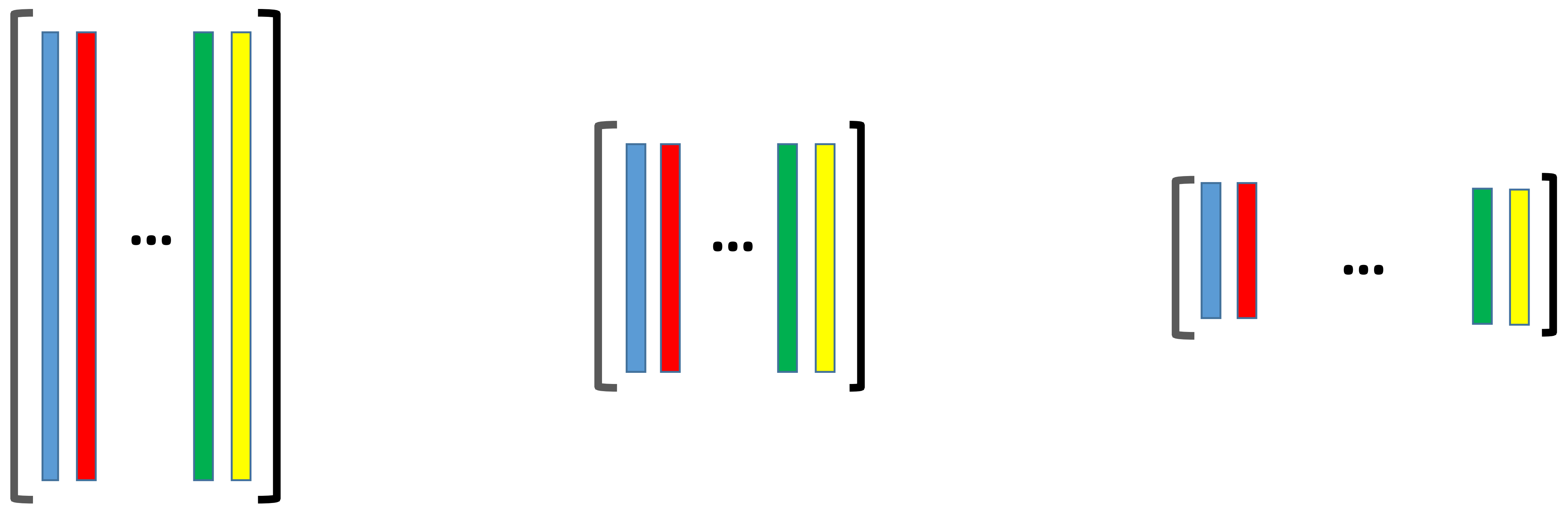
$$x \Longrightarrow \quad \textit{Model} \quad \textit{or Weights}$$

Machine Learning Perspective

$$b \Longrightarrow \textit{Target Values}$$

Goal: Solve x *or find appropriate model*

Based on M & N, there exist three cases:



$M \gg N$ | $M = N$ | $M \ll N$

More equations, less unknowns.

Hence, **over-determined**!

Less equations, more unknowns.

Hence, **under-determined**!

Motivation

What does Gram-Schmidt Process Do?

It orthogonalises a set of vectors!

- A is a nxk matrix
- $k \leq n$ Linearly independent columns
- **Columns are Not orthogonal**
- $\boldsymbol{v_i} \in \boldsymbol{R^n}$, $i = 1..k$

$A = \{v_1, v_2, v_3, \cdots, v_k\}$ → **Gram-Schmidt Process** → $Q = \{u_1, u_2, u_3, \cdots, u_k\}$
R = upper triangle

- **Orthogonal matrix Q**
- **is nxk matrix**
- Orthogonal columns u_i $i = 1..k$, with $k \leq n$
- **and upper triangle R**
- **R is kxk matrix**

$$Span(A) = Span(Q)$$

Note:Q spans the same k-dimensional subspace of R^n as that of A

Applications for Gram-Schmidt Process

1. QR Factorisation:

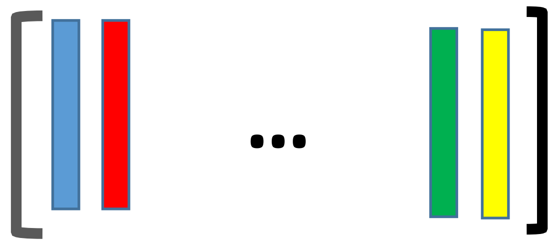
Ref: https://www.quora.com/Why-is-QR-factorization-useful-and-important

For large M and N, the system of equations of the form $Ax = b$ can be solved efficiently if A can be rewritten as:

$A = QR$

where,
Q is an orthogonal matrix (having same col. space of A,
i.e, span (Q) = span (A)) and R is an upper triangular matrix.

2. Feature Selection:



Ref: https://qr.ae/pNKiO5

When $M \ll N$, i.e, when there are more features than examples, features/columns most representative of b can be identified.

$$M \ll N$$

Reviewing finding the solution of x for Ax=b

Matrix A is of dimension $M \times N$.
M Rows, N Columns

Chapter 3. Vector Spaces and Subspaces

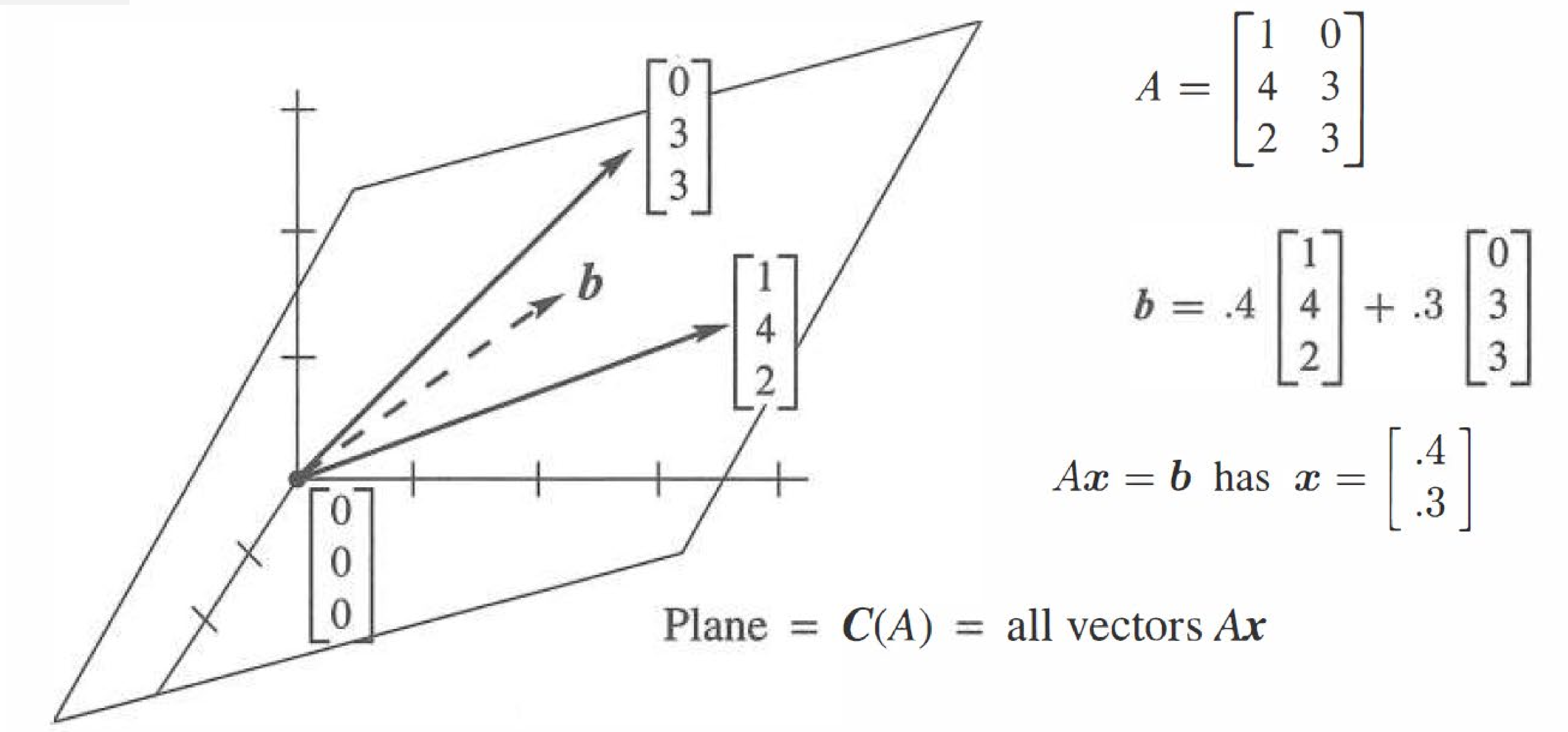


Figure 3.2: The column space $C(A)$ is a plane containing the two columns. $Ax = b$ is solvable when b is on that plane. Then b is a combination of the columns.

In the example above, Matrix A is of dimension 3×2
3 Rows, 2 Columns

If $Ax = b$, then $x \in R^2$, and $b \in R^3$

By b being in a linear combination of columns of A, b lies in the span of column vectors of A.

Note: solving for x is not trivial when
If A does not have orthogonal columns.

QR factorisation, which makes use of GS process, helps decompose A:

$$A = QR$$

where,
Q is an orthogonal matrix (having same col. space of A)
R is an upper triangular matrix.

The problem to find x can then be easily solved by:

$$\begin{aligned} Ax &= b \\ QRx &= b \\ Q^TQRx &= Q^Tb \\ Rx &= Q^Tb \end{aligned}$$

Since R is upper triangle, x can be quickly found by back-substitution.

Note: if b is not in C(A), then the found x will only result in the orthogonal projection of b onto C(A).

Ref: https://www.mathwords.com/b/back_substitution.htm

Ref: Strang, Introduction to Linear Algebra

Reviewing Span

Linear Combinations

Given vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ in $\mathbb{R}^n$ and given scalars $c_1, c_2, \ldots, c_p$, the vector $\mathbf{y}$ defined by

$$\mathbf{y} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$$

is called a **linear combination** of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ with **weights** $c_1, \ldots, c_p$.

One of the key ideas in linear algebra is to study the set of all vectors that can be generated or written as a linear combination of a fixed set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ of vectors.

If $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are in $\mathbb{R}^n$, then the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ is denoted by Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ and is called the **subset of $\mathbb{R}^n$ spanned** (or **generated**) **by** $\mathbf{v}_1, \ldots, \mathbf{v}_p$. That is, Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is the collection of all vectors that can be written in the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

with $c_1, \ldots, c_p$ scalars.

Note that Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ contains every scalar multiple of $\mathbf{v}_1$ (for example), since $c\mathbf{v}_1 = c\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_p$. In particular, the zero vector must be in Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

30 CHAPTER 1 Linear Equations in Linear Algebra

1.3 Vector Equations 27

A Geometric Description of Span $\{\mathbf{v}\}$ and Span $\{\mathbf{u}, \mathbf{v}\}$

Let $\mathbf{v}$ be a nonzero vector in $\mathbb{R}^3$. Then Span $\{\mathbf{v}\}$ is the set of all scalar multiples of $\mathbf{v}$, which is the set of points on the line in $\mathbb{R}^3$ through $\mathbf{v}$ and $\mathbf{0}$. See Fig. 10.

If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in $\mathbb{R}^3$, with $\mathbf{v}$ not a multiple of $\mathbf{u}$, then Span $\{\mathbf{u}, \mathbf{v}\}$ is the plane in $\mathbb{R}^3$ that contains $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{0}$. In particular, Span $\{\mathbf{u}, \mathbf{v}\}$ contains the line in $\mathbb{R}^3$ through $\mathbf{u}$ and $\mathbf{0}$ and the line through $\mathbf{v}$ and $\mathbf{0}$. See Fig. 11.

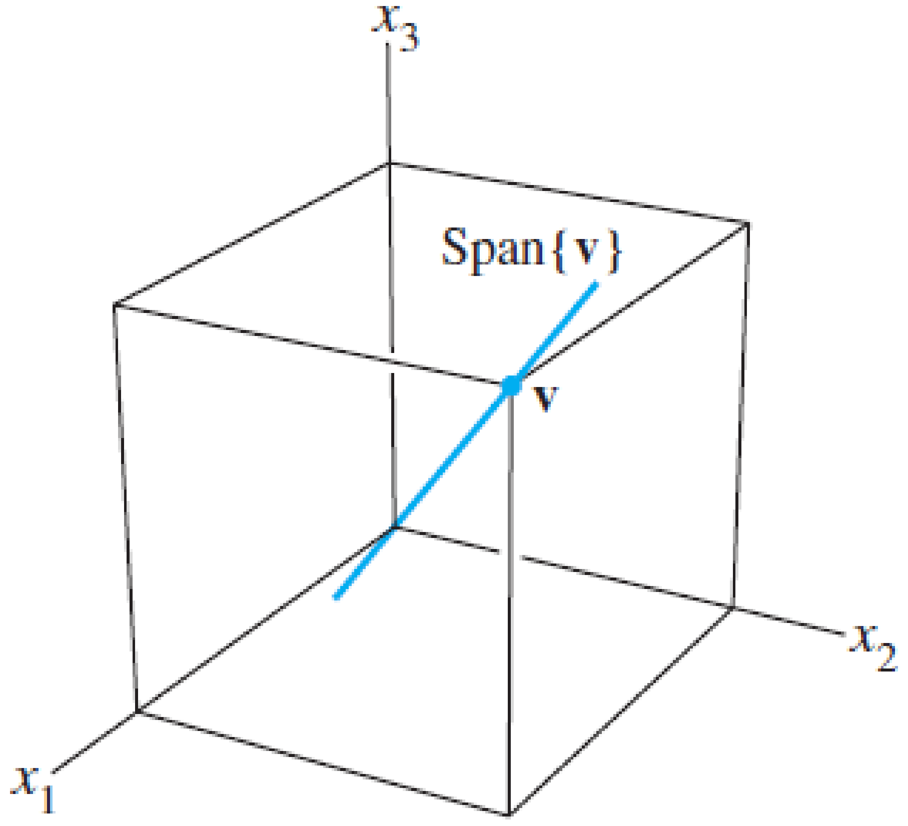


FIGURE 10 Span $\{\mathbf{v}\}$ as a line through the origin.

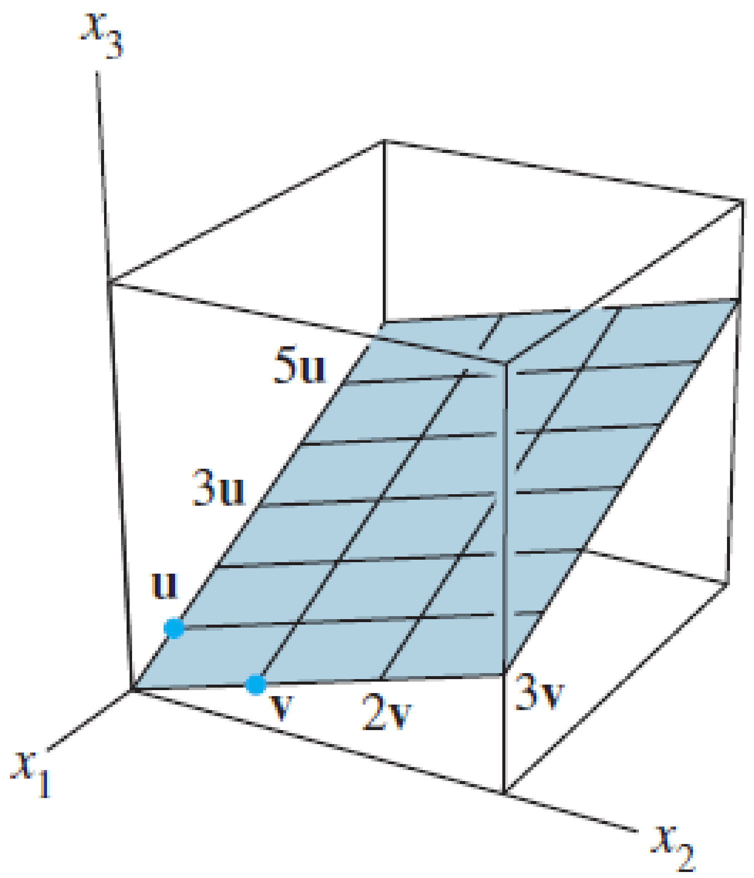


FIGURE 11 Span $\{\mathbf{u}, \mathbf{v}\}$ as a plane through the origin.

Lay, Linear Algebra and its Applications (4th Edition)

4.5 The Dimension of a Vector Space 227

Reviewing Span

EXAMPLE 2 A single linear equation can be treated as a very simple system of equations. Describe all solutions of the homogeneous "system"

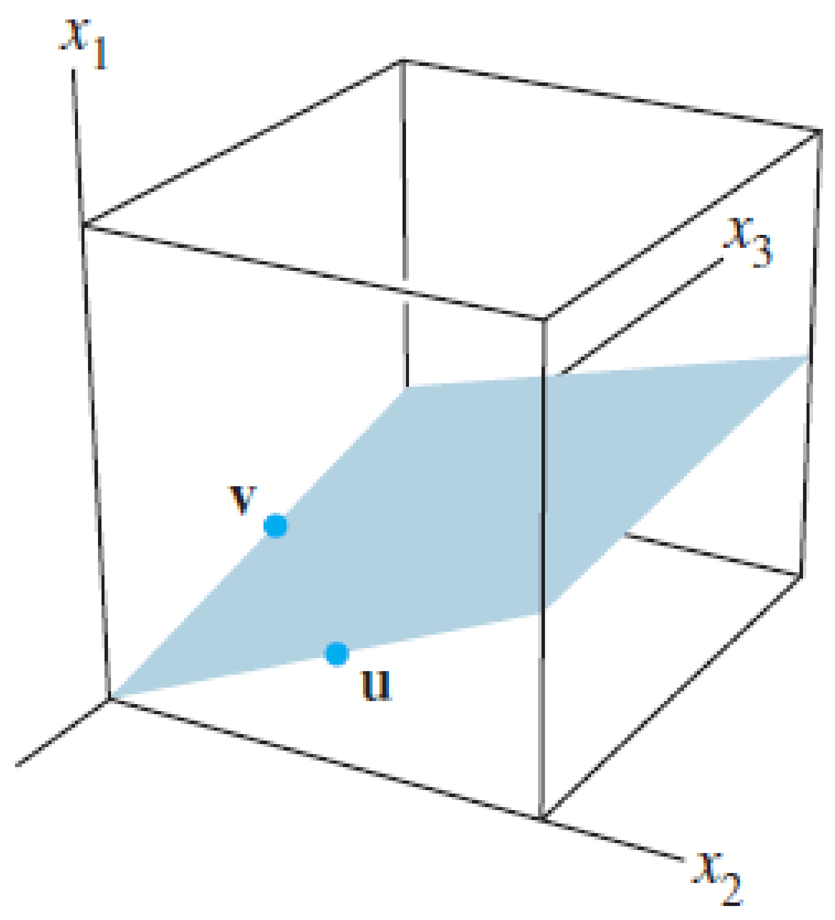
$$10x_1 - 3x_2 - 2x_3 = 0 \qquad (1)$$

SOLUTION There is no need for matrix notation. Solve for the basic variable x_1 in terms of the free variables. The general solution is $x_1 = .3x_2 + .2x_3$, with x_2 and x_3 free. As a vector, the general solution is

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .3x_2 + .2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .3x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} .2x_3 \\ 0 \\ x_3 \end{bmatrix}$$

$$= x_2 \underset{\substack{\uparrow \\ \mathbf{u}}}{\begin{bmatrix} .3 \\ 1 \\ 0 \end{bmatrix}} + x_3 \underset{\substack{\uparrow \\ \mathbf{v}}}{\begin{bmatrix} .2 \\ 0 \\ 1 \end{bmatrix}} \quad \text{(with } x_2, x_3 \text{ free)} \qquad (2)$$

This calculation shows that every solution of (1) is a linear combination of the vectors $\mathbf{u}$ and $\mathbf{v}$, shown in (2). That is, the solution set is Span $\{\mathbf{u}, \mathbf{v}\}$. Since neither $\mathbf{u}$ nor $\mathbf{v}$ is a scalar multiple of the other, the solution set is a plane through the origin. See Fig. 2. ■



Lay, Linear Algebra and its Applications (4th Edition)