Basis and Dimension

Basis

- The vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ form a *basis* for a vector space V if and only if
  - i) $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ are linearly independent
  - ii) $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ span V

- Examples
  - The *standard basis* for $\mathbb{R}^3$ is $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$
  - What about the following vectors?

$$\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}2\\0\\1\end{bmatrix}\right\}, \qquad \left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}\right\}$$

  - In $\mathbb{R}^{2\times 2}$,

$$E_{11} = \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}, \quad E_{12} = \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}, \quad E_{21} = \begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}, \quad E_{22} = \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$$

  - $c_1E_{11} + c_2E_{22} + c_3E_{33} + c_4E_{44} = O \Leftrightarrow c_1 = c_2 = c_3 = c_4 = 0$

Basis

- If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$ is a spanning set for a vector space V, then any collection of m vectors in V, where $m > n$, is linearly dependent

- Proof

Let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ be m vectors in V, where $m > n$. Then, since $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ span V, we have

$$\mathbf{u}_i = a_{i1}\mathbf{v}_1 + a_{i2}\mathbf{v}_2 + \cdots + a_{in}\mathbf{v}_n \quad \text{for} \quad i = 1, 2, \ldots, m$$

A linear combination $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m$ can be written in the form

$$c_1 \sum_{j=1}^{n} a_{1j}\mathbf{v}_j + c_2 \sum_{j=1}^{n} a_{2j}\mathbf{v}_j + \cdots + c_m \sum_{j=1}^{n} a_{mj}\mathbf{v}_j$$

Rearranging the terms, we see that

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m = \sum_{i=1}^{m}\left[c_i\left(\sum_{j=1}^{n} a_{ij}\mathbf{v}_j\right)\right] = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}c_i\right)\mathbf{v}_j$$

Now consider the system of equations

$$\sum_{i=1}^{m} a_{ij}c_i = 0 \qquad j = 1, 2, \ldots, n$$

This is a homogeneous system with more unknowns than equations. Therefore, by Theorem 1.2.1, the system must have a nontrivial solution $(\hat{c}_1, \hat{c}_2, \ldots, \hat{c}_m)^T$. But then

$$\hat{c}_1\mathbf{u}_1 + \hat{c}_2\mathbf{u}_2 + \cdots + \hat{c}_m\mathbf{u}_m = \sum_{j=1}^{n} 0\mathbf{v}_j = \mathbf{0}$$

Hence, $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ are linearly dependent. ■

- If both $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ and $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m\}$ are bases for a vector space V, then $n = m$

- Proof

Let both $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ be bases for V. Since $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ span V and $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ are linearly independent, it follows from Theorem 3.4.1 that $m \leq n$. By the same reasoning, $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ span V and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent, so $n \leq m$. ■

Dimension

- Let V be a vector space. If V has a basis consisting of n vectors, we say that V has *dimension* n
- The subspace $\{\mathbf{0}\}$ of V is said to have dimension 0
- V is said to be *finite dimensional* if there is a finite set of vectors that spans V
- Otherwise, we say that V is infinite dimensional

- Ex 3. Let P be the vector space of all polynomials. P is infinite dimensional.

Let P be the vector space of all polynomials. We claim that P is infinite dimensional. If P were finite dimensional, say, of dimension n, any set of $n+1$ vectors would be linearly dependent. However, $1, x, x^2, \ldots, x^n$ are linearly independent, since $W[1, x, x^2, \ldots, x^n] > 0$. Therefore, P cannot be of dimension n. Since n was arbitrary, P must be infinite dimensional. The same argument shows that $C[a, b]$ is infinite dimensional. ■

Dimension

- If V is a vector space of dimension $n > 0$, then

  i) any set of n linearly independent vectors spans V
  ii) any n vectors that span V are linearly independent

- Proof

To prove (I), suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent and $\mathbf{v}$ is any other vector in V. Since V has dimension n, it has a basis consisting of n vectors and these vectors span V. It follows from Theorem 3.4.1 that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}$ must be linearly dependent. Thus, there exist scalars $c_1, c_2, \ldots, c_n, c_{n+1}$, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n + c_{n+1}\mathbf{v} = \mathbf{0} \tag{1}$$

The scalar c_{n+1} cannot be zero, for then (1) would imply that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly dependent. Hence, (1) can be solved for $\mathbf{v}$:

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n$$

Here, $\alpha_i = -c_i/c_{n+1}$ for $i = 1, 2, \ldots, n$. Since $\mathbf{v}$ was an arbitrary vector in V, it follows that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ span V.

To prove (II), suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ span V. If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly dependent, then one of the $\mathbf{v}_i$'s, say, $\mathbf{v}_n$, can be written as a linear combination of the others. It follows that $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ will still span V. If $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ are linearly dependent, we can eliminate another vector and still have a spanning set. We can continue eliminating vectors in this way until we arrive at a linearly independent spanning set with $k < n$ elements. But this contradicts $\dim V = n$. Therefore, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ must be linearly independent. ∎

Dimension

- If V is a vector space of dimension $n > 0$, then

  i) no set of fewer than n vectors can span V
  ii) any subset of fewer than n linearly independent vectors can be extended to form a basis for V
  iii) any spanning set containing more than n vectors can be pared down to form a basis for V

- Proof

Statement (i) follows by the same reasoning that was used to prove part (I) of Theorem 3.4.3. To prove (ii), suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent and $k < n$. It follows from (i) that $\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ is a proper subspace of V, and hence there exists a vector $\mathbf{v}_{k+1}$ that is in V but not in $\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$. It then follows that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}$ must be linearly independent. If $k + 1 < n$, then, in the same manner, $\{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}\}$ can be extended to a set of $k + 2$ linearly independent vectors. This extension process may be continued until a set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ of n linearly independent vectors is obtained.

To prove (iii), suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_m$ span V and $m > n$. Then, by Theorem 3.4.1, $\mathbf{v}_1, \ldots, \mathbf{v}_m$ must be linearly dependent. It follows that one of the vectors, say, $\mathbf{v}_m$, can be written as a linear combination of the others. Hence, if $\mathbf{v}_m$ is eliminated from the set, the remaining $m - 1$ vectors will still span V. If $m - 1 > n$, we can continue to eliminate vectors in this manner until we arrive at a spanning set containing n elements. ∎

Exercises

11. Let S be the subspace of P_3 consisting of all polynomials of the form $ax^2 + bx + 2a + 3b$. Find a basis for S

Change of Basis

Changing Coordinates in $\mathbb{R}^2$

- The standard basis for $\mathbb{R}^2$ is $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$. Any vector $\boldsymbol{x}$ in $\mathbb{R}^2$ can be expressed as a linear combination

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2$$

- The scalars x_1 and x_2 can be thought of as the *coordinates* of $\boldsymbol{x}$ with respect to the standard basis

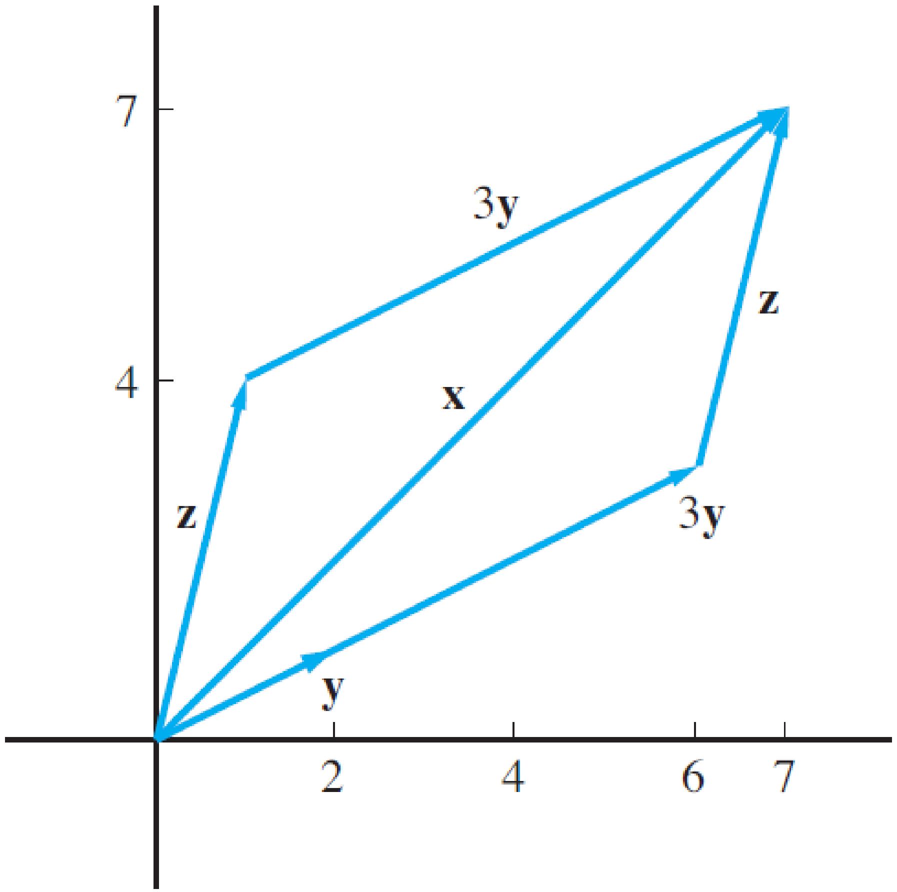
- For any basis $\{\boldsymbol{y}, \boldsymbol{z}\}$ for $\mathbb{R}^2$, a given vector $\boldsymbol{x}$ can be represented uniquely as a linear combination

$$\boldsymbol{x} = \alpha\boldsymbol{y} + \beta\boldsymbol{z}$$

- The vector $(\alpha, \beta)^T$ is the *coordinate vector* of x with respect to $\{\boldsymbol{y}, \boldsymbol{z}\}$

Changing Coordinates in $\mathbb{R}^2$

Ex 1. Let $\boldsymbol{y} = (2, 1)^T$ and $\boldsymbol{z} = (1, 4)^T$. The vectors $\boldsymbol{y}$ and $\boldsymbol{z}$ are linearly independent and hence they form a basis for $\mathbb{R}^2$. The vector $\boldsymbol{x} = (7, 7)^T$ can be written as a linear combination $\boldsymbol{x} = 3\boldsymbol{y} + \boldsymbol{z}$. Thus, the coordinate vector of $\boldsymbol{x}$ with respect to $[\boldsymbol{y}, \boldsymbol{z}]$ is $(3, 1)^T$.



Application: Population Migration

- Suppose that the total population of a large metropolitan area remains relatively fixed
- Each year 6% of the people living in the city move to the suburbs and 2% of the people living in the suburbs move to the city
- Initially, 30% of the population lives in the city and 70% lives in the suburbs
- The changes in population can be determined by matrix multiplications

$$A = \begin{bmatrix} 0.94 & 0.02 \\ 0.06 & 0.98 \end{bmatrix}, \qquad \boldsymbol{x}_0 = \begin{bmatrix} 0.30 \\ 0.70 \end{bmatrix}$$

$$\boldsymbol{x}_{10} = \begin{bmatrix} 0.27 \\ 0.73 \end{bmatrix}, \qquad \boldsymbol{x}_{30} = \begin{bmatrix} 0.25 \\ 0.75 \end{bmatrix}, \qquad \boldsymbol{x}_{50} = \begin{bmatrix} 0.25 \\ 0.75 \end{bmatrix}$$

- As n increases, the sequence of vectors $\boldsymbol{x}_n = A^n \boldsymbol{x}_0$ converges to a limit $\boldsymbol{x} = (0.25, 0.75)^T$
- The limit vector $\boldsymbol{x}$ is called a *steady-state vector* for the process

Application: Population Migration

- Pick $\boldsymbol{u}_1$ to be any multiple of the steady-state vector $\boldsymbol{x}$
- Pick $\boldsymbol{u}_2$ where the effect of multiplying by A is just to scale the vector by a constant factor

$$A\boldsymbol{u}_1 = \begin{bmatrix} 0.94 & 0.02 \\ 0.06 & 0.98 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \boldsymbol{u}_1 \qquad A\boldsymbol{u}_2 = \begin{bmatrix} 0.94 & 0.02 \\ 0.06 & 0.98 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.92 \\ 0.92 \end{bmatrix} = 0.92\boldsymbol{u}_2$$

$$\boldsymbol{x}_0 = \begin{bmatrix} 0.30 \\ 0.70 \end{bmatrix} = 0.25\boldsymbol{u}_1 - 0.05\boldsymbol{u}_2$$

$$\boldsymbol{x}_n = A^n\boldsymbol{x}_0 = 0.25\boldsymbol{u}_1 - 0.05(0.92)^n\boldsymbol{u}_2$$

- A model called a *Markov process*
- The sequence of vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots$ is called a *Markov chain*
- The matrix A is a *stochastic matrix* where its entries are nonnegative and its columns all add up to one

Changing Coordinates

- Once we have decided to work with a new basis, we have the problem of finding the coordinates with respect to that basis

  i) Given a vector $\boldsymbol{x} = (x_1, x_2)^T$, find its coordinates with respect to $\boldsymbol{u}_1 = (3, 2)^T$ and $\boldsymbol{u}_2 = (1, 1)^T$
  ii) Given a vector $c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2$, find its coordinates with respect to $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$

$$\boldsymbol{u}_1 = 3\boldsymbol{e}_1 + 2\boldsymbol{e}_2 \qquad \boldsymbol{u}_2 = \boldsymbol{e}_1 + \boldsymbol{e}_2$$

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 = (3c_1 + c_2)\boldsymbol{e}_1 + (2c_1 + c_2)\boldsymbol{e}_2 = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 = \boldsymbol{x}$$

- Given any coordinate vector $\boldsymbol{c}$ with respect to $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, to find the corresponding coordinate vector $\boldsymbol{x}$ with respect to $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$, we simply multiply U times $\boldsymbol{c}$

$$\boldsymbol{x} = \begin{bmatrix} 3c_1 + c_2 \\ 2c_1 + c_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = [\boldsymbol{u}_1 \quad \boldsymbol{u}_2] \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = U \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

- The matrix U is called the transition matrix from the ordered basis $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ to the standard basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$

Changing Coordinates

- The matrix U is nonsingular since its column vectors are linearly independent

- The transition matrix from $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ to $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ is

$$\boldsymbol{c} = U^{-1}\boldsymbol{x}$$

- Given a vector $\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2$, we need to only multiply by U^{-1} to find its coordinate vector with respect to $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$

- Given two ordered bases $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ and $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ in $\mathbb{R}^2$

$$\boldsymbol{x} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 = d_1\boldsymbol{u}_1 + d_2\boldsymbol{u}_2$$

- If we set $V = [\boldsymbol{v}_1, \boldsymbol{v}_2]$ and $U = [\boldsymbol{u}_1, \boldsymbol{u}_2]$

$$V\boldsymbol{c} = U\boldsymbol{d}$$
$$\boldsymbol{d} = U^{-1}V\boldsymbol{c}$$

$$\begin{tikzcd}
{[\mathbf{v}_1, \mathbf{v}_2]} \arrow[r, "V"] \arrow[dr, "U^{-1}V"'] & {[\mathbf{e}_1, \mathbf{e}_2]} \arrow[d, "U^{-1}"] \\
 & {[\mathbf{u}_1, \mathbf{u}_2]}
\end{tikzcd}$$

Change of Basis for a General Vector Space

- Let V be a vector space and let $E = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$ be an ordered basis for V. If $\boldsymbol{v}$ is any element of V, then $\boldsymbol{v}$ can be written in the form

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n$$

  where $c_1, c_2, \cdots, c_n$ are scalars

- We can associate with each vector $\boldsymbol{v}$ a unique vector $\boldsymbol{c} = (c_1, c_2, \cdots, c_n)^T$ in $\mathbb{R}^n$. The vector $\boldsymbol{c}$ is called the *coordinate vector* of $\boldsymbol{v}$ with respect to the ordered basis E and is denoted $[\boldsymbol{v}]_E$

- The c_i's are called the *coordinates* of $\boldsymbol{v}$ relative to E

Change of Basis for a General Vector Space

- If V is any n-dimensional vector space, it is possible to change from one basis to another by means of an $n \times n$ transition matrix

- Let $E = \{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n\}$ and $F = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ be two ordered bases for V

$$\begin{aligned} \boldsymbol{w}_1 &= s_{11}\boldsymbol{v}_1 + s_{21}\boldsymbol{v}_2 + \cdots + s_{n1}\boldsymbol{v}_n \\ \boldsymbol{w}_2 &= s_{12}\boldsymbol{v}_1 + s_{22}\boldsymbol{v}_2 + \cdots + s_{n2}\boldsymbol{v}_n \\ &\vdots \\ \boldsymbol{w}_n &= s_{1n}\boldsymbol{v}_1 + s_{2n}\boldsymbol{v}_2 + \cdots + s_{nn}\boldsymbol{v}_n \end{aligned}$$

- Let $\boldsymbol{v} \in V$. If $\boldsymbol{x} = [\boldsymbol{v}]_E$,

$$\boldsymbol{v} = x_1\boldsymbol{w}_1 + x_2\boldsymbol{w}_2 + \cdots + x_n\boldsymbol{w}_n = \left(\sum_{j=1}^{n} s_{1j}x_j\right)\boldsymbol{v}_1 + \left(\sum_{j=1}^{n} s_{2j}x_j\right)\boldsymbol{v}_2 + \cdots + \left(\sum_{j=1}^{n} s_{nj}x_j\right)\boldsymbol{v}_n$$

- If $\boldsymbol{y} = [\boldsymbol{v}]_F$, then $y_i = \sum_{j=1}^{n} s_{ij}x_j$, and hence, $\boldsymbol{y} = S\boldsymbol{x} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nn} \end{bmatrix}\boldsymbol{x}$

- The matrix S is referred to as the transition matrix

Exercises

11. Let $E = \{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n\}$ and $F = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ be two ordered bases for $\mathbb{R}^n$, and set

$$U = (\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n) \qquad V = (\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n)$$

Show that the transition matrix from E to F can be determined by calculating the reduced row echelon form of $(V|U)$

Row Space and Column Space

Row Space and Column Space

- If A is an $m \times n$ matrix, each row of A is an n-tuple of real numbers and hence can be considered as a vector in $\mathbb{R}^{1\times n}$
- The m vectors corresponding to the rows of A will be referred to as the *row vectors* of A
- Similarly, each column of A can be considered as a vector in $\mathbb{R}^m$, and we can associate n *column vectors* with the matrix A

- If A is an $m \times n$ matrix, the subspace of $\mathbb{R}^{1\times n}$ spanned by the row vectors of A is called the *row space* of A
- The subspace of $\mathbb{R}^m$ spanned by the column vectors of A is called the *column space* of A

Row Space and Column Space

Ex 5. Let

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

The row space of A is the set of all 3-tuples of the form

$$\alpha(1,0,0) + \beta(0,1,0) = (\alpha, \beta, 0)$$

The column space of A is the set of all vectors of the form

$$\alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \gamma \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

The row space of A is two-dimensional subspace of $\mathbb{R}^{1\times 3}$, and the column space of A is $\mathbb{R}^2$

Rank

- Two row equivalent matrices have the same row space

- Proof

If B is row equivalent to A, then B can be formed from A by a finite sequence of row operations. Thus, the row vectors of B must be linear combinations of the row vectors of A. Consequently, the row space of B must be a subspace of the row space of A. Since A is row equivalent to B, by the same reasoning, the row space of A is a subspace of the row space of B. ■

- The *rank* of a matrix A, denoted $\text{rank}(A)$, is the dimension of the row space of A

Ex 2.

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 2 & -5 & 1 \\ 1 & -4 & -7 \end{bmatrix}, \qquad U = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{bmatrix} \qquad \therefore \text{rank}(A) = 2$$

Consistency Theorem for Linear Systems

- A linear system $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent if and only if $\boldsymbol{b}$ is in the column space of A

Consistency Theorem for Linear Systems

A linear system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ can be written as a linear combination of the column vectors of A.

- Let A be an $m \times n$ matrix. The linear system $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent for every $\boldsymbol{b} \in \mathbb{R}^m$ if and only if the column vectors of A span $\mathbb{R}^m$. The system $\boldsymbol{Ax} = \boldsymbol{b}$ has at most one solution for every $\boldsymbol{b} \in \mathbb{R}^m$ if and only if the column vectors of A are linearly independent

We have seen that the system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in the column space of A. It follows that $A\mathbf{x} = \mathbf{b}$ will be consistent for every $\mathbf{b} \in \mathbb{R}^m$ if and only if the column vectors of A span $\mathbb{R}^m$. To prove the second statement, note that, if $A\mathbf{x} = \mathbf{b}$ has at most one solution for every $\mathbf{b}$, then, in particular, the system $A\mathbf{x} = \mathbf{0}$ can have only the trivial solution, and hence the column vectors of A must be linearly independent. Conversely, if the column vectors of A are linearly independent, $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Now, if $\mathbf{x}_1$ and $\mathbf{x}_2$ were both solutions of $A\mathbf{x} = \mathbf{b}$, then $\mathbf{x}_1 - \mathbf{x}_2$ would be a solution of $A\mathbf{x} = \mathbf{0}$,

$$A(\mathbf{x}_1 - \mathbf{x}_2) = A\mathbf{x}_1 - A\mathbf{x}_2 = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

It follows that $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$, and hence $\mathbf{x}_1$ must equal $\mathbf{x}_2$. ■

- An $n \times n$ matrix A is nonsingular if and only if the column vectors of A form a basis for $\mathbb{R}^n$

Rank-Nullity Theorem

- The dimension of the null space of a matrix is called the *nullity* of the matrix
- If A is an $m \times n$ matrix, then the rank of A plus the nullity of A equals n
- Proof

Let U be the reduced row echelon form of A. The system $A\mathbf{x} = \mathbf{0}$ is equivalent to the system $U\mathbf{x} = \mathbf{0}$. If A has rank r, then U will have r nonzero rows, and consequently the system $U\mathbf{x} = \mathbf{0}$ will involve r lead variables and $n-r$ free variables. The dimension of $N(A)$ will equal the number of free variables. ■

Rank-Nullity Theorem

Ex 3. Let

$$A = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 4 & -3 & 0 \\ 1 & 2 & 1 & 5 \end{pmatrix}$$

Find a basis for the row space of A and a basis for $N(A)$. Verify that $\dim N(A) = n - r$.

The reduced row echelon form of A is given by

$$U = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus, $\{(1, 2, 0, 3), (0, 0, 1, 2)\}$ is a basis for the row space of A, and A has rank 2. Since the systems $A\mathbf{x} = \mathbf{0}$ and $U\mathbf{x} = \mathbf{0}$ are equivalent, it follows that $\mathbf{x}$ is in $N(A)$ if and only if

$$\begin{aligned} x_1 + 2x_2 + 3x_4 &= 0 \\ x_3 + 2x_4 &= 0 \end{aligned}$$

The lead variables x_1 and x_3 can be solved for in terms of the free variables x_2 and x_4:

$$\begin{aligned} x_1 &= -2x_2 - 3x_4 \\ x_3 &= -2x_4 \end{aligned}$$

Let $x_2 = \alpha$ and $x_4 = \beta$. It follows that $N(A)$ consists of all vectors of the form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2\alpha - 3\beta \\ \alpha \\ -2\beta \\ \beta \end{bmatrix} = \alpha \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -3 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

The vectors $(-2, 1, 0, 0)^T$ and $(-3, 0, -2, 1)^T$ form a basis for $N(A)$. Note that

$$n - r = 4 - 2 = 2 = \dim N(A)$$

■

Column Space

- If A is an $m \times n$ matrix, the dimension of the row space of A equals the dimension of the column space of A

- Proof

If A is an $m \times n$ matrix of rank r, the row echelon form U of A will have r leading 1's. The columns of U corresponding to the leading 1's will be linearly independent. They do not, however, form a basis for the column space of A, since, in general, A and U will have different column spaces. Let U_L denote the matrix obtained from U by deleting all the columns corresponding to the free variables. Delete the same columns from A and denote the new matrix by A_L. The matrices A_L and U_L are row equivalent. Thus, if $\mathbf{x}$ is a solution of $A_L\mathbf{x} = \mathbf{0}$, then $\mathbf{x}$ must also be a solution of $U_L\mathbf{x} = \mathbf{0}$. Since the columns of U_L are linearly independent, $\mathbf{x}$ must equal $\mathbf{0}$. It follows from the remarks preceding Theorem 3.6.3 that the columns of A_L are linearly independent. Since A_L has r columns, the dimension of the column space of A is at least r.

We have proved that, for any matrix, the dimension of the column space is greater than or equal to the dimension of the row space. Applying this result to the matrix A^T, we see that

$$\begin{aligned} \dim(\text{row space of A}) &= \dim(\text{column space of } A^T) \\ &\geq \dim(\text{row space of } A^T) \\ &= \dim(\text{column space of } A) \end{aligned}$$

Thus, for any matrix A, the dimension of the row space must equal the dimension of the column space.

- r tells us three things about A and U
  - The number of leading 1's of U
  - The number of nonzero rows of A
  - The number of pivot columns of U

Let A be an $m \times n$ matrix. If the column vectors of A span $\mathbb{R}^m$, then n must be greater than or equal to m, since no set of fewer than m vectors could span $\mathbb{R}^m$. If the columns of A are linearly independent, then n must be less than or equal to m, since every set of more than m vectors in $\mathbb{R}^m$ is linearly dependent. Thus, if the column vectors of A form a basis for $\mathbb{R}^m$, then n must equal m.

$$\therefore \dim(rowsp(A)) \leq m, \qquad \dim(colsp(A)) \leq n \qquad \text{(i)}$$

- Since we only deleted columns to generate A_L and U_L, the number of rows m is unchanged

$$\dim(rowsp(A)) = r \overset{\text{(i)}}{\leq} m \leq \dim(colsp(A)) \overset{\text{(i)}}{\leq} n$$

$$\therefore \dim(rowsp(A)) \leq \dim(colsp(A)) \qquad \text{(ii)}$$

- Applying this result to the matrix A^T,

$$m \geq \dim(rowsp(A)) = \dim(colsp(A^T)) \geq \dim(rowsp(A^T)) = \dim(colsp(A))$$

$$\therefore \dim(rowsp(A)) \geq \dim(colsp(A)) \qquad \text{(iii)}$$

- By (ii) and (iii),

$$\dim(rowsp(A)) = \dim(colsp(A))$$

Column Space

- The row echelon form U tells us only which columns of A to use to form a basis
- We cannot use the column vectors from U, since, in general, U and A have different column spaces

Ex 4.

$$A = \begin{bmatrix} 1 & -2 & 1 & 1 & 2 \\ -1 & 3 & 0 & 2 & -2 \\ 0 & 1 & 1 & 3 & 4 \\ 1 & 2 & 5 & 13 & 5 \end{bmatrix}, \qquad U = \begin{bmatrix} 1 & -2 & 1 & 1 & 2 \\ 0 & 1 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

a_1, a_2, a_5 form a basis for the column space of A

Exercises

11. Let A be an $m \times n$ matrix. Prove that

$$\operatorname{rank}(A) \leq \min(m, n)$$

Thank You