Introduction to Mathematics for Political Science: Linear Spaces

30 August 2018

Suppose we observe some features of N countries related to their level of social and economic development. Each observation (a country) can be represented as a vector, where each element corresponds to a unique feature (life expectancy, per capita gdp, literacy rate, etc). Together, these countries constitute a set. Now suppose, due to poor data quality, we need to predict a country's per capita gdp after observing its life expectancy and literacy rate. By assuming that these data live in a broader *space*, we can extrapolate out of our set and "fill in the blanks." We only observe N countries, but we can consider a broader universe of countries that might have different features than the ones we observe. What allows us to do this extrapolation is the assumption that our data our drawn from a *linear space*, whose elements obey the familiar laws of arithmetic and algebra. Transformations of elements within and between linear spaces is at the core of applied data analysis. We seek to understand how these spaces work today. This lecture will work to unify your self-study of linear algebra with the more abstract notions of sets and spaces we've been studying in the last few days. In the next lecture, we'll attach a metric to these spaces to build "normed linear spaces."

A few preliminaries. As we've been doing so far, we'll use uppercase letters to denote sets and spaces (e.g. A subset of a linear space X will be expressed $S \subset X$). Elements of these sets will be bolded, as they may be multidimensional (e.g. $\boldsymbol{x} = (x_1, ..., x_n)$). $\mathbb{R}$ is the set of real numbers, $\mathbb{R}^2$ two-dimensional real space, and $\mathbb{R}^n$ is n-dimensional real space. We'll start with things we know something about – sets and their elements.

Convex Sets

Definition: A set is *convex* if for all elements $\boldsymbol{x}, \boldsymbol{y} \in S$

$$\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y} \in S$$

for all $\alpha \in [0, 1]$.

The geometric intuition for convexity is straightforward. For a set to be convex, we must be able to connect any two points of the set with a straight line without leaving the set.[1] The line segment $\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y}$ is a *convex combination* of the elements $\boldsymbol{x}$ and $\boldsymbol{y}$.[2] We can generalize this definition to combinations of many elements.

[1] What are a few shapes (or fruits) that are convex? Nonconvex?

[2] What subsets of $\mathbb{R}$ are convex? Is $[0, 1]$ convex? $[0, 1] \cup [2, 3]$? $(0, 1)$?

Definition: A *convex combination* is a finite sum of $\boldsymbol{x}_i \in S$

$$\alpha_1 \boldsymbol{x}_1 + ... + \alpha_n \boldsymbol{x}_n$$

with $\sum_{i=1}^{n} \alpha_n = 1$.

For a set to be convex, it must contain all convex combinations of any points taken from it. Convex combinations are weighted averages of a set of elements.[3] The set of all convex combinations of of elements of a set.

[3] How would one express a sample mean as a convex combination of values?

Definition: The convex hull of a set of elements $\{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\} \in S$ is

$$\{\alpha_1 \boldsymbol{x}_1 + ... + \alpha_n \boldsymbol{x}_n\}$$

for all $\{\alpha_1, ..., \alpha_n | \sum_{i=1}^{n} \alpha_i = 1\}$

Example: What is the convex hull of two points $a, b \in \mathbb{R}$?

Example: What is the convex hull of $\{(0,0), (0,1), (1,1)\} \in \mathbb{R}^2$?

Linear Spaces

Linear spaces are sets whose elements can be combined and scaled while remaining in the set. $\mathbb{R}$ and $\mathbb{R}^n$ are the most common linear spaces, so the properties of linear spaces will be familiar to you – they're the core arithmetic and algebraic rules you've been working with since primary school. $\mathbb{R}$ and $\mathbb{R}^n$ aren't the only linear spaces, though. Formally, a linear space X satisfies the following axioms.

Additivity: For all $\boldsymbol{x}, \boldsymbol{y} \in X$, there is another element $\boldsymbol{x} + \boldsymbol{y} \in X$ with the following properties:

1. Communativity: $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$
2. Associativity: $(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$
3. Zero Element: There exists an element $\boldsymbol{0} \in X$ such that $\boldsymbol{x} + \boldsymbol{0} = \boldsymbol{x}$
4. Negative: For every $\boldsymbol{x} \in X$ there also exists a unique $-\boldsymbol{x} \in X$ such that $\boldsymbol{x} + (-\boldsymbol{x}) = \boldsymbol{0}$.

Homogeneity: For all $\boldsymbol{x} \in X$ and scalars $\alpha, \beta \in \mathbb{R}$, there exists an element $\alpha\boldsymbol{x} \in X$ such that:

5. $(\alpha\beta)\boldsymbol{x} = \alpha(\beta\boldsymbol{x})$
6. $\boldsymbol{1}\boldsymbol{x} = \boldsymbol{x}$

Distributive Property:

7. $\alpha(\boldsymbol{x} + \boldsymbol{y}) = \alpha\boldsymbol{x} + \alpha\boldsymbol{y}$
8. $(\alpha + \beta)\boldsymbol{x} = \alpha\boldsymbol{x} + \beta\boldsymbol{x}$

Example: "Construct" $\mathbb{R}$, starting with the numbers $4, 9$.[4]

[4] In other words, if $4, 9 \in \mathbb{R}$ and $\mathbb{R}$ is a linear space, what other elements must be included in $\mathbb{R}$ in order to satisfy the axioms.

Linear spaces are also referred to as *vector spaces*. Why? Because their elements can always be expressed as vectors.[5]

[5] How would you write a convex combination of $\{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\}$ in vector notation?

Example: "Construct" $\mathbb{R}^2$, starting with the vectors $(2, 1)$ and $(1, 4)$.

I mentioned $\mathbb{R}^1, ..., \mathbb{R}^n$ weren't the only linear spaces. The scalar 0, or any vector of zeros are also linear spaces.[6] The set of all diagonal matrices

[6] See for yourself, verify each of the axioms.

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

is also a vector space.[7] What are some examples of spaces that *are not* vector spaces?[8]

Unlike convex sets, which include all convex combinations of their elements, the axioms defining linear spaces imply that they contain all *linear combinations* of their elements.

Definition: A linear combination is a finite sum of $\boldsymbol{x}_i \in S$

$$\alpha_1 \boldsymbol{x}_1 + ... + \alpha_n \boldsymbol{x}_n$$

with $\boldsymbol{x}_1, ...\boldsymbol{x}_n \in S$ and $\alpha_1, ..., \alpha_n \in \mathbb{R}$.[9]

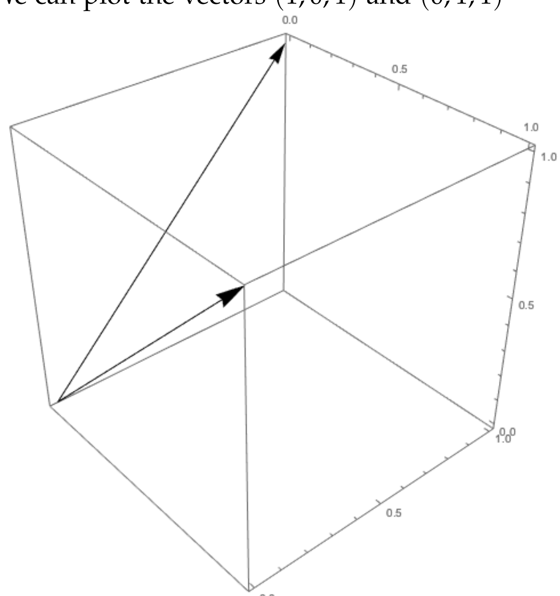
Convex hulls also have cousins in linear spaces.

Definition: A *linear hull* or *span* of a set of elements $\{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\} \in S$ is

$$\{\alpha_1 \boldsymbol{x}_1 + ... + \alpha_n \boldsymbol{x}_n\}$$

for all $\alpha_1, ..., \alpha_n \in \mathbb{R}$

Example: We can plot the vectors $(1, 0, 1)$ and $(0, 1, 1)$

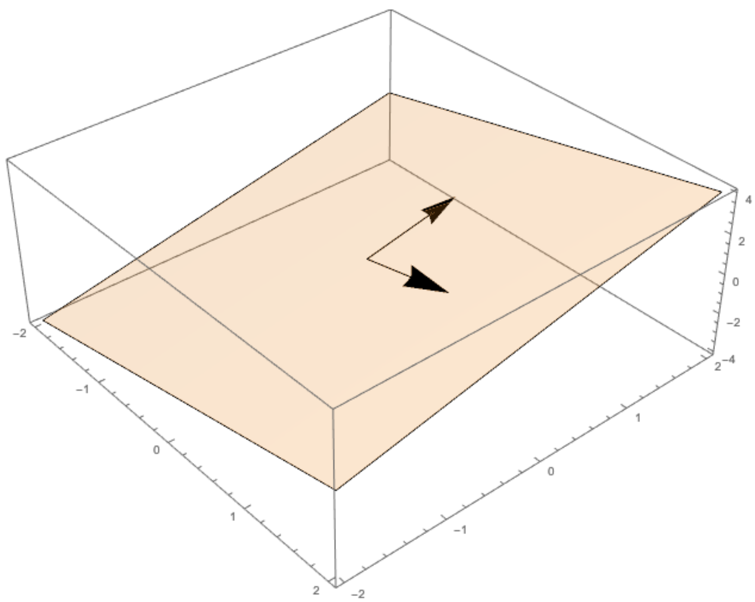


Taking all linear combinations of these vectors gives a separating plane through $\mathbb{R}^3$

[7] Apply the rules of matrix addition and multiplication to see why this is true.

[8] Are the integers ($\mathbb{Z}$) a linear space? What about the rationals ($\mathbb{Q}$)? Why or why not?

[9] Compare this to the definition of a *convex combination*. How do the definitions differ?



Example: What is the linear hull of the vectors $(-1,-1), (1,1) \in \mathbb{R}^2$?

We'll usually think about our data as living in a vector space, where each observation $(x_1, ..., x_n)$ is an element. Note, however, that this requires that certain things be true of the underlying data-generating process.[10]

[10] E.g. data can take on negative values, they can be scaled up or scaled down, etc.

Linear Subspaces

A plane is a linear space. Within a plane are infinitely many lines that pass through the origin. These lines are *linear subspaces* of the plane, because they themselves are linear spaces, and all of their points are contained within the plane.

Definition: A subset S of a linear space X is a *linear subspace* of X if for every $\boldsymbol{x}, \boldsymbol{y} \in S$,

$$\alpha \boldsymbol{x} + \beta \boldsymbol{y} \in S$$

for $\alpha, \beta \in \mathbb{R}$.

Example: Show that linear subspaces are themselves linear spaces.

Example: What are the linear subspaces of $\mathbb{R}^3$?

Linear Independence

We briefly studied linear independence in the context of finding solutions to systems of linear equations of the form

$$A\boldsymbol{x} = \boldsymbol{b}$$

If two rows of A were *linearly dependent*, then $A\boldsymbol{x} = \boldsymbol{b}$ will not have a unique solution and is said to be *underdetermined*.

Now that we have a working understanding of linear spaces, we can think about *linear dependence* in terms of linear spaces and subspaces. When we solve $A\boldsymbol{x} = \boldsymbol{b}$, we are attempting to express $\boldsymbol{b}$ as a linear combination of elements of A. So $\boldsymbol{b}$ must be in the linear space generated by linear combinations of elements of A.

Definition: An element $\boldsymbol{y}$ is *linearly dependent* on a set of vectors $\boldsymbol{x}_1, ..., \boldsymbol{x}_n \in S$ if $\boldsymbol{y}$ is in the linear hull of S, or there exists $\alpha_1, ..., \alpha_n \in \mathbb{R}$ such that

$$\alpha_1 \boldsymbol{x}_1 + ... + \alpha_n \boldsymbol{x}_n = \boldsymbol{y}$$

Example: Consider the system

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

Show that the rows the matrix are linearly dependent. What does this imply about the solution vector (x_1, x_2)?

Example: Now consider

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

What is the linear hull generated by the vectors of A? Is $\boldsymbol{b}$ in this space? What does this imply about the solution vector (x_1, x_2)?

Definition: The *null space* of a set of vectors $\{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\} \in S$ is the set of vectors $\boldsymbol{\alpha}$ such that

$$\alpha_1 \boldsymbol{x}_1 + ... + \alpha_n \boldsymbol{x}_n = 0$$

The zero vector $\boldsymbol{\alpha} = \boldsymbol{0}$ is trivially in the null space. If any nonzero vector is in the nullspace, then the vectors $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$ are linearly dependent.[11]

A *simplex* combines our definitions of convex sets and linear independence, so we introduce them here.

Definition: A *simplex* is the *convex hull* of a set of *linearly indepedent* vectors.[12]

Simplices play important roles in game theory. *Mixed strategies* are convex combinations (probabilities that sum to one) of independent strategies.[13] Multivariate proportions can also be expressed as simplices, see Figure 1.

[11] The proof of this claim is left as an exercise.

[12] Use the definitions of convex hull and linear independence to write a definition that does not rely on prior knowledge of these concepts.

[13] Think of a strategy in "Rock, Paper, Scissors."

Basis and Dimension

A basis is a set of independent vectors from which a given linear space can be constructed. You give me the set of vectors, and I can construct the linear space by applying the additivity and homogeneity axioms. In other words, all other elements of the space are linearly dependent on the basis vectors.

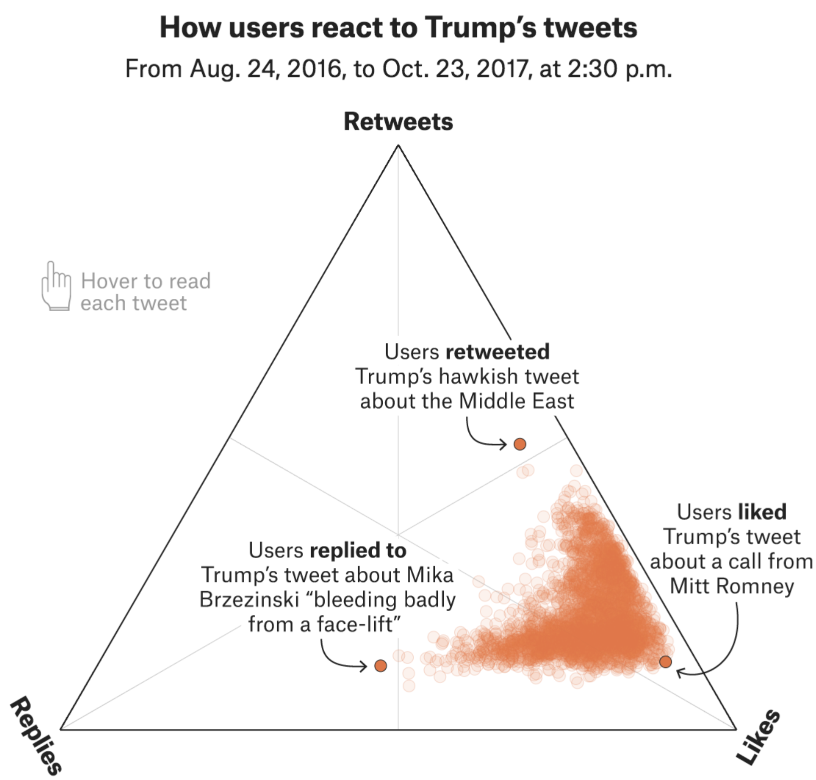


Figure 1: *FiveThirtyEight*, "The Worst Tweeter In Politics Isn't Trump." 24 October 2017.

Definition: A *basis* for a linear space X is a set $\boldsymbol{y}_1, ..., \boldsymbol{y}_n \in S$ such that for all $\boldsymbol{x} \in X$,

$$\alpha_1 \boldsymbol{y}_1 + ... + \alpha_n \boldsymbol{y}_n = \boldsymbol{x}$$

Example: List some bases for $\mathbb{R}^2$.[14]

Example: Are the vectors $(1,1,1), (0,1,1), (0,0,1)$ a basis for $\mathbb{R}^3$?

The easiest basis to visualize and work with is the *standard basis*, or the set of independent *unit vectors*. In the case of $\mathbb{R}^2$, the vectors $(1,0), (0,1)$ are the standard basis.

Definition: The *dimension* of a linear space X is the number of elements in its bases.[15]

Note: The dimension of the linear space occupied by the vectors in a matrix

$$A = \begin{bmatrix} \boldsymbol{x}_1 \\ ... \\ \boldsymbol{x}_n \end{bmatrix}$$

depends not on the number of rows (n) or columns, but on the number of *independent* rows.

[14] Remember, we need vectors from which we can reconstruct by vector addition and scaling all of the 2-dimensional real space.

[15] We now have a more formal definition accompanying our intuition for the dimension $\mathbb{R}^2$, $\mathbb{R}^3$, and $\mathbb{R}^n$.

Affine Set

Axiom 3 distinguishes linear spaces from their cousins, affine sets. Affine sets are like vector spaces, except we don't require that the zero vector, (the "origin"), is included in the space. Geometrically, this includes lines, planes, and *hyperplanes* that may or may not pass through the origin

Definition: An *affine set* is a subset S of a linear space X such that for all $\boldsymbol{x}, \boldsymbol{y} \in S$ and $\alpha \in \mathbb{R}$,

$$\alpha \boldsymbol{x} + (1 - \alpha) \boldsymbol{y} \in S$$

Example: Name all affine subsets of $\mathbb{R}^3$.

Affine sets are really just linear spaces that have been shifted from the zero vector.[16] Affine sets are said to be *parallel* to these linear spaces.

Remark: For each affine set S there exists a unique linear space X such that $S = \boldsymbol{x} + X$ for some shifter $\boldsymbol{x} \in S$.

Definition: Let an affine set S have dimension n. A *hyperplane* through S is a proper affine subset of S with dimension $n - 1$.

Example: What is a hyperplane in $\mathbb{R}^3$? $\mathbb{R}^2$?

[16] Verify that all of the axioms defining linear spaces hold for affine sets, except Axiom 3.

Exercises

1. Prove the following: If a set S is convex, then all sets αS are also convex for all $\alpha \in \mathbb{R}$.[17]

[17] $\boldsymbol{x} \in S \implies \alpha \boldsymbol{x} \in \alpha S$ for all $\alpha \in \mathbb{R}$. Carter 1.166

If S is convex, then for any $\boldsymbol{x}, \boldsymbol{y} \in S$, and $\beta \in [0, 1]$

$$\boldsymbol{z} = \beta \boldsymbol{x} + (1 - \beta)\boldsymbol{y} \in S$$

By the definition of αS, $\alpha \boldsymbol{x}, \alpha \boldsymbol{y}, \alpha \boldsymbol{z} \in \alpha S$. Therefore,

$$\begin{aligned} \alpha \boldsymbol{z} &= \alpha \left(\beta \boldsymbol{x} + (1 - \beta)\boldsymbol{y}\right) \in \alpha S \\ &= \beta \alpha \boldsymbol{x} + (1 - \beta)\alpha \boldsymbol{y} \in \alpha S \end{aligned}$$

for $\beta \in [0, 1]$. Since $\boldsymbol{x}$ and $\boldsymbol{y}$ were chosen arbitrarily, αS is convex. ■

2. Prove that if S_1 and S_2 are subspaces of a linear space X, then $S_1 + S_2$ is also a linear subspace of X.[18]

   If S_1 and S_2 are linear subspaces, then

$$\boldsymbol{z}_1 = \alpha \boldsymbol{x}_1 + \beta \boldsymbol{y}_1 \in S_1$$

$$\boldsymbol{z}_2 = \alpha \boldsymbol{x}_2 + \beta \boldsymbol{y}_2 \in S_2$$

for all $\boldsymbol{x}_1, \boldsymbol{y}_1 \in S_1$, $\boldsymbol{x}_2, \boldsymbol{y}_2 \in S_2$ and $\alpha, \beta \in \mathbb{R}$. By the definition of $S_1 + S_2$, $\boldsymbol{z}_1 + \boldsymbol{z}_2 \in S_1 + S_2$.

$$\begin{aligned} \boldsymbol{z}_1 + \boldsymbol{z}_2 &= \alpha \boldsymbol{x}_1 + \beta \boldsymbol{y}_1 + \alpha \boldsymbol{x}_2 + \beta \boldsymbol{y}_2 \in S_1 + S_2 \\ &= \alpha \left(\boldsymbol{x}_1 + \boldsymbol{x}_2\right) + \beta \left(\boldsymbol{y}_1 + \boldsymbol{y}_2\right) \in S_1 + S_2 \end{aligned}$$

Since $\boldsymbol{x}_1, \boldsymbol{y}_1, \boldsymbol{x}_2, \boldsymbol{y}_2, \alpha, \beta$ were chosen arbitrarily, $S_1 + S_2$ is a linear subspace of X. ■

3. Prove that if S_1 and S_2 are subspaces of a linear space X, then $S_1 + S_2$ is the linear hull of $S_1 \cup S_2$.[19]

   We need to show that for an arbitrary set of $\{x_{11}, ..., x_{1n}\} \in S_1$ and $\{x_{21}, ..., x_{2m}\} \in S_2$,

$$\sum_{i=1}^{n} \alpha_i x_{1i} + \sum_{j=1}^{m} \alpha_j x_{2j} \in S_1 + S_2$$

for $\alpha_i, \alpha_j \in \mathbb{R}$.

We first show that if S is a linear subspace of X, then

$$\sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i \in S$$

for $\{\boldsymbol{x}_i, ..., \boldsymbol{x}_n\} \in S$, $\{\alpha_1, ..., \alpha_n\}$. We proceed by induction. Take $x_1 \in S$. Then $\alpha_1 x_1 \in S$ for all $\alpha_1 \in \mathbb{R}$. Let

$$\boldsymbol{z}_m = \sum_{i=1}^{m} \alpha_i \boldsymbol{x}_i \in S$$

We need to show $z_{m+1} \in S$. By the definition of a linear subspace, this must be true

$$\boldsymbol{z}_m + \alpha_{m+1} \boldsymbol{x}_{m+1} \in S$$

[18] $\boldsymbol{x}_1 \in S_1$ and $\boldsymbol{x}_2 \in X_2$ implies $\boldsymbol{x}_1 + \boldsymbol{x}_2 \in S_1 + S_2$. Carter 1.131

[19] Strang 3.1 #30.

for all $\alpha_{m+1} \in \mathbb{R}$. Therefore, if S is a linear subpace of X,

$$\boldsymbol{z} = \sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i \in S$$

Now take $\boldsymbol{z}_1 \in S_1$ and $\boldsymbol{z}_2 \in S_2$. Since $\boldsymbol{z}_1 + \boldsymbol{z}_2 \in S_1 + S_2$,

$$\sum_{i=1}^{n} \alpha_{1i} \boldsymbol{x}_{1i} + \sum_{j=1}^{m} \alpha_{2j} \boldsymbol{x}_{2j} \in S_1 + S_2$$

as desired. ■

4. Prove the following: Any nonzero vector is in the nullspace of a set S iff ($\Longleftrightarrow$) there exists a linearly dependent vector $y \in S$.

($\Longrightarrow$) Let $\{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\} \in S$. Let $N(S)$ be the nullspace of this set. A vector $\boldsymbol{\alpha} \in N(S)$ if

$$\sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i = 0$$

Since $\boldsymbol{\alpha}$ is nonzero, there is at least one $\alpha_j > 0$. Rearranging, we can write

$$\alpha_j \boldsymbol{x}_j = \sum_{i \neq j}^{n} \alpha_i \boldsymbol{x}_i$$
$$\boldsymbol{x}_j = \frac{1}{\alpha_j} \sum_{i \neq j}^{n} \alpha_i \boldsymbol{x}_i$$

Since $\alpha_j \neq 0$, this implies that there exists a vector $\boldsymbol{x}_j$ that is linearly dependent on $\boldsymbol{x}_{i \neq j}$.

($\Longleftarrow$) Take $\boldsymbol{y}$ linearly dependent on $\{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\} \in S$. Then,

$$\boldsymbol{y} = \sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i$$

for some nonzero $\{\alpha_1, ..., \alpha_n\}$. Rearranging,

$$\sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i + \boldsymbol{y} = 0$$

implying $\boldsymbol{\alpha} = \{\alpha_1, ..., \alpha_n, 1\} \in N(S)$. Since $1 > 0$, $\boldsymbol{\alpha}$ is nonzero. ■