ECON2125/8013

Lecture 7

John Stachurski

Semester 1, 2015

Announcements

- Mid semester exam — date after break requested
- Access to previous exam papers against school policy
- Practice questions with solutions will be posted soon on GitHub

Linear Independence

Important applied questions

- When is a matrix invertible?
- When do regression arguments suffer from collinearity?
- When does a set of linear equations have a solution?
- When is that solution unique?
- How can we approximate complex functions parsimoniously?
- What is the rank of a matrix?

All of these questions closely related to linear independence

Definition

A nonempty collection of vectors $X := \{\mathbf{x}_1, \ldots, \mathbf{x}_K\} \subset \mathbb{R}^N$ is called **linearly independent** if

$$\sum_{k=1}^{K} \alpha_k \mathbf{x}_k = \mathbf{0} \implies \alpha_1 = \cdots = \alpha_K = 0$$

As we'll see, linear independence of a set of vectors determines how large a space they span

Loosely speaking, linearly independent sets span large spaces

Example. Let $\mathbf{x} := (1, 2)$ and $\mathbf{y} := (-5, 3)$

The set $X = \{\mathbf{x}, \mathbf{y}\}$ is linearly independent in $\mathbb{R}^2$

Indeed, suppose α_1 and α_2 are scalars with

$$\alpha_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \alpha_2 \begin{pmatrix} -5 \\ 3 \end{pmatrix} = \mathbf{0}$$

Equivalently,

$$\alpha_1 = 5\alpha_2$$

$$2\alpha_1 = -3\alpha_2$$

Then $2(5\alpha_2) = 10\alpha_2 = -3\alpha_2$, implying $\alpha_2 = 0$ and hence $\alpha_1 = 0$

Example

The set of canonical basis vectors $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ is linearly independent in $\mathbb{R}^N$

Proof: Let $\alpha_1, \ldots, \alpha_N$ be coefficients such that $\sum_{k=1}^N \alpha_k \mathbf{e}_k = \mathbf{0}$

Then

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{pmatrix} = \sum_{k=1}^N \alpha_k \mathbf{e}_k = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

In particular, $\alpha_k = 0$ for all k

Hence $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ linearly independent

As a first step to better understanding linear independence let's look at some equivalences

Take $X := \{\mathbf{x}_1, \ldots, \mathbf{x}_K\} \subset \mathbb{R}^N$

Fact. For $K > 1$ all of following statements are equivalent

1. X is linearly independent
2. No $\mathbf{x}_i \in X$ can be written as linear combination of the others
3. $X_0 \subsetneq X \implies \operatorname{span}(X_0) \subsetneq \operatorname{span}(X)$

- Here $X_0 \subsetneq X$ means $X_0 \subset X$ and $X_0 \neq X$
- We say that X_0 is a **proper subset** of X

As an exercise, let's show that the first two statements are equivalent

The first is

$$\sum_{k=1}^{K} \alpha_k \mathbf{x}_k = \mathbf{0} \implies \alpha_1 = \cdots = \alpha_K = 0 \qquad (\star)$$

The second is

$$\text{no } \mathbf{x}_i \in X \text{ can be written as linear combination of others} \qquad (\star\star)$$

We now show that

- $(\star) \implies (\star\star)$, and
- $(\star\star) \implies (\star)$

To show that $(\star) \implies (\star\star)$ let's suppose to the contrary that

1. $\sum_{k=1}^{K} \alpha_k \mathbf{x}_k = \mathbf{0} \implies \alpha_1 = \cdots = \alpha_K = 0$
2. and yet some $\mathbf{x}_i$ can be written as a linear combination of the other elements of X

In particular, suppose that

$$\mathbf{x}_i = \sum_{k \neq i} \alpha_k \mathbf{x}_k$$

Then, rearranging,

$$\alpha_1 \mathbf{x}_1 + \cdots + (-1)\mathbf{x}_i + \cdots + \alpha_K \mathbf{x}_K = \mathbf{0}$$

This contradicts 1., and hence $(\star\star)$ holds

To show that $(\star\star) \implies (\star)$ let's suppose to the contrary that

1. no $\mathbf{x}_i$ can be written as a linear combination of others
2. and yet $\exists\ \alpha_1, \ldots, \alpha_K$ not all zero with $\alpha_1 \mathbf{x}_1 + \cdots + \alpha_K \mathbf{x}_K = \mathbf{0}$

Suppose without loss of generality that $\alpha_1 \neq 0$

(Similar argument works for any α_j)

Then

$$\mathbf{x}_1 = \frac{\alpha_2}{-\alpha_1} \mathbf{x}_2 + \cdots + \frac{\alpha_K}{-\alpha_1} \mathbf{x}_K$$

This contradicts 1., and hence $(\star)$ holds

Let's show one more part of the proof as an exercise:

X linearly independent $\implies$ proper subsets of X have smaller span

Proof: Suppose to the contrary that

1. X is linearly independent,
2. $X_0 \subsetneq X$ and yet
3. $\mathrm{span}(X_0) = \mathrm{span}(X)$

Let $\mathbf{x}_j$ be in X but not X_0

Since $\mathbf{x}_j \in \mathrm{span}(X)$, we also have $\mathbf{x}_j \in \mathrm{span}(X_0)$

But then $\mathbf{x}_j$ can be written as a linear combination of the other elements of X

This contradicts linear independence

Example. Dropping any of the canonical basis vectors reduces span

Consider the $N = 2$ case

We know that $\mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2\} =$ all of $\mathbb{R}^2$

Removing either element of $\mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2\}$ reduces the span to a line

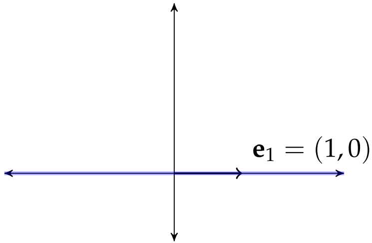

Figure : The span of $\{\mathbf{e}_1\}$ alone is the horizonal axis

Example. As another visual example of linear independence, consider the pair

$$\mathbf{x}_1 = \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix} \quad \text{and} \quad \mathbf{x}_2 = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}$$

The span of this pair is a plane in $\mathbb{R}^3$

But if we drop either one the span reduces to a line

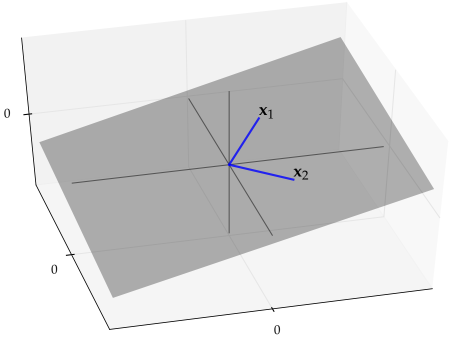

Figure : The span of $\{\mathbf{x}_1, \mathbf{x}_2\}$ is a plane

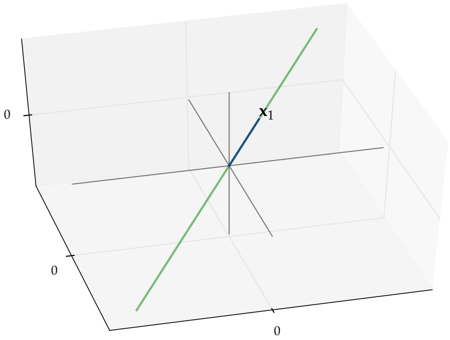

Figure : The span of $\{\mathbf{x}_1\}$ alone is a line

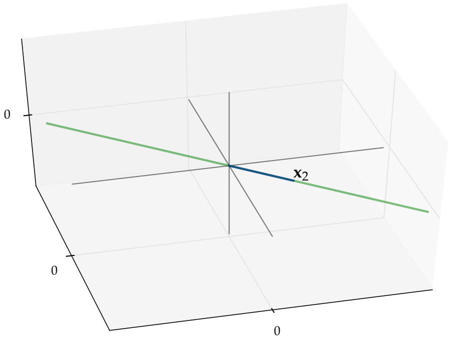

Figure : The span of $\{\mathbf{x}_2\}$ alone is a line

Linear Dependence

If X is not linearly independent then it is called **linearly dependent**

We saw above that

$$\text{linear independence} \iff \text{dropping any elements reduces span}$$

Hence X is linearly dependent when some elements can be removed without changing $\operatorname{span}(X)$

That is,

$$\exists\, X_0 \subsetneq X \;\text{ s.t. }\; \operatorname{span}(X_0) = \operatorname{span}(X)$$

Example. As an example with redundacy, consider $\{\mathbf{x}_1, \mathbf{x}_2\} \subset \mathbb{R}^2$ where

- $\mathbf{x}_1 = \mathbf{e}_1 := (1, 0)$
- $\mathbf{x}_2 = (-2, 0)$

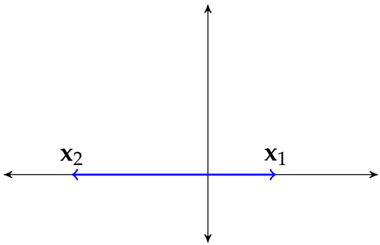

Figure : The vectors $\mathbf{x}_1$ and $\mathbf{x}_2$

We claim that $\text{span}\{\mathbf{x}_1, \mathbf{x}_2\} = \text{span}\{\mathbf{x}_1\}$

Proof: $\text{span}\{\mathbf{x}_1\} \subset \text{span}\{\mathbf{x}_1, \mathbf{x}_2\}$ is clear (why?)

To see the reverse, pick any $\mathbf{y} \in \text{span}\{\mathbf{x}_1, \mathbf{x}_2\}$

By definition,

$$\exists\, \alpha_1, \alpha_2 \;\text{ s.t. }\; \mathbf{y} = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 = \alpha_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

$$\therefore \quad \mathbf{y} = \alpha_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 2\alpha_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = (\alpha_1 - 2\alpha_2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = (\alpha_1 - 2\alpha_2)\mathbf{x}_1$$

The right hand side is clearly in $\text{span}\{\mathbf{x}_1\}$

Hence $\text{span}\{\mathbf{x}_1, \mathbf{x}_2\} \subset \text{span}\{\mathbf{x}_1\}$ as claimed

Implications of Independence

Let $X := \{\mathbf{x}_1, \ldots, \mathbf{x}_K\} \subset \mathbb{R}^N$

Fact. If X is linearly independent, then X does not contain $\mathbf{0}$

Ex. Prove it

Fact. If X is linearly independent, then every subset of X is linearly independent

Sketch of proof: Suppose for example that $\{\mathbf{x}_1, \ldots, \mathbf{x}_{K-1}\} \subset X$ is linearly dependent

Then $\exists\ \alpha_1, \ldots, \alpha_{K-1}$ not all zero with $\sum_{k=1}^{K-1} \alpha_k \mathbf{x}_k = \mathbf{0}$

Setting $\alpha_K = 0$ we can write this as $\sum_{k=1}^{K} \alpha_k \mathbf{x}_k = \mathbf{0}$

Not all scalars zero so contradicts linear independence of X

Fact. If $X := \{\mathbf{x}_1, \ldots, \mathbf{x}_K\} \subset \mathbb{R}^N$ is linearly independent and $\mathbf{z}$ is an N-vector not in $\mathrm{span}(X)$, then $X \cup \{\mathbf{z}\}$ is linearly independent

Proof: Suppose to the contrary that $X \cup \{\mathbf{z}\}$ is linearly dependent:

$$\exists\, \alpha_1, \ldots, \alpha_K, \beta \text{ not all zero with } \sum_{k=1}^{K} \alpha_k \mathbf{x}_k + \beta \mathbf{z} = \mathbf{0} \tag{1}$$

If $\beta = 0$, then by (1) we have $\sum_{k=1}^{K} \alpha_k \mathbf{x}_k = \mathbf{0}$ and $\alpha_k \neq 0$ for some k, a contradiction

If $\beta \neq 0$, then by (1) we have

$$\mathbf{z} = \sum_{k=1}^{K} \frac{-\alpha_k}{\beta} \mathbf{x}_k$$

Hence $\mathbf{z} \in \mathrm{span}(X)$ — contradiction

Unique Representations

Let

- $X := \{\mathbf{x}_1, \ldots, \mathbf{x}_K\} \subset \mathbb{R}^N$
- $\mathbf{y} \in \mathbb{R}^N$

We know that if $\mathbf{y} \in \operatorname{span}(X)$, then exists representation

$$\mathbf{y} = \sum_{k=1}^{K} \alpha_k \mathbf{x}_k$$

But when is this representation unique?

Answer: When X is linearly independent

Fact. If $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_K\} \subset \mathbb{R}^N$ is linearly independent and $\mathbf{y} \in \mathbb{R}^N$, then there is at most one one set of scalars $\alpha_1, \ldots, \alpha_K$ such that $\mathbf{y} = \sum_{k=1}^K \alpha_k \mathbf{x}_k$

Proof: Suppose there are two such sets of scalars

That is,

$$\exists\ \alpha_1, \ldots, \alpha_K \text{ and } \beta_1, \ldots, \beta_K \ \text{ s.t. } \ \mathbf{y} = \sum_{k=1}^K \alpha_k \mathbf{x}_k = \sum_{k=1}^K \beta_k \mathbf{x}_k$$

$$\therefore \quad \sum_{k=1}^K (\alpha_k - \beta_k)\mathbf{x}_k = \mathbf{0}$$

$$\therefore \quad \alpha_k = \beta_k \quad \text{for all} \quad k$$

Exchange Lemma

Here's one of the most fundamental results in linear algebra

Fact. (Exchange lemma) If

1. S is a linear subspace of $\mathbb{R}^N$
2. S is spanned by K vectors,

then any linearly independent subset of S has at most K vectors

Proof: Omitted

Example. If $X := \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\} \subset \mathbb{R}^2$ then X is linearly dependent

- because $\mathbb{R}^2$ is spanned by the two vectors $\mathbf{e}_1, \mathbf{e}_2$

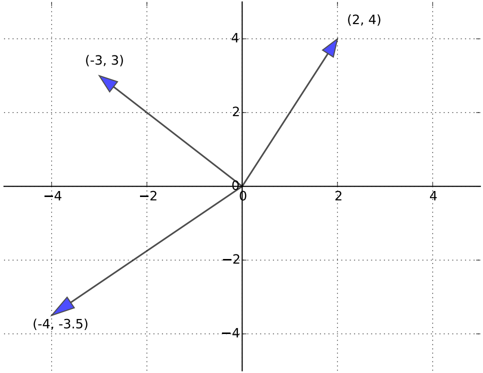

Figure : Must be linearly dependent

Example

Recall the plane

$$P := \{(x_1, x_2, 0) \in \mathbb{R}^3 : x_1, x_2 \in \mathbb{R}\}$$

- flat plane in $\mathbb{R}^3$ where height coordinate = zero

We showed before that $\mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2\} = P$ for

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Therefore any three vectors lying in P are linearly dependent

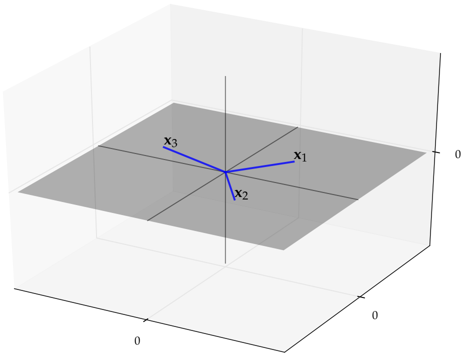

Figure : Any three vectors in P are linearly dependent

When Do N Vectors Span $\mathbb{R}^N$?

In general, linearly independent vectors have a relatively "large" span

- No vector is redundant, so each contributes to the span

This helps explain the following fact:

Let $X := \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ be any N vectors in $\mathbb{R}^N$

Fact. $\operatorname{span}(X) = \mathbb{R}^N$ if and only if X is linearly independent

Example. The vectors $\mathbf{x} = (1, 2)$ and $\mathbf{y} = (-5, 3)$ span $\mathbb{R}^2$

- We already showed $\{\mathbf{x}, \mathbf{y}\}$ is linearly independent

Let's start with the proof that

$$X = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \text{ linearly independent } \implies \operatorname{span}(X) = \mathbb{R}^N$$

Seeking a contradiction, suppose that

1. X is linearly independent
2. and yet $\exists\, \mathbf{z} \in \mathbb{R}^N$ with $\mathbf{z} \notin \operatorname{span}(X)$

But then $X \cup \{\mathbf{z}\} \subset \mathbb{R}^N$ is linearly independent (why?)

This set has $N + 1$ elements

And yet $\mathbb{R}^N$ is spanned by the N canonical basis vectors

Contradiction (of what?)

Next let's show the converse

$$\mathrm{span}(X) = \mathbb{R}^N \implies X = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \text{ linearly independent}$$

Seeking a contradiction, suppose that

1. $\mathrm{span}(X) = \mathbb{R}^N$
2. and yet X is linearly dependent

Since X not independent, $\exists X_0 \subsetneq X$ with $\mathrm{span}(X_0) = \mathrm{span}(X)$

But by 1 this implies that $\mathbb{R}^N$ is spanned by $K < N$ vectors

But then the N canonical basis vectors must be linearly dependent

Contradiction

Bases

Let S be a linear subspace of $\mathbb{R}^N$

A set of vectors $B := \{\mathbf{b}_1, \ldots, \mathbf{b}_K\} \subset S$ is called a **basis of S** if

1. B is linearly independent
2. $\mathrm{span}(B) = S$

Example. Canonical basis vectors form a basis of $\mathbb{R}^N$

Indeed, if $E := \{\mathbf{e}_1, \ldots, \mathbf{e}_N\} \subset \mathbb{R}^N$, then

- E is linearly independent – we showed this earlier
- $\mathrm{span}(E) = \mathbb{R}^N$ – we showed this earlier

Example

Recall the plane

$$P := \{(x_1, x_2, 0) \in \mathbb{R}^3 : x_1, x_2 \in \mathbb{R}\}$$

We showed before that $\mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2\} = P$ for

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Moreover, $\{\mathbf{e}_1, \mathbf{e}_2\}$ is linearly independent (why?)

Hence $\{\mathbf{e}_1, \mathbf{e}_2\}$ is a basis for P

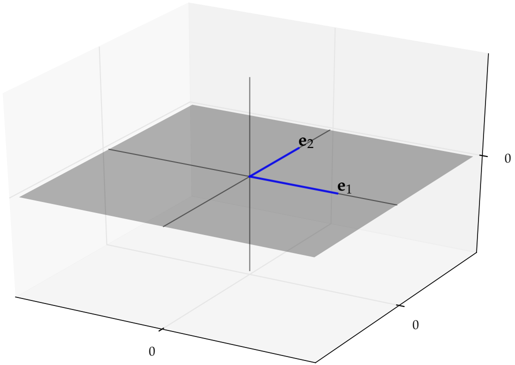

Figure : The pair $\{\mathbf{e}_1, \mathbf{e}_2\}$ form a basis for P

What are the implications of B being a basis of S?

In short, every element of S can be represented uniquely from the smaller set B

In more detail:

- B spans S and, by linear independence, every element is needed to span S — a "minimal" spanning set
- Since B spans S, every $\mathbf{y}$ in S can be represented as a linear combination of the basis vectors
- By independence, this representation is unique

It's obvious given the definition that

Fact. If $B \subset \mathbb{R}^N$ is linearly independent, then B is a basis of $\mathrm{span}(B)$

Example. Let $B := \{\mathbf{x}_1, \mathbf{x}_2\}$ where

$$\mathbf{x}_1 = \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix} \quad \text{and} \quad \mathbf{x}_2 = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}$$

We saw earlier that

- $S := \mathrm{span}(B)$ is the plane in $\mathbb{R}^3$ passing through $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{0}$
- B is linearly independent in $\mathbb{R}^3$ (dropping either reduces span)

Hence B is a basis for the plane S

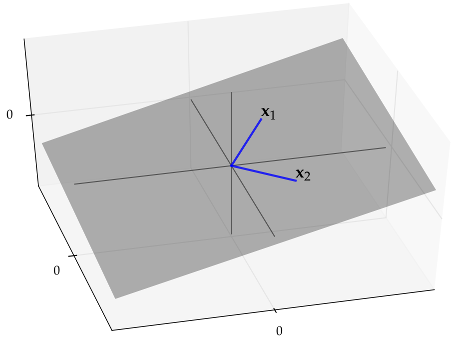

Figure : The pair $\{\mathbf{x}_1, \mathbf{x}_2\}$ is a basis of its span

Fundamental Properties of Bases

Fact. If S is a linear subspace of $\mathbb{R}^N$ distinct from $\{\mathbf{0}\}$, then

1. S has at least one basis, and
2. every basis of S has the same number of elements

Proof of part 2: Let B_i be a basis of S with K_i elements, $i = 1, 2$

By definition, B_2 is a linearly independent subset of S

Moreover, S is spanned by the set B_1, which has K_1 elements

Hence $K_2 \leq K_1$

Reversing the roles of B_1 and B_2 gives $K_1 \leq K_2$

Dimension

Let S be a linear subspace of $\mathbb{R}^N$

We now know that every basis of S has the same number of elements

This common number is called the **dimension** of S

Example. $\mathbb{R}^N$ is N dimensional because the N canonical basis vectors form a basis

Example. $P := \{(x_1, x_2, 0) \in \mathbb{R}^3 : x_1, x_2 \in \mathbb{R}\}$ is two dimensional because the first two canonical basis vectors of $\mathbb{R}^3$ form a basis

Example. In $\mathbb{R}^3$, a line through the origin is one-dimensional, while a plane through the origin is two-dimensional

Dimension of Spans

Fact. Let $X := \{\mathbf{x}_1, \ldots, \mathbf{x}_K\}$ be a $\subset \mathbb{R}^N$ with K elements

The following statements are true:

1. $\dim(\operatorname{span}(X)) \leq K$
2. $\dim(\operatorname{span}(X)) = K \iff X$ is linearly independent

Proof that $\dim(\operatorname{span}(X)) \leq K$

If not then $\operatorname{span}(X)$ has a basis with $M > K$ elements

Hence $\operatorname{span}(X)$ contains $M > K$ linearly independent vectors

This is impossible, given that $\operatorname{span}(X)$ is spanned by K vectors

Now consider the second claim:

1. X is linearly independent $\implies \dim(\text{span}(X)) = K$

Proof: True because the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_K$ form a basis of $\text{span}(X)$

2. $\dim(\text{span}(X)) = K \implies X$ linearly independent

Proof: If not then $\exists\, X_0 \subsetneq X$ such that $\text{span}(X_0) = \text{span}(X)$

By this equality and part 1 of the theorem,

$$\dim(\text{span}(X)) = \dim(\text{span}(X_0)) \leq \# X_0 \leq K - 1$$

Contradiction

Fact. If S a linear subspace of $\mathbb{R}^N$, then

$$\dim(S) = N \iff S = \mathbb{R}^N$$

Useful implications

- The only N-dimensional subspace of $\mathbb{R}^N$ is $\mathbb{R}^N$
- To show $S = \mathbb{R}^N$ just need to show that $\dim(S) = N$

Proof: See course notes