

Basis of the vector space and its set theory background

Zorn's lemma

Table of Contents

- 1, Some review of core concepts needed to understand the **definition of basis**
 - 2, Some set theory background (mainly Zorn's lemma)
 - 3, Going over the proof of “**every vector space has a basis**”
 - 4, A cool application of the previous result
- (you can also learn how one would write mathematical arguments here)

Part 1: Some stuff to recall before understanding the definition of basis of vector spaces

What you need to know to understand the definition of basis

1, The vector space

2, Linear Independence

3, Span of β , where β is some set of vectors

$$\beta = \{v_1, v_2, \dots\}$$

The vector space

*Definitions. A vector space (or linear space) V over a field² F consists of a set on which two operations (called **addition** and **scalar multiplication**, respectively) are defined so that for each pair of elements x, y ,*

The vector space is a set of vectors on which two operations, **Addition** and **Scalar Multiplication** are defined, and satisfies the following 8 axioms. A scalar field is usually \mathbb{R} or \mathbb{C} .

Examples: Polynomial Vector Space, Vector space of functions, Matrices, n -tuples (or what we generally call vectors)

8 Axioms of Vector Space

- (VS 1) For all x, y in V , $x + y = y + x$ (commutativity of addition).
- (VS 2) For all x, y, z in V , $(x + y) + z = x + (y + z)$ (associativity of addition).
- (VS 3) There exists an element in V denoted by 0 such that $x + 0 = x$ for each x in V .
- (VS 4) For each element x in V there exists an element y in V such that $x + y = 0$.
- (VS 5) For each element x in V , $1x = x$.
- (VS 6) For each pair of elements a, b in F and each element x in V ,
 $(ab)x = a(bx)$.
- (VS 7) For each element a in F and each pair of elements x, y in V ,
 $a(x + y) = ax + ay$.
- (VS 8) For each pair of elements a, b in F and each element x in V ,
 $(a + b)x = ax + bx$.

The elements $x + y$ and ax are called the **sum** of x and y and the **product** of a and x , respectively.

Linear Combination

Definitions. Let V be a vector space and S a nonempty subset of V . A vector $v \in V$ is called a **linear combination** of vectors of S if there exist a finite number of vectors u_1, u_2, \dots, u_n in S and scalars a_1, a_2, \dots, a_n in F such that $v = a_1u_1 + a_2u_2 + \dots + a_nu_n$. In this case we also say that v is a linear combination of u_1, u_2, \dots, u_n and call a_1, a_2, \dots, a_n the **coefficients** of the linear combination.

Observe that in any vector space V , $0v = \theta$ for each $v \in V$. Thus the zero vector is a linear combination of any nonempty subset of V .

Or you could interpret that the linear combination is a sum of products of each scalar-vector pair. For example, the following is also a linear combination itself.

$$a_1u_1 + a_2u_2 + \dots + a_nu_n.$$

The definition of linear dependence


Definition. A subset S of a vector space V is called **linearly dependent** if there exist a finite number of distinct vectors u_1, u_2, \dots, u_n in S and scalars a_1, a_2, \dots, a_n , not all zero, such that

$$a_1u_1 + a_2u_2 + \cdots + a_nu_n = 0.$$

In this case we also say that the vectors of S are linearly dependent.

This definition will be important later when we will prove that **every vector space has a basis**.

The linear independence

 **Definition.** A set of vectors $\{v_1, v_2, \dots, v_k\}$ is *linearly independent* if the vector equation

$$x_1 v_1 + x_2 v_2 + \dots + x_k v_k = 0$$

has only the trivial solution $x_1 = x_2 = \dots = x_k = 0$. The set $\{v_1, v_2, \dots, v_k\}$ is *linearly dependent* otherwise.

Note that each indexed “**x**” value associated with a vector is a scalar coming from the field F .

The span of the set of vectors

Definition. Let S be a nonempty subset of a vector space V . The **span** of S , denoted **span**(S), is the set consisting of all linear combinations of the vectors in S . For convenience, we define **span**(\emptyset) = $\{0\}$.

The caveat is that the span is a set of all linear combinations of the vectors in the set. For example, if the field we are working with is infinite, then the span possibly contains infinite elements as well.

Another caveat is that this type of linear combination only makes sense for finitely many vectors. Also, the following is an alternative definition of span. Note that this definition does not always work for infinite dimensional vector spaces.

$$\text{span}(S) = \left\{ \sum_{i=1}^k \lambda_i \mathbf{v}_i \mid k \in \mathbb{N}, \mathbf{v}_i \in S, \lambda_i \in K \right\}.$$

The set of vectors generates (spans) a vector space V

Definition. *A subset S of a vector space V generates (or spans) V if $\text{span}(S) = V$. In this case, we also say that the vectors of S generate (or span) V .*

If the set of all linear combinations of the set of vectors covers the entire vector space, then we say the set of vectors S , **generates** the vector space V .

This is just a terminology you need to know.

Now, the definition of Basis

Definition. A basis β for a vector space V is a linearly independent subset of V that generates V . If β is a basis for V , we also say that the vectors of β form a basis for V .

This definition consists of two important conditions:

- 1, A basis β is a linearly independent set of vectors and a subset of V .**
- 2, A basis β generates (spans) a vector space V .**

This is all that it says.

Some examples of basis

Example 1

Recalling that $\text{span}(\emptyset) = \{0\}$ and \emptyset is linearly independent, we see that \emptyset is a basis for the zero vector space. ♦

Example 2

In F^n , let $e_1 = (1, 0, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, \dots , $e_n = (0, 0, \dots, 0, 1)$; $\{e_1, e_2, \dots, e_n\}$ is readily seen to be a basis for F^n and is called the **standard basis** for F^n . ♦

Example 4

In $P_n(F)$, the set $\{1, x, x^2, \dots, x^n\}$ is a basis. We call this basis the **standard basis** for $P_n(F)$. ♦

Example 5

In $P(F)$, the set $\{1, x, x^2, \dots\}$ is a basis. ♦

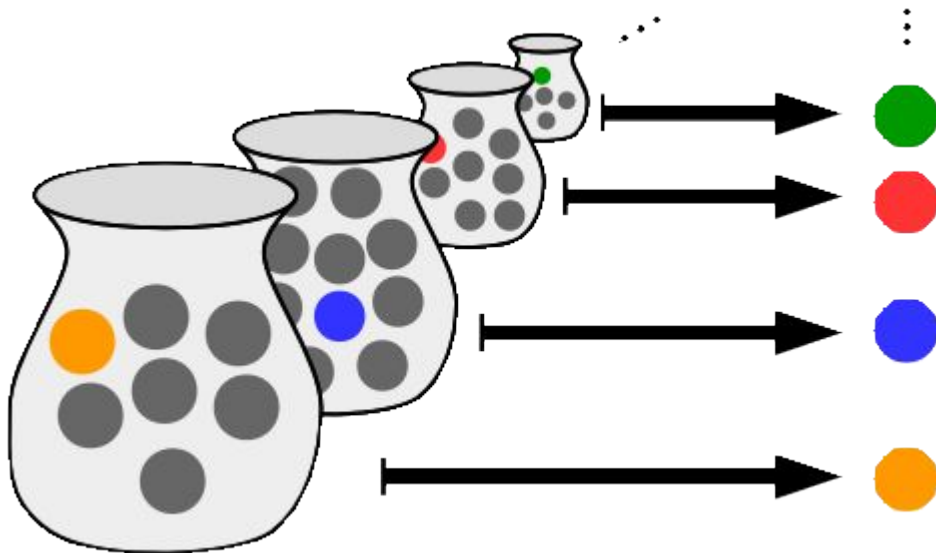
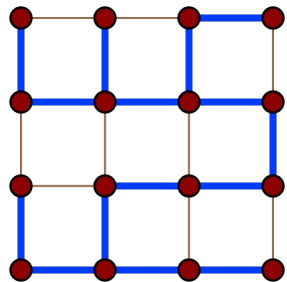
End of review on the linear algebra part

Part 2: Some Set Theory Background

Now, we will cover the basics of set theory we need to know in order to understand the proof of “**every vector space has a basis.**”

Some topics we will cover: Partially ordered set (or poset), Axiom of choice, Well-ordering principle, Zorn’s lemma

$$P = (X, \leq)$$



A little review before covering Axiom of Choice

We need to understand what **Cartesian Product** is.

In **mathematics**, specifically **set theory**, the **Cartesian product** of two **sets** A and B , denoted $A \times B$, is the set of all **ordered pairs** (a, b) where a is in A and b is in B .^[1] In terms of **set-builder notation**, that is

$$A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}.$$

The above is the definition of the cartesian product of two sets. We also generalize this to n many sets and an infinite collection of sets.

Cartesian Product of a collection of n (finite) Many Sets

***n*-ary Cartesian product** [\[edit \]](#)

The Cartesian product can be generalized to the ***n*-ary Cartesian product** over n sets X_1, \dots, X_n as the set

$$X_1 \times \cdots \times X_n = \{(x_1, \dots, x_n) \mid x_i \in X_i \text{ for every } i \in \{1, \dots, n\}\}$$

of ***n*-tuples**. If tuples are defined as **nested ordered pairs**, it can be identified with $(X_1 \times \dots \times X_{n-1}) \times X_n$. If a tuple is defined as a function on $\{1, 2, \dots, n\}$ that takes its value at i to be the i -th element of the tuple, then the Cartesian product $X_1 \times \dots \times X_n$ is the set of functions

$$\{x : \{1, \dots, n\} \rightarrow X_1 \cup \cdots \cup X_n \mid x(i) \in X_i \text{ for every } i \in \{1, \dots, n\}\}.$$

The first definition may be more intuitive. There is an alternative interpretation for an **n-ordered tuple**. We can see each tuple as a function that maps each index between 1 and n , to each corresponding set X_i . Below is the illustration of this alternative interpretation. For some function x , it is indeed some element (x_1, \dots, x_n)

$$\{x : \{1, \dots, n\} \rightarrow X_1 \cup \cdots \cup X_n \mid x(i) \in X_i \text{ for every } i \in \{1, \dots, n\}\}.$$

Cartesian Product for an infinite collection of sets

Consider the following.

It is possible to define the Cartesian product of an arbitrary (possibly **infinite**) **indexed family** of sets. If I is any **index set**, and $\{X_i\}_{i \in I}$ is a family of sets indexed by I , then the Cartesian product of the sets in $\{X_i\}_{i \in I}$ is defined to be

$$\prod_{i \in I} X_i = \left\{ f : I \rightarrow \bigcup_{i \in I} X_i \mid \forall i \in I. f(i) \in X_i \right\},$$

that is, the set of all functions defined on the **index set** I such that the value of the function at a particular index i is an element of X_i . Even if each of the X_i is nonempty, the Cartesian product may be empty if the

This definition is quite important, especially when studying axiom of choice.

Caveat: A cartesian product for an infinite collection of sets is a set of mappings from an arbitrary index set to each index-wise corresponding set.

Axiom of Choice (AoC)

There are roughly two or more definitions of Axiom of Choice. One simple definition of Axiom of Choice is as follows.

Axiom of Choice If $\{A_\alpha : \alpha \in \Lambda\}$ is a collection of nonempty sets, then their product

$$\prod_{\alpha \in \Lambda} A_\alpha$$

is nonempty as well.

Caveat 1: Each set in the collection of sets is assumed to be non-empty.

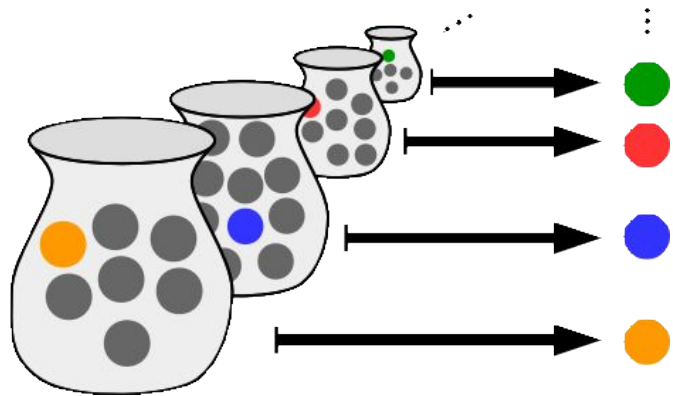
Caveat 2: Translating into a simple language, this just says that a cartesian product of non-empty sets is again non-empty.

This is intuitively very obvious, but extremely useful in set theory.

Axiom of Choice (AoC)

A **choice function** (also called selector or selection) is a function f , defined on a collection X of nonempty sets, such that for every set A in X , $f(A)$ is an element of A . With this concept, the axiom can be stated:

Axiom — For any set X of nonempty sets, there exists a choice function f that is defined on X and maps each set of X to an element of that set.



Formally, this may be expressed as follows:

$$\forall X \left[\emptyset \notin X \implies \exists f: X \rightarrow \bigcup_{A \in X} A \quad \forall A \in X (f(A) \in A) \right].$$

Axiom of Choice states the existence of such choice function given a collection of non-empty sets.

Caveat: X is a collection of non-empty sets, not the set to which a function maps.

Partially ordered set (poset)

We need to know what a partial order is first. The following is its definition.

Definition A *partial order* on a set P is a binary relation \leq on P such that

1. For all $p \in P$, $p \leq p$.
2. For all $p, q \in P$, if $p \leq q$ and $q \leq p$ then $p = q$.
3. For all $p, q, r \in P$, if $p \leq q$ and $q \leq r$ then $p \leq r$.

If any two elements of P are *comparable*, i.e., for all $p, q \in P$, either $p \leq q$ or $q \leq p$, then we will say that the order relation \leq is *total* or *linear*.

For any poset, it is **not** necessary that all elements are comparable. However, if all elements in a poset are comparable, we say the order relation is total or linear. We call such set a **totally ordered set**.

Some important definitions regarding poset

Let (P, \leq) be a partially ordered set:

- ▶ $p \in P$ is *maximal* if there is no $q \in P$ with $p \lessneq q$.
- ▶ $q \in P$ is *minimal* if there is no $p \in P$ with $p \lessneq q$.
- ▶ $Q \subseteq P$ is *bounded* if there is some $r \in P$ such that $q \leq r$ for all $q \in Q$.
- ▶ $C \subseteq P$ is a *chain* if the restriction of the order relation \leq to C is total.

A partially ordered set is a set P with an order relation. In general, the order relation is a partial order.

In a chain C , the restriction of the order relation to C is a total order. In other words, all elements in C are comparable to each other.

Zorn's lemma

Zorn's Lemma Every nonempty partially ordered set in which every chain is bounded has a maximal element.

Here is what it says:

For any non-empty partially ordered set (P, \leq) , if every chain C in P is bounded in a poset, then a poset P has a maximal element.

- ▶ $Q \subseteq P$ is *bounded* if there is some $r \in P$ such that $q \leq r$ for all $q \in Q$.
- ▶ $C \subseteq P$ is a *chain* if the restriction of the order relation \leq to C is total.

The proof of Zorn's lemma is truly out of scope, (possibly it is what we do in grad school), so I do not do the proof of Zorn's lemma here.

The Zorn's lemma is useful for lots of things in mathematics. For example, we will use this lemma to prove that “every vector space has a basis.”

Well order and well-ordering principle

Here, we introduce the well-ordering principle in the context of set theory. First, the definition of well order on a set X is as follows.

Definition A total (linear) ordered set is a *well-order* if all of its nonempty subsets have a minimal element.

Now, here comes what well-ordering principle states.

Well-ordering Principle Every set is well-orderable.

Note: The well-ordering principle implies we can put a well-order on any given set. This implies even \mathbb{R} (a set of real numbers) is well-ordered. Of course, an ordinary order relation does not work. We all do not know how a well order on \mathbb{R} looks like, but we know it exists according to this principle.

Food for thought: Well ordering principle is equivalent to Axiom of Choice.

End of Set Theory Part

Part 3: Every vector space has a basis

We will go over the proof, and we will use the Zorn's lemma. The Zorn's lemma is as follows.

Zorn's Lemma Every nonempty partially ordered set in which every chain is bounded has a maximal element.

We consider some arbitrary vector space V . We will consider two cases.

Case 1: It is a zero vector space. Recall that the zero vector space is a vector space whose only element is a zero vector.

Then, the empty set is a basis for a zero vector space, following the convention.

Case 2: The vector space is not a zero vector space.

This part takes some effort.

Proof of “Every vector space has a basis”

If $V = \{\mathbf{0}\}$, then the empty set is a basis for V . Now, suppose that $V \neq \{\mathbf{0}\}$. Let P be the set consisting of all linearly independent subsets of V . Since V is not the zero vector space, there exists a nonzero element \mathbf{v} of V , so P contains the linearly independent subset $\{\mathbf{v}\}$. Furthermore, P is partially ordered by set inclusion (see inclusion order). Finding a maximal linearly independent subset of V is the same as finding a maximal element in P .

It is straightforward to verify that P is a partially ordered set where its order relation is a set inclusion “ \subseteq ”

Notice that not every element in P is comparable. For example, consider $A = \{\mathbf{v}_1, \mathbf{v}_2\}$, $B = \{\mathbf{v}_3\}$, where A and B are linearly independent subsets of V . then A and B are not comparable since the set inclusion does not happen.

Also, consider $C = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, and $D = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6, \mathbf{v}_7\}$, where C and D are linearly independent subsets of V .

Then we have $C \subseteq D$. Notice that the order relation satisfies.

Proof of “Every vector space has a basis” (continued)

In order to apply Zorn’s lemma, we need to show every chain of the poset P , (the set of all linearly independent subsets of V), is **bounded** in P .

Below shows if the chain is the empty set, then we can just take $\{v\}$ where v is non-zero

To apply Zorn’s lemma, take a chain T in P (that is, T is a subset of P that is totally ordered). If T is the empty set, then $\{v\}$ is an upper bound for T in P . Suppose then that T is non-empty. We need to show that T has an upper bound, that is, there exists a linearly independent subset B of V containing all the members of T .

Q, What would be an upper bound of a chain T ?

A, The union of all sets in the chain T

Let’s call this union “ B .”

We need to show that B exists in the poset P as an upper bound of a chain T . In other words, B is also **a linearly independent subset of V .**

Proof of “Every vector space has a basis” (continued)

We show by contradiction that B is a linearly independent subset of V .

Suppose otherwise, that B is not linearly independent. Then there exists vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in B$ and **scalars** a_1, a_2, \dots, a_k , not all zero, such that

$$a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \cdots + a_k \mathbf{v}_k = \mathbf{0}.$$

Since B is the union of all the sets in T , there are some sets $S_1, S_2, \dots, S_k \in T$ such that $\mathbf{v}_i \in S_i$ for every $i = 1, 2, \dots, k$. As T is totally ordered, one of the sets S_1, S_2, \dots, S_k must contain the others, so there is some set S_i that contains all of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$. This tells us there is a linearly dependent set of vectors in S_i , contradicting that S_i is linearly independent (because it is a member of P).

Since a chain T is a totally ordered set, there is a chain relation that looks like below

$S_1 \subseteq S_2 \subseteq S_3 \subseteq S_4 \dots$ where each S_i is an element in a chain T

Since we are considering a **finite** linear combination here, this chain relation **terminates** at some point, and hence there is a set S that contains all other elements in T .

Proof of “Every vector space has a basis” (continued)

We have just shown that every chain T in a poset P where P is a set of linearly independent subsets of V , is bounded. Notice that the hypothesis of Zorn's lemma has been checked as well. Recall the Zorn's lemma below.

Zorn's Lemma Every nonempty partially ordered set in which every chain is bounded has a maximal element.

Now, Zorn's lemma states that there is a **maximal** linearly independent subset of V in a poset P .

Now, our final task is to show that B is indeed a basis of V .

Since we know that this maximal element in P is linearly independent, we only need to show that this maximal element of P **spans** or **generates** V .

Proof of “Every vector space has a basis” (continued)

Let's call this maximal linearly independent set **B**.

We will use one extremely important fact from linear algebra. It is as follows.

Theorem 1.7. Let S be a linearly independent subset of a vector space V , and let v be a vector in V that is not in S . Then $S \cup \{v\}$ is linearly dependent if and only if $v \in \text{span}(S)$.

Now, let's start the proof. We will prove that a maximal linearly independent set **B** spans or generates V by **contradiction**.

Suppose for the contradiction that **B** does **not** span V .

Then there exists some vector v in V such that v not in $\text{span}(B)$.

The theorem above implies that $\{v\} \cup B$ is a linearly independent subset of V , which is larger than **B**, contradicting the maximality of **B**, thus contradicting the Zorn's lemma. Therefore, **B** is a desired basis for V .

We finished the proof, how is it useful?

We can simply assume that given any vector space, there is a basis for it.

We can also notice that the basis of V is any maximal linearly independent subset of V .

One great thing is that this also generalizes to an **infinite** dimensional vector spaces. It is generally far more difficult to deal with infinite cases in mathematics. Almost everything is far more trivial in finite cases.

Now, we move to the final part of this slide.

Part 4: cool application of “Every vector space has a basis”

We will solve the following problem to demonstrate how this fact can be useful.

8. Recall that the *distance* between two real numbers is defined as the function $d: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ given by $d\langle x, y \rangle = |x - y|$. We will say that a subset $A \subseteq \mathbb{R}$ is *distance-injective* if for all $x, y, z, w \in A$, if $d\langle x, y \rangle = d\langle z, w \rangle > 0$, then $\{x, y\} = \{z, w\}$.

Prove that there exists an uncountable, distance-injective subset of \mathbb{R} .

What this question says:

**In a distance-injective subset A of \mathbb{R} ,
consider taking two pairs of two points in A , $\{x, y\}$ and $\{z, w\}$.**

Then if $|x - y| = |z - w|$, then these two pairs are the same.

**For example, such subset cannot contain 1, 3 and 5 at the same time.
since $|5 - 3| = |3 - 1|$ and $\{1, 3\} \neq \{3, 5\}$, not distance-injective.**

Distance-injective question (continued)

8. Recall that the *distance* between two real numbers is defined as the function $d: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ given by $d\langle x, y \rangle = |x - y|$. We will say that a subset $A \subseteq \mathbb{R}$ is *distance-injective* if for all $x, y, z, w \in A$, if $d\langle x, y \rangle = d\langle z, w \rangle > 0$, then $\{x, y\} = \{z, w\}$.

Prove that there exists an uncountable, distance-injective subset of \mathbb{R} .

To solve this question, you can assume the following:

1, \mathbb{R} is uncountable.

2, The basis of the vector space \mathbb{R} over the field of rational numbers \mathbb{Q} is uncountable. This follows from the fact that \mathbb{Q} is countable and the set of all finite subsets of a countable set is again countable. (you don't need to know these for now)

You don't need to worry about the definition of uncountability for now.

By Zorn's lemma, since every vector space has a basis, let's take the **basis for the vector space \mathbb{R} over the field of rational numbers \mathbb{Q} .**

Distance-injective question (continued)

8. Recall that the *distance* between two real numbers is defined as the function $d: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ given by $d\langle x, y \rangle = |x - y|$. We will say that a subset $A \subseteq \mathbb{R}$ is *distance-injective* if for all $x, y, z, w \in A$, if $d\langle x, y \rangle = d\langle z, w \rangle > 0$, then $\{x, y\} = \{z, w\}$.

Prove that there exists an uncountable, distance-injective subset of \mathbb{R} .

Recall that we picked a basis of \mathbb{R} over a field of rational numbers.

We need to show that this basis is distance-injective.

We use the fact that the basis is linearly independent.

Consider two pairs of two distinct points, whose distance between the points is equal. Specifically, take $\{x, y\}$ and $\{z, w\}$ such that $|x - y| = |z - w|$

Distance-injective question (continued)

8. Recall that the *distance* between two real numbers is defined as the function $d: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ given by $d\langle x, y \rangle = |x - y|$. We will say that a subset $A \subseteq \mathbb{R}$ is *distance-injective* if for all $x, y, z, w \in A$, if $d\langle x, y \rangle = d\langle z, w \rangle > 0$, then $\{x, y\} = \{z, w\}$.

Prove that there exists an uncountable, distance-injective subset of \mathbb{R} .

Consider two pairs of two distinct points, whose distance between the points is equal. Specifically, take $\{x, y\}$ and $\{z, w\}$ such that $|x - y| = |z - w|$

Case 1: WLOG, suppose $x < y$ and $z < w$

then we have $y - x = w - z$. Then this contradicts a linear independence of the basis since we have $y - x - w + z = 0$

Case 2: WLOG suppose $x = z$, then $w = y$, then we are done.

Case 3: WLOG suppose $y = z$ and $x < y$ and $z < w$.

Then we have $y - x = w - z = w - y$, then $2y - x - w = 0$, contradiction.

Distance-injective question (continued)

Hence, we have shown that the basis for the vector space of \mathbb{R} over the field of rational numbers is a distance-injective subset of \mathbb{R} . Hence, we solved the problem.