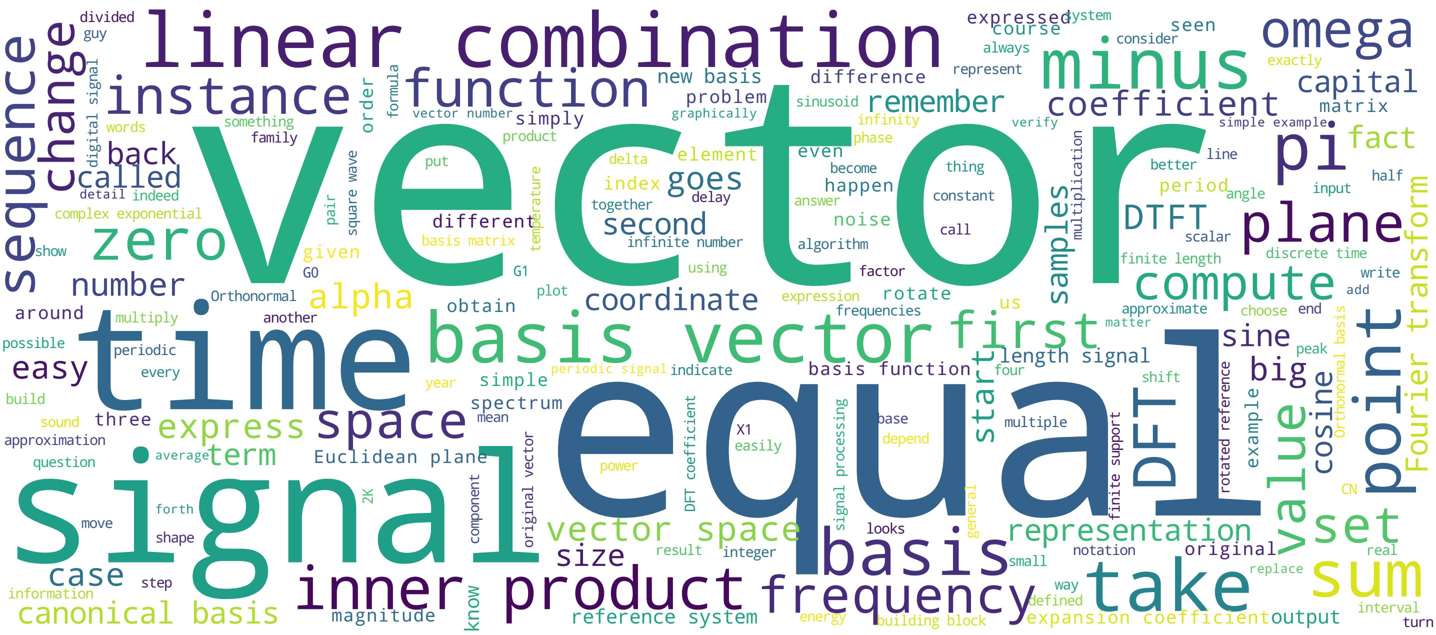
linear combination is the basic operation in vector spaces:

$$\mathbf{g} = \alpha\,\mathbf{x} + \beta\,\mathbf{y}$$

can we find a set of vectors $\{\mathbf{w}^{(k)}\}$ so that we can write *any* vector as a linear combination of the $\{\mathbf{w}^{(k)}\}$?



Search MOOC

Video

Bases

linear combination is the basic operation in vector spaces:

$$\mathbf{g} = \alpha\,\mathbf{x} + \beta\,\mathbf{y}$$

can we find a set of vectors $\{\mathbf{w}^{(k)}\}$ so that we can write *any* vector as a linear combination of the $\{\mathbf{w}^{(k)}\}$?

1

The axioms of a vector space tell us how to combine vectors together. So that a linear combination of vector is the basic operation that we perform in vector spaces. We take two or more vectors, X and Y, we rescale them via scalar multiplication, and then we combine them together by a vector addition. So when we combine vectors together to obtain new vectors in the space, a natural question is whether we can find a minimal set of vectors so that we can express any vector in the space as a linear combination of this basis vector. So this set of building blocks will be called the basis. You can think back of the Lego block analogy that we drew before. The question would be, can we find a set of minimal building blocks with which we can build any other piece in the set? A brief remark about notation when we want to indicate a family of vectors and we want to be able to index them, then we will use a superscript in parenthesis here that indicates the cardinal number of the vector in the family.

Notes

Summary

0m 00s

The canonical $\mathbb{R}^2$ basis

$$\mathbf{e}^{(0)} = \begin{bmatrix}1\\0\end{bmatrix} \qquad \mathbf{e}^{(1)} = \begin{bmatrix}0\\1\end{bmatrix}$$

$$\underbrace{\begin{bmatrix}x_0\\x_1\end{bmatrix}}_{\mathbb{R}^2} = x_0 \begin{bmatrix}1\\0\end{bmatrix} + x_1 \begin{bmatrix}0\\1\end{bmatrix}$$

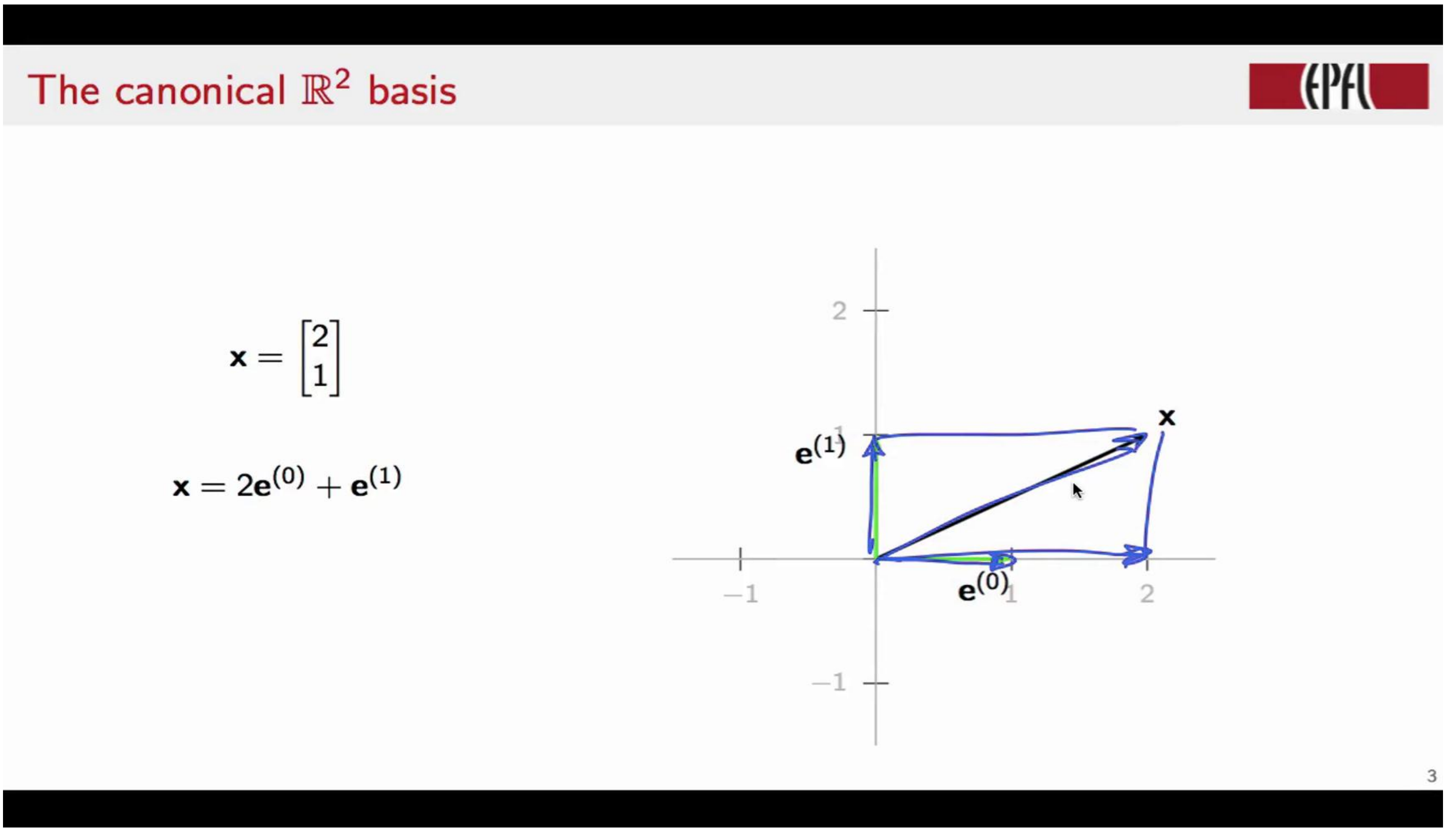
2

Let's start with a very simple example and introduce the canonical basis for the Euclidean plane. You remember, any vector in the Euclidean plane is a point on the plane, and this can be expressed as a pair of coordinates, X0 and X1 on the plane. Now, it's easy to see that if you take two vectors shaped like so, E0 will be the pair onE0, and E1 will be the pair zero one, then any vector in R2 can be expressed as a linear combination of E0 and E1. And the coefficients, the scalars that you will have to use in this linear combination are actually the values themselves in the original vector. The fact that the scalars in the linear combination coincide with the values of the elements of the vector is why we call this basis the canonical basis.

Notes

Summary

1m 05s



As a graphical example, consider the vector X of coordinates two and one. So this guy here, and we can see very easily that we can express this as the linear combination of two times the horizontal basis vector, canonical basis vector, plus one time the vertical canonical basis vector. So if we sum this two times zero plus this, we have the usual parallelogram rule and we get back our X.

Notes

Summary

1m 52s

Another $\mathbb{R}^2$ basis

$$\mathbf{v}^{(0)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \mathbf{v}^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \alpha_0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\alpha_0 = x_0 - x_1, \quad \alpha_1 = x_1$$

4

However, there are other bases that we could choose for the Euclidean plane. For instance, the basis given by the vectors 1, 0 and 1, 1 also represents a basis for the plane. And any vector of coordinates X0 and X1 can be expressed as a linear combination of V 0 and V 1. How do we find the coefficients in the linear combination? Well, we have a system of two equations and two unknowns here's alpha 0 and alpha 1 and we get that alpha 0 will be equal to X 0 minus X 1, and alpha 1 will be equal to X 1.

Notes

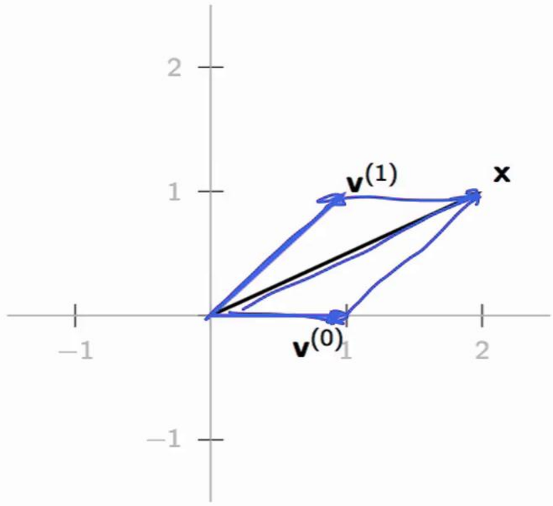
Summary

2m 20s

Another $\mathbb{R}^2$ basis

$$\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\mathbf{x} = \underset{1}{\mathbf{v}^{(0)}} + \underset{1}{\mathbf{v}^{(1)}}$$



5

So again, graphically, the same vector of coordinate two and one can be expressed as a linear combination of V 0 and V 1. And now the coefficients are different because they will be 1 and 1 and you can see, indeed, if you sum this guy plus this guy, again, with a parallelogram rule, you get back X. There are, in fact, an infinite number of bases for R 2, but not all pairs of vectors represent the basis for the plane.

Notes

Summary

2m 55s

But this is not a basis for $\mathbb{R}^2$...

$$\mathbf{g}^{(0)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \mathbf{g}^{(1)} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x_0 \\ x_1 \end{bmatrix} \neq \alpha_0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha_1 \begin{bmatrix} -1 \\ 0 \end{bmatrix} \text{ whenever } x_1 \neq 0$$

6

Here's a counter example. If you take G0, the vector of coordinates 1 and 0, so equal to the first canonical basis vector, and then you take G1 as minus 1, 0, then it's easy to show that it's not possible to express any point on the plane as a linear combination of G0 and G1. As a matter of fact, you cannot express any vector for which the second component is not equal to zero.

Notes

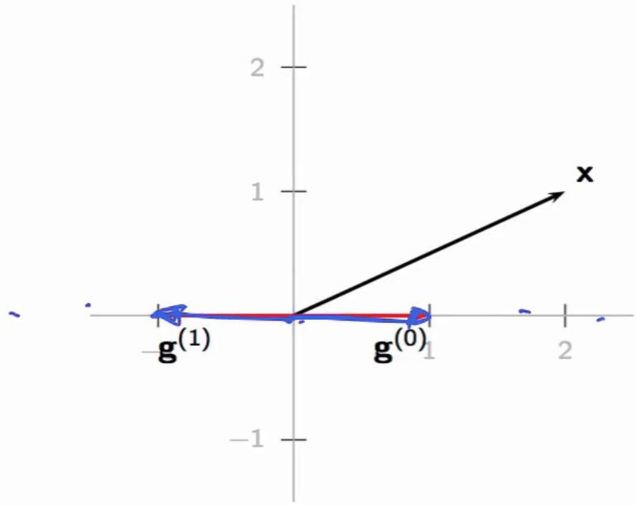
Summary

3m 27s

Another $\mathbb{R}^2$ basis

$$\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\mathbf{x} \neq \alpha_1 \mathbf{g}^{(0)} + \alpha_2 \mathbf{g}^{(1)}$$



7

Graphically, that's easy to see because you can see that G0 and G1 are actually collinear. They're in the same line on the plane. And so they can only capture 1 dimensional component of the plane. So any linear combination of G0 and G1 will be a point on this line, but you will not be able to reach the other points on the plane. We express this by saying that the vectors are not linearly independent.

Notes

Summary

3m 51s

What about infinte-dimensional spaces?

$$\mathbf{x} = \sum_{k=0}^{\infty} \alpha_k \mathbf{w}^{(k)}$$

8

Okay, so the Euclidean plane, Euclidean space, and in general, finite dimensional spaces are well behaved animals. And so it's easy to believe that there will always be a basis for this basis. But what about infinite dimensional spaces? We have seen, for instance, the space of infinite length signals. What is the basis for that? Can we write linear combinations of an infinite number of basis vectors?

Notes

Summary

4m 19s

A basis for $\ell_2(\mathbb{Z})$

$$\mathbf{e}^{(k)} = \begin{bmatrix} \vdots \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

k-th position, $k \in \mathbb{Z}$

9

So with some precautions, like for instance, in the case of square summable sequences, the answer is yes. And the canonical basis for L2 of Z, for instance, is a collection of an infinite number of vectors in which there's only one nonzero component which is equal to one. Even more interestingly, we can develop bases for function vector spaces.

Notes

Summary

4m 44s

What about function vector spaces?

$$f(t) = \sum_k \alpha_k h^{(k)}(t)$$

10

Take for instance, the space of square integral functions over an interval, you can express any function in that vector space as a linear combination of basis vectors. What are these basis vectors? Again, there exists n infinite number of bases for this space, but one of the most famous ones is the Fourier basis.

Notes

Summary

5m 05s

A basis for the functions over an interval?

the Fourier basis for $[-1, 1]$:

$$\frac{1}{\sqrt{2}},\ \cos\pi t,\ \sin\pi t,\ \cos 2\pi t,\ \sin 2\pi t,\ \cos 3\pi t,\ \sin 3\pi t,\ \ldots$$

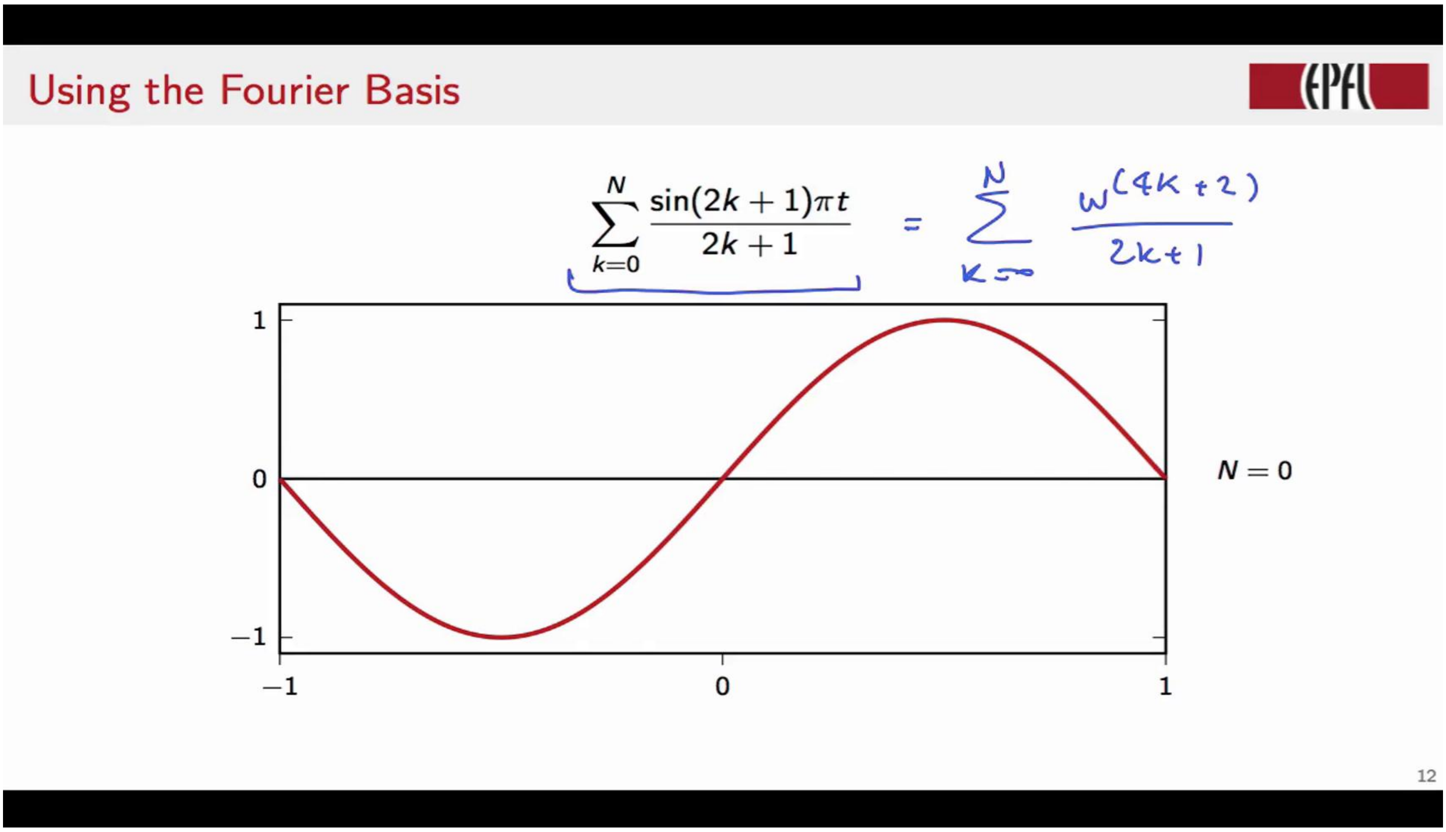
11

In this case for, say, the interval minus 1,1. We can order the basis vectors and we have that basis vector zero is the constant one over square root two, basis vector one is cosine of Pi t, basis vector two is sine of Pi t, and so on, and so forth. Now, can we really represent any square integral function over the minus 1, 1 interval as a linear combination of these basis functions? The answer is yes, and this leads to some surprising results. For instance, we can approximate a discontinuous function such as a square wave as a linear combination of continuous functions, a combination of sines.

Notes

Summary

5m 24s



Consider the following sum here for K that goes from zero to capital N of sine of 2K plus 1 Pi T over 2K plus 1. If you consider the order of the basis functions that we showed in the previous slide, this is equal to the sum for K that goes to zero to capital N of a basis function W of 4K plus 2 divided by 2K plus 1. So it is a linear combination of basis functions, and the basis function of index 4K plus 2 is normalized by the factor of 1 over 2K plus 1. So in this figure here, you see what happens when N is equal to zero. So you have just the first term in the sum is just the sign of period Pi T.

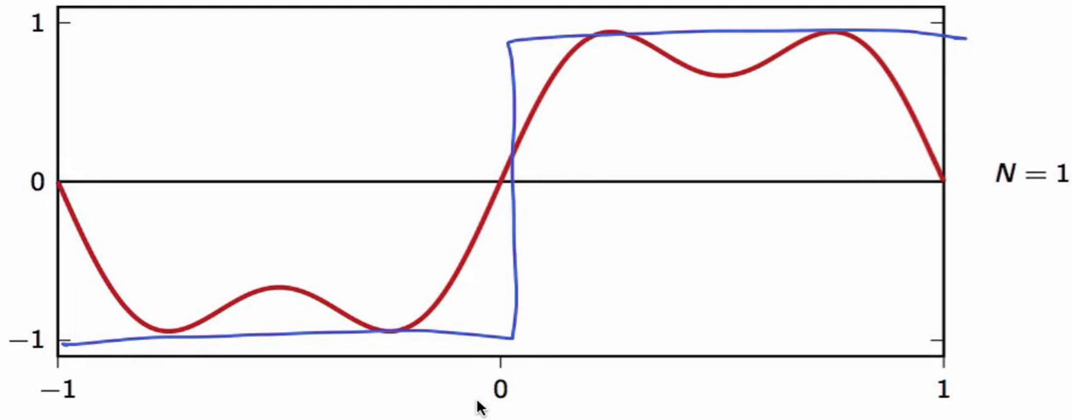
Notes

Summary

6m 05s

Using the Fourier Basis

$$\sum_{k=0}^{N} \frac{\sin(2k+1)\pi t}{2k+1}$$



12

As you increase the number of terms, so N equal to one. You see that the shape starts to converge to that of a square wave, which, to remind you, will look exactly like this. So 1 from 0 to 1 and minus 1 from 0 to minus 1.

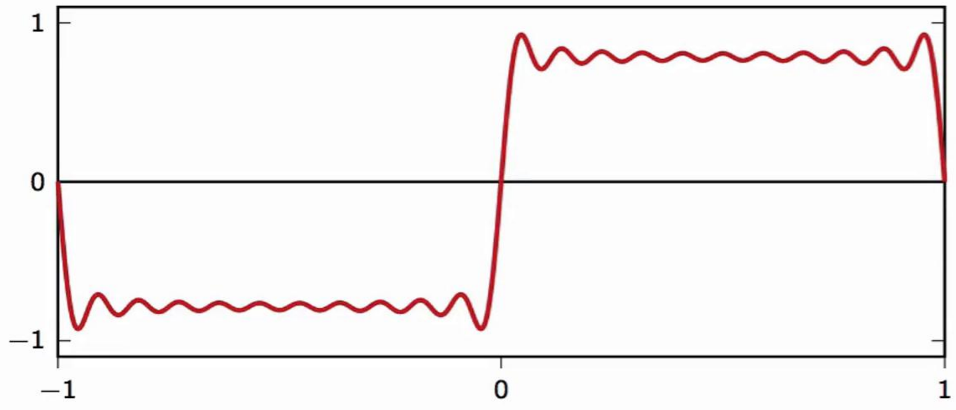
Notes

Summary

6m 55s

Using the Fourier Basis

$$\sum_{k=0}^{N} \frac{\sin(2k+1)\pi t}{2k+1}$$



$N = 10$

12

So we increment the number of terms, and the more we add. The more we approximate the function that we want to approximate.

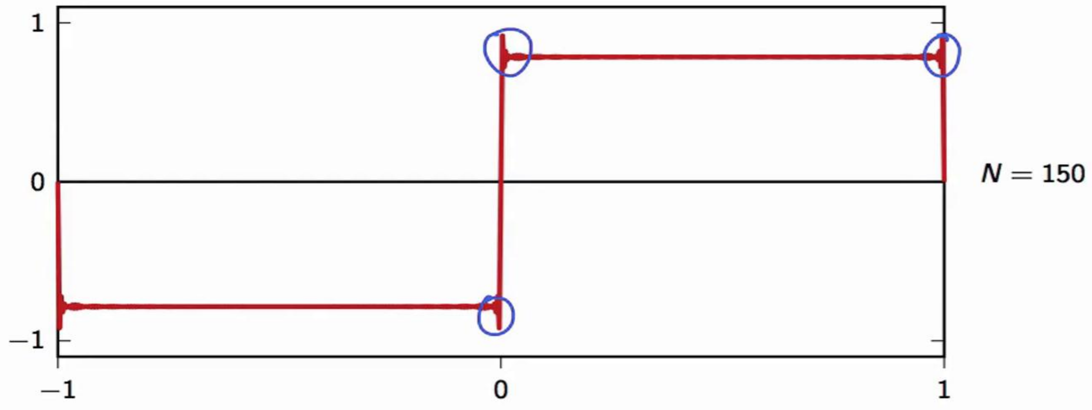
Notes

Summary

7m 13s

Using the Fourier Basis

$$\sum_{k=0}^{N} \frac{\sin(2k+1)\pi t}{2k+1}$$



Now, of course, you can see that as the number of terms grows. The approximation is better and better. But something is going on here at the transition point. This wiggle. That you see around the transition point is called the Gibb's phenomenon, and it's due to the fact that the convergence of this sum to the actual square wave only happens in the mean square sense and not point wise. We will see this in more detail when we study filter design.

Notes

Summary

7m 21s

Bases: formal definition

Given:

- a vector space H
- a set of K vectors from H: $W = \{\mathbf{w}^{(k)}\}_{k=0,1,\ldots,K-1}$

W is a basis for H if:

1. we can write for *all* $\mathbf{x} \in H$:
$$\mathbf{x} = \sum_{k=0}^{K-1} \alpha_k \mathbf{w}^{(k)}, \quad \alpha_k \in \mathbb{C}$$
2. the coefficients α_k are unique

13

Okay, so after the examples, let's formalize the concept of basis for a vector space H. We take a set of K vectors from H and we call the set W. W Will be a basis if two things happen. First, we need to be able to express any vector in the space as a linear combination of vectors from W. And two, the coefficients in this linear combination must be unique.

Notes

Summary

7m 54s

Bases: formal definition

Unique representation implies linear independence:

$$\sum_{k=0}^{K-1} \alpha_k \mathbf{w}^{(k)} = 0 \quad \Longleftrightarrow \quad \alpha_k = 0, \; k = 0, 1, \ldots, K-1$$

14

Uniqueness of the representation is equivalent to linear independence of the vector in the basis. In other words, if we take a linear combination of basis vectors and we obtain the null vector, this implies that the coefficients in the expansion are identically zero and vice versa.

Notes

Summary

8m 19s

Special bases

Orthogonal basis:

$$\langle \mathbf{w}^{(k)}, \mathbf{w}^{(n)} \rangle = 0 \text{ for } k \neq n$$

Orthonormal basis:

$$\langle \mathbf{w}^{(k)}, \mathbf{w}^{(n)} \rangle = \delta[n-k]$$

(we can always orthonormalize a basis via the Gram-Schmidt algorithm)

15

Of all the possible basis for a space, the most important ones are orthogonal basis. In orthogonal basis, the basis vectors are mutually orthogonal. And if we are particularly lucky, these vectors will be also unit norm, which gives us an Orthonormal basis. We can express the key property of an Orthonormal basis in compact form like so. The inner product between any two vectors in the basis is equal to the delta of the difference between the indices. So in other words, if the vectors are not the same vector, the inner product will be equal to zero. If it's the self inner product of a basis vector, the inner product will be equal to one. The good news is that we have an algorithm called the Gramm Schmidt Orthonormalisation procedure that will allow us to turn any basis for a space into an ortho normal basis if we need it.

Notes

Summary

8m 38s

Basis expansion

$$\mathbf{x} = \sum_{k=0}^{K-1} \alpha_k \mathbf{w}^{(k)}$$

how do we find the α's ?

Orthonormal bases are the best:

$$\alpha_k = \langle \mathbf{w}^{(k)}, \mathbf{x} \rangle$$

16

So if we have a basis, the question is, how do we find the expansion coefficients, the scalars that we have to use in this representation for a vector that is a linear combination of basis vectors? In general, the answer could be quite involved, but if we have an Orthonormal basis, then the procedure is extremely simple. If we take the inner product between the basis vector number K and the original vector, we directly obtain the expansion coefficient number K. You can verify this very easily by replacing here X with this expression and then considering the Orthonormality of the basis.

Notes

Summary

9m 29s

Change of basis

$$\mathbf{x} = \sum_{k=0}^{K-1} \alpha_k \mathbf{w}^{(k)} = \sum_{k=0}^{K-1} \beta_k \mathbf{v}^{(k)}$$

if $\{\mathbf{v}^{(k)}\}$ is orthonormal:

$$\beta_h = \langle \mathbf{v}^{(h)}, \mathbf{x} \rangle$$

$$= \langle \mathbf{v}^{(h)}, \sum_{k=0}^{K-1} \alpha_k \mathbf{w}^{(k)} \rangle$$

$$= \sum_{k=0}^{K-1} \alpha_k \langle \mathbf{v}^{(h)}, \mathbf{w}^{(k)} \rangle$$

17

Another common question when using basis is how to change the representation of a vector from one basis to another. Take a vector X and consider its representation as a linear combination of basis vectors from a basis W of K. Now, the question is to find the expansion coefficients of the same vector using a different basis V of K. If V of K is Orthonormal, the solution is extremely elegant. The expansion coefficient number H in the new basis can be obtained as we have seen in the previous slide, as the inner product between the basis vector number H in the new basis and the original vector. Now, let's replace the expression for X by the linear combination of basis vectors from the original basis. So we get this expression here, and now we exploit the distributivity of the inner product to obtain this expression here, where we can see that each new expansion coefficient is actually a linear combination of the previous expansion coefficients.

Notes

Summary

10m 04s

Change of basis

$$\beta_h = \sum_{k=0}^{K-1} \alpha_k \boxed{\langle \mathbf{v}^{(h)}, \mathbf{w}^{(k)} \rangle}$$

$$= \sum_{k=0}^{K-1} \alpha_k \, c_{hk}$$

$$= \underbrace{\begin{bmatrix} c_{00} & c_{01} & \cdots & c_{0(K-1)} \\ & & \vdots & \\ c_{(K-1)0} & c_{(K-1)1} & \cdots & c_{(K-1)(K-1)} \end{bmatrix}} \begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_{K-1} \end{bmatrix}$$

18

What is remarkable however, is that the factors used in this linear combination do not depend on the vectors that we are manipulating. These are just inner products between vectors in the original basis and the new basis. As a matter of fact, we can write this sum here as a matrix vector multiplication, and this matrix here contains entries that are all the possible cross products between vectors in the original and in the new basis. In other words, given a starting basis and an orthogonal arrival basis, you can compute this change of basis matrix once and for all. And every time you have to change the reference system for a vector, you just need to compute a matrix vector multiplication. We will see very soon that this is really what's behind the discrete Fourier transform algorithm for finite length signals.

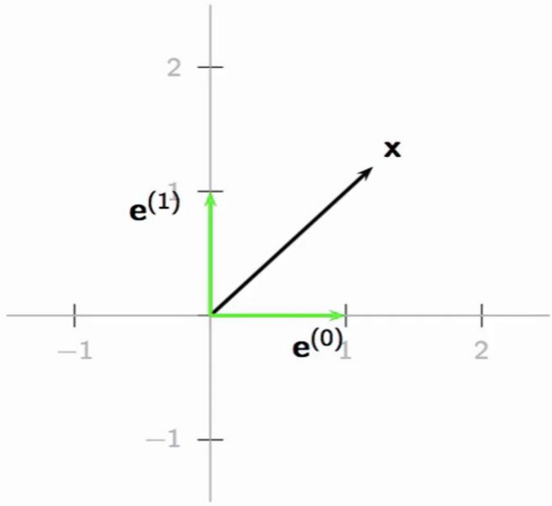
Notes

Summary

11m 08s

Change of basis: example

- canonical basis $E = \{\mathbf{e}^{(0)}, \mathbf{e}^{(1)}\}$
- $\mathbf{x} = \alpha_0 \mathbf{e}^{(0)} + \alpha_1 \mathbf{e}^{(1)}$



19

In the meantime, let's just show a simple example using an elementary rotation of the Euclidean plane. Let's start with a canonical basis, E0, E1. We know this basis to be Orthonormal. Any vector in the plane can be expressed as a linear combination of the two canonical vectors.

Notes

Summary

11m 58s

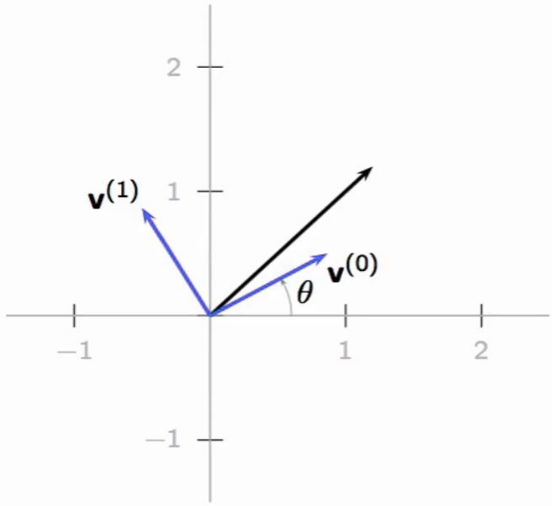
Change of basis: example

- canonical basis $E = \{\mathbf{e}^{(0)}, \mathbf{e}^{(1)}\}$
- $\mathbf{x} = \alpha_0 \mathbf{e}^{(0)} + \alpha_1 \mathbf{e}^{(1)}$
- new basis $V = \{\mathbf{v}^{(0)}, \mathbf{v}^{(1)}\}$ with

$$\mathbf{v}^{(0)} = \begin{bmatrix} \cos\theta & \sin\theta \end{bmatrix}^T$$

$$\mathbf{v}^{(1)} = \begin{bmatrix} -\sin\theta & \cos\theta \end{bmatrix}^T$$

- $\mathbf{x} = \beta_0 \mathbf{v}^{(0)} + \beta_1 \mathbf{v}^{(1)}$



19

Now we rotate the plane, so we rotate the reference system by considering a new basis, V0 and V1 with V0 equal to cosine of thetasine of theta, and V1 minus sine of theta, cosine of theta. Theta is the angle that we rotate the plane by. We want to find the coordinates of a vector in this new rotated reference system. It is easy to verify with basic trigonometry that the new basis is also Orthonormal, so we can use the previous result and try to compute this change of basis matrix.

Notes

Summary

12m 15s

Change of basis: example

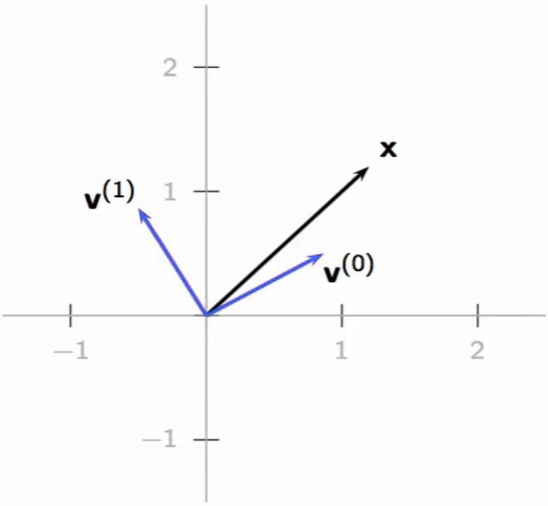
- new basis is orthonormal:

$$c_{hk} = \langle \mathbf{v}^{(h)}, \mathbf{e}^{(k)} \rangle$$

- in compact form:

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \mathbf{R}\boldsymbol{\alpha}$$

- $\mathbf{R}$: rotation matrix
- key fact: $\mathbf{R}^T\mathbf{R} = \mathbf{I}$



20

It will be a two by two matrix where the entries are going to be given by the inner product between every vector in the rotated reference system and a vector in the canonical system. Now, because this vector is very simple, they're just 1, 0 and 0, 1, this inner products are going to be very easy to compute and the change of basis matrix has this form here. The coefficients in the rotated reference system will be given by the product between the change of basis matrix and the coefficients in the original system. Well, we were starting rom the canonical reference system, so these are also the coordinates of the point in the plane. R is called the rotation matrix, and this is what you would use, for instance, in computer graphics, to rotate a set of coordinates to rotate an object on the plane. A key fact that depends on the orthonormality of the rotated basis is that if we take the rotationmatrix and multiply it by its transpose, we obtain the identity matrix.

Notes

Summary

12m 49s