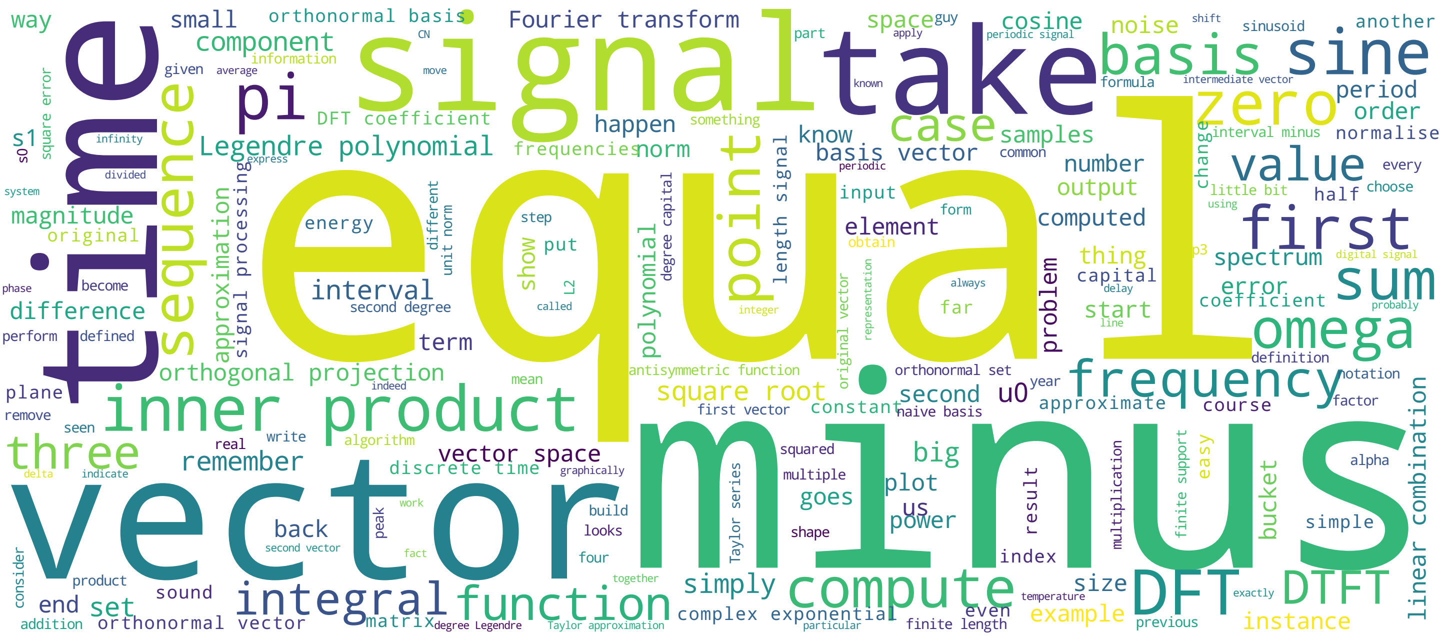
Example: polynomial approximation

- vector space $P_N[-1,1] \subset L_2[-1,1]$
- $\mathbf{p} = a_0 + a_1 t + \ldots + a_{N-1} t^{N-1}$
- a self-evident, naive basis: $\underbrace{\mathbf{s}^{(k)} = t^k,}$ $\quad k = 0, 1, \ldots, N-1$
- naive basis is not orthonormal



Search MOOC

Video

EPFL

Example: polynomial approximation

goal: approximate $\mathbf{x} = \sin t \in L_2[-1,1]$ over $P_3[-1,1]$

- build orthonormal basis from naive basis
- project $\mathbf{x}$ over the orthonormal basis
- compute approximation error
- compare error to Taylor approximation (well known but not optimal over the interval)

11

Okay, now let's try and make things a little bit more interesting by considering function vector spaces. In particular, we will try to approximate elements of L2 of minus 1,1, the space of square integral functions over the interval minus 1,1, with elements of a very particular subspace called PN of minus 1,1, which is a space of polynomials of degree capital N minus one over the interval. A generic element of the subspace will have the following form, just a polynomial of degree capital N minus one. You can see that like all polynomials, this is a linear combination of increasing powers of the variable t. A self-evident, naive basis if you want for this basis is the set of all powers of t up to degree capital n minus one. However, the naive basis is not orthonormal. It's clear that we can express any polynomial as a linear combination of vectors sk, but these vectors are not orthogonal nor unit norm. To make things more concrete, let's choose a candidate vector for approximation, and we will choose x equal to sine of t over the interval minus one, one. We will try to approximate it with a second degree polynomial. So with an element of the subspace p3 of minus 1,1.

Notes

Summary

0m 00s

Example: polynomial approximation

goal: approximate $\mathbf{x} = \sin t \in L_2[-1, 1]$ over $P_3[-1, 1]$

- build orthonormal basis from naive basis
- project $\mathbf{x}$ over the orthonormal basis
- compute approximation error
- compare error to Taylor approximation (well known but not optimal over the interval)

11

Clearly, second degree polynomials are square integral functions over the interval, so this is a legitimate subspace. The way we're going to perform the approximation is first, try to build an orthonormal basis for p3 of minus 1, 1. Because once we have an orthonormal basis, the best approximation will be simply the orthogonal projection of sine of t over this basis. To project x over this basis, all we will need to do is take the inner product between sine of t and the elements of the basis. We will then compute the approximation error and compare that to what is probably the most common approximation of the sine function over an interval, which is Taylor approximation. The Taylor approximation is well known and easy to compute, but we will show that it is not optimal in the mean square error sense.

Notes

Summary

1m 25s

Building an orthonormal basis

Gram-Schmidt orthonormalization procedure:

$$\{\mathbf{s}^{(k)}\} \longrightarrow \{\mathbf{u}^{(k)}\}$$

original set — orthonormal set

Algorithmic procedure: at each step k

1. $\mathbf{p}^{(k)} = \mathbf{s}^{(k)} - \sum_{n=0}^{k-1} \langle \mathbf{u}^{(n)}, \mathbf{s}^{(k)} \rangle \, \mathbf{u}^{(n)}$
2. $\mathbf{u}^{(k)} = \mathbf{p}^{(k)} / \|\mathbf{p}^{(k)}\|$

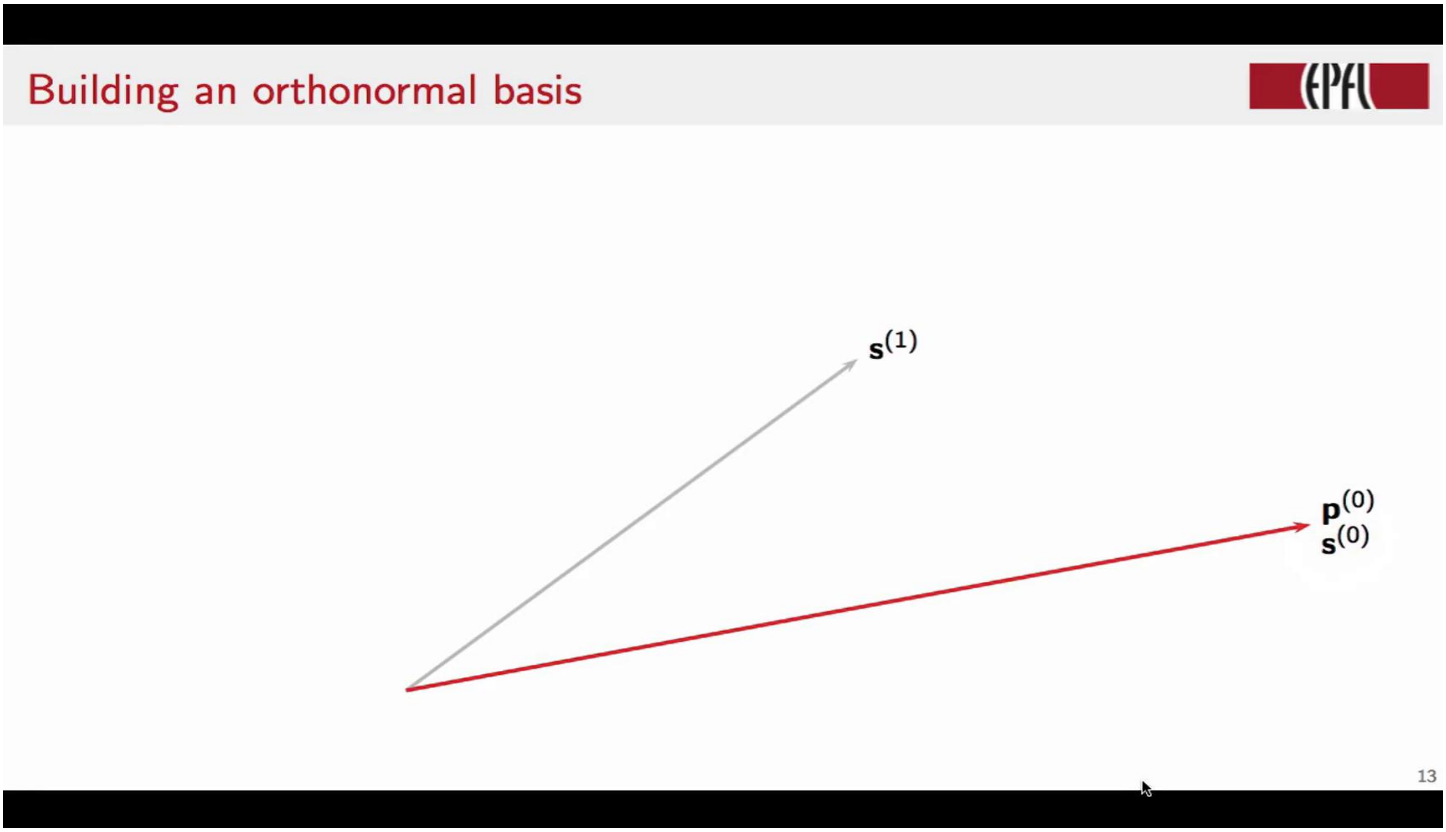
12

Okay, so the first step is to build an orthonormal basis for PN of minus one, one. We will use the Gram-Schmidt orthonormalization procedure, start with a naive basis and arrive at an orthonormal set. The Gram-Schmidt algorithm is an iterative procedure that takes one vector at the time from the original set and incrementally produces an orthonormal set. Each step has two sub-steps. The first is to take the next vector in line, say s of k, and remove from it all components that are linear, namely in the same direction as all the other vectors that you have produced so far for the orthonormal set. Then once you have whittled away all the components that are dependent on the previous vectors, you normalise the vector and obtain a new member of the orthonormal set.

Notes

Summary

2m 18s

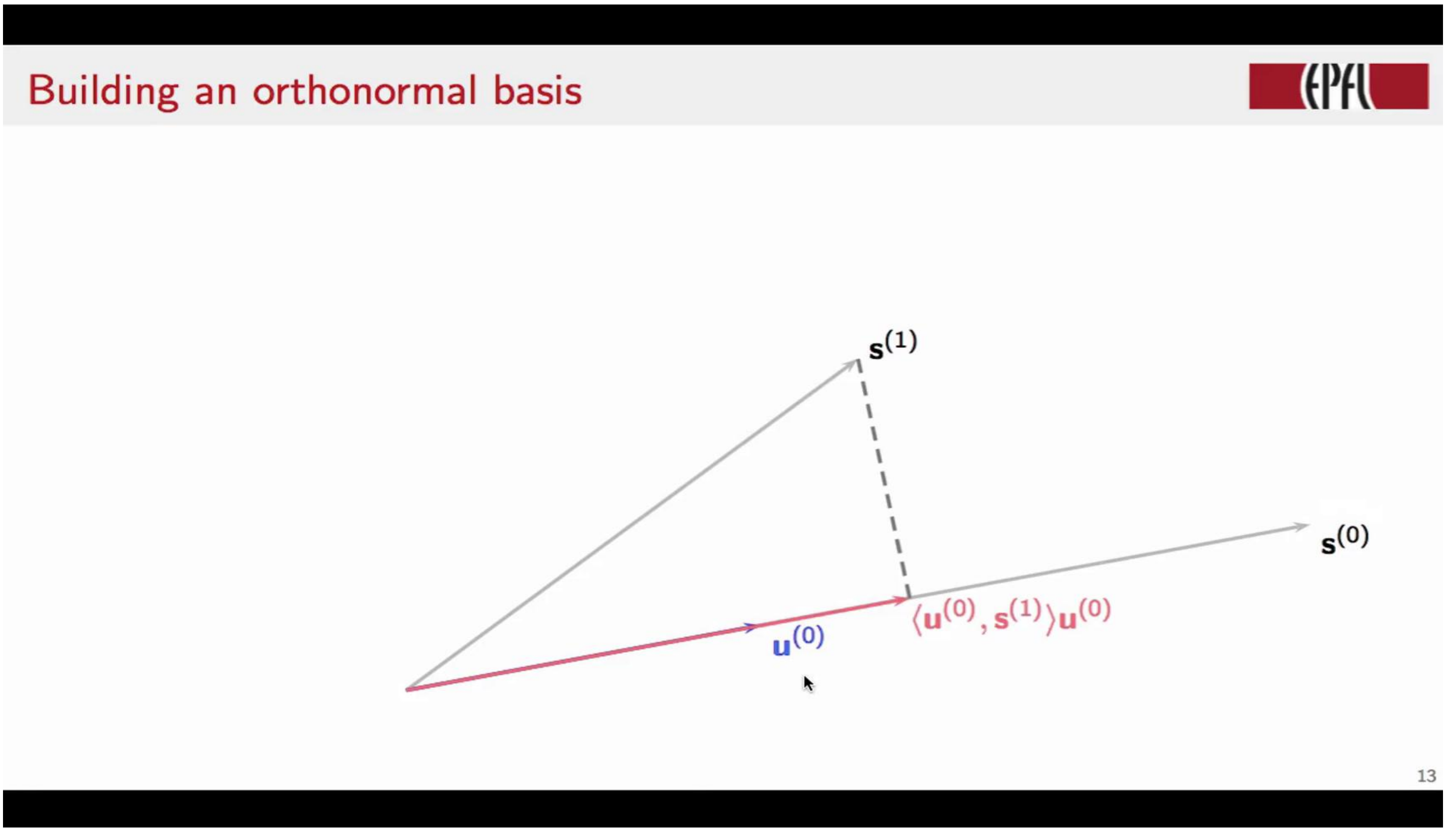


Let's look at it graphically for a very simple example. Suppose you have two vectors on the plane which do form a basis because they're pointing in different directions, but they do not form an orthonormal basis. The first step takes the first vector. Since it's the first vector, you don't have any other vector that you have computed so far. All you need to perform is the second sub-step of the procedure, just normalise the length to one. So the intermediate vector is identical to the original vector and the final vector you put in the bucket of orthonormal vector will be the rescaled version of s0 normalised to unit norm.

Notes

Summary

3m 13s



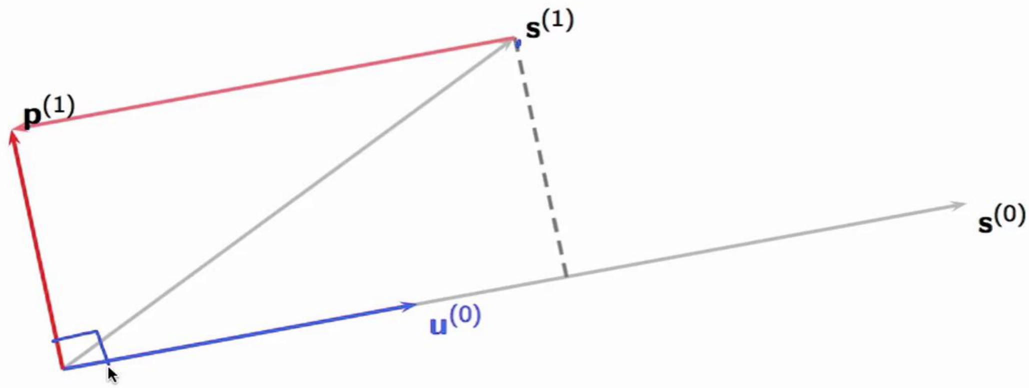
Then you take the second vector. You can see that s1 and u0 are not orthogonal, so they have some directionality in common. The idea is to remove from s1, all components that could be captured by u0. In order to do so, you project s1 over u0. This is the orthogonal projection of s1 over u0, okay, this orange vector. Then you remove it from s1. You take away the part that could have been captured by the other vector.

Notes

Summary

3m 53s

Building an orthonormal basis



13

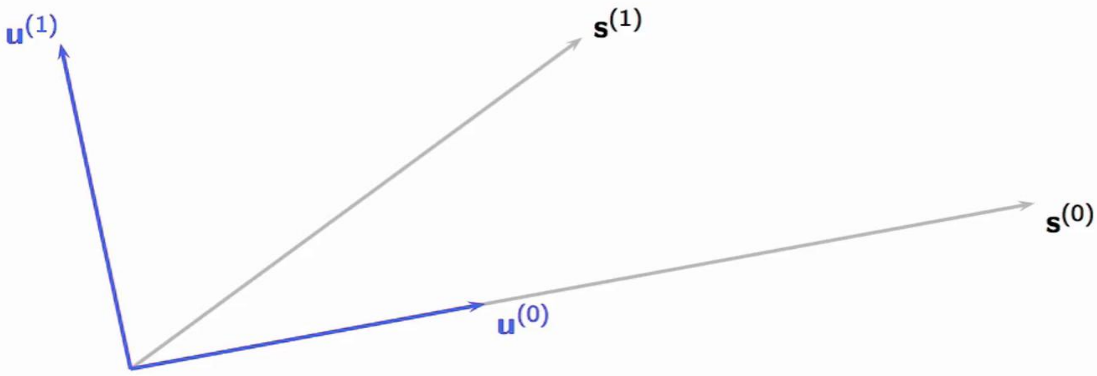
If you subtract this orthogonal projection from s1, you obtain p1, which is now, of course, orthogonal to u0. The only thing that's left to do is to normalise p1 to unit norm.

Notes

Summary

4m 22s

Building an orthonormal basis



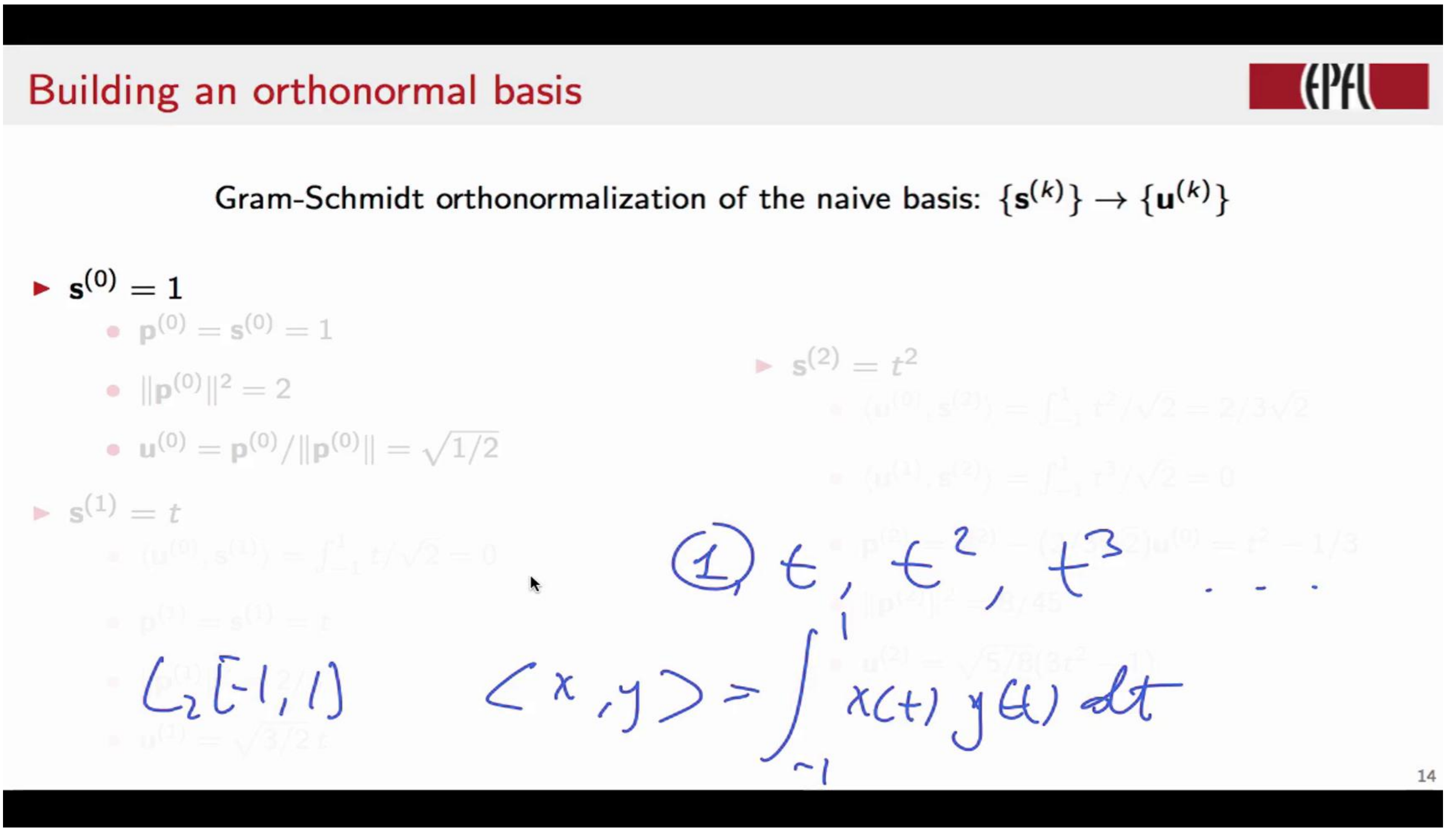
13

Here you have converted s1 and s0 to two orthonormal vectors.

Notes

Summary

4m 33s



If we now apply the orthonormalisation procedure to the naive basis for p3 of minus 1,1, the only things that we have to adapt in our computations, well, first of all, remember, the basis is 1, t, t-squared, t-cubed and so on, and the inner product between two vectors in l2 of minus 1, 1 is defined as the integral from minus 1 to 1 of x of t, y of t and dt, alright. With this in mind, let's start with the orthonormalisation procedure. The first vector in the bucket of non-orthogonal vector will be this guy, one. S0 is equal to one.

Notes

Summary

4m 38s

Building an orthonormal basis

Gram-Schmidt orthonormalization of the naive basis: $\{\mathbf{s}^{(k)}\} \rightarrow \{\mathbf{u}^{(k)}\}$

- $\mathbf{s}^{(0)} = 1$
  - $\mathbf{p}^{(0)} = \mathbf{s}^{(0)} = 1$
  - $\|\mathbf{p}^{(0)}\|^2 = 2$
  - $\mathbf{u}^{(0)} = \mathbf{p}^{(0)}/\|\mathbf{p}^{(0)}\| = \sqrt{1/2}$
- $\mathbf{s}^{(1)} = t$
  - $\langle \mathbf{u}^{(0)}, \mathbf{s}^{(1)} \rangle = \int_{-1}^{1} t/\sqrt{2} = 0$
  - $\mathbf{p}^{(1)} = \mathbf{s}^{(1)} = t$
  - $\|\mathbf{p}^{(1)}\|^2 = 2/3$
  - $\mathbf{u}^{(1)} = \sqrt{3/2}\, t$
- $\mathbf{s}^{(2)} = t^2$
  - $\langle \mathbf{u}^{(0)}, \mathbf{s}^{(2)} \rangle = \int_{-1}^{1} t^2/\sqrt{2} = 2/3\sqrt{2}$
  - $\langle \mathbf{u}^{(1)}, \mathbf{s}^{(2)} \rangle = \int_{-1}^{1} t^3/\sqrt{2} = 0$
  - $\mathbf{p}^{(2)} = \mathbf{s}^{(2)} - (2/3\sqrt{2})\mathbf{u}^{(0)} = t^2 - 1/3$
  - $\|\mathbf{p}^{(2)}\|^2 = 8/45$
  - $\mathbf{u}^{(2)} = \sqrt{5/8}(3t^2 - 1)$

14

The intermediate vector because there are no vectors computed so far in the bucket of orthonormal vectors, will be equal to s0. The norm is the integral of one between minus one and one equal to two. Therefore, the first orthonormal vector in our basis will be the original vector, normalised by the norm, which is square root of two. U0 will be square root of 1 over 2. Second vector in the bucket of non-normalised vector will be t, so the polynomial of degree one. We take the inner product between this guy and the vector that we have computed before. Remember, that is square root of one-half. The inner product between these two vectors is simply the integral between minus one and one of t divided by square root of two. Well, t is an antisymmetric function over the interval, so the integral between minus one and one will be equal to zero. What that means is that u0 and s1 are already orthogonal. P1, therefore, will be equal to s1. We now compute the norm. It's easy to show that it's equal to two-three, and therefore the second orthonormal vector in our bucket will be equal to square root of three over two times t. Third vector that we want to normalise is t-square.

Notes

Summary

5m 26s

Building an orthonormal basis

Gram-Schmidt orthonormalization of the naive basis: $\{\mathbf{s}^{(k)}\} \to \{\mathbf{u}^{(k)}\}$

- $\mathbf{s}^{(0)} = 1$
  - $\mathbf{p}^{(0)} = \mathbf{s}^{(0)} = 1$
  - $\|\mathbf{p}^{(0)}\|^2 = 2$
  - $\mathbf{u}^{(0)} = \mathbf{p}^{(0)}/\|\mathbf{p}^{(0)}\| = \sqrt{1/2}$
- $\mathbf{s}^{(1)} = t$
  - $\langle \mathbf{u}^{(0)}, \mathbf{s}^{(1)} \rangle = \int_{-1}^{1} t/\sqrt{2} = 0$
  - $\mathbf{p}^{(1)} = \mathbf{s}^{(1)} = t$
  - $\|\mathbf{p}^{(1)}\|^2 = 2/3$
  - $\mathbf{u}^{(1)} = \sqrt{3/2}\, t$
- $\mathbf{s}^{(2)} = t^2$
  - $\langle \mathbf{u}^{(0)}, \mathbf{s}^{(2)} \rangle = \int_{-1}^{1} t^2/\sqrt{2} = 2/3\sqrt{2}$
  - $\langle \mathbf{u}^{(1)}, \mathbf{s}^{(2)} \rangle = \int_{-1}^{1} t^3/\sqrt{2} = 0$
  - $\mathbf{p}^{(2)} = \mathbf{s}^{(2)} - (2/3\sqrt{2})\mathbf{u}^{(0)} = t^2 - 1/3$
  - $\|\mathbf{p}^{(2)}\|^2 = 8/45$
  - $\mathbf{u}^{(2)} = \sqrt{5/8}(3t^2 - 1)$

14

We take now the inner product between u0 and t-square. This time the inner product is not zero, it's actually two over three square root of two. So there is a directionality in common. We take the inner product between t-squared and u1 that we just computed the previous step. Again, an antisymmetric function, so the integral over symmetric interval will be zero. Finally p2, this intermediate vector will be equal to s2 minus the component that has in common with u0. So in this case, minus one-three. We normalise this by computing the norm, and we find that u2 is equal to square root of five over eight times three t-squared minus one. As you can see, the procedure is a little bit laborious and not particularly interesting.

Notes

Summary

6m 45s

Legendre polynomials

The Gram-Schmidt algorithm leads to an orthonormal basis for $P_N([-1,1])$

$$\mathbf{u}^{(0)} = \sqrt{1/2}$$
$$\mathbf{u}^{(1)} = \sqrt{3/2}\,t$$
$$\mathbf{u}^{(2)} = \sqrt{5/8}(3t^2 - 1)$$
$$\mathbf{u}^{(3)} = \ldots$$

15

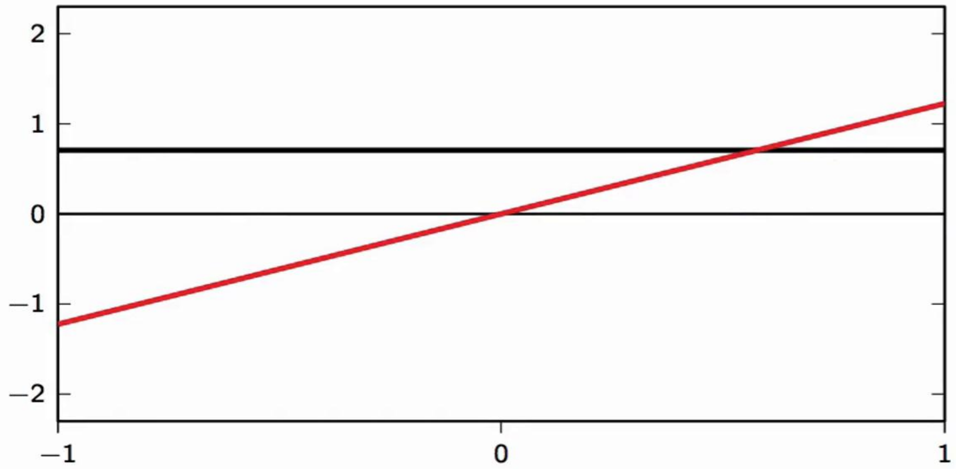
But what we do get in the end is a set of orthogonal polynomials for the interval minus one, one. These polynomials are known as the Legendre polynomials, and of course you can compute the Legendre polynomials of arbitrary degree.

Notes

Summary

7m 36s

Legendre Polynomials



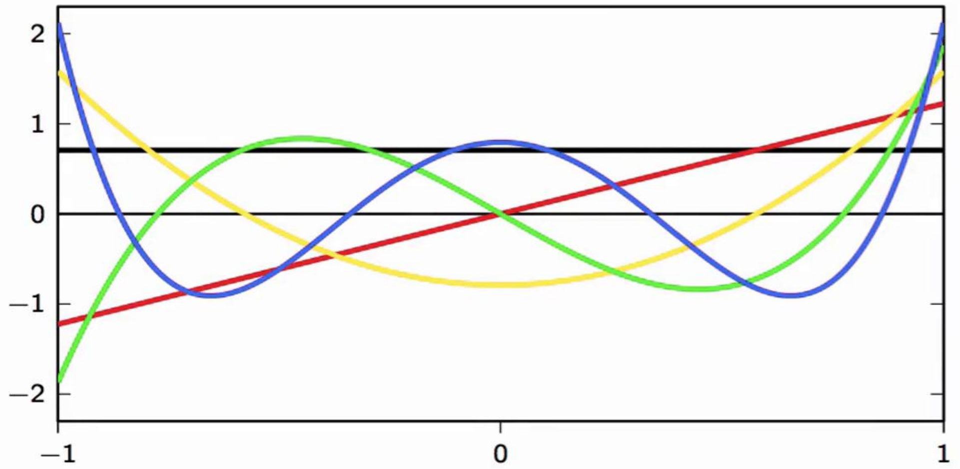
16

Here we're going to show the first five. So degree zero is the first basis vector, degree one and degree two.

Notes

Summary

7m 51s



16

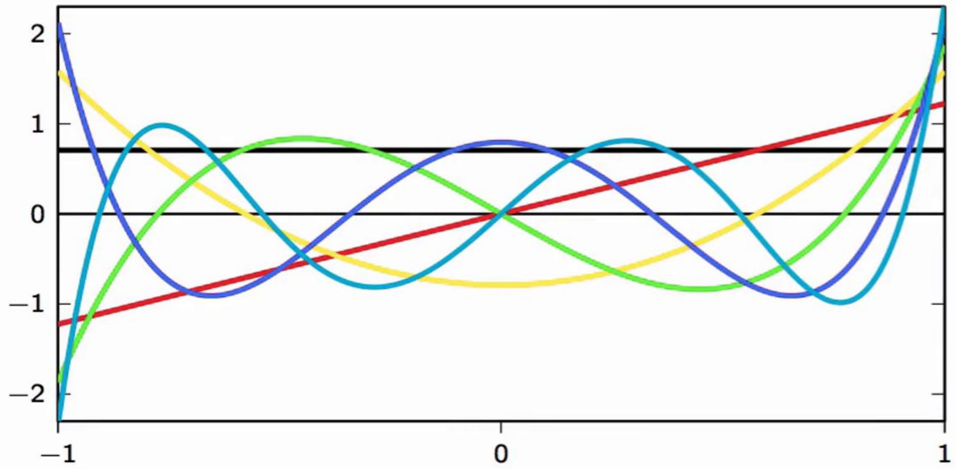
This is what we're going to use in the approximation example that follows. But if you're interested, this is the third degree, Legendre polynomial, the fourth degree, and the fifth degree, and so on and so forth.

Notes

Summary

7m 59s

Legendre Polynomials



16

Notes

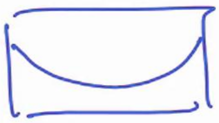
Summary

8m 08s

Orthogonal projection over $P_3[-1,1]$

$$\alpha_k = \langle \mathbf{u}^{(k)}, \mathbf{x} \rangle = \int_{-1}^{1} u_k(t) \sin t \, dt$$

- $\alpha_0 = \langle \sqrt{1/2}, \sin t \rangle = 0$
- $\alpha_1 = \langle \sqrt{3/2}\, t, \sin t \rangle \approx 0.7377$
- $\alpha_2 = \langle \sqrt{5/8}(3t^2 - 1), \sin t \rangle = 0$



17

Back to our approximation problem. We're trying to approximate sine of t over the subspace p3 of minus 1,1. We know that the best approximation is going to be given by the orthogonal projection of sine of t over the set of Legendre polynomials up to degree two. We compute these expansion coefficients as the inner product between the Legendre polynomials and the function that we want to approximate. Let's begin. Alpha zero is the inner product within the first Legendre polynomial sine of t. The first Legendre polynomial is just a constant, so we're taking the integral between minus one, one of sine of t, times a constant which is equal to zero because sine of t is antisymmetric. The first coefficient, alpha one, is now the integral of square root of three over two times t times sine of t between minus 1 and 1, and it's equal to approximately 0.7377. The last coefficient is the inner product between the second degree Legendre polynomial and sine of t. Now, we could compute this, but it's enough to see that we're computing the integral of the product between an antisymmetric function and a symmetric function. Remember, the second degree Legendre polynomial is a parable like this. The product will be an antisymmetric function, and the integral over a symmetric interval will be equal to zero.

Notes

Summary

8m 12s

Approximation

Using the orthogonal projection over $P_3[-1,1]$:

$$\sin t \to \alpha_1 \mathbf{u}^{(1)} \approx 0.9035\, t$$

Using Taylor's series:

$$\sin t \approx t$$

18

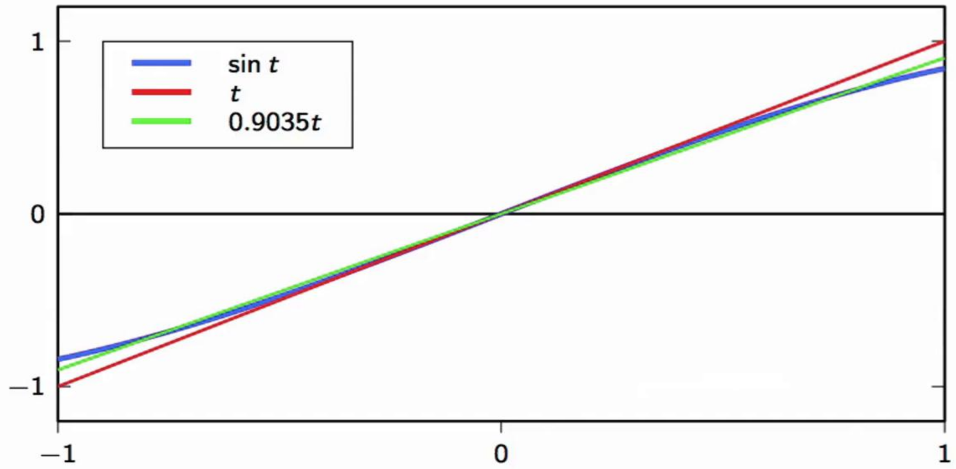
Finally, the orthogonal projection of sine of t over p3 of minus 1,1 consists of only 1 term, alpha 1 times u of 1. If you go back to the definition of why 1, you can see this is equal to 0.9035t. This is similar but not quite identical to the standard approximation for the sine around zero given by Taylor's series, which is sine of t approximately equal to t. So in both cases, we're approximating the sine with a straight line, but the slope is a bit different. Let's see how this affects the approximation result.

Notes

Summary

9m 37s

Sine approximation



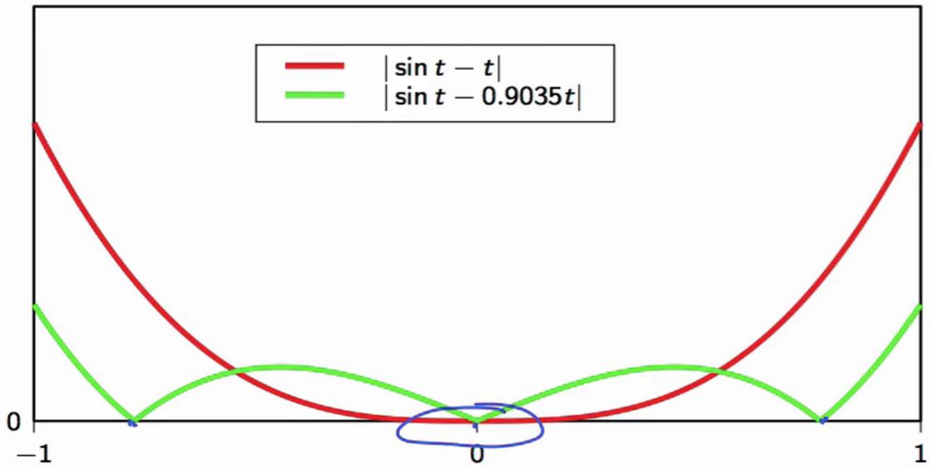
If we plot the sine, the Taylor approximation and the orthogonal projection approximation, we don't really see a lot of difference, but we can concentrate on the error between the original function and the approximates.

Notes

Summary

10m 16s

Approximation error



20

So here you can see what happens. In red, we have the error of the Taylor series approximation, and we can see that the error is very small in an interval around zero, but then tends to grow large as we move away from zero. This is because the Tailor approximation is a local approximation. The orthogonal projection on the other hand, minimises the global distance, the mean square error between the approximation and the original vector. Although the error is not as small around the origin, the energy of the error is lower considered over the entire interval. It's a global optimisation of the approximation.

Notes

Summary

10m 33s

Error norm

Orthogonal projection over $P_3[-1, 1]$:

$$\| \sin t - \alpha_1 \mathbf{u}^{(1)} \| \approx 0.0337$$

Taylor series:

$$\| \sin t - t \| \approx 0.0857$$

21

We can compute the distance between the sine function and the approximate. In the case of the Legendre approximation for the sine, we have the dominant square error of the approximation, norm of the error is 0.0337, whereas in the case of Taylor series, we have that it's 0.0857. So we have a factor of approximately two and a half between the two means square errors.

Notes

Summary

11m 10s