

MathOverflow is a question and answer site for professional mathematicians. Join them; it only takes a minute:

[Sign up](#)

Here's how it works:

Anybody can ask a question

Anybody can answer

The best answers are voted up and rise to the top

## Visualizing Orthogonal Polynomials

Recently I was introduced to the concept of [Orthogonal Polynomials](#) through the `poly()` function in the R programming language. These were introduced to me in the concept of polynomial transformations in order to do a linear regression. Bear in mind that I'm an economist and, as should be obvious, am not all that smart (choice of profession has an odd signaling characteristic). I'm really trying to wrap my head around what Orthogonal Polynomials are and how, if possible, to visualize them. Is there any way to visualize orthogonal polynomials vs. simple polynomials?

[orthogonal-polynomials](#)

asked Sep 15 '10 at 20:11



[jd long](#)  
138 1 5

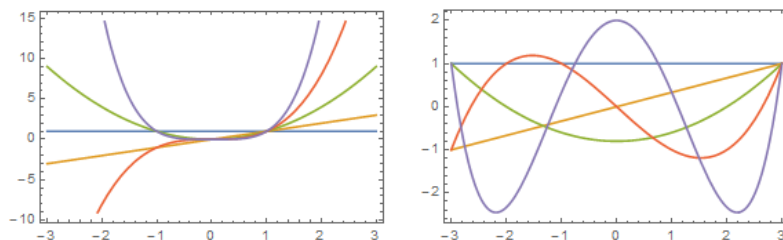
### 4 Answers

Helge presented the *continuous* case in his answer; for the purposes of data fitting in statistics, one usually deals with *discrete* orthogonal polynomials. Associated with a set of abscissas  $x_i$ ,  $i = 1 \dots n$  is the discrete inner product

$$\langle f, g \rangle = \sum_{i=1}^n w(x_i) f(x_i) g(x_i)$$

where  $w(x)$  is a weight function, a function that associates a "weight" or "importance" to each abscissa. A frequently occurring case is one where the  $x_i$  are equispaced,  $x_{i+1} - x_i = h$  where  $h$  is a constant, and the weight function is  $w(x) = 1$ ; for this special case, special polynomials called Gram polynomials are used as the basis set for polynomial fitting. (I won't be dealing with the nonequispaced case in the rest of this answer, but I'll add a few words on it if asked).

Let's compare a plot of the regular monomials  $x^k$  to a plot of the Gram polynomials:



On the left, you have the regular monomials. The "bad" thing about using them in data fitting is that for  $k$  high enough,  $x^k$  and  $x^{k+1}$  are high-indistinguishable, and this spells trouble for data-fitting methods since the matrix associated with the linear system describing the fit is dangerously close to becoming singular.

On the right, you have the Gram polynomials. Each member of the family does not resemble its predecessor or successor, and thus the underlying matrix used for fitting is a lot less likely to be close to singularity.

This is the reason why discrete orthogonal polynomials are of interest in data fitting.

edited May 25 at 17:04



[J. M.](#)  
128 7

answered Sep 16 '10 at 0:35



[J. M.](#)  
2,180 14 27

- 1 A few small notes: 1. in the limit of infinitely many abscissas (the *discrete* becomes *continuous*), the Gram polynomial becomes the Legendre polynomial. 2. The orthogonal polynomials associated with nonequispaced abscissas are easily constructed through the so-called Stieltjes procedure. Different abscissas correspond to different basis sets. The reason the equispaced case is much nicer is that one can factor out the spacing  $h$  from the appropriate equations. 3. Any good old book on the difference calculus should have some mention on the Gram polynomials. – J. M. Sep 16 '10 at 5:40
- 2 Thanks for taking the time to explain this. The issue of multicollinearity (singular matrix) is exactly the context where I was using the `poly()` function which creates orthogonal polynomials. Thank you for grabbing hold of my context and sharing an answer that really gets at my specific domain of interest. – jd long Sep 16 '10 at 18:13
- jd: You're very much welcome. I have to say I wish someone had explained it like that to me when I was first learning about this stuff. – J. M. Sep 17 '10 at 1:03

Start [here](#)

answered Sep 15 '10 at 20:26



[Piero D'Ancona](#)  
4,731 1 14 30

That was actually a really helpful resource. Thank you. – jd long Sep 15 '10 at 20:34

I am not sure of your math-background, so I am trying to keep it simple, without oversimplifying some ideas. First off polynomials are nice for various reasons, e.g. a polynomial of degree  $n$  has at most  $n$  zeros. However, there are still many polynomials and it makes sense to choose VERY nice ones: orthogonal polynomials.

To choose orthogonal polynomials, one has a problem at hand, which comes with a way to measure functions  $f : \mathbb{R} \rightarrow \mathbb{R}$  by an expression of the form

$$\mathcal{E}(f) = \int_{-\infty}^{\infty} f(x)^2 w(x) dx,$$

where  $w(x) > 0$  is a weight that satisfy  $\int w(x) dx = 1$ . One should think of  $\mathcal{E}$  as an energy.

Now the orthogonal polynomial of degree  $n$  can be defined as the polynomial

$P_n(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ , where  $a_{n-1}, \dots, a_0$  are real numbers, that minimizes  $\mathcal{E}(P_n)$ . It is this minimization property that is responsible for some of the power of orthogonal polynomials.

At this point let me also say that it is through the weight  $w$  that your problem enters the definition of orthogonal polynomials. And that one also has orthonormal polynomials, which satisfy  $p_n = 1$ .

These are given by  $p_n = \frac{1}{\sqrt{\mathcal{E}(P_n)}} P_n$ .

answered Sep 15 '10 at 21:38



[Helge](#)  
2,598 14 20

I think of the space of polynomials on  $\mathbb{R}$  as a set of graphs arranged round the real line like the pages of a book round the axis. Polynomials which have almost the same graph are close to each other; then orthogonal polynomials are those which fall at right angles in the picture, and the linear combinations generate the space just as the basis vectors of  $\mathbb{R}^n$  generate  $\mathbb{R}^n$ .

answered Sep 15 '10 at 21:22



[Tom Smith](#)  
792 1 8 17

PS If you have had a decent linear algebra course you will be familiar with the concept of a vector space as a collection of abstract vectors satisfying certain axioms. You can see that the set of polynomials on  $\mathbb{R}$  satisfies these axioms and so is a vector space. An orthogonal set of polynomials then generates the whole space in roughly the same way that an orthogonal basis for an ordinary vectors space does. – Tom Smith Sep 15 '10 at 21:26