Polynomial regression using trapezoidal rule for computing Legendre coefficients

Demetris T. Christopoulos\textsuperscript{1,2}

\textsuperscript{1}National and Kapodistrian University of Athens, Department of Economics
\textsuperscript{2}dchristop@econ.uoa.gr, dem.christop@gmail.com

December 2, 2013

Abstract

We are presenting a method for computing the Fourier coefficients of a given polynomial regression by using the trapezoidal rule for numerical integration. As function basis we use the orthogonal Legendre polynomials. The results are accurate and stable compared to Forsythe’s method.

MSC2000. Primary 62J05, Secondary 65D99

Keywords. Basis function, regression, orthogonal polynomials, trapezoidal rule, numerical integration

1 Polynomial regression

The polynomial regression technique is based on the OLS computation of the coefficients in the formal truncated series expansion of degree \( m \)

\[
y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \ldots + \beta_m x_i^m + \epsilon_i \sim \text{iid}(0, \sigma^2), \quad i = 1, 2, \ldots, n \tag{1}
\]

or in matrix form

\[
y = X \cdot \beta + \epsilon, \quad \epsilon | X \sim \text{iid} \left(0, \sigma^2 I_{m+1}\right) \tag{2}
\]

which has the well known OLS solution which is given by the pseudo-inverse Moore-Penrose matrix after \cite{4} and \cite{5}

\[
\hat{\beta} = \left(X' X\right)^{-1} X' y = X^+ y \tag{3}
\]
The problem is that matrix $X'X$ can be proved, that for equidistant $x_i$, see [2], [6] and for increasing $n$ is approximately the Hilbert matrix, one of the most famous ill-conditioned matrices. Early computations, see [3] has sown that only for a degree up to 10 we could have satisfactory outputs. Although the situation has been better now due to arbitrary precision arithmetic computations, the computational effort is still big enough. If we manage to diagonalize $X'X$ then our task is much more easy computationally. This process has been done by Forsythe, see [2] & [3], where a recursive method for defining orthogonal polynomials was introduced. The concept of polynomial orthogonality there had the sense of discrete orthogonality, i.e. two polynomial $\phi_k(x), \phi_l(x)$ are said to be orthogonal over a set of abscissae $\{x_1, x_2, \ldots, x_n\}$ if the next vanishing equation holds

$$\sum_{i=1}^{n} \phi_k(x_i)\phi_l(x_i) = 0$$

This is nothing else than the zero common Euclidean real inner product

$$\langle \phi_k, \phi_l \rangle = \phi_k' \phi_l = 0$$

By using this procedure instead of directly computing $\beta$ coefficients of we compute the coefficients of next truncated series

$$y_i = \beta_0\phi_0(x_i) + \beta_1\phi_1(x_i) + \beta_2\phi_2(x_i) + \ldots + \beta_m\phi_m(x_i) + \epsilon_i, \epsilon_i \sim iid(0, \sigma^2)$$

or in matrix form

$$y = \Phi \cdot \beta + \epsilon, \epsilon | \Phi \sim iid(0, \sigma^2 I_{m+1})$$

where the design matrix is

$$\Phi = \begin{bmatrix} \phi_0 & \phi_1 & \ldots & \phi_m \end{bmatrix} = \begin{bmatrix} 1 & \phi_1(x_1) & \ldots & \phi_m(x_1) \\ 1 & \phi_1(x_2) & \ldots & \phi_m(x_2) \\ \vdots & \vdots & \ldots & \vdots \\ 1 & \phi_1(x_n) & \ldots & \phi_m(x_n) \end{bmatrix}$$

Our coefficients are now simply the well known Fourier coefficients

$$\beta_j = \frac{\langle y, \phi_j \rangle}{\langle \phi_j, \phi_j \rangle} = \frac{\sum_{i=1}^{n} y_i \phi_j(x_i)}{\sum_{i=1}^{n} \phi_j(x_i)^2}$$

If we take orthonormal polynomials, see [7], i.e. if it holds that

$$\langle \phi_k, \phi_l \rangle = \phi_k' \phi_l = \delta_{kl} = \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases}$$
then $\Phi' \Phi = I_{m+1}$ and the coefficients are simply

$$\beta_j = \Phi' y = \langle y, \phi_j \rangle = \sum_{i=1}^{n} y_i \phi_j(x_i), \ j = 0, \ldots, m$$

(11)

The concept of orthogonality or orthonormality is a linear algebraic term and is independent of the chosen function basis representation. It is an elementary exercise, see [1] page 61, that starting from the linearly independent set of monomials $\{x^i, i = 0, 1, 2, \ldots, m, \forall m \in \mathbb{N}\}$ and by using Gram-Schmidt orthonormalisation process we can end to the normalised Legendre polynomials

$$P_k(x) = \sqrt{\frac{m+1}{2}} P_k(x) = \sqrt{\frac{m+1}{2}} \frac{1}{2^k k!} \frac{d^k}{dx^k} (x^2 - 1)^k$$

(12)

The above polynomials are orthonormal in the interval $[-1, 1]$

$$\int_{-1}^{1} P_k(x) P_l(x) = \delta_{kl}$$

(13)

Now we can expand every function in a Legendre series expansion

$$f(x) = \sum_{i=1}^{m} c_k P_k(x) dx$$

(14)

with the Fourier coefficients given by

$$\beta_k = \langle f(x), P_k(x) \rangle = \int_{-1}^{1} f(x) P_k(x) dx$$

(15)

If we follow the guides of [6] and make the linear transformation

$$T_2 : [a, b] \to [-1, 1], \quad T_2(x) = \frac{2x - a - b}{b - a}$$

(16)

in order to convert our initial range $[a, b]$ to the $[-1, 1]$, where many orthogonal polynomials are defined, then the Forsythe polynomials are just a scale version of Legendre polynomials. So, the norm we have chosen does not play any other role except for the simplicity of computations. If we choose the $l_2$ norm, then we can proceed like Forsythe and construct a set of orthogonal polynomials for solving our polynomial regression problem.

The discrete case is

$$y_i = f(x_i) = \beta_0 P_0(x_i) + \beta_1 P_1(x_i) + \beta_2 P_2(x_i) + \ldots + \beta_m P_m(x_i)$$

(17)
It is obvious to think about computing the Fourier coefficients \([15]\) by a numerical approximation of the relevant integral. For the equidistant case with \(x_{j+1} - x_j = 1\) we have that

\[
\beta_k = \langle f(x), \mathcal{P}_k(x) \rangle \approx \sum_{i=1}^{n} y_i \mathcal{P}_k(x_i)
\]  

(18)

Now we have approximated the integral via the orthogonal rule. We can also use trapezoidal method in order to increase the accuracy. By comparing \([18]\) and \([11]\) we see that the latter is just the left orthogonal Riemannian approximation for the continuous case \([15]\). If we had use the simple Legendre polynomials \(P_k(x)\) then our coefficients could be

\[
\beta_k = \frac{\langle f(x), P_k(x) \rangle}{\langle P_k(x), P_k(x) \rangle}
\]  

(19)

For the equidistant case \(x_{j+1} - x_j = h\) we have that

\[
\beta_k \approx \frac{\sum_{i=1}^{n} y_i P_k(x_i) h}{\sum_{i=1}^{n} P_k^2(x_i) h} = \frac{\sum_{i=1}^{n} y_i P_k(x_i)}{\sum_{i=1}^{n} P_k^2(x_i)}
\]  

(20)

which is just \([9]\) for Fourier coefficients.

## 2 Trapezoidal estimation of Fourier coefficients

Our task is to compute the integrals of our Fourier coefficients, \([18]\) for normalised or \([19]\) for simple orthogonal polynomials by using the trapezoidal rule of numerical integration. We shall constraint in the equidistant case, since we have closed formulas using less arithmetic operations.

For the case of simple orthogonal polynomials we have that

\[
\beta_k \approx \frac{y_1 P_k(x_1) + 2 \sum_{i=2}^{n-1} y_i P_k(x_i) + y_n P_k(x_n)}{P_k^2(x_1) + 2 \sum_{i=1}^{n} P_k^2(x_i) h + P_k^2(x_n)}
\]  

(21)

If we use orthonormal polynomials we have the estimation

\[
\beta_k \approx \frac{h}{2} \left( y_1 \mathcal{P}_k(x_1) + 2 \sum_{i=2}^{n-1} y_i \mathcal{P}_k(x_i) + y_n \mathcal{P}_k(x_n) \right)
\]  

(22)

The total sum of squares is almost identical for the two cases and for the simple OLS regression by mean of \([3]\) with \(X = \Phi\).
3 A numerical example

Let us consider the known function:

\[ f : [-\pi, \pi] \to \mathbb{R}, \quad f(x) = \sin(3x) \cos(5x)e^{-x} + 3 \sin(\pi x)e^{x/2} \]  \hspace{1cm} (23)

at an equal spaced grid \( x_i, i = 0, \ldots, 628 \). The graph of the function is presented in Figure 1. This function is a smooth function, \( f \in C^\infty \), it has 6 local maxima and 5 local minima inside the interval \([-\pi, \pi]\). It is a rather complicated function, for example it has one local minimum and maximum in the small interval \([-1, -0.5]\), so the task of recovering this shape is difficult.

By using floating point arithmetic with 32 digits of accuracy we can obtain the next Taylor polynomial of 30th degree

\[ T_{30}(x) = +0.0000016359x^{30} - 0.0000097947x^{29} - 0.00013156x^{28} + 0.0001336926x^{27} + 0.0000097947x^{26} - 0.001477436x^{25} + 0.00710316x^{24} + 0.0131083743x^{23} - 0.015365948x^{22} - 0.091642089x^{21} + 0.168431320x^{20} + 0.488498913x^{19} - 1.270259183x^{18} - 1.866729762x^{17} + 6.956422865x^{16} + 4.386886795x^{15} + 27.7967780x^{14} - 2.36993754x^{13} + 78.47606337x^{12} - 23.60696807x^{11} - 151.3009259x^{10} + 86.58066121x^9 + 184.9106325x^8 - 142.4655789x^7 - 125.9202816x^6 + 121.2621435x^5 + 33.94478037x^4 - 54.82504110x^3 + 1.71238898x^2 + 12.4247796x \]

Although the above polynomial is not identical to the initial function outside approximately the interval \([-1.5, 1.5]\) it is a representation that carries a lot of information about the function since it can give the derivatives until the 30th order. So, our task is (i) to recover as many as possible coefficients of the above series expansion and (ii) to approximate the functional data with the smallest possible error.

In order to avoid multicollinearity problems due to lower accuracy we are using 32 digits in our arithmetic operations and we are transforming to the interval \([-1, 1]\) both \( x_i \) & \( y_i \) data. After finishing our coefficient computations we are performing the inverse \( T_2 \)-transform and return to our initial data scale.

Results are presented at Table 1 while the sum of squares for both cases, the transformed to \([-1, 1]\) and the initial, are given at Table 2. The relevant plots of all Legendre series are indistinguishable from the original data, see Figure 2.
As a benchmark to our effort we shall compare our results with those obtained by using [2] method as has been implemented in FORTRAN 90 by [8]. We find that under double precision arithmetic, i.e. with 16 digits accuracy, the solution divergences very fast from the true series expansion. The [2] polynomial coefficients after inverse transforming to the initial domain are presented at Table 3.

4 Discussion

The times for computing the coefficients were \((P_k(x), \mathcal{P}_k(x), P_k^{OLS}(x)) = (16.895, 11.013, 24.820)\) CPU seconds in a typical Intel Core i5 CPU with 4 GB RAM memory and by using Maple program. We observe that the use of orthonormal polynomials is reducing the computational time. If we decrease our accuracy to 16 digits in order to be compatible with FORTRAN we obtain similar results, see Table 4. Thus our methods still found converged and suitable outputs compared to the [2] orthogonal polynomial method.

References

[1] M.R. Spiegel, Theory and Problems of Fourier Analysis with Applications to Boundary Value Problems, McGraw-Hill New York, 1974

[2] G. E. Forsythe, Generation and use of orthogonal polynomials for data-fitting with a digital computer, J. Soc. Indust. Appl. Math. 5, 7488, 1957

[3] M. Asche and G.E. Forsythe, SWAC Experiments on the Use of Orthogonal Polynomials for Data Fitting, J. ACM, ACM, 5, 9-21, 1958

[4] E. H. Moore, On the reciprocal of the general algebraic matrix. Bulletin of the American Mathematical Society 26 (9): 394395, 1920

[5] R. Penrose, A generalized inverse for matrices. Proceedings of the Cambridge Philosophical Society 51: 406413, 1955

[6] M. Shacham and N. Brauner, Minimizing the Effects of Collinearity in Polynomial Regression, Industrial & Engineering Chemistry Research, 36, 4405-4412, 1997
[7] T.R. Ten Have, C.J. Kowalski, E. D. Schneiderman, PC program for obtaining orthogonal polynomial regression coefficients for use in longitudinal data analysis, *American Journal of Human Biology*, 4(3): 403-416, 1992

[8] J.P. Moreau Programs In FORTRAN language, [http://jean-pierre.moreau.pagesperso-orange.fr/Fortran/approx1_f90.txt](http://jean-pierre.moreau.pagesperso-orange.fr/Fortran/approx1_f90.txt), last visited 11/29/2013
Figure 1: The plot of the function used at numerical example
Figure 2: Legendre series approximation
| $k$ | $P_k(x)$ | $\mathcal{P}_k(x)$ | $P_k^{OLS}(x)$ | $T_k(x)$ |
|-----|---------|----------------|----------------|---------|
| 0   | -0.0001255282 | 0.0018148468 | -0.002035062 | 0.0     |
| 1   | 12.33649  | 12.44636 | 12.50049 | 12.42478 |
| 2   | 1.852769  | 1.629086 | 1.821539 | 1.712389 |
| 3   | -54.36178 | -55.30185 | -56.14278 | -54.82504 |
| 4   | 32.78326  | 34.56230 | 32.97494 | 33.94478 |
| 5   | 121.7605  | 124.2576 | 128.1177 | 121.2621 |
| 6   | -122.8261 | -127.6625 | -122.5218 | -125.9203 |
| 7   | -148.0910 | -151.0391 | -159.2726 | -142.4656 |
| 8   | 180.7955  | 187.3529 | 178.6751 | 184.9106 |
| 9   | 98.54291  | 100.3063 | 110.2045 | 86.58066 |
| 10  | -147.9384 | -153.1944 | -144.4110 | -151.3009 |
| 11  | -36.87822 | -37.37939 | -44.82297 | -23.60697 |
| 12  | 76.54964  | 79.28914 | 73.50992 | 78.47606 |
| 13  | 7.061322  | 7.061628 | 10.78721 | -2.236994 |
| 14  | -26.87537 | -27.85941 | -25.26188 | -27.70968 |
| 15  | -0.09374941 | -0.04223660 | -1.330764 | 4.386887 |
| 16  | 6.680829  | 6.933201 | 6.111261 | 6.956423 |
| 17  | -0.3086695 | -0.3285492 | -0.01411589 | -1.866730 |
| 18  | -1.204759 | -1.251863 | -1.065779 | -1.270259 |
| 19  | 0.08555660 | 0.08970893 | 0.03517729 | 0.4884989 |
| 20  | 0.1590441 | 0.1654770 | 0.1352121 | 0.1684313 |
| 21  | -0.01253891 | -0.01309178 | -0.006412513 | -0.09164209 |
| 22  | -0.01527912 | -0.01591658 | -0.01241044 | -0.0153695 |
| 23  | 0.001141524 | 0.001189852 | 0.0006253121 | 0.01310837 |
| 24  | 0.001042632 | 0.001087308 | 0.0008050851 | 0.0007170316 |
| 25  | -0.00006491531 | -0.00006761426 | -0.00003629592 | -0.001477436 |
| 26  | -0.00004797304 | -0.00005007409 | -0.00003507896 | 0.00004194580 |
| 27  | 0.000002124622 | 0.000002212277 | 0.000001186322 | 0.00001336926 |
| 28  | 0.000001335500 | 0.000001395001 | 0.0000009223072 | -0.00001315592 |
| 29  | -0.00000003068878 | -0.0000003195140 | -0.0000001691387 | -0.000009794687 |
| 30  | -0.00000001699460 | -0.0000001776155 | -0.0000001106804 | 0.000001635893 |
Table 2: Total Sum of Squares

| Domain                  | $P_k(x)$ | $\mathcal{P}_k(x)$ | $P_k^{OLS}(x)$ |
|------------------------|----------|--------------------|----------------|
| $[-1, 1] \times [-1, 1]$ | 218.869  | 219.183            | 219.095        |
| $[-\pi, \pi] \times [y_{min}, y_{max}]$ | 13132.6  | 13177.4            | 13169.9        |
Table 3: Forsythe orthogonal polynomial coefficients

| $k$ | $P_k^{Forsythe}(x)$ | $T_k(x)$ |
|-----|---------------------|----------|
| 0   | 0.05094099          | 0.0      |
| 1   | 0.9398571           | 12.42478 |
| 2   | 0.1309315           | 1.712389 |
| 3   | 0.1168369           | -54.82504|
| 4   | 0.1049183           | 33.94478 |
| 5   | -0.2143114          | 121.2621 |
| 6   | -0.04026457         | -125.9203|
| 7   | -0.08311646         | -142.4656|
| 8   | -0.006660081        | 184.9106 |
| 9   | 0.02541863          | 86.58066 |
| 10  | 0.001024787         | -151.3009|
| 11  | 0.002504286         | -23.60697|
| 12  | -0.003783044        | 78.47606 |
| 13  | 0.001420267         | -2.236994|
| 14  | -0.0002494057       | -27.70968|
| 15  | -0.0002088982       | 4.386887 |
| 16  | 0.0003358103        | 6.956423 |
| 17  | -0.0002421829       | -1.866730|
| 18  | 0.00008259028       | -1.270259|
| 19  | 0.00001781517       | 0.4884989|
| 20  | -0.00004319505      | 0.1684313|
| 21  | 0.00002519070       | -0.09164209|
| 22  | 0.00000009195903    | -0.01536595|
| 23  | -0.000007925483     | 0.01310837|
| 24  | 0.000002842523      | 0.0007170316|
| 25  | 0.0000009651675     | -0.001477436|
| 26  | -0.0000007763782    | 0.00004194580|
| 27  | -0.00000005683953   | 0.0001336926|
| 28  | 0.0000001149406     | -0.00001315592|
| 29  | -0.0000001691387    | -0.000009794687|
| 30  | -0.0000001106804    | 0.000001635893|
Table 4: Legendre coefficients for 16 digits accuracy

| $k$ | $P_k(x)$ | $\mathcal{P}_k(x)$ | $P_{k}^{OLS}(x)$ | $T_k(x)$ |
|-----|----------|----------------------|-------------------|---------|
| 0   | -0.0001255282 | 0.001814486 | -0.002035063 | 0.0 |
| 1   | 12.33649 | 12.44636 | 12.50049 | 12.42478 |
| 2   | 1.852769 | 1.629086 | 1.821539 | 1.712389 |
| 3   | -54.36178 | -55.30185 | -56.14278 | -54.82504 |
| 4   | 32.78326 | 34.56230 | 32.97494 | 33.94478 |
| 5   | 121.7605 | 124.2576 | 128.1177 | 121.2621 |
| 6   | -122.8261 | -127.6625 | -122.5218 | -125.9203 |
| 7   | -148.0910 | -151.0391 | -159.2726 | -142.4656 |
| 8   | 180.7955 | 187.3529 | 178.6751 | 184.9106 |
| 9   | 98.54291 | 100.3063 | 110.2045 | 86.58066 |
| 10  | -147.9384 | -153.1944 | -144.4110 | -151.3009 |
| 11  | -36.87822 | -37.37939 | -44.82297 | -23.60697 |
| 12  | 76.54964 | 79.28914 | 73.50992 | 78.47606 |
| 13  | 7.061322 | 7.061628 | 10.78721 | -2.236994 |
| 14  | -26.87537 | -27.85941 | -25.26188 | -27.70968 |
| 15  | -0.09374941 | -0.04223660 | -1.330764 | 4.386887 |
| 16  | 6.680829 | 6.933201 | 6.111261 | 6.956423 |
| 17  | -0.3086695 | -0.3285492 | -0.0141583 | -1.866730 |
| 18  | -1.204759 | -1.251863 | -1.065779 | -1.270259 |
| 19  | 0.08555660 | 0.08970893 | 0.03517728 | 0.4884989 |
| 20  | 0.1590441 | 0.1654770 | 0.1352121 | 0.1684313 |
| 21  | -0.01253891 | -0.01309178 | -0.00641251 | -0.09164209 |
| 22  | -0.01527912 | -0.01591658 | -0.01241044 | -0.01536595 |
| 23  | 0.001141524 | 0.001189852 | 0.0006253120 | 0.01310837 |
| 24  | 0.0001042632 | 0.001087308 | 0.0008050850 | 0.0007170316 |
| 25  | -0.00006491531 | -0.00006761426 | -0.00003629592 | -0.001477436 |
| 26  | -0.00004797304 | -0.00005007409 | -0.00003507895 | 0.00004194580 |
| 27  | 0.000002124622 | 0.000002212277 | 0.000001186322 | 0.00001336926 |
| 28  | 0.000001335500 | 0.000001395001 | 0.0000009223070 | -0.00001315592 |
| 29  | -0.00000003068878 | -0.00000003195140 | -0.00000001691387 | -0.000009794687 |
| 30  | -0.00000001699460 | -0.00000001776155 | -0.00000001106803 | 0.000001635893 |