INTERPOLATION AND FUNCTION APPROXIMATION

Max Gunzburger, Janet Peterson, and Sachin Shanbhag

Department of Scientific Computing, Florida State University

Spring 2013

1. Introduction

1.1. Function approximation

Function approximation is the task of constructing, for a given function, a simpler function so that the difference between the two functions is small and to then provide a quantifiable estimate for the size of the difference.

Why would one want to do this? Consider the evaluation of the integral

$$\int_0^1 e^{x^2} dx.$$

The antiderivative of e^{x^2} cannot be expressed in terms of simple functions, i.e., in terms of powers of x, trigonometric functions, exponential functions, ... Suppose we could find a function $p(x)$ that is "close" to e^{x^2} over the interval $[0,1]$ and which has an easy to define antiderivative $F(x)$. Then, we can use $p(x)$ as a surrogate for $f(x)$ and then approximate the integral of $f(x)$ by the integral of $p(x)$, i.e., we have

$$\int_0^1 e^{x^2} dx \approx \int_0^1 p(x) dx = F(1) - F(0).$$

EXAMPLE. In calculus, one learns a particular way to define a simple function $p(x)$: use the Maclaurin series, i.e., the Taylor series about the point $x = 0$, for e^{x^2} which is given by

$$e^{x^2} = 1 + x^2 + \frac{1}{2}x^4 + \frac{1}{6}x^6 + \cdots = \sum_{i=0}^{\infty} \frac{1}{i!} x^{2i}.$$

We approximate e^{x^2} by keeping only the first $n+1$ terms in the series:

$$e^{x^2} \approx p(x) = \sum_{i=0}^{n} \frac{1}{i!} x^{2i}.$$

We then replace the integrand e^{x^2} by $\sum_{i=0}^{n} x^{2i}$ to obtain the approximation

$$\int_0^1 e^{x^2} dx \approx \sum_{i=0}^{n} \int_0^1 \frac{1}{i!} x^{2i} dx = \sum_{i=0}^{n} \frac{1}{i!(2i+1)}.$$

For $n = 2$, we then have

$$\int_0^1 e^{x^2} dx \approx 1 + \frac{1}{3} + \frac{1}{10} = \frac{43}{30} = 1.43333\underline{3}\ldots.$$

Question: how big an error are we making by replacing the integral $\int_0^1 e^{x^2} dx$ by $F(1) - F(0)$, i.e., how close is $\frac{43}{30}$ to the correct answer? We do not know the exact answer but MATLAB has a very good method for approximating integrals. The MATLAB approximation is

$$\int_0^1 e^{x^2} dx \approx 1.46265.$$

Thus, the error in our approximation is about $1.46265 - 1.43333 = 0.02932$.

What about keeping fewer or more terms in the Taylor series? We have that

$$\int_0^1 e^{x^2} dx \approx \begin{cases} 1 + \frac{1}{3} = \frac{4}{3} \approx 1.3333333 & \text{for } n = 1 \\ 1 + \frac{1}{3} + \frac{1}{10} + \frac{1}{42} = \frac{51}{35} \approx 1.4571429 & \text{for } n = 3 \\ 1 + \frac{1}{3} + \frac{1}{10} + \frac{1}{42} + \frac{1}{216} = \frac{22039}{15120} \approx 1.4617726 & \text{for } n = 4 \end{cases}$$

The corresponding errors are given in the table below, from which we observe that the approximation seems to get better as we increase n.

n	error	relative error $= \frac{\text{error}}{\text{exact answer}}$
0	0.46265	32%
1	0.12932	8.8%
2	0.02932	2.0%
3	0.00551	0.38%
4	0.00088	0.06%

□

What if we do not have[1] MATLAB? Could we get some idea about the size of the error we make by approximating the integral using Taylor polynomials? We can indeed obtain an estimate for that error, specifically by estimating the remainder term in the Taylor formula

$$e^{x^2} = \sum_{i=0}^{n} \frac{1}{i!} x^{2i} + \text{remainder term.}$$

1.2. Interpolation

We are given data pairs $\{x_i, f_i\}_{i=1}^m$, where x_i are distinct points in an interval $[a, b]$ and f_i are real numbers. We are asked to determine a function $p(x)$ that *interpolates* this data, i.e., that satisfies the m *interpolation conditions*

$$p(x_i) = f_i \qquad \text{for } i = 1, 2, \ldots, m,$$

in which case $p(x)$ is referred to as the *interpolant* of the data $\{f_i\}_{i=1}^m$ at the *interpolation points* $\{x_i\}_{i=1}^m$. Of course, there are an infinity of such functions so that it is necessary to narrow down the types of functions one considers for $p(x)$. In these notes, we will restrict attention to polynomial functions.

There are two means for producing the data pairs $\{x_i, f_i\}_{i=1}^m$. First, this data can be obtained from a given function $f(x)$, i.e., we *select* m points $x_i \in [a, b]$ and then let

$$f_i = f(x_i) \qquad \text{for } i = 1, 2, \ldots, m.$$

[1] Of course, for this simple problem, if we have MATLAB we would just go ahead and use it to approximate the integral and not bother with the Taylor polynomial approximation! But here we are trying to learn how functions are approximated, which includes learning something about how MATLAB does it.

In this case, interpolation is simply another means to do function approximation, i.e., the interpolant $p(x)$ is a presumably simpler function, e.g., a polynomial, that can be used as a surrogate for $f(x)$.

The second means for obtaining the data pairs $\{x_i, f_i\}_{i=1}^m$ is from observations, e.g., from measurements; thus, we do not know the function $f(x)$ but only know its value a m points. The construction of the interpolant $p(x)$, being a function of x, allows us to have some idea of what occurs at points other than the given points $\{x_i\}_{i=1}^m$.

EXAMPLE. For example, suppose we measure the ambient temperature every hour over a one-day period; then, we could have $m = 24$ and, in units of hours, $x_i = i$ and f_i is the measured temperature at hour i. We can then use the interpolant $p(x)$ corresponding to the data pairs $\{x_i, f_i\}_{i=1}^m$ to get an idea what the temperature was at other times during the one-day period; for example, the temperature at 1:30PM would roughly be $p(13.5)$. □

In general, for the second case, we cannot choose the points x_i either because someone else provides the data or because there are constraints on how they are chosen. In the first case, we are totally free to choose the set of interpolation points $\{x_i\}_{i=1}^m$. A natural question is then: are some choices of points better than others, i.e., do some point sets $\{x_i\}_{i=1}^m$ result in more accurate approximations to the given function $f(x)$? This is a fundamental question that we will address in these notes.

1.3. Measuring errors

We have used phrases such as the "difference between two functions" or the "error in an approximation of a function." We want to quantify these notions, i.e., precisely say what we mean by these phrases, specifically, say how we are going to measure such differences and errors. In our context, we are given a function $f(x)$ and then we will consider several ways to construct a function $p(x)$ that is hopefully "close" to $f(x)$; we want to quantify what it means to be "close," or equivalently, quantify what it means for the difference or error function $E(x) = f(x) - p(x)$ to be "small." There are many ways to measure the closeness of two functions; we use two such ways in these notes.

Continuous norms. Given the error function $E(x) = f(x) - p(x)$ defined on the interval $[a, b]$, we can measure its size using the $L^2(a, b)$ or *root-mean square* norm

$$\|E\|_2 = \Big(\int_a^b |E(x)|^2 \, dx \Big)^{1/2},$$

provided $E(x)$ is square integrable. Alternately, we will use the $L^\infty(a, b)$ or *uniform* or *maximum* or *pointwise* norm

$$\|E\|_\infty = \max_{x \in [a,b]} |E(x)|,$$

provided $E(x)$ is continuous.

Note that, for a continuous function $g(x)$,

$$\int_a^b |g(x)|^2 \, dx \le (b - a) \max_{x \in [a,b]} |f(x)|^2 = (b - a) \big(\max_{x \in [a,b]} |g(x)| \big)^2$$

so that

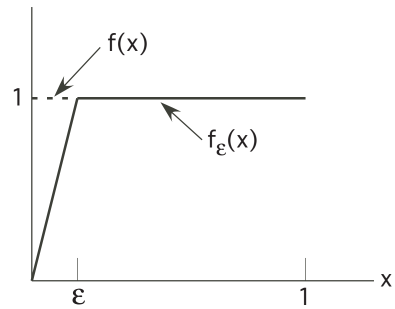
$$\|E\|_2 \le (b-a)^{1/2}\|E\|_\infty.$$

This implies that so long as the interval $[a,b]$ is finite, i.e, we have $-\infty < a < b < \infty$, small errors as measured in the $L^\infty(a,b)$ norm imply that the error is also small if it is measured in the $L^2(a,b)$ norm. The converse, however, is not true, i.e., close approximation in the $L^2(a,b)$ norm does not imply close approximation in the $L^\infty(a,b)$ norm.

EXAMPLE. For example, consider the function $f(x)=1$ on the interval $[0,1]$ and the continuous function $f_\varepsilon(x)$ as depicted in the figure, where $0<\varepsilon<1$. We have that

$$\|f(x)-f_\varepsilon(x)\|_2 = \varepsilon^{1/2} \qquad \text{and} \qquad \|f(x)-f_\varepsilon(x)\|_\infty = 1.$$

If $\varepsilon \ll 1$, then $\|f(x)-f_\varepsilon(x)\|_2$ is small but $\|f(x)-f_\varepsilon(x)\|_\infty = 1$, i.e., small $L^2(0,1)$ norm does not imply small $L^\infty(0,1)$ norm.



□

Small $L^\infty(a,b)$ norm implies that the function is small at every point in the interval $[a,b]$ whereas small $L^2(a,b)$ norm only implies that the function is small in some averaged sense. In this example, we see that $f_\varepsilon(x)$ is a good approximation to the function $f(x)=1$ if the error is measured in the $L^2(0,1)$ norm, but is a bad approximation if the error is measured in the $L^\infty(0,1)$ norm. This is why *it is very important to specify how one chooses to measure error and how one interprets the smallness of lack thereof of errors.*

Discrete norms. As we have seen, sometimes the value of a function is known only at a set of points $\{x_i\}_{i=1}^m$ so that we cannot evaluate either the $L^2(a,b)$ or $L^\infty(a,b)$ norms of that function. In this case, we make use the discrete norms

$$|||g|||_2 = \Big(\sum_{i=1}^m |g(x_i)|^2\,dx\Big)^{1/2} \qquad \text{and} \qquad |||g|||_\infty = \max_{x_i,\, i=1,\dots,m} |g(x_i)|.$$

1.4. Best approximation

We have touched on two means (Taylor polynomials and interpolants) for approximating a given function $f(x)$ by a simpler function $p(x)$. Suppose we restrict attention to polynomial interpolants so that both cases involve polynomial approximations. One may ask if either of these approaches produces a polynomial that is the best possible approximation of the given function, where now we know two ways of measuring what "best" is. Specifically, we ask if one can construct a polynomial $p_2^*(x)$ of a given degree n such that

$$\|f(x)-p_2^*(x)\|_2 \le \|f(x)-p(x)\|_2 \qquad \text{for any and all polynomials } p(x) \text{ of degree } n \text{ or less}$$

or construct a polynomial $p^*_\infty(x)$ of a given degree n such that

$$\|f(x)-p^*_\infty(x)\|_\infty \le \|f(x)-p(x)\|_\infty \qquad \text{for any and all polynomials } p(x) \text{ of degree } n \text{ or less.}$$

For obvious reasons, polynomials such as $p^*_2(x)$ and $p^*_\infty(x)$ are referred to as *best approximations* of the function $f(x)$; $p^*_2(x)$ is referred to as the *best* $L^2(a,b)$ or *best least-squares approximation* of $f(x)$ whereas $p^*_\infty(x)$ is referred to as the *best* $L^\infty(a,b)$ or *best uniform* or *best pointwise approximation* of $f(x)$. Because we have that

$$\begin{aligned}\|f(x)-p^*_\infty(x)\|_\infty &= \min_{p(x) \text{ of degree } n \text{ or less}} \|f(x)-p(x)\|_\infty \\ &= \min_{p(x) \text{ of degree } n \text{ or less}} \max_{x\in[a,b]} |f(x)-p(x)|,\end{aligned}$$

$p^*_\infty(x)$ is also referred to as the *best min-max approximation* of $f(x)$. Note that, in general, $p^*_2(x) \neq p^*_\infty(x)$ so that what is "best" depends on how we choose to measure the error.

Best approximations can also be defined for discrete data, i.e., when all that is known about a function are its values at the points x_i, $i = 1,\ldots,m$. In this case, we seek a polynomial $p^*_2(x)$ of a given degree n such that

$$\begin{aligned}&\sum_{i=1}^m |f(x_i)-p^*_2(x_i)|^2 \\ &\qquad \le \sum_{i=1}^m |f(x_i)-p(x_i)|^2 \quad \text{for any and all polynomials } p(x) \text{ of degree } n \text{ or less}\end{aligned}$$

or a polynomial $p^*_\infty(x)$ of a given degree n such that

$$\begin{aligned}&\max_{x_i,\, i=1,\ldots,m} |f(x_i)-p^*_\infty(x_i)|_\infty \\ &\quad \le \max_{x_i,\, i=1,\ldots,m} |f(x_i)-p(x_i)|_\infty \quad \text{for any and all polynomials } p(x) \text{ of degree } n \text{ or less.}\end{aligned}$$

1.5. The Weierstrass theorem

For the most part, in these notes we deal with polynomial approximation. One of the most important theorems in approximation theory, indeed, in all of mathematics, is the Weierstrass approximation theorem:

THEOREM [The Weierstrass theorem]. *Given any function $f(x)$ that is continuous on the interval $[a,b]$ and given any $\varepsilon > 0$, there exists a polynomial $p(x)$ such that*

$$\|f(x)-p(x)\|_\infty < \varepsilon.$$

Thus, no matter how small one chooses ε, there exists a polynomial whose value at every point in the interval $[a,b]$ is within ε of the value of $f(x)$ at that point. Polynomial interpolation and best approximation is motivated by the desire to construct polynomials that do well at approximating continuous functions.