Note on Weighted Residual Methods

Makoto Nakajima, UIUC
January 2006

1 Introduction

We study how to approximate a function with one variable, using the so-called weighted residual method. They are also called the projection methods.

The problem is the same as the Finite Element Methods. We want to approximate a function which is defined over a compact space and which does not have a parametric form, using another function with finite number of parameters.

Finite Element Methods divide the domain into sub-domains and approximate the original function in each sub-domain using simple functions, like linear functions or low order polynomials.

On the other hand, Weighted Residual Methods use a polynomial over the original domain and choose the coefficients of the polynomial such that the weighted error between the approximating function and the original function is minimized. The methods are called Weighted Residuals Methods because we minimize the weighted residuals.

Naturally, continuity and high order of differentiability is usually obtained in this method. Also the method can be intuitively extended to the multivariate case. Lastly, because the properties of polynomials are well known, we can have some results like the maximum error using some family of polynomials. However, since an approximating function is usually oscillating around the original function, we usually lose concavity or monotonicity of the original function.

2 The Problem

Let's limit ourself to the one-dimensional case for now. So the problem is like the following. We are going to start from a simple example and use the example to motivate more general treatment.

Problem 1
We want to approximate a continuous function $f(x)$ which is defined over $[\underline{x}, \overline{x}]$.

3 A Simple Example

Suppose we want to approximate a function $f(x)$ defined over $[\underline{x}, \overline{x}]$. Suppose we are using the following fourth order polynomial with coefficients $\theta = \{\theta_i\}_{i=0}^{3}$

$$\tilde{f}(x, \theta) = \theta_3 x^3 + \theta_2 x^2 + \theta_1 x + \theta_0$$

How we can choose the set of coefficients θ? One natural way is the following. First of all, we define a residual function $R(x, \theta)$:

$$R(x, \theta) = f(x) - \tilde{(}x, \theta)$$

Using the residual function, we can choose the set of parameters θ using the following formula:

$$\theta = \arg\min \int_{\underline{x}}^{\overline{x}} [R(x,\theta)]^2 dx$$

Taking the first order conditions with respect to θ_i, we can obtain:

$$0 = \int_{\underline{x}}^{\overline{x}} R(x,\theta) \frac{\partial R(x,\theta)}{\partial \theta_i} dx$$

We have four first order conditions ($i = 0, 1, 2, 3$) and four coefficients. Therefore, all we need to do (which might not be a trivial thing to do though) is to find the four parameters which satisfy the four equations above simultaneously.

4 The General Framework

4.1 The General Set-Up

Suppose we want to approximate a function $f(x)$ defined over $[\underline{x}, \overline{x}]$. Given a degree of polynomials n_p, a family of polynomials $\{\varphi_i\}_{i=0}^{n_p}$, and weights to each of the polynomials $\theta = \{\theta_i\}_{i=0}^{n_p}$, we can define an approximating function as follows:

$$\tilde{f}(x,\theta) = \sum_{i=0}^{n_p} \theta_i \varphi_i(x)$$

Suppose that the true function $f(x)$ satisfies the following equation:

$$F(f(x)) = 0$$

The approximating function $\tilde{f}(x,\theta)$ is expected to satisfy the equation above as much as possible. Now, let's define the residual function $R(x,\theta)$ as follows:

$$R(x,\theta) = F(\tilde{f}(x,\theta))$$

What we want to find is the parameters θ which achieves $R(x,\theta) = 0$.

If we can evaluate the $f(x)$ function easily, the residual function can be defined as:

$$R(x,\theta) = f(x) - \tilde{f}(x,\theta)$$

Finally, choose a set of test functions $\{g_i(x,\theta)\}_{i=0}^{n_p}$ and a weighting function $w(x)$ and we can obtain $n_p + 1$ equations which are used to pin down $n_p + 1$ coefficients:

$$0 = \int_{\underline{x}}^{\overline{x}} R(x,\theta) g_i(x,\theta) w(x) dx$$

Now the questions that we need to answer are:

1. How to choose $\{\varphi_i\}_{i=0}^{n_p}$?
2. How to choose $\{g_i(x,\theta)\}_{i=0}^{n_p}$
3. How to choose $w(x)$?

4.2 Orthogonal Polynomials

4.2.1 Definition

Since we will need to be able to identify coefficients associated with different order of polynomials, a desirable property for the family of polynomials $\{\varphi_i\}_{i=0}^{n_p}$ is that the polynomials are *different* to one another. Since the families of orthogonal polynomials satisfy the feature, they are usually the choice for weighted residual methods. We will see how the property of families of orthogonal polynomials is useful in our problem. First of all, let's begin by defining a weighting function.

Definition 1 (Weighting function)
A weighting function $w(x)$ on $[\underline{x}, \overline{x}]$ is any function that is positive almost everywhere and has a finite integral on $[\underline{x}, \overline{x}]$.

The weighting function is used to define a family of orthogonal polynomials:

Definition 2 (Orthogonal polynomials)
A family of polynomials $\{\varphi_i\}_{i=0}^{n_p}$ is mutually orthogonal with respect to a weighting function $w(x)$ if and only if:

$$\int_{\underline{x}}^{\overline{x}} \varphi_m(x)\varphi_n(x)w(x)dx$$

for $\forall i \neq j$.

4.2.2 Example of orthogonal polynomials 1: Chebyshev polynomials

1. **Closed-form definition:** $\varphi_i(x) = cos(i\ cos^{-1}x)$
2. **Recursive-form definition:** $\varphi_{i+1}(x) = 2x\varphi_i(x) - \varphi_{i-1}(x)$
3. **Domain:** $[-1, 1]$
4. **Weighting function:** $w(x) = (1 - x^2)^{-\frac{1}{2}}$

4.2.3 Example of orthogonal polynomials 2: Legendre polynomials

1. **Closed-form definition:** $\varphi_i(x) = \frac{(-1)^i d^i}{2^i i!\ dx^i}[(1 - x^2)^i]$
2. **Recursive-form definition:** $\varphi_{i+1}(x) = \frac{2i+1}{i+1}\varphi_i(x) - \frac{i}{i+1}\varphi_{i-1}(x)$
3. **Domain:** $[-1, 1]$
4. **Weighting function:** $w(x) = 1$

4.2.4 Example of orthogonal polynomials 2: Laguerre polynomials

1. **Closed-form definition:** $\varphi_i(x) = \frac{e^x d^i}{i!\ dx^i}[x^i e^{-x}]$
2. **Recursive-form definition:** $\varphi_{i+1}(x) = \frac{2i+1-x}{i+1}\varphi_i(x) - \frac{i}{i+1}\varphi_{i-1}(x)$
3. **Domain:** $[0, \infty)$
4. **Weighting function:** $w(x) = e^{-x}$

4.2.5 Example of orthogonal polynomials 1: Hermite polynomials

1. **Closed-form definition:** $\varphi_i(x) = (-1)^i e^{x^2} \frac{d^i e^{-x^2}}{dx^i}$
2. **Recursive-form definition:** $\varphi_{i+1}(x) = 2x\varphi_i(x) - 2i\ \varphi_{i-1}(x)$
3. **Domain:** $(-\infty, \infty)$
4. **Weighting function:** $w(x) = e^{-x^2}$

4.2.6 Evaluation of Orthogonal Polynomials

Usually it is easier to evaluate an orthogonal polynomial using the recursive form rather than the closed form. For example, below is how to evaluate Chebyshev polynomial of order n_p at x_0:

Algorithm 1 (Evaluating Chebyshev polynomial of order $\boldsymbol{n_p}$)

1. *Suppose we want to evaluate Chebyshev polynomial of order n_p ($\varphi_{n_p}(x)$).*
2. *It's easy to compute the first two orders of Chebyshev polynomials as follows:*

$$\varphi_0(x_0) = 1$$

$$\varphi_1(x_0) = x_0$$

3. *The Chebyshev polynomial of order i can be computed using the values of Chebyshev polynomials of order $i-1$ and $i-2$ and the following recursive formula:*

$$\varphi_i(x_0) = 2x_0\varphi_{i-1}(x_0) - \varphi_{i-2}(x_0)$$

4. *We only need to apply the formula up to the order n_p to evaluate Chebyshev polynomial of order n_p*

4.3 Methods of Weighted Residuals

Suppose we chooses some family of polynomials as $\{\varphi_i(x)\}_{i=0}^{n_p}$. We then need to decide how to choose $g_i(x,\theta)$ and $w(x)$ to specify the equations which are used to find the set of parameters θ for the approximating function:

$$0 = \int_{\underline{x}}^{\overline{x}} R(x,\theta) g_i(x,\theta) w(x) dx$$

There are four ways to choose $g_i(x,\theta)$ and $w(x)$:

1. **Least Squares** uses the following function for $g_i(x)$ and $w(x)$:

$$g_i(x) = \frac{\partial R(x,\theta)}{\partial \theta_i}$$

$$w(x) = 1$$

Equivalently, least squares method is characterized by:

$$\theta = \arg\min \int_{\underline{x}}^{\overline{x}} R^2(x,\theta) dx$$

2. **Collocation** imposes that the residual function $R(x, \theta)$ takes zero at $n_p + 1$ points over $[\underline{x}, \overline{x}]$. Let's denote these points as $\{x_i\}_{i=0}^{n_p}$. These points are called the collocation points. Then, the collocation method is characterized by the following $g_i(x)$ and $w(x)$ functions:

$$g_i(x) = \delta(x - x_i)$$
$$w(x) = 1$$

where $\delta(x - x_i)$ is the Dirac's delta function. In other words, collocation implies that:

$$R(x_i, \theta) = 0 \qquad \forall i = 0, 1, ..., n_p$$

A key to an easy calculation is to choose wisely the collocation points. The Chebyshev polynomials are extensively used with collocation methods, because of some nice property that we will learn.

Notice that the collocation method is similar to the finite element method, in the sense that the original function and the approximating function takes the same value only at finite number of collocation points.

3. **Galerkin** method uses the polynomials used to approximate the function as the test functions. Specifically

$$g_i(x) = \varphi_i(x)$$

As for the weighting function, if we are using a family of orthogonal polynomials for the base functions, the weighting function associated with the orthogonal polynomials is the default choice.

It is shown that, if we use a family of orthogonal polynomials and associated weighting function, and the residual function is the simple one like below, the solution to the Galerkin method becomes a very simple one:

$$R(x, \theta) = f(x) - \tilde{f}(x, \theta)$$

Remember that the Galerkin method takes the following form for each of $i = 0, 1, ..., n_p$:

$$0 = \int_{\underline{x}}^{\overline{x}} \left[f(x) - \sum_{i=0}^{n_p} \theta_i \varphi_i(x) \right] \varphi_i(x) w(x) dx$$

Using the orthogonality property of $\varphi_i(x)$, the above expression can be simplified to:

$$0 = \int_{\underline{x}}^{\overline{x}} \left[f(x) - \theta_i \varphi_i(x) \right] \varphi_i(x) w(x) dx$$

Solving for θ_i and we can obtain:

$$\theta_i = \frac{\int_{\underline{x}}^{\overline{x}} f(x) \varphi_i(x) w(x) dx}{\int_{\underline{x}}^{\overline{x}} \varphi_i^2(x) w(x) dx}$$

It's easy to solve because each equation only contains θ_i and not $\theta_{j \neq i}$. However, still we need to be able to compute nontrivial integral. A standard way to do the integration is to use numerical integration.

4. **Regression** is a special form of the collocation method. For the collocation method, the number of polynomials $n_p + 1$ is equal to the number of points $\{x_i\}_{i=0}^{n_p}$. But we can have more collocation points than the number of polynomials. In other words, collocation is the exactly identified case of the regression.