Class notes: Advanced Topics in Macroeconomics

Topic: Weighted Residual Methods

Date: October 8, 2018

In class, we worked through the chapter on weighted residual methods in Marimon and Scott (1999). We started with students trying to think about how one would go about solving a functional equation. Interestingly, the discussion lead to someone describing a least-squares version of the weighted residual method.

We then set out the problem more generally as in the chapter. The problem is to find $d : I\!R^m \rightarrow I\!R^n$ that satisfies a functional equation $F(d) = 0$, where $F : C_1 \rightarrow C_2$ and C_1 and C_2 are function spaces. As an example, think of d as decision or policy variables and F as first-order conditions from some maximization problem. The goal here is to find an approximation $d^n(x;\theta)$ on $x \in \Omega$ which depends on a finite-dimensional vector of parameters $\theta = [\theta_1, \theta_2, \ldots, \theta_n]'$. Weighted residual methods assume that d^n is a finite linear combination of known functions, $\psi_i(x)$, $i = 0, \ldots, n$, called *basis functions*:

$$d^n\left(x;\theta\right) = \psi_0\left(x\right) + \sum_{i=1}^{n} \theta_i \psi_i\left(x\right). \tag{1}$$

The functions $\psi_i(x)$, $i = 0, \ldots, n$ are typically simple functions. Standard examples of basis functions include simple polynomials (for example, $\psi_0(x) = 1$, $\psi_i(x) = x^i$), orthogonal polynomials (for example, Chebyshev polynomials), and piecewise linear functions.

We worked through examples with piecewise linear approximations. More specfically, we used the following basis functions:

$$\psi_i\left(x\right) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}} & \text{if } x \in [x_{i-1}, x_i] \\ \\ \frac{x_{i+1} - x}{x_{i+1} - x_i} & \text{if } x \in [x_i, x_{i+1}] \\ 0 & \text{elsewhere.} \end{cases} \tag{2}$$

We do not need to have the points x_i, $i = 1, \ldots, n$ equally spaced. For example, if we want to represent a function that has large gradients or kinks in certain places – say, because inequality constraints bind – then we can cluster points in those regions. In regions where the function is near-linear, we do not need many points.

We worked with the *residual equation*:

$$R\left(x;\theta\right)=F\left(d^{n}\left(x;\theta\right)\right)$$

and discussed how to choose θ so that $R(x;\theta)$ is close to zero for all x. Weighted residual methods get the residual close to zero in the weighted integral sense. That is, we choose θ so that

$$\int_{\Omega}\phi_{i}\left(x\right)R\left(x;\theta\right)dx=0,\quad i=1,\ldots,n,$$

where $\phi_i(x)$, $i=1,\ldots,n$ are *weight functions*. Note that $\phi_i(x)$ and $\psi_i(x)$ can be different functions. Alternatively, the weighted integral can be written

$$\int_{\Omega}w\left(x\right)R\left(x;\theta\right)dx=0,\tag{3}$$

where $w(x)=\sum_i\omega_i\phi_i(x)$ and (3) must hold for any nonzero weights ω_i, $i=1,\ldots,n$. Therefore, instead of setting $R(x;\theta)$ to zero for all $x\in\Omega$, the method sets a weighted integral of R to zero.

We discussed different choices of of weight functions, for example: determining the coefficients $\theta_1,\ldots,\theta_n$.

- **Least Squares**: $\phi_i(x)=\partial R(x;\theta)/\partial\theta_i$. This set of weights can be derived by calculating the first-order derivatives for the following optimization problem:

$$\min_{\theta}\int_{\Omega}R\left(x;\theta\right)^{2}\,dx.$$

- **Collocation**: $\phi_i(x)=\delta(x-x_i)$, where δ is the Dirac delta function. This set of weights implies that the residual is set to zero at n points $x_1,\ldots,x_n$ called the *collocation points*: $R(x_i;\theta)=0$, $i=1,\ldots,n$. If the basis functions are chosen from a set of orthogonal polynomials with collocation points given as the roots of the nth polynomial in the set, the method is called *orthogonal collocation.*

- **Galerkin**: $\phi_i(x)=\psi_i(x)$. In this case, the set of weight functions is the same as the basis functions used to represent d. Thus, the Galerkin method forces the residual to be orthogonal to each of the basis functions. As long as the basis functions are chosen from a complete set of functions, then equation (1) represents the exact solution, given that enough terms are included. The Galerkin method is motivated by the fact that a continuous function is zero if it is orthogonal to every member of a complete set of functions.

To illustrate weighted residual methods, we worked through a a simple problem in which the coefficients θ_i, $i = 1, \ldots, n$ of (1) satisfy a linear system of equations (that is, $A\theta = b$, where A and b do not depend on θ), namely,

$$F(d)(x) = d'(x) + d(x) = 0. \tag{4}$$

If we use simple polynomials for the d^n, that is, x^i, $i = 1, \ldots, n$, then the approximation is:

$$d^n(x;\theta) = 1 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \ldots + \theta_n x^n. \tag{5}$$

Note that $\psi_0(x) = 1$ so that the boundary condition at $x = 0$ is satisfied. The task is to find the coefficients θ_i, $i = 1, \ldots, n$ by applying a weighted residual method with one of the possible sets of weights. In each case, we solve a linear system of equations for θ, $A\theta = b$.

Let's start with least squares. In this case, the problem is to find θ that minimizes the integral of the squared residual. The residual can be found by substituting equation (5) into equation (4). The first-order conditions of the minimization of the squared residual imply that $\theta_1, \ldots, \theta_n$ satisfy

$$\int_0^{\bar{x}} \frac{\partial R(x;\theta)}{\partial \theta_i} R(x;\theta)\, dx = 0, \quad i = 1, \ldots, n,$$

where the residual and its derivative are given by

$$R(x;\theta) = 1 + \sum_{i=1}^{n} \theta_i \{ i x^{i-1} + x^i \},$$

$$\frac{\partial R(x;\theta)}{\partial \theta_i} = i x^{i-1} + x^i.$$

Suppose that $n = 3$ and $\bar{x} = 6$. Then the following system of equations is solved for θ:

$$\left\{ \int_0^6 \begin{bmatrix} 1+x \\ 2x+x^2 \\ 3x^2+x^3 \end{bmatrix} \begin{bmatrix} 1+x & 2x+x^2 & 3x^2+x^3 \end{bmatrix} dx \right\} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = -\int_0^6 \begin{bmatrix} 1+x \\ 2x+x^2 \\ 3x^2+x^3 \end{bmatrix} dx$$

or, more simply,

$$\begin{bmatrix} 114.0 & 576.0 & 3067.2 \\ 576.0 & 3139.2 & 17496.0 \\ 3067.2 & 17496.0 & 100643.7 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = \begin{bmatrix} -24 \\ -108 \\ -540 \end{bmatrix}.$$

More generally, we can use the fact that

$$R\left(x;\theta\right) = \left(C\vec{x} + e\right)' \theta + 1,$$

where $\vec{x} = [x, x^2, \ldots, x^n]'$, $e = [1, 0, \ldots, 0]'$, and

$$C = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 3 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & n & 1 \end{bmatrix}.$$

Since the residual R is linear in θ, the derivatives with respect to θ are given by $C\vec{x} + e$. Thus, the system of equations to be solved to compute the coefficients θ for the least squares method is given by

$$\left\{ \int_0^{\bar{x}} \left(C\vec{x} + e\right)\left(C\vec{x} + e\right)' \, dx \right\} \theta = - \int_0^{\bar{x}} \left(C\vec{x} + e\right) \, dx$$

or, more succinctly, $A\theta = b$ with

$$A = CMC' + ePC' + CP'e' + \bar{x}ee',$$
$$b = -CP' - \bar{x}e,$$

and

$$M = \int_0^{\bar{x}} \vec{x}\vec{x}' \, dx = \begin{bmatrix} \bar{x}^3/3 & \bar{x}^4/4 & \cdots & \bar{x}^{n+1}/\left(n+1\right) \\ \bar{x}^4/4 & \bar{x}^5/5 & \cdots & \bar{x}^{n+2}/\left(n+2\right) \\ \vdots & \vdots & \vdots & \vdots \\ \bar{x}^{n+2}/\left(n+2\right) & \bar{x}^{n+3}/\left(n+3\right) & \cdots & \bar{x}^{2n+1}/\left(2n+1\right) \end{bmatrix},$$

$$P = \int_0^{\bar{x}} \vec{x} \, dx = \begin{bmatrix} \bar{x}^2/2 \\ \bar{x}^3/3 \\ \vdots \\ \bar{x}^{n+1}/\left(n+1\right) \end{bmatrix}.$$

In Figure 3 of the chapter, I plot the approximate function d^n for $n = 3$ and the exact solution $exp(-x)$. If I had used $n = 5$, then the two lines would be visually indistinguishable.

Next, consider collocation. In this case, the problem is to find θ so that the residual is equal to 0 at n points in $[0, \bar{x}]$: $x_1, \ldots, x_n$. Suppose that the x_i are evenly spaced on $[0, 6]$ and that $n = 3$, so that $x_1 = 0$, $x_2 = 3$, and $x_3 = 6$. Then, θ must satisfy the following system of equations:

$$\begin{bmatrix} 1 & 0 & 0 \\ 4 & 15 & 54 \\ 7 & 48 & 324 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}.$$

More generally, we can solve $A\theta = b$ with $(C\vec{x} + e)'$ defined above evaluated at x_i in the ith row of A and b set to a vector of -1's:

$$\begin{bmatrix} (C\vec{x} + e)' \,|_{x=x_1} \\ (C\vec{x} + e)' \,|_{x=x_2} \\ \vdots \\ (C\vec{x} + e)' \,|_{x=x_n} \end{bmatrix} \theta = \begin{bmatrix} -1 \\ -1 \\ \vdots \\ -1 \end{bmatrix}.$$

In Figure 4 of the chapter, I plot the approximate function d^n and the exact solution. If I choose $n = 5$, the two lines are nearly indistinguishable. However, for $n = 3$, the approximation is not as good as the least squares approximation.

Finally, consider the Galerkin variant of the method. In this case, the problem is to find $\theta_1, \ldots, \theta_n$ that satisfy

$$\int_0^{\bar{x}} x^i R\,(x; \theta)\; dx = 0, \quad i = 1, \ldots, n. \tag{6}$$

Again, consider $n = 3$ and $\bar{x} = 6$. For these choices, the equations in (6) are given by

$$\left\{ \int_0^6 \begin{bmatrix} x \\ x^2 \\ x^3 \end{bmatrix} [\, 1 + x \quad 2x + x^2 \quad 3x^2 + x^3 \,]\; dx \right\} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = -\int_0^6 \begin{bmatrix} x \\ x^2 \\ x^3 \end{bmatrix} dx. \tag{7}$$

Note that we have written these equations in the form $A\theta = b$. If we compute the integrals in equation (7), then the system of equations becomes

$$\begin{bmatrix} 90.0 & 468.0 & 2527.2 \\ 396.0 & 2203.2 & 12441.6 \\ 1879.2 & 10886.4 & 63318.9 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = \begin{bmatrix} -18 \\ -72 \\ -324 \end{bmatrix}.$$

For general n and $\bar{x}$, the coefficients solve $A\theta = b$, where A and b are the following functions:

$$A = MC' + P'e'$$

$$b = -P'$$

with M, C, P, and e as defined above. In Figure 5 of the chapter, I plot the approximate function d^n and the exact solution. The results are similar to those obtained with the least squares method. Again, if I choose $n = 5$, then the approximate and exact solutions are visually indistinguishable.

Next we will work with basis functions that are nonzero on only small regions of the domain of x. The resulting representations of d^n will be piecewise functions (for example, piecewise linear, piecewise quadratic). In the terminology of numerical analysts, we will be applying a *finite element method.* Finite element methods use basis functions that are only nonzero on small regions of the domain of x (for example, the tent functions drawn in Figure 2).

The idea behind the finite element method is to break up the domain of x into smaller pieces, use low-order polynomials to get good local approximations for the function d, and then piece the local approximations together to get a good global approximation. In effect, one can think of the finite element method as a piecewise application of a weighted residual method. Thus, to apply a finite element method, we first divide the domain into smaller nonoverlapping subdomains. On each of the subdomains, we construct a local approximation to the function d. For the problem in (4), Ω is one-dimensional, and therefore, division of Ω means coming up with some partition, say, $[x_1, x_2, \ldots, x_n]$ on $I\!R$. Each subinterval $[x_i, x_{i+1}]$ is called an *element.*

Suppose, for example, that we want to represent d as a piecewise linear function; that is, over each element, we assume that the approximation is of the form $a + bx$. Suppose also that we want the function d to be continuous on the whole domain Ω. How would we construct basis functions $\psi_i(x)$ so that we can write d^n as in (1)?

The first step is to assign *nodes* on the element. For the finite element method, nodes are points on an element that are used to define the geometry of the element and to uniquely define the order of the polynomial being used to approximate the true solution over the element. Since we are assuming that an element is some interval $[x_i, x_{i+1}]$, two nodes – in particular, the two endpoints x_i and x_{i+1} – are needed to define the geometry. And only two points are needed to uniquely define a linear function. Therefore, the nodes on a one-dimensional element with linear bases are the two endpoints of the element.

The second step in constructing the basis functions is to assume that the undetermined coefficients are equal to the approximate solution at the nodal points. Assume that the numbering of elements and nodes is such that element i is the interval $[x_i, x_{i+1}]$: the first element is $[x_1, x_2]$, the second element is $[x_2, x_3]$, and so on. Assume also that the approximate solution on element i, $d_i^n(x;\theta)$, satisfies $d_i^n(x_i) = \theta_i$ and $d_i^n(x_{i+1}) = \theta_{i+1}$. In other words, assume that the undetermined coefficients represent the solution at the nodes. The approximation of d on element i, d_i^n, is therefore uniquely given by

$$d_i^n\left(x;\theta\right) = \theta_i \psi_i\left(x\right) + \theta_{i+1}\psi_{i+1}\left(x\right), \quad x \in \left[x_i, x_{i+1}\right],$$

where the basis functions are given by equation (2) and drawn in Figure 2. Since elements are connected to each other at nodal points on the element boundaries, this choice of basis functions guarantees that the approximation is continuous across elements. Notice also that any linear function (and, hence, any continuous piecewise linear d^n) can be represented with the basis functions given in (2).

Let the approximate solution to (4) be of the form

$$d^n\left(x;\theta\right) = \sum_{i=1}^{n} \theta_i \psi_i\left(x\right),$$

with $\psi_i(x)$, $i = 1, \ldots, n$ given by (2). To impose the boundary condition $d^n(0;\theta) = 1$, we need to set θ_1 to one. Let's apply a Galerkin method. Therefore, the weight functions are given by the bases $\psi_i(x)$, $i = 1, \ldots, n$.

Suppose that there are three elements with nodes at 0, 1, 3, and 6. Then the residual equation is given by

$$R\left(x;\theta\right) = \sum_{i=1}^{4} \theta_i \left(\psi_i'\left(x\right) + \psi_i\left(x\right)\right)$$

$$= \begin{cases} \theta_1 \; (-x) + \theta_2 \; (1 + x) & \text{if } x \in [0, 1] \\ \theta_2 \; \left(1 \; - \frac{1}{2}x\right) + \theta_3 \; \left(\frac{1}{2}x\right) & \text{if } x \in [1, 3] \\ \theta_3 \; \left(\frac{5}{3} - \frac{1}{3}x\right) + \theta_4 \; \left(-\frac{2}{3} + \frac{1}{3}x\right) & \text{if } x \in [3, 6]. \end{cases} \tag{8}$$

If we substitute the residual (8) into the weighted integral (3) with $\phi_i(x) = \psi_i(x)$, then we get the following system of equations:

$$\begin{aligned} &\Bigg\{ \int_0^1 \begin{bmatrix} 1-x \\ x \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} -x & 1+x & 0 & 0 \end{bmatrix} dx \\ &+ \int_1^3 \begin{bmatrix} 0 \\ \frac{3}{2} - \frac{1}{2}x \\ -\frac{1}{2} + \frac{1}{2}x \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 - \frac{1}{2}x & \frac{1}{2}x & 0 \end{bmatrix} dx \\ &+ \int_3^6 \begin{bmatrix} 0 \\ 0 \\ 2 - \frac{1}{3}x \\ -1 + \frac{1}{3}x \end{bmatrix} \begin{bmatrix} 0 & 0 & \frac{5}{3} - \frac{1}{3}x & -\frac{2}{3} + \frac{1}{3}x \end{bmatrix} dx \Bigg\} \begin{bmatrix} 1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \end{aligned}$$

or if we compute the integrals,

$$\begin{bmatrix} -1/6 & 2/3 & 0 & 0 \\ -1/3 & 1 & 5/6 & 0 \\ 0 & -1/6 & 5/3 & 1 \\ 0 & 0 & 0 & 3/2 \end{bmatrix} \begin{bmatrix} 1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{9}$$

Note that we need to drop the first equation because we have to impose that $\theta_1 = 1$ for the boundary condition to be satisfied. Recall that the integral equation can be written as in (3), where in this case, $w(x) = \sum_i \omega_i \psi_i(x)$. The function $w(x)$ must satisfy the homogeneous counterpart of the boundary condition $d(0) = 1$, that is, $w(0) = 0$. For those familiar with the calculus of variations, w is like the variation of the solution and thus must satisfy the homogeneous counterparts of boundary conditions for d. Enforcing the condition $w(0) = 0$ is equivalent to dropping the first equation in (9). Therefore, the system of equations reduces to

$$\begin{bmatrix} 1 & 5/6 & 0 \\ -1/6 & 5/3 & 1 \\ 0 & 0 & 3/2 \end{bmatrix} \begin{bmatrix} \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 0 \\ 0 \end{bmatrix},$$

with three equations and three unknowns. In Figure 7 of the chapter, I plot the finite element approximation and the exact solution. By construction, the approximate function is piecewise linear.

Next class, we'll discuss extensions to problems in economics with more states and more nonlinearities.