Federal Reserve Bank of Minneapolis
Research Department

May 1997

Application of Weighted Residual Methods to Dynamic Economic Models †

Ellen R. McGrattan
Federal Reserve Bank of Minneapolis

† This paper is based on lectures given at the 7th Annual Summer School of the European Economic Association held September 3-13, 1996, at the European University Institute. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1. Introduction

A necessary intermediate step for many problems in economics is the solution of a functional equation. Typically, we are seeking decision functions that satisfy a set of Euler conditions or a value function that satisfies Bellman's equation. In many cases, we cannot derive analytical solutions for these functions and instead must rely on numerical methods. In this paper, I will show how to apply weighted residual and finite element methods to this type of problem.

The basic idea behind weighted residual methods is to approximate an unknown function by a linear combination of known *basis* functions. In many applications, the basis functions are polynomials. The coefficients on each basis function are the objects to be determined in order to obtain an approximate solution. These coefficients are found by setting the residual of the equation being solved equal to zero in an average sense. In other words, a weighted integral of the residual is set equal to zero.

The finite element method can be viewed as an piecewise application of the weighted residual method. With the finite element method, the first step in solving the functional equation is to subdivide the domain of the state space into nonintersecting subdomains called *elements*. The domain is subdivided because the method relies on fitting low-order polynomials on subdomains of the state space rather than high-order polynomials on the entire state space. The local approximations are then pieced together to get a global approximation. As the dimensionality of the problem increases, higher order functions can be used where needed, with fewer elements.

The primary goal of this paper is to illustrate the application of weighted residual and finite element methods by way of examples. I start with a simple differential equation because the coefficients to be computed satisfy a linear system of equations. In this case, I can work through the example without the use of a computer. I then apply the methods to a deterministic growth model and a stochastic growth model – two standard models in economics. In the growth examples, the coefficients to be computed satisfy nonlinear systems of equations. Fortunately, in the case of the finite element method, these nonlinear equations are exploitably sparse.

2. The General Procedure

The problem is to find $d : \mathbb{R}^m \to \mathbb{R}^n$ that satisfies some functional equation $F(d) = 0$, where $F : C_1 \to C_2$ and C_1 and C_2 are function spaces. The goal here is to find some approximation $d^h(x;\theta)$ 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^h is a finite linear combination of known functions, called basis functions,

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

where $\psi_i(x)$, $i = 0, \ldots, n$, are known and are typically simple functions (e.g., polynomials, splines, piecewise linear functions, etc.).

If I substitute the approximate solution d^h into the functional equation $F(d) = 0$, I get the residual equation,

$$R(x;\theta) = F(d^h(x;\theta)),$$

that I want to get 'close to zero' for all x. Weighted residual methods get the residual close to zero in the weighted integral sense. That is, θ is chosen so that

$$\int_\Omega \phi_i(x) R(x;\theta) dx = 0, \quad i = 1, \ldots, n,$$

where $\phi_i(x)$ and $\psi_i(x)$ can be different functions. Alternatively, the weighted integral can be written

$$\int_\Omega w(x) R(x;\theta) dx = 0 \tag{2}$$

where $w(x) = \sum_i \omega_i \phi_i(x)$ and (2) must hold for any nonzero weights ω_i, $i = 1, \ldots, n$. Therefore, instead of setting $R(x;\theta)$ equal to zero for all $x \in \Omega$, the method sets a weighted integral of R equal to zero.

I will consider three specific sets of weights and, hence, three different ways of determining the coefficients θ_1, ..., θ_n.

1. **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:

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

2. **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 equal to zero at n points $x_1, \ldots, x_n$ called the collocation points. 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, then the method is called *orthogonal collocation.*

3. **Galerkin**: $\phi_i(x) = \psi_i(x)$. In this case, the set of weights 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 (1) represents the exact solution, given that enough terms are included. The motivation for the Galerkin method is the fact that a continuous function is zero if it is orthogonal to every member of a complete set of functions.

Another way to classify weighted residual methods is by the choice of the basis functions for d^h. That is, weighted residual methods can be divided into two categories: spectral methods and finite element methods. Spectral methods use basis functions that are smooth (i.e., C^∞) and nonzero on most of the domain of x (e.g., sets of polynomials defined on Ω). Finite element methods use basis functions that are only nonzero on small regions of the domain of x. As a result, the approximate functions (e.g., d^h) are piecewise functions. In fact, one could think of the finite element method as a piecewise application of a weighted residual method. The idea of the finite element method is to use low-order polynomials to get good local approximations for the function and then to piece them together to get a good global approximation.

The first step in applying a finite element method is to discretize the domain. For example, if Ω is one-dimensional, then the discretization is a partition on an interval of $\mathbb{R}$, say, $[x_1, x_2, \ldots, x_n]$. Each subinterval $[x_i, x_{i+1}]$ is an *element*. If Ω is two-dimensional, then the finite element discretization consists of simple two-dimensional (nonoverlapping) subdomains, such as triangles, rectangles, or quadrilaterals, which are referred to as elements. In most economic applications, a reasonable choice for the element shape is a rectangle in two-dimensional problems, a cube in three-dimensional problems, and so on. Over each element, the approximation to d is represented as a linear combination of basis functions (typically low-order polynomials). The undetermined coefficients represent the approximate solution to $d(x)$ at a finite number of preselected points called *nodes* that appear on the boundary and in the interior of the element.

Suppose, for example, that we wanted to represent d as piecewise linear function; that is, over each element, we assume that the approximation is of the form $a + bx$. How would we construct basis functions $\psi_i(x)$ so as to write d^h 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 I am 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, $d_i^h(x)$, satisfies: $d_i^h(x_i) = \theta_i$ and $d_i^h(x_{i+1}) = \theta_{i+1}$; in other words, we can assume that the undetermined coefficients represent the solution at the nodes. The approximation of d on element i, d_i^h, is therefore uniquely given by

$$d_i^h(x) = \theta_i \psi_i(x) + \theta_{i+1}\psi_{i+1}(x), \quad x \in [x_i, x_{i+1}], \tag{3}$$

where the basis functions are given by

$$\psi_i(x) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}} & \text{if } x \in [x_{i-1}, x_i] \\ \\ \frac{x_{i+1} - x}{x_{i+1} - x_i} & \text{if } x \in [x_i, x_{i+1}] \\ 0 & \text{elsewhere} \end{cases} \tag{4}$$

$i = 1, \ldots, n$, and the nodal points of element i are the endpoints x_i and x_{i+1}. 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^h) can be represented with the basis functions given in (4).

In the examples in Sections 3-5, I will focus on only the simplest finite element approximations that are used in practice. For the one-dimensional examples, I will use linear and quadratic basis functions. For the two-dimensional examples, I will use rectangular elements with linear and quadratic basis functions. Extensions to higher dimensions and higher order polynomials are relatively straightforward. (See, for example, Hughes 1987.) Extensions to non-rectangular element shapes require additional work but are not as useful in economic problems as in engineering problems, which sometimes involve irregularly shaped domains. (See, for example, Hughes 1987 and Reddy 1993.)

In general, it must be the case that the finite element approximations satisfy certain conditions to guarantee convergence of the method as the number of elements is increased. The approximate solution should be continuous and differentiable over each element – with the order of differentiability such that all terms in the residual equation are nonzero. The polynomials used to represent the approximate solution on an element should be complete, and the approximation should be continuous across element boundaries.

Although the finite element method – as I will apply it – can be viewed as a Galerkin weighted residual method, there is a lot of structure put on the bases and the element geometries. In fact, one criticism that is sometimes raised against spectral methods is that, unlike with the finite element method, the selection of basis and weighting functions seems arbitrary. This is why the finite element method and spectral weighted residual methods are typically treated separately in textbooks. (See, for example, Reddy 1993.) However, this does not mean that there are not good and bad choices to be made in the case of spectral methods. Applications typically lead one to choose certain functions over others. For example, although the set of simple polynomials, 1, x, $\ldots$, x^n, seems a natural choice, it may be better to use a set of orthogonal polynomials. The reason for this is that x^n and x^{n+1} take on similar values when n is large. As a result, there may be problems with ill-conditioned matrices when computing θ in equation (2). There may also be other benefits of using certain classes of functions. For example, approximation with Chebyshev polynomials has the nice property that coefficients on higher order polynomials fall off rapidly. Therefore, using a high value of n may not be necessary in many examples.

To illustrate weighted residual and finite element methods, it helps to start with a simple problem in which the undetermined coefficients satisfy a linear system of equations. Once I have tackled the simple problem, I will then apply the methods to the standard growth model.[1]

[1] See Aiyagari and McGrattan (1996), Braun and McGrattan (1993), and Chari, Kehoe and McGrattan (1996, 1997) for other examples.

3. Solution of a First-order Differential Equation

Consider the following first-order problem: find f such that $f'(x) + f(x) = 0$, with $f(0) = 1$ and $x \in [0, \bar{x}]$. The solution to this problem is $f(x) = e^{-x}$. Using the notation of Section 2, we have $d = f$ with $d(0) = 1$ and

$$F(f)(x) = f'(x) + f(x). \tag{5}$$

In this section, we consider various choices of basis functions (ψ_i) and weighting functions (ϕ_i), so as to illustrate the weighted residual methods described in Section 2.

Example 1. Suppose I start with the simplest set of basis functions: x^i, $i = 1, \ldots, n$. Then the approximation to $f(x)$ is given by

$$f^h(x; \theta) = 1 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \ldots \theta_n x^n. \tag{6}$$

Note that I have chosen $\psi_0(x) = 1$ in order to satisfy the boundary condition at $x = 0$. The coefficients θ_i, $i = 1, \ldots, n$ are found by applying a weighted residual method with one of three possible sets of weights. In each case, a linear system of equations, $A\theta = b$, will be solved for θ.

1a. 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 (6) in (5). The first order conditions of the minimization of the squared residual imply that θ_1, ..., θ_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 $n = 3$ 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, written 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, I can use the fact that

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

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, the derivatives 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

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

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

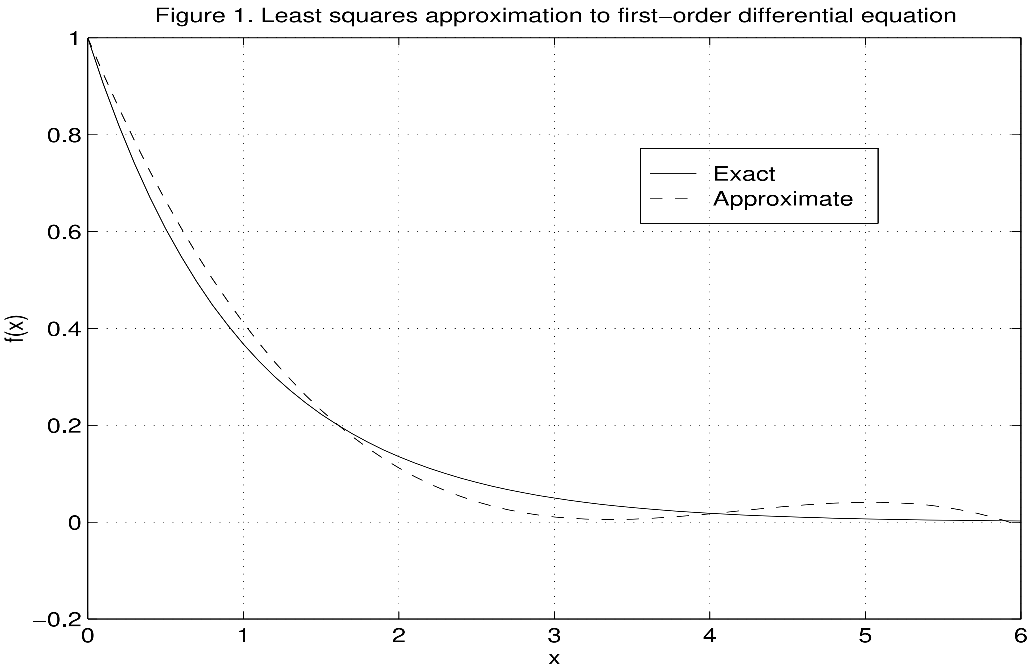
$$A = CMC' + ePC' + CP'e' + \bar{x}ee',$$

$$b = -CP' - \bar{x}e,$$

and

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

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



In Figure 1, I have plotted the approximate function f^h for $n = 3$ and the exact solution e^{-x}. If I had used $n = 5$, then the two lines would be visually indistinguishable.

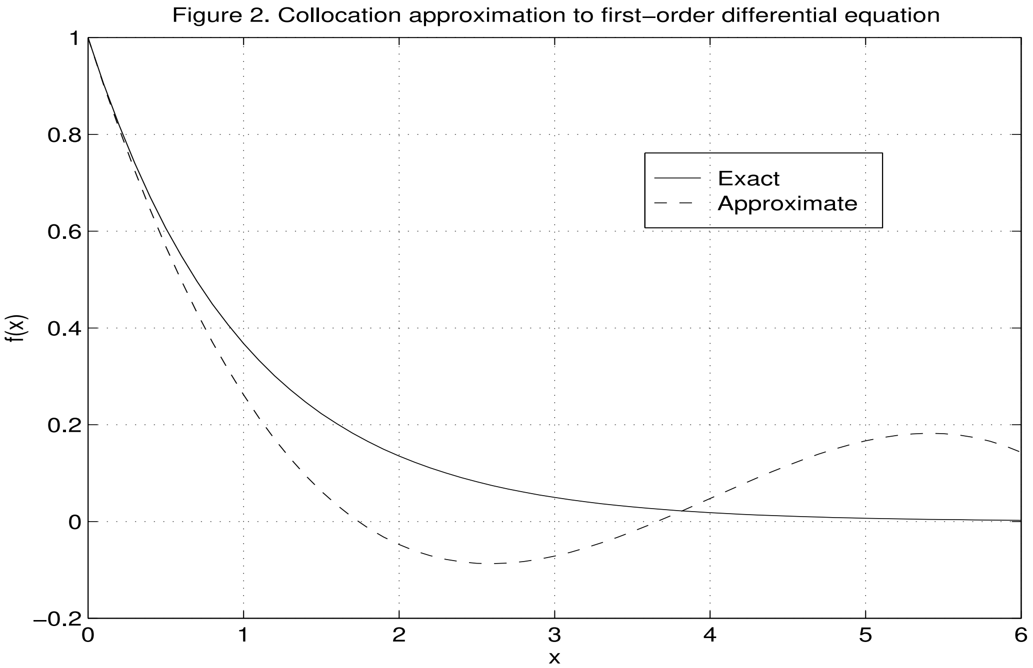
1b. Collocation. In this case, the problem is to find θ so that the residual is equal to 0 at n points in $[0, \bar{x}]$: x_1, …, x_n. Suppose that the x_i are evenly spaced on $[0, 6]$ and $n = 3$, so that $x_1 = 0$, $x_2 = 3$ and $x_3 = 6$. Then, θ must satisfy the following system of equations:

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

More generally, I can solve $A\theta = b$ with $(C\vec{x} + e)'$ evaluated at x_i in the ith row of A and b set equal 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 2, I have plotted the approximate function f^h 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. This is due in part to our choices of



basis functions and collocation points. Below I show that significant improvements can be made by using orthogonal collocation.

1c. Galerkin. In this case, the problem is to find θ_1, ..., θ_n that satisfy

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

Again, consider $n = 3$ and $\bar{x} = 6$. For these choices, I solve the following system of equations:

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

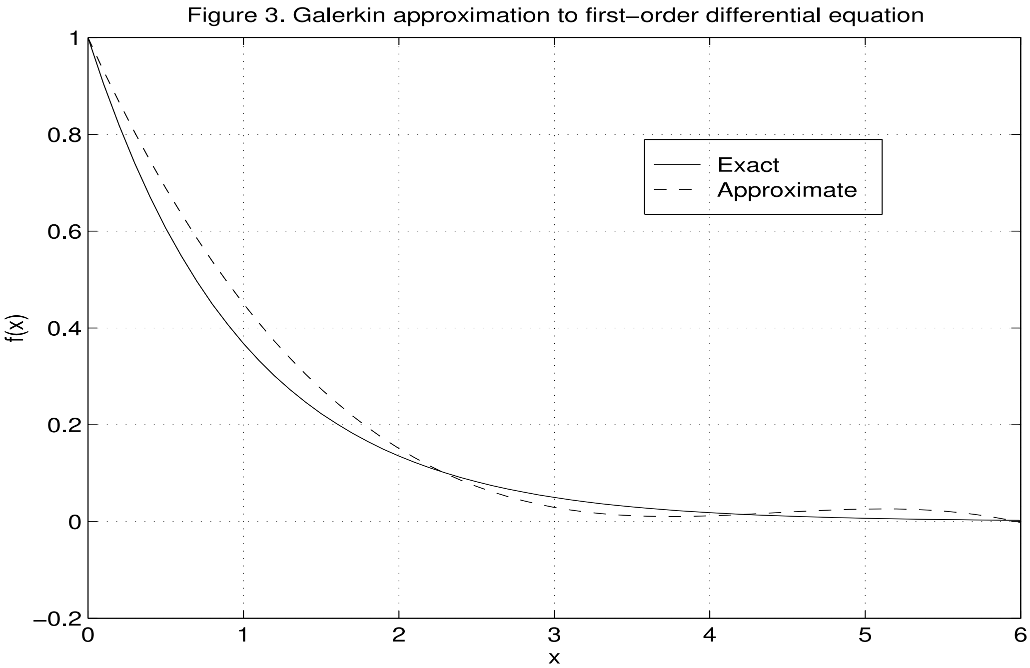
If we compute the integrals, then the system becomes

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

More generally, I can solve $A\theta = b$, where A and b are the following functions of $\bar{x}$ and n:

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

$$b = -P'$$



with M, C, P, and e as defined above. In Figure 3, I have plotted the approximate function f^h 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.

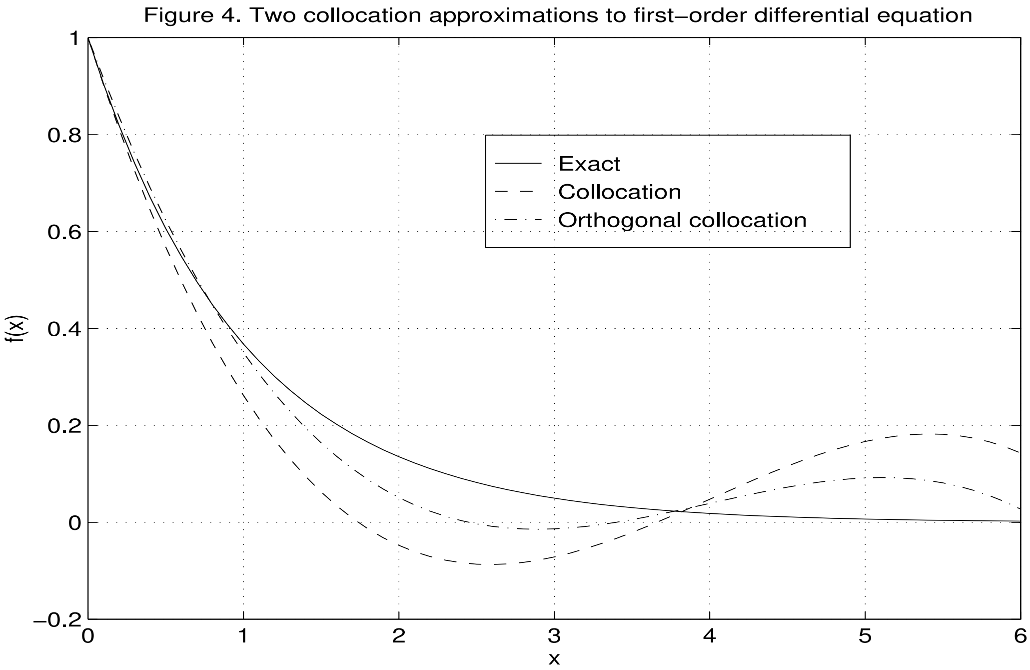
In Example 1, I assumed that the basis functions were the set of simple polynomials, $\{1, x, \ldots, x^n\}$. As I noted in Section 2, numerical problems can arise with this set of bases when the order of the approximating function (i.e., n) gets large. Suppose instead that I use a set of orthogonal polynomials. In particular, suppose I use Chebyshev polynomials. Chebyshev polynomials are defined on $[-1, 1]$ and are given recursively as follows: $p_0(x) = 1$, $p_1(x) = x$, and

$$p_{i+1}(x) = 2xp_i(x) - p_{i-1}(x), \quad i \geq 1. \tag{7}$$

Note also that $p_i(x) = \cos(i \arccos x)$.

Example 2. Let the approximate solution to (5) be of the following form:

$$f^h(x; \theta) = 1 + x \sum_{i=1}^{n} \theta_i p_{i-1}(2x/\bar{x} - 1), \tag{8}$$



where p_i is the ith Chebyshev polynomial. Therefore, $\psi_i(x) = xp_{i-1}(2x/\bar{x} - 1)$ are the basis functions defined on $[0, \bar{x}]$. Suppose I apply the method of orthogonal collocation. In this case, the collocation points are the n roots of the nth Chebyshev polynomial. Suppose that $n = 3$ and $\bar{x} = 6$. Then, the collocation points are $(3(\cos(5\pi/6)+1), 3, 3(\cos(\pi/6)+1))$ (which is equal to (0.4,3,5.6)). The system of equations to be solved is, therefore,

$$\begin{bmatrix} 1.4 & -1.1 & 0.2 \\ 4.0 & 1.0 & -4.0 \\ 6.6 & 7.6 & 9.8 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}. \tag{9}$$

The solution to (9) is plotted in Figure 4 along with the solution from Example 1b and the exact solution. Notice that the approximation using orthogonal polynomials is closer to the exact solution at all points in the domain.

Next I consider an application of the finite element method with linear basis functions. To apply a finite element method, I must first discretize the domain. For $\Omega = [0, \bar{x}]$, I choose some partition $[x_1, \ldots, x_n]$ with $x_1 = 0$ and $x_n = \bar{x}$. I will refer to the subintervals $[x_i, x_{i+1}]$, $i = 1, \ldots, n-1$ as elements and to the indices i as nodes.

Example 3. Let the approximate solution to (5) be of the form

$$f^h(x;\theta) = \sum_{i=1}^{n} \theta_i \psi_i(x),$$

with $\psi_i(x)$, $i, \ldots, n$, given by (4). If I apply a Galerkin method, then the weighting function has the same representation as f^h, i.e.,

$$w(x) = \sum_{i=1}^{n} \omega_i \psi_i(x),$$

where $w(x)$ satisfies the homogeneous counterpart of the boundary condition $f(0) = 1$, i.e., $w(0) = 0$.[2]

3a. Three elements. Suppose there are 3 elements with nodes at 0, 1, 3, and 6. The weighted integral form of the residual is given by

$$\begin{bmatrix} 0 & \omega_2 & \omega_3 & \omega_4 \end{bmatrix} \left\{ \int_0^1 \begin{bmatrix} 1-x \\ x \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} -x & x+1 & 0 & 0 \end{bmatrix} dx \right.$$

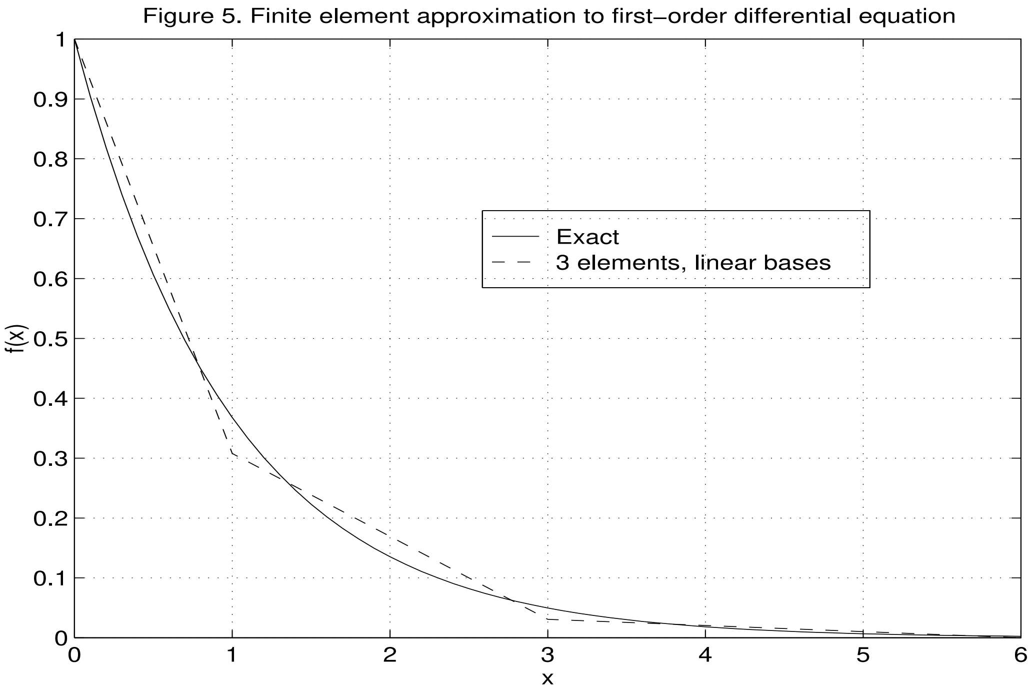
$$+ \int_1^3 \begin{bmatrix} 0 \\ \frac{3}{2} - \frac{1}{2}x \\ \frac{1}{2}x - \frac{1}{2} \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 - \frac{1}{2}x & \frac{1}{2}x & 0 \end{bmatrix} dx$$

$$\left. + \int_3^6 \begin{bmatrix} 0 \\ 0 \\ \frac{1}{2} - \frac{1}{3}x \\ \frac{1}{3}x - 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & \frac{1}{6} - \frac{1}{3}x & \frac{1}{3}x - \frac{2}{3} \end{bmatrix} dx \right\} \begin{bmatrix} 1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} = 0,$$

or, if we compute the integrals,

$$\begin{bmatrix} 0 & \omega_2 & \omega_3 & \omega_4 \end{bmatrix} \begin{bmatrix} -1/6 & 2/3 & 0 & 0 \\ -1/3 & 1 & 5/6 & 0 \\ 0 & -1/6 & 5/3 & 1 \\ 0 & 0 & 0 & 3/2 \end{bmatrix} \begin{bmatrix} 1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} = 0.$$

[2] 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 f.



The assumption that this equation has to hold for any nonzero ω_2, ω_3, and ω_4 implies that θ must satisfy

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

In Figure 5, I have plotted the finite element approximation and the exact solution. By construction, the approximate function is piecewise linear. As in the case of Examples 1-3, with five degrees of freedom it is difficult to distinguish the approximate and exact solutions when they are plotted.

3b. m elements, linear bases. It turns out that the same calculations are repeated for each element and, therefore, constructing the linear system of equations for θ can be made very simple. Construction of the approximation on any element can be standardized by working with a *master* element and introducing notation that is local rather than global. Consider this master element. Assume that it has length ℓ_e, that node 1 is placed at 0 and that node 2 is placed at ℓ_e. To construct the approximation f^h on the master element, I use a linear combination of two functions:

$$\psi_1^e(\bar{x}) = 1 - \frac{\bar{x}}{\ell_e} \qquad \text{and} \qquad \psi_2^e(\bar{x}) = \frac{\bar{x}}{\ell_e}$$

where $\bar{x}$ lies in $[0,\ell_e]$. If we were mapping element e given by $[x_e, x_{e+1}]$ to the master element, then the relationship between the *local coordinate* $\bar{x}$ and the *global coordinate* x (which lies somewhere between x_e and x_{e+1}) is given by $\bar{x} = x - x_e$.

Let

$$\Psi(\bar{x}) = [\, 1 - \tfrac{\bar{x}}{\ell_e} \quad \tfrac{\bar{x}}{\ell_e} \,]$$

$$\frac{d\Psi(\bar{x})}{d\bar{x}} = [\, -\tfrac{1}{\ell_e} \quad \tfrac{1}{\ell_e} \,].$$

Define K^e to be the following integral:

$$\begin{aligned} K^e &= \int_0^{\ell_e} \left\{ \Psi(\bar{x})' \left(\Psi(\bar{x}) + \frac{d\Psi(\bar{x})}{d\bar{x}} \right) \right\} d\bar{x} \\ &= \int_0^{\ell_e} \left\{ \begin{bmatrix} 1 - \frac{\bar{x}}{\ell_e} \\ \frac{\bar{x}}{\ell_e} \end{bmatrix} \left([\, 1 - \tfrac{\bar{x}}{\ell_e} \quad \tfrac{\bar{x}}{\ell_e} \,] + [\, -\tfrac{1}{\ell_e} \quad \tfrac{1}{\ell_e} \,] \right) \right\} d\bar{x} \\ &= \begin{bmatrix} \frac{\ell_e}{3} - \frac{1}{2} & \frac{\ell_e}{6} + \frac{1}{2} \\ \frac{\ell_e}{6} - \frac{1}{2} & \frac{\ell_e}{3} + \frac{1}{2} \end{bmatrix}. \end{aligned}$$

Then, one can show that the residual equation is given by

$$[\, 0 \quad \omega_2 \quad \cdots \quad \omega_{m+1} \,] \begin{bmatrix} \begin{bmatrix} K^1 \end{bmatrix} & & & & 0 \\ & \begin{bmatrix} K^2 \end{bmatrix} & & & \\ & & \begin{bmatrix} K^3 \end{bmatrix} & & \\ & & & \ddots & \\ 0 & & & & \begin{bmatrix} K^m \end{bmatrix} \end{bmatrix} \begin{bmatrix} 1 \\ \theta_2 \\ \vdots \\ \theta_{m+1} \end{bmatrix} = 0. \tag{10}$$

With the exceptions of (1,1) and $(m+1, m+1)$, the diagonal elements of the matrix multiplying θ in (10) are the sum of two numbers. For example, diagonal element (2,2) is found by summing $K^1(2,2)$ and $K^2(1,1)$, (3,3) is found by summing $K^2(2,2)$ and $K^3(1,1)$, and so on. The residual equation holds for all nonzero ω_i, $i \geq 2$. Therefore, I have a system of m equations of the form $A\theta = b$ to solve for θ in which A and b depend only on the element lengths ℓ_e, $e = 1, \ldots, m$.

An important advantage of the finite element method over the spectral methods is that the matrix A is sparse. For example, for $m = 6$, A is a 6×6 matrix with more zeros than nonzeros (20 versus 16). Furthermore, the matrix is tridiagonal and, therefore, has an "exploitable" structure. Saad (1996) describes iterative methods for solving large sparse systems of equations. Although the examples given in this paper have a small number of unknowns and therefore can be solved without resorting to sparse equation solvers, problems such as that described in Chari, Kehoe, and McGrattan (1997) require solving systems as large as $25{,}000 \times 25{,}000$. In such cases, the memory limitations make standard (non-sparse) matrix inversion routines infeasible to use.

In Example 3, I assumed that the approximate solution on each element was a linear function. There are two possible ways to improve the approximation: increase the number of elements or increase the order of the polynomial approximation on each element. We next consider an approximation with quadratic basis functions. In this case, a typical element will have 3 nodes: two at the endpoints and one interior. The two nodes at the endpoints are needed to define the geometry of the element (which is the same as in the case of linear bases). The interior point is needed to uniquely determine the order of the polynomial; that is, there are three degrees of freedom if the approximate solution is of the form

$$f_e^h(x) = a + bx + cx^2. \tag{11}$$

In order to allow flexibility, I do not necessarily assume that the third nodal point is at the midpoint of the element. However, because I want to avoid repeating the same calculations, we should calculate the functions and residuals for the master quadratic element just as I did in the linear case in Example 3. Consider the master element. Assume it has length ℓ_e. Let $\bar{x}$ be the local coordinate of the element which has node 1 located at x_1^e. Then, $\bar{x} = x - x_1^e$ where x is the global coordinate. Locally, assume that the nodes are located at $\bar{x} = 0$, $\bar{x} = \alpha\ell_e$, and $\bar{x} = \ell_e$, with $0 < \alpha < 1$.

As in the linear case, I can write the approximation on element e as a linear combination of basis functions that have the property that they are equal to 1 at one node and 0 at the other nodes, i.e.,

$$f_e^h(\bar{x}) = \theta_1^e\psi_1^e(\bar{x}) + \theta_2^e\psi_2^e(\bar{x}) + \theta_3^e\psi_3^e(\bar{x})$$

where the ψ_i^e's are constructed so as to satisfy $f_e^h(0) = \theta_1^e$, $f_e^h(\alpha\ell_e) = \theta_2^e$, and $f_e^h(\ell_e) = \theta_3^e$. The only functions that satisfy these conditions are given by

$$\begin{aligned} \psi_1^e(\bar{x}) &= \left(1 - \frac{\bar{x}}{\ell_e}\right)\left(1 - \frac{\bar{x}}{\alpha\ell_e}\right) \\ \psi_2^e(\bar{x}) &= \frac{\bar{x}}{\alpha(1-\alpha)\ell_e}\left(1 - \frac{\bar{x}}{\ell_e}\right) \\ \psi_3^e(\bar{x}) &= -\frac{\alpha\bar{x}}{(1-\alpha)\ell_e}\left(1 - \frac{\bar{x}}{\alpha\ell_e}\right). \end{aligned} \tag{12}$$

Example 4. Let the approximate solution be piecewise quadratic. Assume that the interior nodes on each element are at the midpoints of the elements (so that $\alpha = 1/2$.) Let

$$\Psi(\bar{x}) = [\,\psi_1^e(\bar{x}) \quad \psi_2^e(\bar{x}) \quad \psi_3^e(\bar{x})\,]$$

where $\psi_i^e(\bar{x})$ is given in (12), and define K^e to be the following integral:

$$\begin{aligned} K^e &= \int_0^{\ell_e} \left\{\Psi(\bar{x})'\left(\Psi(\bar{x}) + \frac{d\Psi(\bar{x})}{d\bar{x}}\right)\right\} d\bar{x} \\ &= \begin{bmatrix} \frac{2}{15}\ell_e - \frac{1}{2} & \frac{1}{15}\ell_e + \frac{2}{3} & -\frac{1}{30}\ell_e - \frac{1}{6} \\ \frac{1}{15}\ell_e - \frac{2}{3} & \frac{8}{15}\ell_e & \frac{1}{15}\ell_e + \frac{2}{3} \\ -\frac{1}{30}\ell_e + \frac{1}{6} & \frac{1}{15}\ell_e - \frac{2}{3} & \frac{2}{15}\ell_e + \frac{1}{2} \end{bmatrix}. \end{aligned}$$

Then, one can show that, if there are m elements, then the residual equation is given by

$$\begin{bmatrix} 0 & \omega_2 & \cdots & \omega_{2m+1} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} K^1 & \\ & \end{bmatrix} & & & & 0 \\ & \begin{bmatrix} K^2 & \\ & \end{bmatrix} & & & \\ & & \begin{bmatrix} K^3 & \\ & \end{bmatrix} & & \\ & & & \ddots & \\ 0 & & & & \begin{bmatrix} K^m & \\ & \end{bmatrix} \end{bmatrix} \begin{bmatrix} 1 \\ \theta_2 \\ \vdots \\ \theta_{2m+1} \end{bmatrix} = 0. \tag{13}$$

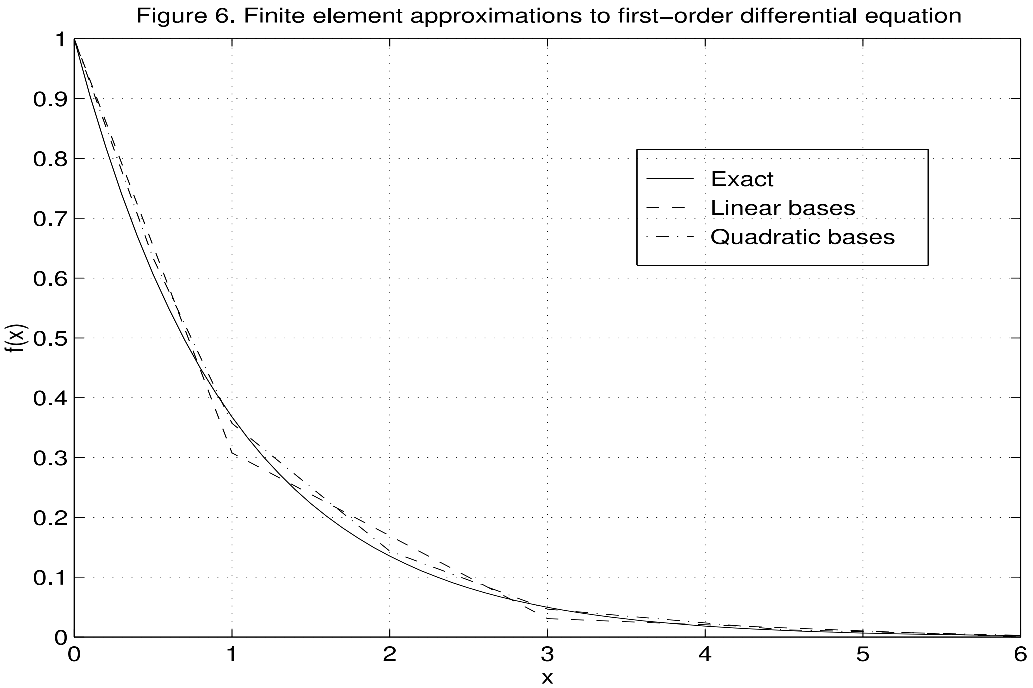
Diagonal element (3,3) is the sum of $K^1(3,3)$ and $K^2(1,1)$, diagonal element (5,5) is the sum of $K^2(3,3)$ and $K^3(1,1)$, and so on.

Note that there is a relationship between the θ_i's and the θ_i^e's. The θ_i^e denotes the ith unknown coefficient on element e, $i = 1, 2, 3$, whereas θ_i denotes the unknown coefficient for node i, $i = 1, \ldots, 2m+1$. If I label nodes consecutively, then $\theta_2 = \theta_2^1$, $\theta_3 = \theta_3^1 = \theta_1^2$, $\theta_4 = \theta_2^2$, $\theta_5 = \theta_3^2 = \theta_1^3$, and so on.

Since equation (13) must hold for all nonzero weights, I have a $2m \times 2m$ system of equations to solve for θ. If $m = 3$ and the elements are the same as in Example 3, namely [0,1], [1,3], and [3,6], then the approximate solution satisfies

$$\begin{bmatrix} 8/15 & 11/15 & 0 & 0 & 0 & 0 \\ -3/5 & 2/5 & 4/5 & -7/30 & 0 & 0 \\ 0 & -8/15 & 16/15 & 4/5 & 0 & 0 \\ 0 & 1/10 & -8/15 & 2/3 & 13/15 & -4/15 \\ 0 & 0 & 0 & -7/15 & 8/5 & 13/15 \\ 0 & 0 & 0 & 1/15 & -7/15 & 9/10 \end{bmatrix} \begin{bmatrix} \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \\ \theta_6 \\ \theta_7 \end{bmatrix} = \begin{bmatrix} 3/5 \\ 2/15 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

In Figure 6, the approximate solution is plotted (the line marked "quadratic bases") along with the exact solution and the approximate solution with linear basis functions found in Example 3. Notice that the fit is better with quadratic functions. However, for this example, the number of unknowns doubled.



4. Solution of the Discrete-time, Deterministic Growth Model

We now turn to the deterministic growth model:

$$\max_{c_t} \sum_{t=0}^{\infty} \beta^t u(c_t) \qquad (14)$$
$$\text{subject to} \;\; c_t + k_{t+1} = f(k_t),$$

where c_t is consumption at date t, k_t is the capital stock at date t, $u(\cdot)$ is the utility function, $f(\cdot)$ is the production function, and $\beta < 1$ is a discount factor. If I solve the Euler equation, then the functional equation to be solved is given by

$$F(c)(k) = u'\left(c(k)\right) - \beta u'\left(c\left(f(k) - c(k)\right)\right) \, f'\left(f(k) - c(k)\right)$$

and our boundary condition is given by $c(0) = 0$.

Example 5. Let $u(c) = \ln(c)$ and $f(k) = \lambda k^\alpha$. In this case, the functional equation is

$$F(c)(k) = \frac{1}{c(k)} - \frac{\beta\alpha\lambda\left(\lambda k^\alpha - c(k)\right)^{\alpha-1}}{c\left(\lambda k^\alpha - c(k)\right)}$$

or, alternatively,

$$F(c)(k) = c\left(\lambda k^\alpha - c(k)\right) - \beta\alpha\lambda\left(\lambda k^\alpha - c(k)\right)^{\alpha-1} c(k).$$

The solution for consumption in this case is

$$c(k) = (1 - \beta\alpha)\lambda k^{\alpha}.$$

Suppose we want to obtain an approximate solution of the form

$$c^h(k;\theta) = \theta_1 k + \theta_2 k^2 + \ldots + \theta_n k^n$$

which satisfies the boundary condition at $k = 0$. The residual equation is therefore

$$R(k;\theta) = \sum_{i=1}^{n} \theta_i (\lambda k^{\alpha} - \sum_{j=1}^{n} \theta_j k^j)^i - \beta\alpha\lambda(\lambda k^{\alpha} - \sum_{j=1}^{n} \theta_j k^j)^{\alpha-1} \sum_{j=1}^{n} \theta_j k^j.$$

To apply weighted residual methods, I have to compute integrals of the form

$$\int_0^{\bar{k}} \phi_i(k) R(k;\theta)\, dk, \quad i = 1, \ldots, n. \tag{15}$$

where $\bar{k}$ is the upper bound of the domain for the capital stock. Since the residual R is a nonlinear function of θ, it makes sense to do numerical integration. If I apply Gaussian quadrature (which is typically done), then (15) is replaced by

$$\sum_{l} \omega_l \phi_i(k_l) R(k_l;\theta), \quad i = 1, \ldots, n \tag{16}$$

where ω_l are called the quadrature weights and the grid points k_l are called the quadrature abscissas. (See Press, et al. 1986 for the quadrature formulas and a description of how they are derived.) Depending on the specific quadrature rule (e.g., Legendre, Chebyshev, etc.) that is used, ω_l's and k_l's will differ, but the calculations of R and ϕ will look the same no matter what quadrature rule is used.

The final step is to solve the system of equations in (16). In this case, the system is nonlinear; the problem is to find θ such that $G(\theta) = 0$, where G has the same dimension as θ. Applying Newton's method to $G(\theta) = 0$ means iterating on

$$\theta^{j+1} = \theta^j - \left[\left. \frac{\partial G(\theta)}{\partial \theta} \right|_{\theta=\theta^j} \right]^{-1} G(\theta^j), \quad j = 1, 2, \ldots$$

with some initial guess θ^0, where θ^j is the vector of unknown coefficients at the jth iteration. Notice that as we iterate, we solve a sequence of problems of the following form: find θ such that $A\theta = b$, where A is the Jacobian matrix $\partial G/\partial\theta$ evaluated at θ^j and b is the function itself, $G(\theta^j)$.

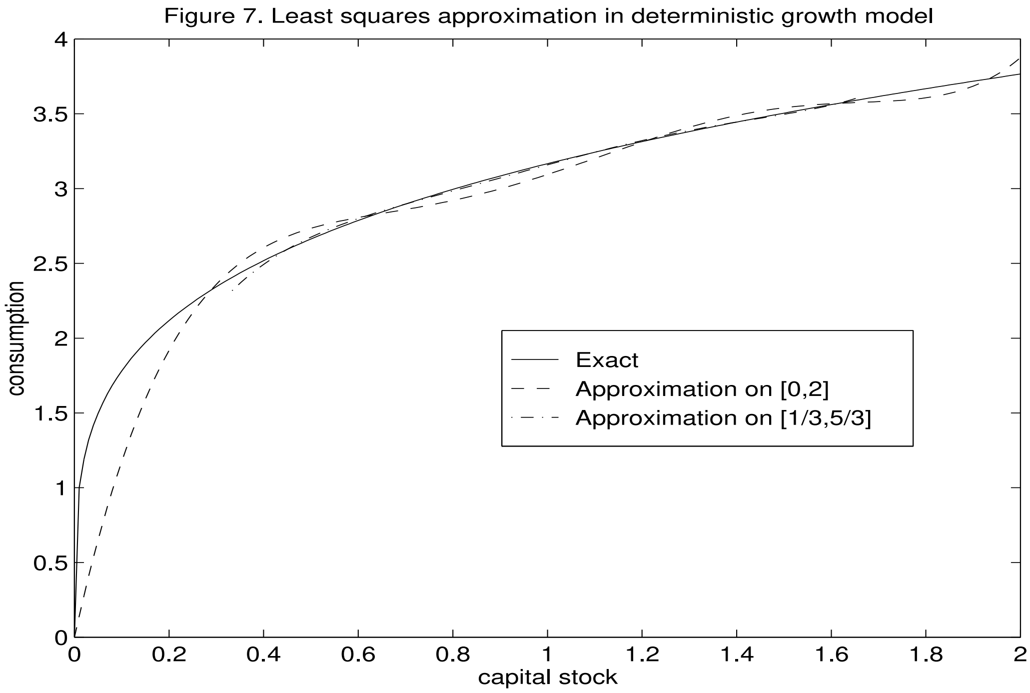
For the three weighted residual applications below, assume that $\alpha = 0.25$, $\beta = 0.96$, $\lambda = 1/(\alpha\beta)$, and $\bar{k} = 2$. For this set of parameters, the steady state capital stock is equal to 1. Assume also that the quadrature rule is Legendre with 20 quadrature abscissas used to approximate the integral in (15).

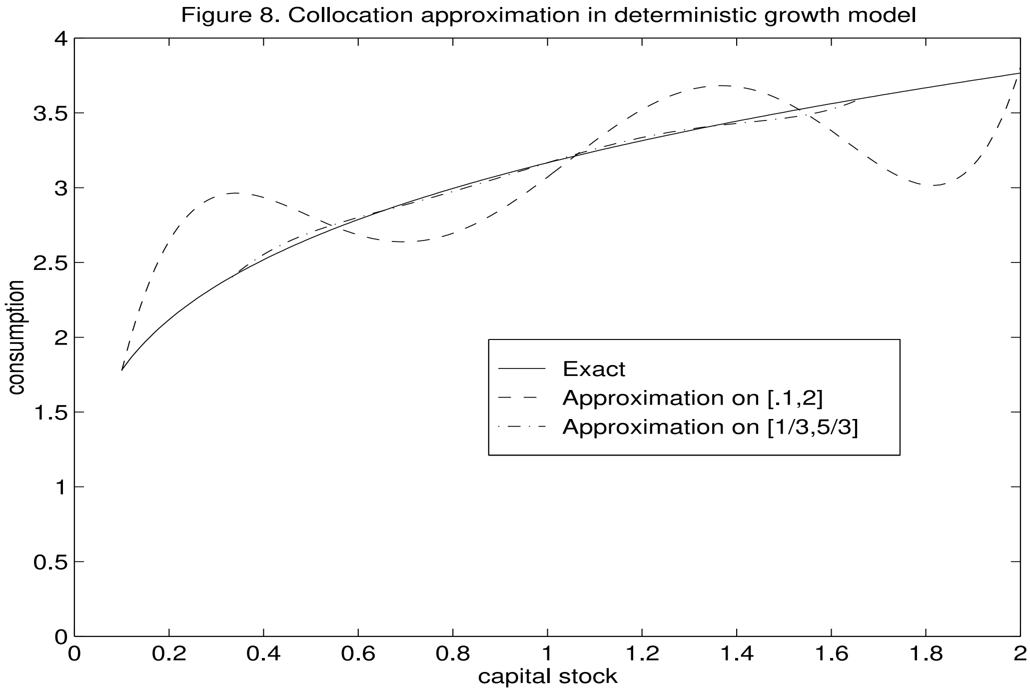
5a. Least squares. To apply least squares, I set $\phi_i(k) = \partial R(k;\theta)/\partial\theta_i$, where the derivative of the residual is given by

$$\frac{\partial R(k;\theta)}{\partial\theta_l} = -k^l \left\{ \sum_{i=1}^{n} i\theta_i \tilde{k}^{i-1} - \beta\alpha\lambda(\alpha-1)\tilde{k}^{\alpha-2} \sum_{j=1}^{n} \theta_j k^j + \beta\alpha\lambda\tilde{k}^{\alpha-1} \right\} + \tilde{k}^l,$$

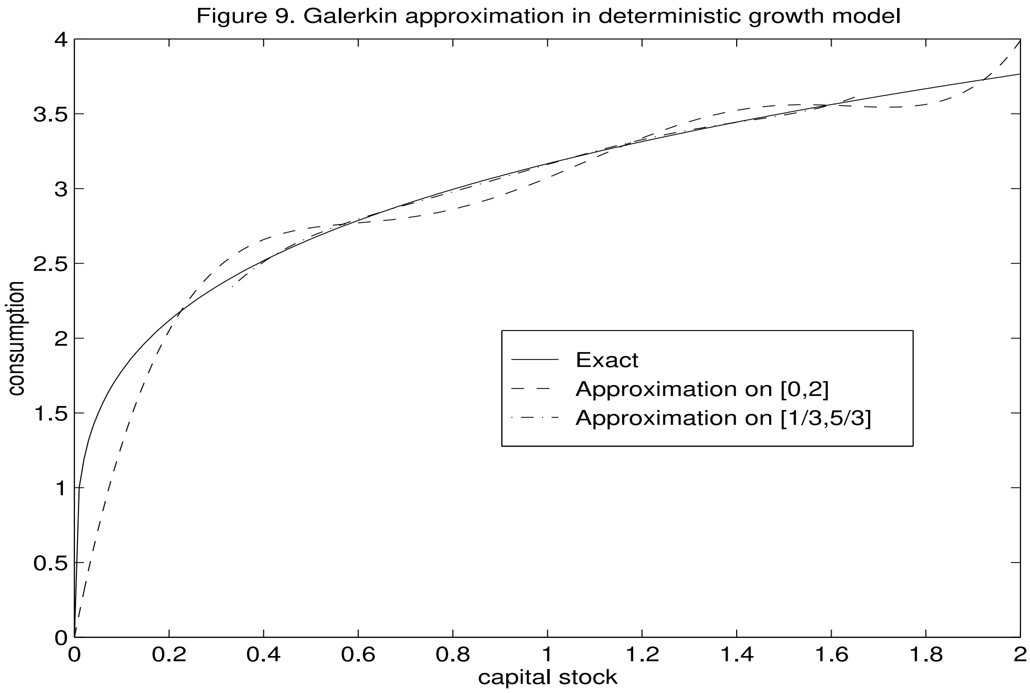
$l = 1, \ldots, n$, and $\tilde{k} = \lambda k^\alpha + \sum_j \theta_j k^j$. In Figure 7, I have plotted the approximate solution c^h for $n = 5$ along with the exact solution. Since the derivative of the true function is infinite at $k = 0$ and relatively small for high values of k, it is necessary to add more polynomials in order to completely resolve the solution at all capital stocks. I have also plotted the result for a more restricted grid on the capital stocks, namely [1/3,5/3]. This is the grid that Judd (1992) uses when evaluating weighted residual methods for the deterministic growth model. For both approximations, I have assumed that $n = 5$. Notice that although the approximation on [1/3,5/3] is very close to the true solution, the exact solution is very smooth – almost linear.

5b. Collocation. To apply the collocation method, I set $\phi_i(k) = \delta(k - k_i)$, where k_i, $i = 1, \ldots, n$ are collocation points in $[0,\bar{k}]$. In Figure 8, I have plotted two approximations: one with 5 evenly spaced collocation points between 0.1 and 2 and one with 5 evenly spaced collocation points between 1/3 and 2/3. The problem of fitting functions with steep gradients becomes acute in this case, which is why I avoided the region of capital





stocks below 0.1. Even so, the approximation on [0.1,2] is not very accurate. It is clear that some better choices for basis functions and collocation points are needed to make this method competitive with least squares. On [1/3,2/3], I find that the approximation is not quite as good as that for least squares but is not too different from the exact solution. Here again, the fit is good because the exact solution is very smooth on [1/3,2/3].



5c. Galerkin. To apply the Galerkin method, I set $\phi_i(k) = k^i$, $i = 1, \ldots, n$. In Figure 9, I have plotted approximate functions on [0,2] and [1/3,2/3] along with the exact solution. The results here are similar to the least squares method.

Because we need to include more polynomials, which, in the case of k^i, $i = 1, \ldots, n$, become similar to each other as n gets large, it makes sense to use a class of orthogonal polynomials. Judd (1992) uses a representation for consumption of the form

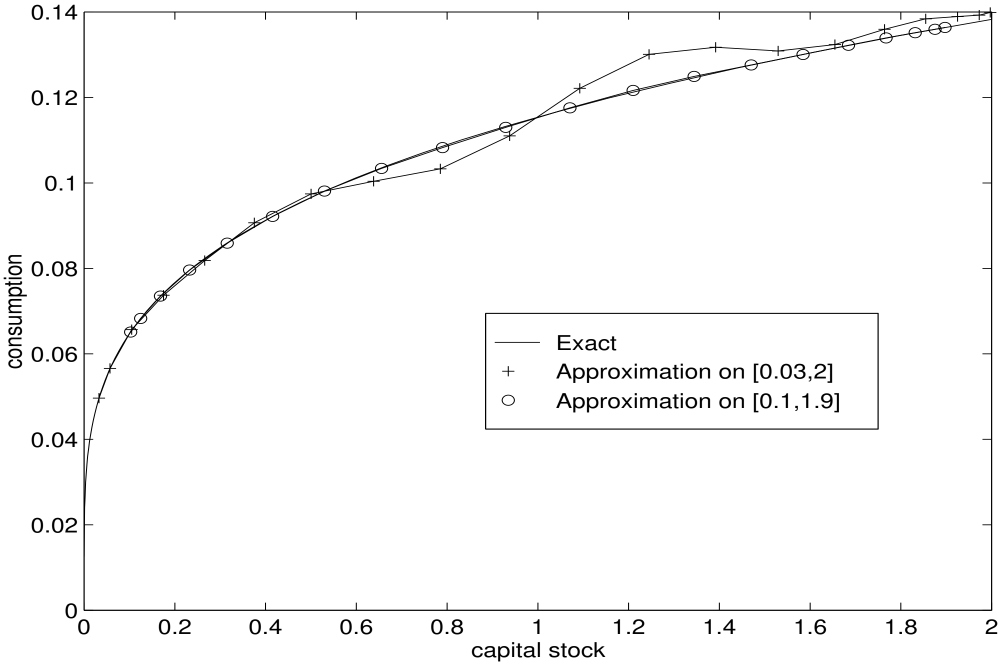
$$c^h(k;\theta) = \sum_{i=1}^{n} \theta_i \psi_i(k) \tag{17}$$

where $\psi_i(k) = p_{i-1}(2(k - \underline{k})/(\bar{k} - \underline{k}) - 1)$, $\underline{k}$ is a lower bound on the capital stocks, and $p_i(x)$ is the ith Chebyshev polynomial defined in (7).[3]

Example 6. In this case, assume that $u(c) = c^{1-\tau}/(1-\tau)$ and $f(k) = \lambda k^\alpha + (1-\delta)k$. Let $\tau = 5$, $\alpha = 0.25$, $\delta = 0.025$, $\beta = 0.99$, and $\lambda = (1 - \beta(1-\delta))/(\alpha\beta)$ (so that the steady state capital is equal to 1). Let c^h take the form of (17) with $n = 10$. In Figure 10, I have

[3] Note that this approximation would not satisfy the boundary condition at $c(0) = 0$ if $\underline{k} = 0$ for any θ. However, if we made a slight modification, namely, $\psi_i(k) = kp_{i-1}(2k/\bar{k} - 1)$ defined on $[0, \bar{k}]$, then the boundary condition is satisfied for all possible choices of θ.

Figure 10. Galerkin approximation with Chebyshev bases in the deterministic growth model



plotted the approximate solutions for $\underline{k} = 0.03$, $\bar{k} = 2$ (marked '+'), and $\underline{k} = 0.1$, $\bar{k} = 1.9$ (marked 'o') along with the exact solution.[4] The location of the points marked by '+' or 'o' are the quadrature abscissas.

It is clear from Figure 10 that more polynomials would be required for a good approximation on [0.03,2]. This is due to the fact that we are trying to approximate a very steep part of the function and a very flat part of the function using the same basis functions. When I restrict the domain to be [0.1,1.9], there is a significant improvement in the approximation over this region of the state space. The approximation is visually indistinguishable from the exact solution. In this restricted region of the domain, the function does not have any large gradients.

Suppose instead of using Chebyshev polynomials, we applied the Galerkin method with piecewise linear basis functions as is done for the finite element method.

Example 7. Assume that $u(\cdot)$, $f(\cdot)$, and the parameterization is the same as in Example

[4] What I am calling the 'exact solution' here is actually a finite element approximation with a large number of elements. Although this itself is an approximation, doubling the number of elements leaves Figure 10 unchanged.

6. Let $x_1 = 0$, x_{11}=2, and $x_i = x_{i-1} + 0.005 \exp(.574(i-2))$. This partition implies that there are 10 elements with lengths that increase exponentially. Thus, there will be more points near the origin where the function has a large (infinite in this case) gradient. To compute the weighted integral, I used a Legendre quadrature rule with two quadrature points per element. On an element of length ℓ_e, the Legendre quadrature rule with two quadrature points implies the following weights and abscissas for (16): $\omega_l = \ell_e/2$, $l = 1, 2$, and $k_1 = k_e + 0.211\ell_e$, $k_2 = k_e + 0.789\ell_e$, where k_e is the first endpoint of the element.

In Figure 11, I have plotted the finite element approximation along with the exact solution. Because the finite element method is a piecewise application of a weighted residual method, it is possible to get a more accurate approximation over the entire [0,2] domain – we are not using the same basis functions in the very steep region and the very flat region of the consumption function.

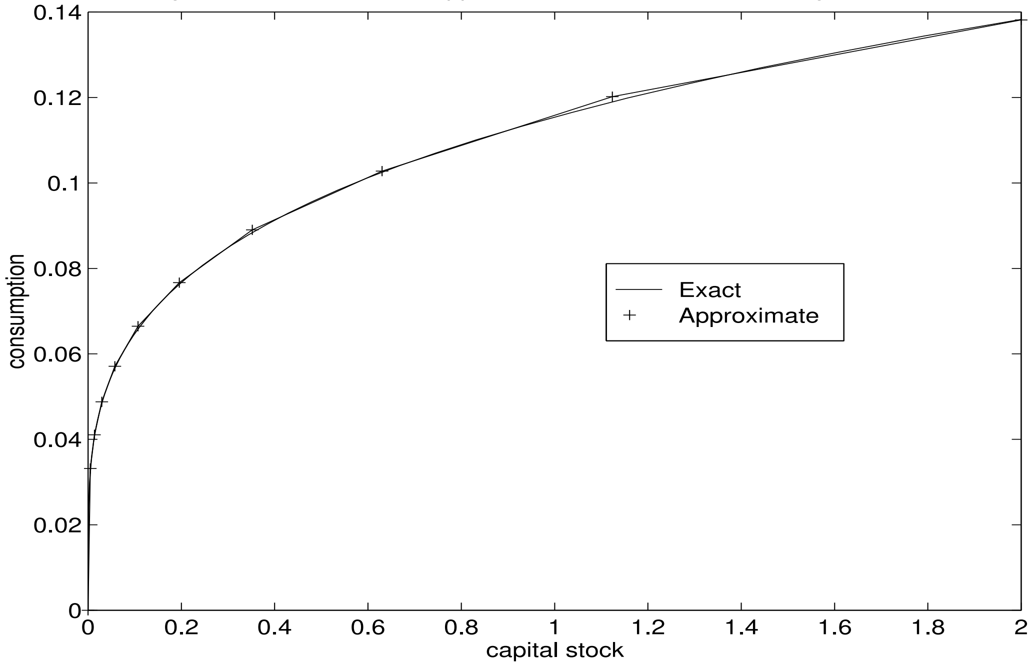
To obtain the approximation in Figure 11, the main computational task was the inversion of a 10×10 matrix. Sixty-eight elements of the 100 elements in this matrix were zeros and the structure of the matrix was band diagonal. As the number of unknowns becomes large, it becomes expensive and, in some cases, infeasible to invert the matrix without utilizing inversion routines that exploit the fact that the matrix is band diagonal.

5. Solution of the Discrete-time, Stochastic Growth Model[5]

Suppose instead of the deterministic growth model, we want to calculate the decision functions for the stochastic growth model where decisions depend on the capital stock and a stochastic shock. The stochastic growth model assumes that output in period t can be allocated either to current consumption c_t or to current investment i_t. The consumption-savings decision is assumed to be optimal in that the preferences of households are maxi-

[5] Interested readers should also consult Judd (1992) for more details on spectral methods as applied to this problem and McGrattan (1996) for more details on the finite element method as applied to this problem.

Figure 11. Finite element approximation in the deterministic growth model



mized. The preferences are given by

$$E\left[\sum_{t=0}^{\infty}\beta^t\, u(c_t)\,\middle|\, k_0\right], \qquad 0 < \beta < 1, \tag{18}$$

where k_t is the capital stock at t and k_0 is known. The maximization of (18) is done subject to the feasibility constraints

$$c_t + k_{t+1} - (1-\delta)k_t = \lambda_t k_t^{\alpha}, \qquad 0 < \alpha < 1,\ 0 \leq \delta \leq 1, \tag{19}$$

the nonnegativity constraints $c_t \geq 0$, $k_t \geq 0$ for all $t \geq 0$ and subject to the process for the technology shock,

$$\ln\lambda_t = \rho\ln\lambda_{t-1} + \varepsilon_t, \qquad -1 < \rho < 1, \tag{20}$$

where ε_t is a serially uncorrelated, normally distributed random variable with mean zero and variance σ^2. Since ε is normally distributed, it does not have a compact support. Let $z = \tanh(\ln(\lambda))$, which has a compact support [-1,1]. Then I can rewrite (20) as follows:

$$z_t = \tanh(\rho\tanh^{-1}(z_{t-1}) + \sqrt{2}\sigma\nu_t),$$

where $\nu_t = \varepsilon_t/(\sqrt{2}\sigma)$.

Since the stochastic shock takes on a continuum of values, we need to solve a two-dimensional problem. The representation of the approximate solution would then be

$$c^h(k, z; \theta) = \sum_{i=1}^{n} \theta_i \psi_i(k, z).$$

A simple set of basis functions would be all products of the elements of $\{1, k, k^2, \ldots, k^{n_k}\}$ and $\{1, z, z^2, \ldots, z^{n_z}\}$. Alternatively, one could use all products of the elements of two sets of orthogonal polynomials. In either case, however, the number of unknowns starts to add up quickly, especially if a large number of polynomials are needed to approximate consumption at both high and low values of the capital stock.

One way to keep the problem tractable is to use the set of complete polynomials rather than all products of terms in $\{k^i\}_{i=0}^{n_k}$ and $\{z^i\}_{i=0}^{n_z}$ (e.g., use bases $\{1, k, z, k^2, kz, z^2\}$ rather than $\{1, k, z, kz, k^2, z^2, k^2z, kz^2, k^2z^2\}$). Using the set of complete polynomials allows one to approximate higher order functions but limits the number of unknown coefficients.[6] A second way to keep the problem tractable is to apply a finite element method. As we saw in earlier examples, the system of equations to be solved for the unknown coefficients θ is typically very sparse. Therefore, for big problems, we do not need as much storage as a typical spectral method and we can apply algorithms for solving sparse systems of equations.

Consider application of the finite element method to the stochastic growth model. The first step to be taken is to write out the residual equation using the first order condition for the problem in (18):

$$R(k, z; \theta) = c^h(k, z; \theta)^{-\tau} - \frac{\beta}{\sqrt{\pi}} \int_{-\infty}^{\infty} c^h(\tilde{k}, \tilde{z}; \theta)^{-\tau} \left(\alpha \tilde{k}^{\alpha-1} \sqrt{\frac{1+\tilde{z}}{1-\tilde{z}}} + 1 - \delta\right) e^{-\nu^2}\, d\nu = 0, \quad (21)$$

where

$$\tilde{k} = k^\alpha \sqrt{(1+z)/(1-z)} + (1-\delta)k - c^h(k, z; \theta),$$

$$\tilde{z} = \tanh(\rho \tanh^{-1}(z) + \sqrt{2}\sigma\nu),$$

[6] See Judd (1992) for a comparison of complete polynomials and tensor products in the case of the stochastic growth model.

$c^h(0, z; \theta) = 0$, ν is distributed normally with mean zero and variance $1/2$, and the domain for the state space is $\Omega = [0, \bar{k}] \times [-1, 1]$. If I apply a Gauss-Hermite quadrature rule when computing the integral in (21), then the residual equation becomes

$$R(k, z; \theta) \simeq c^h(k, z; \theta)^{-\tau} - \frac{\beta}{\sqrt{\pi}} \sum_{l=1}^{m_\nu} c^h(\tilde{k}, \tilde{z}_l; \theta)^{-\tau} \Big(\alpha \tilde{k}^{\alpha-1} \sqrt{\frac{1+\tilde{z}_l}{1-\tilde{z}_l}} + 1 - \delta \Big) \omega_l, \tag{22}$$

where $\tilde{z}_l = \tanh(\rho \tanh^{-1}(z) + \sqrt{2}\sigma\nu_l)$ and ν_l, ω_l, $l = 1, \ldots, m_\nu$ are the abscissas and weights for an m_ν-point quadrature rule. (For the quadrature formulas, see Press, et al. (1986).)

The second step in applying the finite element method is to discretize the domain. In this problem, the domain is two dimensional: $\Omega = [0, \bar{k}] \times [-1, 1]$. Suppose we divide the domain into smaller rectangular subdomains which do not overlap.

I will consider two types of approximations over the rectangular elements: linear and quadratic. Suppose the representation for consumption on some element e is linear,

$$c_e^h(k, z) = a + bk + cz + dkz. \tag{23}$$

Since there are four unknowns, we require an element with four nodes. If we place the four nodes at the corners of the rectangularly shaped element, then we can uniquely define the geometry of the element, and we can use the values of the solution at the four nodes to pin down the constants in (23). That is, as in the one-dimensional case, we can rewrite the approximation in (23) so that $c_e^h(k, z; \theta) = \sum_i \theta_i^e \psi_i^e(k, z)$, $i = 1, \ldots, 4$, where the basis functions are such that ψ_i^e is 1 at node i and zero at the other three nodes on the element.

Before giving formulas for the basis functions, it is convenient to first consider a mapping from global coordinates (k, z) to local coordinates (ξ, η) defined on a master element. This is done for convenience, since the master element has a fixed set of coordinates, while each element in Ω has a different set of coordinates. Thus, we can construct basis functions once but use them for each element. Consider functions $\xi(k)$

and $\eta(z)$ that map a typical element $[k_i, k_{i+1}] \times [z_j, z_{j+1}]$ to the square $[-1,1] \times [-1,1]$; i.e., $\xi(k) = (2k - k_i - k_{i+1})/(k_{i+1} - k_i)$ and $\eta(z) = (2z - z_j - z_{j+1})/(z_{j+1} - z_j)$. Assume that the four nodes of the master element are $(-1,-1)$, $(1,-1)$, $(1,1)$, and $(-1,1)$ using the local coordinates. In this case, the basis functions are constructed so that $c_e^h(\xi, \eta; \theta) = \sum_i \theta_i^e \psi_i^e(\xi, \eta)$ with $\theta_1^e = c_e^h(-1,-1;\theta)$, $\theta_2^e = c_e^h(1,-1;\theta)$, $\theta_3^e = c_e^h(1,1;\theta)$, and $\theta_4^e = c_e^h(-1,1;\theta)$. These restrictions imply that

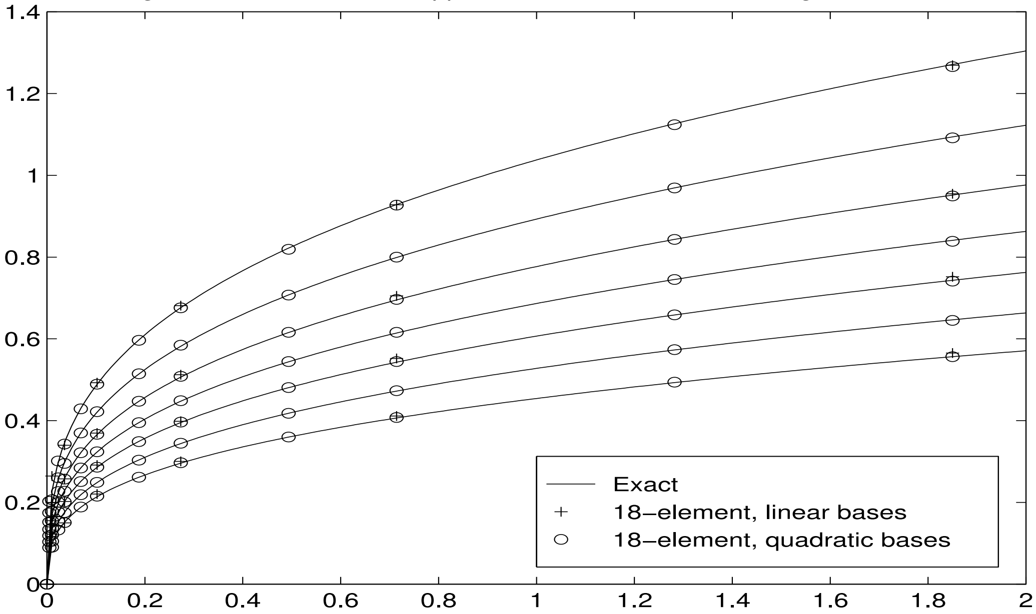
$$\begin{aligned} c_e^h(\xi, \eta; \theta) = & \frac{1}{4}(1-\xi)(1-\eta)\,\theta_1^e + \frac{1}{4}(1+\xi)(1-\eta)\,\theta_2^e \\ & + \frac{1}{4}(1+\xi)(1+\eta)\,\theta_3^e + \frac{1}{4}(1-\xi)(1+\eta)\,\theta_4^e. \end{aligned} \tag{24}$$

To attain a more accurate approximation, one can increase the number of elements while retaining linear basis functions or use higher-order polynomials. Consider, for example, quadratic functions in two dimensions. One simple way to construct these functions would be to take the product of one-dimensional quadratic polynomials. A unique set of coefficients for the polynomial requires that there be nine nodes and, hence, nine interpolation functions. In this case, the approximation on the master element $[-1,1] \times [-1,1]$ is given by

$$\begin{aligned} c_e^h(\xi, \eta; \theta) = & \frac{1}{4}\xi(\xi-1)\eta(\eta-1)\,\theta_1^e + \frac{1}{4}\xi(\xi+1)\eta(\eta-1)\,\theta_2^e + \frac{1}{4}\xi(\xi+1)\eta(\eta+1)\,\theta_3^e \\ & + \frac{1}{4}\xi(\xi-1)\eta(\eta+1)\,\theta_4^e + \frac{1}{2}(1-\xi^2)\eta(\eta-1)\,\theta_5^e + \frac{1}{2}\xi(\xi+1)(1-\eta^2)\,\theta_6^e \\ & + \frac{1}{2}(1-\xi^2)\eta(\eta+1)\,\theta_7^e + \frac{1}{2}\xi(\xi-1)(1-\eta^2)\,\theta_8^e + (1-\xi^2)(1-\eta^2)\,\theta_9^e. \end{aligned} \tag{25}$$

Example 8. Let $\tau = 1$, $\delta = 0$, $\beta = 0.95$, $\alpha = 0.33$, $\rho = 0.95$, and $\sigma = 0.1$. Assume that the partition on z is given by [−0.391, −0.123, 0.123, 0.391] and that the partition on k is given by [0, 0.010, 0.036, 0.102, 0.273, 0.714, 1.85]. I set the number of quadrature points on each element equal to 9, i.e., 3 points for integration with respect to the capital stock and 3 points for integration with respect to the technology shock. For integration over ν, I set the number of quadrature points, m_ν, equal to 10.

Figure 12. Finite element approximations in the stochastic growth model



In Figure 12, I have plotted the approximate piecewise linear solution (marked '+') along with the exact solution. Even though there are only 18 elements, it is hard to distinguish the two.

Example 9. Suppose we use the same parameterization as Example 8, but instead of linear basis functions, we use the quadratic functions in (25). In Figure 12, the solution is marked by 'o'. Notice that the fit is slightly better – however, since the coarse piecewise linear approximation was very accurate, there was not much room for improvement.

6. Conclusions

In this paper, I have applied weighted residual methods to several standard examples. I have applied both spectral methods and finite element methods.

One advantage of spectral methods is that they are very simple to implement once the choice of bases and weights has been made. The disadvantages are most obvious when we move away from simple one-dimensional linear problems to higher-order problems and highly nonlinear problems. For example, when we consider bigger problems, the system of equations to be solved (the $A\theta = b$ that we solve for our parameters θ) is dense with spectral methods, since each element of θ will in general affect every residual equation. That is not to say that you do not encounter the curse of dimensionality with the finite element method. You do. But with the finite element method, A is typically sparse and its structure can be exploited. When we have highly nonlinear problems, we also need to abandon the simpler spectral methods. The finite element method, on the other hand, can handle large gradients in the functions. Also, given the fact that the finite element solution is piecewise, it has no problem with kinks in the policy or value functions – say, because of inequality constraints binding.

In summary, it is probably best to use spectral methods when first starting out. They will be much faster than many standard methods (see Taylor and Uhlig (1990)). However, when the problems get too big or too nonlinear, it may be necessary to turn to the finite element method.

REFERENCES

Given that weighted residual and finite element methods have been applied for three or more decades, it is not hard to find textbooks on the subject. Since the finite element method has wider use, it is more likely that you will find discussions of weighted residual methods (or projection methods as they are sometimes called) in the first few chapters of an introductory text on finite element methods. Below I give references for a small subset of books, papers, and software relevant to the subject. These are materials that I have found useful.

Background reading

[1] Burden, Richard L. and J. Douglas Faires. 1989. *Numerical Analysis.* Boston: PWS-Kent.

(It is always helpful to have one or two undergraduate-level numerical analysis books on your shelf. This is one that I like.)

[2] Golub, Gene H. and James M. Ortega. 1992. *Scientific Computing and Differential Equations: An Introduction to Numerical Methods.* New York: Academic Press.

(This is an easy-to-read introduction to numerical methods for computing differential equations. There is some discussion of collocation and Galerkin weighted residual – which they refer to as projection methods – but no discussion of the finite element method.)

[3] Golub, Gene H. and Charles F. Van Loan. 1989. *Matrix Computations.* Baltimore: Johns Hopkins.

(This is always a handy reference for anyone doing computations that involve linear algebra.)

[4] Hughes, Thomas J.R. 1987. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis.* Englewood Cliffs: Prentice-Hall

(This book is useful only after you have looked at an introductory textbook on the finite element method. It goes into greater detail on solving three-dimensional problems – three is the maximum dimension that many engineers would ever consider. However, the key thing to extract from discussions of 3-D problems is how to choose element bases so as to avoid adding a lot of elements to the vector θ.)

[5] Judd, Kenneth. 1992. Projection methods for solving aggregate growth models, *Journal of Economic Theory* 58(2): 410-52.

(In this paper, Judd provides a detailed discussion of weighted residual methods – especially those that use Chebyshev polynomial bases.)

[6] McGrattan, Ellen. 1996. Solving the Stochastic Growth Model with a Finite Element Method. *Journal of Economic Dynamics and Control* 20: 19-42.

(In this paper, I apply standard finite element techniques to solving the stochastic growth model. I did this for my class project with Yousef Saad at the University of Minnesota. The title reflects the fact that I tried to give an alternative to methods described in Taylor and Uhlig 1990.)

[7] Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. 1986. *Numerical recipes: The art of scientific computing.* Cambridge: Cambridge University Press.

(This is a useful book to have amongst your references – although it gives very little details about the methods I described here.)

[8] Reddy, J.N. 1993. *An introduction to the finite element method.* New York: McGraw-Hill.

(This book is an easy-to-read introduction to the finite element method. It gives a good summary of weighted residual methods in Chapter 2 and some history of the use of these methods.)

[9] Saad, Yousef. 1996. *Iterative Methods for Sparse Linear Systems.* Boston: PWS.

(This is the book that was written from Saad's class notes at the University of Minnesota. I used them in note form when I took his class and found them to be extremely useful. And presumably the book will be an improvement over the notes – fewer typos.)

[10] Taylor, John B. and Harald Uhlig. 1990. Solving nonlinear stochastic growth models: A comparison of alternative solution methods. *Journal of Business and Economic Statistics* 8: 1-17.

(This paper, along with the others that accompany it in the JBES 1990 issue, serves as a nice survey of methods that have been applied to solving a standard problem in economics.)

Software

[11] IMSL

(This is a package of FORTRAN routines for purchase. For those who like having a slick manual and well-tested algorithms for doing basic operations, this is the package for you.)

[12] LAPACK

(This is a set of well-tested FORTRAN routines available for free. For those who don't have any dough, this is the package for you. I started using these routines – previously LINPACK – as a graduate student and never felt the need to change to IMSL. There is a manual available – see the 'readme' files for details.)

[13] SPARSKIT2

(This package is used for solving large, sparse linear systems. It was written by Yousef Saad and accompanies his book. I took his class that covered the material in the book and have found the routines in SPARSKIT2 to be very useful.)

Specific applications

[14] Aiyagari, S. Rao and Ellen R. McGrattan. 1996. The optimal quantity of debt. Staff Report #203, Federal Reserve Bank of Minneapolis.

(In this paper, individuals hit borrowing constraints, which implies that nonlinearities in their decision functions can be important. There is a technical appendix that accompanies the paper that describes in gory detail how to apply the finite element method and how it does for some test cases.)

[15] Braun, Richard A. and Ellen R. McGrattan. 1993. The macroeconomics of war and peace. *NBER Macroeconomics Annual 1993*. Cambridge: MIT Press.

(In this paper, we tried to model aggregate changes in output and employment during World War II. Because the shocks in our model are large, we applied a finite element method to solve the nonlinear decision functions. Linear approximations are not valid for this problem.)

[16] Chari, V.V., Patrick Kehoe, and Ellen R. McGrattan. 1996. Sticky price models of the business cycle: Can the contract multiplier solve the persistence problem? Staff Report #217, Federal Reserve Bank of Minneapolis.

(In this paper, the finite element method was used to test whether or not our approximation was good. For many of the examples, however, the state space is simply too big to apply the FEM in the standard way. So we compared the linear and nonlinear solutions for a subset of cases and found that the linear solution was very accurate.)

[17] Chari, V.V., Patrick Kehoe, and Ellen R. McGrattan. 1997. The poverty of nations: A quantitative investigation. Staff Report #204, Federal Reserve Bank of Minneapolis.

(In this paper, the model that we use is a variant of the stochastic growth model. Because we assume that countries can at some time period have very little capital and in other time periods have a lot of capital, it is necessary to solve the nonlinear decision functions. We apply a finite element method to do so.)