1 Week 12: 13 November 2003

Today we will begin the numerical solution of two–point boundary value problems (BVPs)

$$\begin{aligned} y' &= f(t,y) \qquad a < t < b \\ g(y(a),y(b)) &= 0 \end{aligned}$$

where $y : [a,b] \to \mathbf{R}^d$ is a d–dimensional vector function. Classical examples include the standard linear second–order equations of mathematical physics (transformed to first–order systems)

$$a(t)y'' + b(t)y' + c(t)y = f(t), \qquad 0 < t < \infty$$

subject to linear boundary conditions such as

$$y(0) = 0, \qquad y(t) \quad \text{is bounded as} \quad t \to \infty,$$

or the linear system in standard separated form

$$\begin{aligned} y' &= R(t)y(t) + f(t) \qquad a < t < b, \\ Ay(a) + By(b) &= g \end{aligned}$$

where $y(t)$ is a d–dimensional vector. Useful references for this subject include [AMR95, Kel92, SR94].

2 Theory

The existence, uniqueness and stability theory of the BVP is different from corresponding theories for the IVP, because the boundary conditions impose global requirements on the solution. Thus standard IVP conditions on the niceness of f guarantee nothing about the BVP. Thus even linear BVPs resemble linear algebraic eigenvalue problems $Ax = \lambda x$ more than they do IVPs or linear systems of equations $Ax = b$; the theory is therefore more complicated. A single condition on A guarantees that $Ax = b$ can always be uniquely solved for any b, while $Ax = \lambda x$ has no solution for most λ and the solution is not unique even when it does exist. For example,

$$y'' = 1$$

always has a unique solution $y_0 + y_1 t + t^2/2$ satisfying boundary conditions

$$y(a) = \alpha, \quad y(b) = \beta,$$

but

$$y'' + y = 1, \qquad y(a) = y(b) = 0$$

has no solution at all if the length of the interval is a multiple of π.

Thus we will not present a general theory of existence, uniqueness and stability for BVPs as we did for IVPs. Instead, we will develop methods to compute solutions and check a posteriori when possible.

3 Quasilinearization

Most methods for the general nonlinear BVP rely on one of two paths: solve a sequence of linearized BVPs by numerical methods ("quasilinearization"), or discretize the BVP directly and solve a large system of nonlinear equations. Both approaches have many followers, and often both arrive at the same final equations. We focus mainly on the first approach, because it gives us more scope for analysis of the BVP and thus tends to produce more effective numerical methods.

The nonlinear BVP can be linearized by assuming an approximate solution u and perturbing it to become the exact solution $y = u + e$. Then e must satisfy

$$\begin{aligned} u' + e' &= f(t, u + e) = f(t, u) + Df(t, u)e + O(e^2) \\ g(u(a) + e(a), u(b) + e(b)) &= g(u(a), u(b)) + Ae(a) + Be(b) + O(e^2) = 0 \end{aligned}$$

where

$$A = \frac{\partial g}{\partial u(a)} \qquad \text{and} \qquad B = \frac{\partial g}{\partial u(b)}.$$

If u is a good approximation, we can neglect quadratic terms in e and solve

$$\begin{aligned} e' &= Df(t, u)e + f(t, u) - u' \\ Ae(a) + Be(b) &= -g(u(a), u(b)) \end{aligned}$$

for the update e, improve $u \leftarrow u + e$, and repeat until satisfied. It is not a coincidence that this linearized problem has exactly the form of Newton's method, with a residual on the right–hand side of each equation and the Jacobian operator acting on the correction e. Indeed, quasilinearization is simply Newton's method in function space. Like the finite–dimensional Newton's method, it can be improved and modified in various ways for more assured convergence. For example, starting values can be obtained by homotopy methods, Broyden-type methods with fewer Jacobians can be used, and we can move a different distance along the Newton direction to guarantee that the residual will be reduced at each Newton step.

4 Linear Theory

We can now concentrate on numerical methods for the linear two–point BVP

$$\begin{aligned} y' &= P(t)y + f(t) \\ Ay(a) + By(b) &= g \end{aligned}$$

where $P(t)$ is a matrix function, f is a vector function, A and B are n by n matrices and g is a constant n–vector.

Most such methods are built on the "shooting" viewpoint which constructs the BVP solution by solving an equation for $y(a)$ and then solving an IVP for y by any of the methods previously discussed in this course. To do this, we need to represent the solution of the IVP

$$\begin{aligned} y' &= P(t)y + f(t) \\ Ay(a) + By(b) &= g \end{aligned}$$

as a functional of $y(a)$. Thus we introduce the fundamental matrix and solve the IVP by variation of parameters.

Suppose first that $P(t) = P$ is constant and $f = 0$. Then the homogeneous IVP

$$y' = Py, \qquad y(a) \quad \text{given}$$

is solved by the matrix exponential

$$y(t) = e^{tP} e^{-aP} y(a)$$

where the matrix exponential is defined by the convergent power series

$$e^{tP} = \sum_{n=0}^{\infty} \frac{(tP)^n}{n!}.$$

(The series definition is bad for computation in most cases, and there are many alternative formulas some of which are numerically useful [ML78].) The inhomogeneous problem

$$y' = Py + f, \qquad y(a) \quad \text{given}$$

then succumbs to the variation of parameters formula

$$y(t) = e^{(t-a)P} y(a) + \int_a^t e^{(t-s)P} f(s) ds.$$

Example The second–order linear constant–coefficient IVP

$$\begin{aligned} y'' &= f \\ y(a) &= y_0 \\ y'(a) &= y_0' \end{aligned}$$

can be transformed into the first–order system

$$\begin{aligned} y' &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} y + \begin{bmatrix} 0 \\ f \end{bmatrix} = Py + f \\ y(a) &= \begin{bmatrix} y_0 \\ y_0' \end{bmatrix}. \end{aligned} \tag{1}$$

Since $P^2 = 0$, the series for the matrix exponential terminates and we get

$$e^{tP} = I + tP.$$

Thus the inhomogeneous IVP solution is

$$y(t) = (I + tP)(I + aP)^{-1} y(a) + \int_a^t (I + (t-s)P) f(s) ds.$$

Exercise 1 *Convert the fourth-order scalar linear inhomogeneous constant-coefficient IVP*

$$y'''' - k^4 y = q(t),$$

$$y(0) = y_0, \quad y'(0) = y_0', \quad y''(0) = y_0''(0), \quad y'''(0) = y_0'''(0)$$

into a first-order system $u' = Pu + f(t)$*, diagonalize the matrix* P*, and thus find the general solution* $u(t)$ *of the inhomogeneous IVP.*

For nonconstant matrix coefficients $P(t)$, the natural matrix exponential formula

$$e^{\int_a^t P(s)ds} y(a)$$

immediately fails, essentially because $P(t)$ no longer commutes with $P(s)$ for $t \neq s$. The derivative of the exponential series contains products in all possible orders of $P(t)$ with $\int_a^t P(s)ds$ which do not commute, so terms can no longer be collected to yield $P(t)e^{\int_a^t P(s)ds}$.

Instead, we introduce an analytical tool called the *fundamental matrix* $Y(t)$ which satisfies

$$\begin{aligned} Y'(t) &= P(t)Y(t) \\ Y(a) & \quad \text{is nonsingular.} \end{aligned}$$

In terms of Y, the solution of the homogeneous IVP is

$$y(t) = Y(t)Y(a)^{-1}y(a)$$

and the inhomogeneous IVP is solved by the variation of parameters formula

$$y(t) = Y(t)Y(a)^{-1}y(a) + \int_a^t Y(t)Y(s)^{-1}f(s)ds$$

where we have replaced e^{tP} by $Y(t)$. It turns out that if $Y(a)$ is nonsingular and P is continuous, then $Y(t)$ is nonsingular for any t. This follows from Abel's formula

$$\det Y(t) = e^{\int_a^t \mathrm{tr} P(s)ds} \det Y(a)$$

where $\mathrm{tr}A = \sum_i A_{ii}$ is the trace of a square matrix A.

Computing a fundamental matrix Y for a given IVP is of course much more difficult than computing the solution y directly, since it requires n IVP solves, one for each column of Y, so it rarely constitutes a useful numerical object. But computing the result of applying the fundamental matrix to a vector is equivalent to solving an IVP, so Y can be used as a conveniently organized shorthand which represents the solution y as an explicit linear function of the initial values $y(a)$. Similarly, when solving a linear system $Ax = b$, the inverse matrix A^{-1} provides a useful theoretical representation of the solution $x = A^{-1}b$ but not often a practical computational tool.

Thus representing the BVP solution in terms of its unknown initial values by the variation of parameters formula gives a n by n linear system to be solved for the initial values, after which y is easily computed by solving an IVP: indeed, the boundary conditions require

$$\begin{aligned} Ay(a) + By(b) &= Ay(a) + B(Y(b)Y(a)^{-1}y(a) + \int_a^b Y(b)Y(s)^{-1}f(s)ds) \\ &= (A + BY(b)Y(a)^{-1})y(a) + B\int_a^b Y(b)Y(s)^{-1}f(s)ds \\ &= g \end{aligned}$$

or

$$(AY(a) + BY(b))Y(a)^{-1}y(a) = g - B\int_a^b Y(b)Y(s)^{-1}f(s)ds.$$

Thus we see that solvability of the BVP depends on the n by n "boundary condition matrix"

$$D = AY(a) + BY(b);$$

the BVP is solvable for arbitrary right–hand sides and boundary values iff D is invertible.

Exercise 2 *Find a fundamental matrix $Y(t)$ and a boundary condition matrix D for the n–dimensional BVP*

$$\begin{aligned} y' &= \Lambda y + f(t) \\ Ay(a) + By(b) &= g \end{aligned}$$

where Λ is a constant diagonal matrix, and find a condition on Λ, A and B which is equivalent to the invertibility of D. Show that the BVP is uniquely solvable for arbitrary data (f, g) iff D is invertible.

Supposing from now on that D is invertible, we can put the solution into the "Green function" form

$$y(t) = Y(t)D^{-1}g + \int_a^b G(t,s)f(s)ds$$

by straightforward manipulation. This form is analogous to writing the solution of a linear system $Ax = b$ via the inverse matrix $x = A^{-1}b$; the Green function plays the role of the matrix entries $(A^{-1})_{ij}$ and summation is replaced by integration.

The calculation goes as follows:

$$\begin{aligned} y(t) &= Y(t)Y(a)^{-1}y(a) + \int_a^t Y(t)Y(s)^{-1}f(s)ds \\ &= Y(t)D^{-1}(g - B\int_a^b Y(b)Y(s)^{-1}f(s)ds) \\ &\quad + \int_a^t Y(t)Y(s)^{-1}f(s)ds \\ &= Y(t)D^{-1}g + \int_a^t Y(t)(I - D^{-1}BY(b))Y(s)^{-1}f(s)ds \end{aligned}$$

$$
\begin{aligned}
& - \int_t^b Y(t) D^{-1} B Y(b) Y(s)^{-1} f(s) ds \\
= \; & Y(t) D^{-1} g + \int_a^t Y(t) D^{-1} A Y(a) Y(s)^{-1} f(s) ds \\
& - \int_t^b Y(t) D^{-1} B Y(b) Y(s)^{-1} f(s) ds.
\end{aligned}
$$

Thus if we define the Green function by

$$
G(t,s) = \begin{cases} Y(t) D^{-1} A Y(a) Y(s)^{-1} & a < s < t < b \\ -Y(t) D^{-1} B Y(b) Y(s)^{-1} & a < t < s < b \end{cases}
$$

then we have the Green function formula above.

References

[AMR95] U. M. Ascher, R. M. M. Mattheij, and R. D. Russell. *Numerical solution of boundary value problems for ordinary differential equations.* SIAM, Philadelphia, 1995.

[Kel92] H. B. Keller. *Numerical Methods for Two–Point Boundary–Value Problems.* Dover, 1992.

[ML78] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix. *SIAM Rev.*, 20:801–836, 1978.

[SR94] Page Starr and Vladimir Rokhlin. On the numerical solution of two-point boundary value problems. II. *Comm. Pure Appl. Math.*, 47(8):1117–1159, 1994.