**Scientific Computing and Optimisation**

Matrix-vector formulations of finite-difference methods

Matthew Hennessy

Consider the 1D linear Poisson equation

$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} + q(x) = 0 \tag{1}$$

with the Dirichlet boundary conditions

$$u(a) = \alpha, \quad u(b) = \beta. \tag{2}$$

In the videos and demos, we've shown two ways of solving this problem using finite differences. The first is to discretise the system to arrive at an algebraic system of the form $\boldsymbol{F}(\boldsymbol{u}) = \boldsymbol{0}$, which is then solved using SciPy's `root` function. Alternatively, the algebraic system can be written in matrix-vector form as $\mathbf{A}^{DD}\boldsymbol{u} = -\boldsymbol{b}^{DD} - (\Delta x)^2\boldsymbol{q}$ and solved with NumPy. In these notes, we provide more explanation of how to formulate the algebraic system in matrix-vector form.

Suppose that we want to solve (1) using the finite-difference method. We let $x_i = a + i\Delta x$ be the set of grid points, where $\Delta x = (b-a)/N$ and $i = 0, 1, \ldots, N$. The solution approximation at $x_i$ is given by $u_i$ and the source term at $x_i$ is $q_i = q(x_i)$. After discretising (1) using the central difference formula, the algebraic system that must be solved is given by

$$u_0 = \alpha, \tag{3a}$$

$$\frac{u_{i+1} - 2u_i + u_{i-1}}{(\Delta x)^2} + q_i, \quad i = 1, 2, \ldots, N-1, \tag{3b}$$

$$u_N = \beta. \tag{3c}$$

This is a system of $N+1$ equations for $N+1$ unknowns. We can reduce the size of the system by eliminating $u_0$ and $u_N$ to obtain

$$\frac{u_2 - 2u_1 + \alpha}{(\Delta x)^2} + q_1 = 0, \tag{4a}$$

$$\frac{u_{i+1} - 2u_i + u_{i-1}}{(\Delta x)^2} + q_i = 0, \quad i = 2, \ldots, N-2, \tag{4b}$$

$$\frac{\beta - 2u_{N-1} + u_{N-2}}{(\Delta x)^2} + q_{N-1} = 0. \tag{4c}$$

This is a system of $N-1$ unknowns for the solution components $u_i$ at the $N-1$ interior grid points. To see how to convert (4) into a matrix-vector system, let's assume that $N = 5$ so there are 4 interior grid points and hence 4 unknowns. Then (4) becomes, after some re-ordering of terms,

$$-2u_1 + u_2 = -\alpha - (\Delta x)^2 q_1, \tag{5a}$$

$$u_1 - 2u_2 + u_3 = -0 - (\Delta x)^2 q_2, \tag{5b}$$

$$u_2 - 2u_3 + u_4 = -0 - (\Delta x)^2 q_3, \tag{5c}$$

$$u_3 - 2u_4 = -\beta - (\Delta x)^2 q_4. \tag{5d}$$

This is now a $4\times 4$ linear system of algebraic equations. Any linear system of algebraic equations can be written in matrix-vector form. We group the solution components $u_1$, $u_2$, $u_3$, and $u_4$ into a vector $\boldsymbol{u}$ such that $\boldsymbol{u} = (u_1, u_2, u_3, u_4)^T$. The matrix $\mathbf{A}^{DD}$ has entries $A_{jk}^{DD}$. The entry $A_{jk}^{DD}$ is equal to the coefficient of $u_k$ in the $j$-th algebraic equation. This means that rows of the matrix correspond to different algebraic equations, and columns of the matrix correspond to different solution components. For example, the second row of the matrix contains information about the second algebraic equation in (5), and the third column of the matrix contains information about the third solution component $u_3$. To be more specific, the second row of the matrix contains the coefficients of all of the $u_i$ in the second equation of (5), and the third column of the matrix contains all of the coefficients of $u_3$ across all of the equations. With this in mind, we can examine the left-hand side of (5) to find that the matrix $\mathbf{A}^{DD}$ is given by

$$\mathbf{A}^{DD} = \begin{pmatrix} -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \end{pmatrix}. \tag{6}$$

To check this is correct, we calculate $\mathbf{A}^{DD}\boldsymbol{u}$:

$$\mathbf{A}^{DD}\boldsymbol{u} = \begin{pmatrix} -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} -2u_1 + u_2 \\ u_1 - 2u_2 + u_3 \\ u_2 - 2u_3 + u_4 \\ u_3 - 2u_4 \end{pmatrix}, \tag{7}$$

which matches the left-hand side of (5). Therefore, we can re-write (5) as

$$\begin{pmatrix} -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = - \begin{pmatrix} \alpha \\ 0 \\ 0 \\ \beta \end{pmatrix} - (\Delta x)^2 \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix}. \tag{8}$$

By defining the boundary condition vector $\boldsymbol{b}^{DD} = (\alpha, 0, 0, \beta)^T$ and the source term vector $\boldsymbol{q} = (q_1, q_2, q_3, q_4)^T$, the linear system in (8) can finally be written as in matrix-vector form:

$$\mathbf{A}^{DD}\boldsymbol{u} = -\boldsymbol{b}^{DD} - (\Delta x)^2 \boldsymbol{q}. \tag{9}$$

Although the derivation here was for $N = 5$ and two Dirichlet boundary conditions, the same approach can be used to create matrix-vector systems for any number of grid points and any combination of boundary conditions.