Chapter 1.2: Linear Systems; Solving ODEs numerically

Linear Systems. Linear systems, which is one of the two main topics we will be concerned with in this class, are ubiquitous in all fields of science and engineering. They can be used to solve problems from optimization to ODEs (Ordinary Differential Equations), PDEs (Partial Differential Equations), model the airflow around the wing of an airplane, understand the stock market, etc.

Solving ODEs Numerically. One very common example of a problem where linear systems are central is finding the solution to ODEs. We will use this example to illustrate how a problem can be numerically solved by the use of linear algebra, even though it might not be clear a priori that it makes sense to do so.

Suppose one is trying to find a numerical solution to the equation

\[ u''(x) + a \cdot u'(x) + b \cdot u(x) = f(x), \] for all \( x \) in \([0, 1]\),

\( u(0) = u(1) = 0 \).

The second set of conditions, \( u(0) = u(1) = 0 \), represents the “boundary conditions”, and ensures that the solution is unique (and thus that the problem is well-defined). Generally, \( f \) is taken to be a “nice” function (for example, continuous on \([0, 1]\)). In the above, we will think of \( a \) and \( b \) as constants.

Thus, given \( a, b, \) and \( f \), one must produce \( u \), or rather one must produce \( u \) numerically. What we understand by that is the following: we must produce an approximation for \( u \) at a large set of points equidistributed in the interval \([0, 1]\). In other words, we must do the following.

Given a partition of the interval \([0, 1]\) in \( m \) points equally spaced at distance \( h = 1/m \) (generally, \( m \) will be large), we define

\[ x_0 = 0, x_1 = h, x_2 = 2h, \ldots, x_i = i \cdot h, \ldots, x_m = m \cdot h = 1. \]

To numerically solve the problem, we must find a vector of large length \( m \), call it \([u_0, u_1, \ldots, u_m]\) such that \( u_0 = u(0) = 0, u_m = u(1) = 0, \) and \( u_i \approx u(x_i) \) at all points \( x_i \) in the partition with \( 1 \leq i \leq m - 1 \). The accuracy of the approximation will depend on the number of points taken, i.e., the more points, the better the approximation.

How might one go about doing something like this? Somehow, we must find ways to replace \( u'' \) and \( u' \) with linear combinations of values \( u \) takes “nearby”. The answer is given by the Taylor series approximation:

\[ u(x + h) = u(x) + h \cdot u'(x) + \frac{h^2}{2} \cdot u''(x) + \frac{h^3}{6} \cdot u'''(x) + \ldots, \]

where we could basically take as many terms in the expansion as we want, but for all practical purposes we will not take more than a few.

Apply the above to \( x = x_i \) and \( h \), respectively \(-h\), and cut off terms after \( h^3 \) to get

\[ u(x_{i+1}) = u(x_i + h) \approx u(x_i) + h \cdot u'(x_i) + \frac{h^2}{2} \cdot u''(x_i) + \frac{h^3}{6} \cdot u'''(x_i), \]

\[ u(x_{i-1}) = u(x_i - h) \approx u(x_i) - h \cdot u'(x_i) + \frac{h^2}{2} \cdot u''(x_i) - \frac{h^3}{6} \cdot u'''(x_i); \]
now subtract the two approximations to get

\[ (u(x_{i+1}) - u(x_{i-1})) \approx 2hu'(x_i). \]

Note that the terms indeed do cancel on the right hand side, and in fact the error in the right hand side will be in the \(O(h^3)\) term (which we have ommitted!) Since \(h = 1/m\) with \(m\) very large, \(h\) is very small, and hence \(h^3\) is much, much smaller than \(h\).

Thus, we will be able to replace the first derivative \(u'(x_i)\) using the equivalent linear expression

\[ u'(x_i) = \frac{u(x_{i+1}) - u(x_{i-1})}{2h}. \]  

Similarly, we will obtain a linear expression for the second derivative \(u''(x_i)\), which—you may check—follows immediately from the Taylor series expansion:

\[ u''(x_i) = \frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2}. \]  

Armed with these two expressions, we can now set up an approximate linear system

\[
\frac{u_{i+1} - 2u_{i} + u_{i-1}}{h^2} + a \cdot \frac{u_{i+1} - u_{i-1}}{2h} + b \cdot u_{i} = f_{i}, \quad 1 \leq i \leq (m - 1), u_0 = 0 = u_m.
\]

Here \(f_i = f(x_i)\), and is one of the givens.

This mimics the ODE, and in fact the Taylor series approximation assures us that the approximation \(u_i \approx u(x_i)\) will be very good when \(h\) is small enough (or, equivalently, \(m\) is large enough).

From the \(i\)th equation we see that the only non-zero coefficients are those of \(u_{i-1}, u_i, \) and \(u_{i+1}.\) This means that the matrix is banded, that is, it only has non-zero entries in a band around the main diagonal (any system arising from a homogeneous, constant-coefficient ODE is going to be banded; in this case the band length is 3, as the system is tridiagonal). Rewrite the equation:

\[
\begin{pmatrix}
\frac{1}{h^2} - \frac{a}{h} \\
\frac{1}{h^2} - \frac{a}{h} \\
\vdots \\
0
\end{pmatrix}
\begin{pmatrix}
u_{i-1} \\
u_i \\
u_{i+1}
\end{pmatrix}
+ \begin{pmatrix}
\frac{1}{h^2} + b \\
\frac{1}{h^2} + b \\
\vdots \\
\frac{1}{h^2} + b
\end{pmatrix}
\begin{pmatrix}
u_{i+1} \\
u_i \\
u_{i-1}
\end{pmatrix}
= f_i;
\]

now we see that the matrix looks like

\[ A_h = \begin{bmatrix}
\frac{1}{h^2} - \frac{a}{h} & 0 & \cdots & 0 \\
\frac{1}{h^2} - \frac{a}{h} & \frac{1}{h^2} + b & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{1}{h^2} - \frac{a}{h} & -\frac{2}{h^2} + b
\end{bmatrix}, \]

The rows of this matrix are independent, and a very well-known theorem from linear algebra tells us that the system will have exactly one solution for each right hand side.

**Theorem 1.** Let \(A\) be an \(n \times n\) matrix with real (or complex) entries. The following are equivalent:

1. \(A^{-1}\) exists (the matrix is invertible or non-singular);
2. there is no \(y \neq 0\) such that \(Ay = 0\) (meaning, if \(Ay = 0\) then \(y = 0\));
3. the columns of $A$ are linearly independent;
4. the rows of $A$ are linearly independent;
5. for any column vector $b$ of length $n$, there exists precisely one column vector $y$ such that $Ay = b$.

For Homework 1 you will have to set up, or to set up and solve, systems of linear equations (the former for an ODE, the latter for a mass-spring system, which you can also read more about in Section 1.2 of the textbook).

Example. Let us take a look at a concrete example, for a simpler ODE, where no term in $u''$ appears:

$$u'(x) + u(x) = x(1 - x), \quad \text{for all } x \in [0, 1], \quad u(0) = u(1) = 0,$$

and suppose for simplicity we choose $m = 5$ (although this will not necessarily give a very good numerical solution, since $m$ is not large enough).

As $m = 5$, $h = 1/5$, and so by (1), we will replace

$$u'(x_i) = \frac{u_{i+1} - u_{i-1}}{2 \cdot \frac{1}{5}} = \frac{5}{2} (u_{i+1} - u_{i-1}), \quad \text{for all } 1 \leq i \leq (m - 1).$$

Note that we shall have $f_i = f(x_i) = x_i$, as given.

The linear system is then given by the equations

$$\frac{5}{2} (u_2 - u_0) + u_1 = x_1 = \frac{1}{5} \left(1 - \frac{1}{5}\right),$$
$$\frac{5}{2} (u_3 - u_1) + u_2 = x_2 = \frac{2}{5} \left(1 - \frac{2}{5}\right),$$
$$\frac{5}{2} (u_4 - u_2) + u_3 = x_3 = \frac{3}{5} \left(1 - \frac{3}{5}\right),$$
$$\frac{5}{2} (u_5 - u_3) + u_4 = x_4 = \frac{4}{5} \left(1 - \frac{4}{5}\right),$$

which, when we recall $u_0 = u_5 = 0$, becomes

$$\frac{5}{2} u_2 + u_1 = \frac{4}{25},$$
$$\frac{5}{2} (u_3 - u_1) + u_2 = \frac{6}{25},$$
$$\frac{5}{2} (u_4 - u_2) + u_3 = \frac{6}{25},$$
$$\frac{5}{2} (-u_3) + u_4 = \frac{4}{25}.$$

This is the same as solving the system $A \cdot u = b$, where $u = [u_1, u_2, u_3, u_4]^T$ and $b = [4/25, 6/25, 6/25, 4/25]^T$, while

$$A = \begin{pmatrix} 1 & \frac{5}{2} & 0 & 0 \\ -\frac{5}{2} & 1 & \frac{5}{2} & 0 \\ 0 & -\frac{5}{2} & 1 & \frac{5}{2} \\ 0 & 0 & -\frac{5}{2} & 1 \end{pmatrix}.$$