2003/11/12	**M174 : Numerical Methods in Science and Engineering**
Steven Pav	SPAV@MATH.UCSD.EDU

§6.5 Iterative Solutions Continued

We are trying to solve

$$\mathsf{A}\boldsymbol{x} = \boldsymbol{b} \tag{1}$$

iteratively by starting with some $\boldsymbol{x}^{(0)}$, then defining the iteration implicitly:

$$\mathsf{Q}\boldsymbol{x}^{(k+1)} = (\mathsf{Q} - \mathsf{A})\,\boldsymbol{x}^{(k)} + \boldsymbol{b},$$

for some chosen matrix Q.

We defined the error vector as

$$\boldsymbol{e}^{(k)} = \boldsymbol{x}^{(k)} - \boldsymbol{x},$$

and found that

$$\boldsymbol{e}^{(k+1)} = \left(\mathsf{I} - \mathsf{Q}^{-1}\mathsf{A}\right)\boldsymbol{e}^{(k)}.$$

We then defined the vector norm as

$$\|\boldsymbol{x}\|_2 = \sqrt{\boldsymbol{x}^\top \boldsymbol{x}} = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}}.$$

We defined the matrix norm *subordinate* to this vector norm as

$$\|\mathsf{A}\|_2 = \max_{\boldsymbol{x}\neq \boldsymbol{0}} \frac{\|\mathsf{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}.$$

Using the property that for any $\boldsymbol{x}$,

$$\|\mathsf{A}\boldsymbol{x}\|_2 \le \|\mathsf{A}\|_2 \|\boldsymbol{x}\|_2,$$

we see that

$$\left\|\boldsymbol{e}^{(k+1)}\right\|_2 = \left\|\left(\mathsf{I} - \mathsf{Q}^{-1}\mathsf{A}\right)\boldsymbol{e}^{(k)}\right\|_2 \le \left\|\mathsf{I} - \mathsf{Q}^{-1}\mathsf{A}\right\|_2 \left\|\boldsymbol{e}^{(k)}\right\|_2.$$

Thus our iteration converges if

$$\left\|\mathsf{I} - \mathsf{Q}^{-1}\mathsf{A}\right\|_2 < 1.$$

There are a number of alternative characterizations of the matrix norm. One is

$$\mathsf{A} = \max_{1\le i\le n} |\lambda_i|,$$

where λ_i are the n *singular values* of the matrix. Singular values are the eigenvalues of the matrix if it is nonsingular. Thus we have the theorem:

Theorem 1. An iterative solution scheme converges for any starting $\boldsymbol{x}^{(0)}$ if and only if all eigenvalues of $\mathsf{I} - \mathsf{Q}^{-1}\mathsf{A}$ are less than 1 in absolute value.

Another way of saying this is "the spectral radius of $\mathsf{I} - \mathsf{Q}^{-1}\mathsf{A}$ is less than 1."

An easier condition to check is *diagonal dominance*. The matrix A is said to be diagonally dominant if for $j = 1, 2, \dots, n$,

$$|a_{jj}| > \sum_{i=1, i \neq j}^{n} |a_{ji}| \,.$$

Thus in a diagonally dominant matrix, a diagonal element is larger, in absolute value, than the sum of the absolute values of the rest of the row.

A diagonally dominant matrix plus the right choice of Q gives the right eigenvalues, as the following theorem asserts:

Theorem 2. If the matrix A is diagonally dominant, then the Jacobi and Gauss Seidel iterative solution schemes converge for any starting $\boldsymbol{x}^{(0)}$.

The following material does not appear on the second midterm:

§7.1 First and Second Degree Splines

Splines are used to approximate complex functions and shapes. A spline is a function consisting of simple functions glued together. In this way a spline is different from a polynomial interpolation, which consists of a single well defined function that approximates a given shape; splines are normally piecewise polynomial.

To define a spline on $[a, b]$, we partition the interval into *knots* $t_0, t_1, \dots, t_n$. The spline is a polynomial on each subinterval $[t_i, t_{i+1}]$. We define splines of degree 1:

Definition 3 (Splines of Degree 1). A function S is a spline of degree 1 on $[a, b]$ if

1. The domain of S is $[a, b]$.
2. S is continuous on $[a, b]$.
3. There is a partition $a = t_0 < t_1 < \dots < t_n = b$ such that on $[t_i, t_{i+1}]$, S is a linear polynomial.

A degree 1 spline is defined entirely by its value at the knots. That is, given

t	t_0	t_1	$\dots$	t_n
y	y_0	y_1	$\dots$	y_n

there is only one degree 1 spline with these values at the knots and linear on each given subinterval.

For a spline with this data, the linear polynomial on each subinterval is defined as

$$S_i(x) = y_i + \frac{y_{i+1} - y_i}{t_{i+1} - t_i}\,(x - t_i)\,.$$

Note that if $x \in [t_i, t_{i+1}]$, then $x - t_i > 0$, but $x - t_{i-1} \le 0$. Thus if we wish to evaluate $S(x)$, we search for the largest i such that $x - t_i > 0$, then evaluate $S_i(x)$.

First Degree Spline Accuracy

The book has a fancy "theorem" with no content. I claim we can immediately see that if $p(x)$ is the linear polynomial interpolating $f(x)$ at the endpoints of the interval $[a, b]$, then for $x \in [a, b]$,

$$|f(x) - p(x)| \le \max\{|f(x) - f(a)|, |f(x) - f(b)|\}.$$

Thus $|f(x) - p(x)|$ is no larger than the "maximum variation" of $f(x)$ on this interval.

In particular, if $f'(x)$ exists and is bounded by M_1 on $[a, b]$, then

$$|f(x) - p(x)| \le \frac{M_1}{2}(b - a).$$

Similarly, if $f''(x)$ exists and is bounded by M_2 on $[a, b]$, then

$$|f(x) - p(x)| \le \frac{M_2}{8}(b - a)^2.$$

Note this is better than polynomial interpolation, which may get worse as the number of nodes is increased.

Second Degree Splines

Piecewise quadratic splines, or *second degree splines* are defined similarly to the degree 1 splines:

Definition 4 (Splines of Degree 2). A function Q is a spline of degree 2 on $[a, b]$ if

1. The domain of Q is $[a, b]$.
2. Q is continuous on $[a, b]$.
3. Q' is continuous on $[a, b]$.
4. There is a partition $a = t_0 < t_1 < \ldots < t_n = b$ such that on $[t_i, t_{i+1}]$, Q is a polynomial of degree at most 2.

Example 5. The following is a quadratic splne:

$$Q(x) = \begin{cases} -x & x \le 0, \\ x^2 - x & 0 \le x \le 2, \\ -x^2 + 7x - 1 & 2 \le x. \end{cases}$$

Unlike degree 1 splines, degree 2 splines are *not* defined entirely by their values at the knots. We consider why that is. The spline $Q(x)$ is defined by its piecewise polynomials,

$$Q_i(x) = a_i x^2 + b_i x + c_i.$$

Thus there are $3n$ parameters to define $Q(x)$.

For each of the n subintervals, the data

t	t_0	t_1	$\ldots$	t_n
y	y_0	y_1	$\ldots$	y_n

give two equations regarding $Q_i(x)$, namely that $Q_i(t_i)$ must equal y_i and $Q_i(t_{i+1})$ must equal y_{i+1}. This is $2n$ equations. The condition on continuity of Q' gives a single equation for each of the $n-1$ internal nodes. This totatls $3n-1$ equations, but $3n$ unknowns. This system is underdetermined.

Thus some additional user-chosen condition is required to determine the quadratic spline. One might choose, for example, $Q'(a) = 0$, or $Q''(a) = 0$, or some other condition.

Computing Second Degree Splines

Suppose the data

t	t_0	t_1	$\ldots$	t_n
y	y_0	y_1	$\ldots$	y_n

are given. Let $z_i = Q_i'(t_i)$, and suppose that the additional condition to define the quadratic spline is given by specifying z_0. We want to be able to compute the form of $Q_i(x)$.

Because $Q_i(t_i) = y_i, Q_i'(t_i) = z_i, Q_i'(t_{i+1}) = z_{i+1}$, we see that we can define

$$Q_i(x) = \frac{z_{i+1} - z_i}{2\left(t_{i+1} - t_i\right)}\left(x - t_i\right)^2 + z_i\left(x - t_i\right) + y_i.$$

Use this at t_{i+1} :

$$\begin{aligned} y_{i+1} = Q_i(t_{i+1}) &= \frac{z_{i+1} - z_i}{2\left(t_{i+1} - t_i\right)}\left(t_{i+1} - t_i\right)^2 + z_i\left(t_{i+1} - t_i\right) + y_i, \\ y_{i+1} - y_i &= \frac{z_{i+1} - z_i}{2}\left(t_{i+1} - t_i\right) + z_i\left(t_{i+1} - t_i\right), \\ y_{i+1} - y_i &= \frac{z_{i+1} + z_i}{2}\left(t_{i+1} - t_i\right). \end{aligned}$$

Thus we can determine, from the data alone, z_{i+1} from z_i:

$$z_{i+1} = 2\frac{y_{i+1} - y_i}{t_{i+1} - t_i} - z_i.$$