Linear Algebraic Systems

Lecture 2
Indirect methods: Jacobi, Gauss-Zeidel and SOR methods

By contrast with direct methods which provide you with a finite algorithm which leads to a formula for the exact answers, *indirect* or, as they are sometimes known, *iterative* methods consist of a number of steps, designed to get you pretty close to the answer. They consist of a number of steps:

1. *Guess* an initial iteration
$$\mathbf{x} = \mathbf{x}^{(0)}. \tag{1}$$
where (0) is a counter, *not* a power!

2. Substitute $\mathbf{x}^{(0)}$ in to the original equation
$$A\mathbf{x} = \mathbf{b}, \tag{2}$$
in some way, in such a way as to generate a new iteration $\mathbf{x}^{(1)}$.

3. Continue (or *iterate*) this procedure, generating $\mathbf{x}^{(n+1)}$ from $\mathbf{x}^{(n)}$.

4. Stop when (by some criterion or other) the iteration $\mathbf{x}^{(n)}$ is good enough.

The iteration procedure must:
Ensure that $\mathbf{x}^{(n+1)}$ is closer to the true solution than $\mathbf{x}^{(n)}$ - convergent.
Efficient in the sense that the amount of work (computer time and memory) is balanced between the effort to produce the next iterate and the improvement of that iterate as opposed to the present one.

An iterative procedure (unlike a direct procedure) will usually *not* reach the exact result in a finite set of computations. In order to define 'closer to the exact solution' one needs a procedure for defining the size of a vector or matrix.

Review of Norms

Definition

A *norm* of $\mathbf{x}$, $||x||$, where $\mathbf{x}$ is a real or complex n vector (strictly $\mathbf{x} \in \mathcal{R}_n$ or $\mathcal{C}_n$) satisfies

$||x|| > 0$, and $\mathbf{x} = 0 \Leftrightarrow ||\mathbf{x}|| = 0$.

For λ real or complex, $||\lambda x|| = |\lambda| \cdot ||x||$.

$||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$.

The norm is not unique. Specific examples are the so-called l_p norms:

$$||\mathbf{x}||_p = \left\{ \sum_{i=1}^{n} |x_i|^p \right\}^{1/p} \tag{3}$$

with p specified. Thus $p = 1$ yields the *abstract* norm :

$$||\mathbf{x}||_1 = \sum_{i=1}^{n} |x_i|, \tag{4}$$

$p = 2$ yields the *Euclidean* norm (or length as usually defined):

$$||\mathbf{x}||_2 = \left\{ \sum_{i=1}^{n} |x_i|^2 \right\}^{1/2}, \tag{5}$$

and $p = \infty$ yields the *maximum* norm:

$$||\mathbf{x}||_\infty = \max|x_i| \tag{6}$$

Example:

$$\mathbf{x} = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}.$$

Then

$$\begin{aligned} ||\mathbf{x}||_1 = &\quad 1+2+4 &&= 7 \\ ||\mathbf{x}||_2 = &\quad \left(1^2+2^2+4^2\right)^{1/2} &&= \sqrt{21} \\ ||\mathbf{x}||_\infty = &\quad \max\{1,2,4\} &&= 4 \end{aligned} \tag{7}$$

There is an extensive literature about norms, which we shall not need to go into. The important point is that we can choose the one which is most convenient for us to check whether some difference is decreasing, and that whatever norm we choose, more or less, we'll get the same answer to that question. We can also write down norms for matrices. These must satisfy:

$||A|| \geq 0$, and $||A|| = 0 \Leftrightarrow A = 0$;

$||\lambda A|| = |\lambda| ||A||$;

$||A + B|| \leq ||A|| + ||B||$;

$||AB|| \leq ||A|| \cdot ||B||$. (The *Cauchy-Schwartz* inequality)

As with vector norms there are various choices, but in any application one should ensure that one is using a *consistent* choice:

$$||A\mathbf{x}|| \leq ||A||\ ||\mathbf{x}|| \quad \forall A, \mathbf{x} \tag{8}$$

(The Cauchy-Schwartz inequality again). The norm of a matrix of order $m \times n$ could be constructed by regarding it as a vector in $\mathcal{R}_{mn}$. For example, the Frobenius norm is defined by

$$||A||_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2}. \tag{9}$$

A common choice of matrix norm makes it *subordinate* to the vector norm:

$$||A|| = \max_{\mathbf{x} \neq 0} \frac{||A\mathbf{x}||}{||\mathbf{x}||}. \tag{10}$$

This choice of matrix norm makes the identity matrix I have unit norm: $||I|| = 1$, which surprisingly is *not* true for a general norm. Some examples:

$$\begin{aligned} ||A||_1 &= \max_j \left\{ \sum_{i=1}^{n} |a_{ij}| \right\}, \text{(maximum column sum)} \\ ||A||_\infty &= \max_i \left\{ \sum_{j=1}^{n} |a_{ij}| \right\}, \text{(maximum row sum)} \\ ||A||_2 &= \left\{ \rho(A^* A) \right\}^{1/2}, \end{aligned} \tag{11}$$

where A^* is the *Hermitian conjugate* (= complex transpose) of A, and $\rho(A)$ is the *spectral radius* of B:

$$\rho(A) = \max_i |\lambda_i(A)|$$

This is the maximum out of the moduli of the eigenvalues.

Matrix norms that are subordinate to vector norms satisfy the submultiplicative property,

$$||AB|| \leq ||A||||B||. \tag{12}$$

Common indirect methods

Recall that we are trying to solve the general equation $A\mathbf{x} = \mathbf{b}$. We've got as far as a guess $\mathbf{x}^{(n)}$. How close are we? One way of keeping tabs on this is to define the *residual*

$$\mathbf{r}^{(n)} = \mathbf{b} - A\mathbf{x}^{(n)}. \tag{13}$$

If the residual is zero we've found the solution. Otherwise we want $||\mathbf{r}^{(n)}||$ to keep going down. The three most common iterative methods use this idea. Let us summarise what goes on in the iterative part of the solution process in these cases.

Determine a step

$$\Delta^{(n)} = \mathbf{x}^{(n+1)} - \mathbf{x}^{(n)},$$

where

$$M\Delta^{(n)} = \mathbf{r}^{(n)}.$$

The only problem now is to identify the matrix M.

There are a number of common choices:

1. *Jacobi's* method chooses the diagonal of A. Thus

$$\begin{cases} m_{ij} &= 0 \quad i \neq j \\ m_{ii} &= a_{ii} \quad i = j \end{cases} \tag{14}$$

2. The *Gauss-Seidel* method chooses the upper triangular part of A:

$$\begin{cases} m_{ij} &= 0 \quad i > j \\ m_{ij} &= a_{ij} \quad i \leq j \end{cases} \tag{15}$$

3. The SOR/SUR method is the same as Gauss-Seidel, but includes a parameter ω which improves convergence (but which must be determined from A). SOR stands for **S**uccessive **O**ver **R**elaxation:

$$\begin{cases} m_{ij} &= 0 \quad & i > j \\ m_{ij} &= a_{ij} \quad & i < j \\ m_{ii} &= a_{ii}/\omega \quad & i = j \end{cases} \tag{16}$$

Numerical procedure

The basic procedure is as follows: Guess $\mathbf{x}_0$ and set M. At this stage $n = 0$. This is the iterative loop.

1. Calculate $\mathbf{r}^{(n)} = \mathbf{b} - A\mathbf{x}^{(n)}$
2. $M\Delta^{(n)} = \mathbf{r}^{(n)}$
3. Calculate $\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} + \Delta^{(n)}$
4. Test if $\mathbf{x}^{(n+1)}$ is good enough.

   If NO, put $n = n + 1$ and go back to the beginning of the loop.

   If YES FINISH.

We decide to stop iterating if

$$\frac{||\Delta^{(n+1)}||}{||\mathbf{x}^{(n+1)}||} < \text{Tolerance}, \tag{17}$$

or

$$\frac{||r^{(n+1)}||}{||\mathbf{b}||} < \text{Tolerance}. \tag{18}$$

Convergence condition and Condition Number

Let $\hat{\mathbf{x}}$ be the *exact* solution of the equation $A\mathbf{x} = \mathbf{b}$. Then

$$\mathbf{r}^{(n)} = \mathbf{b} - A\mathbf{x}^{(n)} = A(\hat{\mathbf{x}} - \mathbf{x}^{(n)}) \tag{19}$$

Now, using the Cauchy-Schwartz inequality eq.(8)

$$||\mathbf{r}^{(n)}|| \ \leq \ ||A||\,||\hat{\mathbf{x}} - \mathbf{x}^{(n)}||, \tag{20}$$

$$||\hat{\mathbf{x}} - \mathbf{x}^{(n)}|| \;\leq\; ||A^{-1}||\;||\mathbf{r}^{(n)}||. \tag{21}$$

and thus

$$\frac{||\mathbf{r}^{(n)}||}{||A||} \;\leq\; ||\hat{\mathbf{x}} - \mathbf{x}^{(n)}|| \leq ||A^{-1}||\;||\mathbf{r}^{(n)}||. \tag{22}$$

An exactly analogous argument, using the result $A\hat{\mathbf{x}} = \mathbf{b}$, yields:

$$\frac{||\mathbf{b}||}{||A||} \;\leq ||\hat{\mathbf{x}}|| \leq ||A^{-1}||\;||\mathbf{b}||. \tag{23}$$

Combining the results of eqs.(22,23) yields an inequality which bounds the relative error of the solution $\mathbf{x}^{(n)}$:

$$\frac{1}{||A||\;||A^{-1}||}\frac{||\mathbf{r}^{(n)}||}{||\mathbf{b}||} \;\leq\; \frac{||\hat{\mathbf{x}} - \mathbf{x}^{(n)}||}{||\hat{\mathbf{x}}||} \;\leq\; ||A||\;||A^{-1}||\frac{||\mathbf{r}^{(n)}||}{||\mathbf{b}||}. \tag{24}$$

Thus if the ratio

$$\frac{||\mathbf{r}^{(n)}||}{||\mathbf{b}||} \leq \text{tolerance} \tag{25}$$

we have a constraint on how close the estimated solution is to the real one. Probably the constraint is factors too high, actually. How close the solution actually is partly depends on the value of $||A||\;||A^{-1}||$. This is known as the *condition number* of the matrix. If the condition number is large then even if $||\mathbf{r}^{(n)}||$ is small, we may not be close to the solution. The problem would then be *ill-conditioned*. The magnitude of the condition number depends on the chosen norm. If matrix A is symmetric and we take $||A||_2$, then

$$||A||\;||A^{-1}|| = \frac{\max|\lambda_i(A)|}{\min|\lambda_i(A)|}. \tag{26}$$

Note that of the determinant gets very small, the matrix becomes almost singular and the condition number becomes large. Singular matrices will

have no solution (usually); nearly singular matrices will have very sensitive solutions.

Convergence Rate

We can also investigate the *convergence rate* of an iterative method. Let us examine what happens to

$$\mathbf{y}^{(n)} = \hat{\mathbf{x}} - \mathbf{x}^{(n)}. \tag{27}$$

Then

$$\mathbf{y}^{(n+1)} = \hat{\mathbf{x}} - \mathbf{x}^{(n+1)}, \tag{28}$$

and thus

$$\mathbf{y}^{(n+1)} = (I - M^{-1}A)\mathbf{y}^{(n)} = R\mathbf{y}^{(n)}. \tag{29}$$

The matrix $R = I - M^{-1}A$ is called the *iteration matrix*. The idea is that if R makes the vector $\mathbf{r}^{(n)} = A\mathbf{y}^{(n)}$ smaller in some sense, then the solution method is stable. Thus

$$\mathbf{y}^{(n)} = R^n\mathbf{y}^{(0)} = R^nA^{-1}r^{(0)}, \tag{30}$$

or, using the Cauchy-Schwartz inequality:

$$||\mathbf{y}^{(n)}|| \leq ||R^n||\ ||A^{-1}||\ ||r^{(0)}|| \leq ||R||^n\ ||A^{-1}||\ ||r^{(0)}||. \tag{31}$$

Thus the iterative method will converge if and only if $||R|| < 1$. A necessary and sufficient condition for this is that all the eigenvalues of R have modulus less than one. In the SOR/SUR method we choose the coordinate ω in such a way as to minimise the largest eigenvalue of the matrix R. This will ensure the most rapid convergence.