AN ITERATION

In part as motivation, we consider an iteration method for solving a system of linear equations which has the form

$$x - Ax = b$$

In this, A is an $n \times n$ matrix and $b \in \mathbb{R}^n$. Systems of this form arise in a number of applications. One way in which such systems are solved is as follows. First rewrite the system as

$$x = b + Ax \qquad (1)$$

Choose and initial guess $x^{(0)}$. Then define a sequence of iterates $x^{(1)}, x^{(2)}, \ldots$ by

$$x^{(k+1)} = b + Ax^{(k)}, \quad k = 0, 1, 2, \ldots \qquad (2)$$

Do these iterates converge to x? Subtract (2) from (1), obtaining

$$x - x^{(k+1)} = Ax - Ax^{(k)} = A\left(x - x^{(k)}\right)$$

$$x - x^{(k+1)} = A\left(x - x^{(k)}\right), \quad k \geq 0$$

By recursion,

$$x - x^{(k)} = A^k\left(x - x^{(0)}\right), \quad k \geq 0$$

This error depends on both A^k and the initial error $x - x^{(0)}$,

$$\left\|x - x^{(k)}\right\| \leq \left\|A^k\right\| \left\|x - x^{(0)}\right\|$$

Generally we want a method which works for all possible initial guesses, in part because rounding error will eventually bring a randomness into the error. Consequently, in order to determine convergence of $x^{(k)}$ to x, we need to know whether $A^k \to 0$ as $k \to \infty$.

THEOREM

Let A be an $n \times n$ matrix. Then

$$A^k \to 0 \quad \text{as} \quad k \to \infty$$

if and only if

$$r_\sigma(A) < 1$$

PROOF: There are a number of ways to prove this; and in the book, I use the Jordan canonical form for A. To simplify things for this class presentation, we make the simplifying assumption that the Jordan canonical form for A is a diagonal matrix: there is a nonsingular matrix P for which

$$P^{-1}AP = D = \text{diag}\,[\lambda_1, ..., \lambda_n]$$

We can rewrite this as

$$A = PDP^{-1}$$

Then

$$\begin{aligned} A^k &= \left(PDP^{-1}\right)^k \\ &= \left(PDP^{-1}\right)\left(PDP^{-1}\right)\cdots\left(PDP^{-1}\right) \\ &= PD^kP^{-1} \end{aligned}$$

Then $A^k \to 0$ if and only if $D^k \to 0$.

For a diagonal matrix $D = \text{diag}\,[\lambda_1, ..., \lambda_n]$,

$$D^k = \text{diag}\left[\lambda_1^k, ..., \lambda_n^k\right]$$

Thus $D^k \to 0$ if and only if

$$\lambda_i^k \to 0 \quad \text{as} \quad k \to 0$$

for $i = 1, 2, ..., n$. This is true if and only if

$$r_\sigma(A) \equiv \max_{1 \le i \le n} |\lambda_i| < 1$$

COROLLARY

Suppose $\|A\| < 1$. Then $A^k \to 0$ as $k \to \infty$, because

$$r_\sigma(A) \le \|A\| < 1$$

SOLVABILITY OF A LINEAR SYSTEM

Return to the linear system

$$(I - A)\,x = x - Ax = b$$

Is this system uniquely solvable? Equivalently, is $I - A$ an invertible matrix? The answer to this question turns out to be of fundamental importance in the numerical analysis of linear algebra problems.

THEOREM

Let A be an $n \times n$ matrix. If $r_\sigma(A) < 1$, then $(I - A)^{-1}$ exists; and moreover, it can be expressed as the convergent infinite series

$$(I - A)^{-1} = I + A + A^2 + A^3 + \cdots$$

Conversely, if this series is convergent, then

$$r_\sigma(A) < 1$$

PROOF: Assume $r_\sigma(A) < 1$. Then $I - A$ must be nonsingular. Otherwise there would be a vector $x \neq 0$ for which

$$\begin{aligned} (I - A)\,x &= 0 \\ Ax &= x \end{aligned}$$

and this says 1 is an eigenvalue of A, contrary to assumption.

To prove the series converges, look at its partial sums

$$B_k = I + A + A^2 + A^3 + \cdots + A^k$$

for $k = 1, 2, 3, ...$ Then

$$\begin{aligned} (I - A)B_k &= (I - A)\left(I + A + \cdots + A^k\right) \\ &= I - A^{k+1} \\ B_k &= (I - A)^{-1}\left(I - A^{k+1}\right) \end{aligned}$$

Then the partial sums B_k converge to $(I - A)^{-1}$ because $A^{k+1} \to 0$.

If instead we assume the series

$$I + A + A^2 + A^3 + \cdots$$

is convergent, then the general term A^k must converge to the zero matrix 0 as $k \to \infty$. That in turn implies

$$r_\sigma(A) < 1$$

This result about

$$(I - A)^{-1} = I + A + A^2 + A^3 + \cdots$$

is called the *Geometric Series Theorem*. Note its resemblance to the well-known geometric series

$$\frac{1}{1 - z} = 1 + z + z^2 + z^3 + \cdots, \qquad |z| < 1$$

This in not an accidental resemblance. We can do similar things with series such as that for e^z, leading to

$$e^A = 1 + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots$$

THEOREM

Let A be an $n \times n$ matrix with $\|A\| < 1$. Then $(I - A)^{-1}$ exists, with

$$(I - A)^{-1} = I + A + A^2 + A^3 + \cdots$$

Moreover,

$$\left\|(I - A)^{-1}\right\| \leq \frac{1}{1 - \|A\|}$$

PROOF: The only thing to be proven is the bound. Recall that

$$\begin{aligned} B_k &= I + A + A^2 + A^3 + \cdots + A^k \\ &= (I - A)^{-1}\left(I - A^{k+1}\right) \end{aligned}$$

Then

$$\begin{aligned} \|B_k\| &\leq \|I\| + \|A\| + \left\|A^2\right\| + \left\|A^3\right\| + \cdots + \left\|A^k\right\| \\ &\leq 1 + \|A\| + \|A\|^2 + \cdots + \|A\|^k \\ &= \frac{1 - \|A\|^{k+1}}{1 - \|A\|} \leq \frac{1}{1 - \|A\|} \end{aligned}$$

$$\|B_k\| \leq \frac{1}{1 - \|A\|}$$

Also,

$$\begin{aligned} \left| \left\| (I - A)^{-1} \right\| - \|B_k\| \right| &\leq \left\| (I - A)^{-1} - B_k \right\| \\ &= \left\| (I - A)^{-1} A^{k+1} \right\| \\ &\leq \left\| (I - A)^{-1} \right\| \|A\|^{k+1} \to 0 \end{aligned}$$

Combining these results proves the bound

$$\left\| (I - A)^{-1} \right\| \leq \frac{1}{1 - \|A\|}$$

AN APPROXIMATION RESULT

Suppose we are considering a system

$$Ax = b$$

where A is some nonsingular $n \times n$ matrix. What happens to the solvability of this system if we change the matrix A by a "small amount" to a new square matrix B? And what is the relation of the solution of

$$Bz = b$$

to the original solution x?

THEOREM

Let A and B be two $n \times n$ matrices, and suppose A is nonsingular. Moreover, assume

$$\|A - B\| < \frac{1}{\left\|A^{-1}\right\|}$$

Then the matrix B is also nonsingular, with

$$\left\|B^{-1}\right\| \leq \frac{\left\|A^{-1}\right\|}{1 - \left\|A^{-1}\right\| \|A - B\|}$$

For the solutions of $Ax = b$ and $Bz = b$,

$$\|x - z\| \leq \left\|B^{-1}\right\| \|A - B\| \|x\|$$

This theorem tells us that if A is invertible, then all "nearby" matrices B are also invertible.

PROOF: Write

$$\begin{aligned} B &= A - (A - B) \\ &= A\left[I - A^{-1}(A - B)\right] \end{aligned}$$

Examine the matrix $I - A^{-1}(A - B)$.

$$\left\| A^{-1}(A - B) \right\| \leq \left\| A^{-1} \right\| \|A - B\| < 1$$

Then by the Geometric Series Theorem, $\left[I - A^{-1}(A - B)\right]^{-1}$ exists, and

$$\begin{aligned} \left\| \left[I - A^{-1}(A - B)\right]^{-1} \right\| &\leq \frac{1}{1 - \left\| A^{-1}(A - B) \right\|} \\ &\leq \frac{1}{1 - \left\| A^{-1} \right\| \|A - B\|} \end{aligned}$$

Returning to

$$B = A\left[I - A^{-1}(A - B)\right]$$

we have B is the product of invertible matrices, and therefore it is itself invertible:

$$B^{-1} = \left[I - A^{-1}(A - B)\right]^{-1} A^{-1}$$

$$\left\|B^{-1}\right\| \le \left\|\left[I - A^{-1}(A - B)\right]^{-1}\right\| \left\|A^{-1}\right\|$$

The bound for $\left\|B^{-1}\right\|$ follows from this formula.

For the equations $Ax = b$ and $Bz = b$, write

$$\begin{aligned} x - z &= A^{-1}b - B^{-1}b \\ &= \left(A^{-1} - B^{-1}\right) b \\ &= B^{-1}(B - A) A^{-1}b \\ &= B^{-1}(B - A)\, x \\ \|x - z\| &\le \left\|B^{-1}\right\| \|A - B\| \|x\| \end{aligned}$$

EXAMPLE

Let

$$B = \begin{bmatrix} c & 1 & 0 & 0 & 0 \\ 1 & c & 1 & 0 & 0 \\ 0 & 1 & c & 1 & 0 \\ 0 & 0 & 1 & c & 1 \\ 0 & 0 & 0 & 1 & c \end{bmatrix}$$

Approximate this by the matrix $A = cI$. Then

$$B - A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Using the row norm,

$$\|B - A\|_\infty = 2$$

Also,

$$A^{-1} = \frac{1}{c} I, \qquad \left\| A^{-1} \right\|_\infty = \left| \frac{1}{c} \right|.$$

The condition

$$\| A - B \|_\infty < \frac{1}{\left\| A^{-1} \right\|}$$

becomes

$$2 < |c|$$

Then B is invertible if this is true; and then the bound

$$\left\| B^{-1} \right\|_\infty \le \frac{\left\| A^{-1} \right\|_\infty}{1 - \left\| A^{-1} \right\|_\infty \| A - B \|_\infty}$$

becomes

$$\left\| B^{-1} \right\|_\infty \le \frac{\left| \frac{1}{c} \right|}{1 - 2 \left| \frac{1}{c} \right|} = \frac{1}{|c| - 2} \qquad (3)$$

This also tells us something about the eigenvalues of B. Let $Bx = \lambda x$ for some $x \neq 0$. Then

$$|\lambda| \leq \|B\| = |c| + 2 \tag{4}$$

Multiply both sides of $Bx = \lambda x$ by both B^{-1} and λ^{-1} to get

$$\lambda^{-1} x = B^{-1} x$$

Therefore the eigenvalues of B^{-1} are the reciprocals of those of B. Recall that

$$r_\sigma(B^{-1}) \leq \left\| B^{-1} \right\|_\infty$$

Then (3) implies

$$\frac{1}{|\lambda|} \leq \frac{1}{|c| - 2}$$

$$|\lambda| \geq |c| - 2$$

With (4),

$$|c| - 2 \leq |\lambda| \leq |c| + 2$$

for all eigenvalues λ of the matrix B.