Fixed point iterations

The goal: find the solution of a fixed point equation

$X = G(X)$, where $X \in R^n$, $G : R^n \to R^n$

The method: choose $X^{(0)}$, then for $k = 0, 1, 2, \dots$ compute

$X^{(k+1)} = G(X^{(k)})$

Example: Consider the following system of nonlinear equations

$$\begin{aligned} x_1 &= 1 + 0.2 \sin(x_1 - 2x_2) \\ x_2 &= \sqrt{x_1 + x_2 + 4} \end{aligned}$$

We have

$$X = \begin{bmatrix} x_1 \\ \\ x_2 \end{bmatrix}, \qquad G(X) = \begin{bmatrix} g_1(x_1, x_2) \\ \\ g_2(x_1, x_2) \end{bmatrix} = \begin{bmatrix} 1 + 0.2 \sin(x_1 - 2x_2) \\ \\ \sqrt{x_1 + x_2 + 4} \end{bmatrix}.$$

Theorem 1 – contraction mapping theorem:

Let $D \subset R^n$, D be closed, $G : D \to R^n$. Assume

1. if $X \in D$, then $G(X) \in D$
2. the mapping G is a *contraction* on D: there exists $q < 1$ such that

$$\|G(X) - G(Y)\| \le q \, \|X - Y\| \qquad \forall \; X, Y \in D \tag{1}$$

Then

- there exists a unique $X^* \in D$ such that $X^* = G(X^*)$ and the fixed point iterations converge to X^* for any choice of $X^{(0)} \in D$,
- $X^{(k)}$ satisfies the a-priori error estimate $\|X^{(k)} - X^*\| \le \frac{q^k}{1-q} \, \|X^{(1)} - X^{(0)}\|$ and the a-posteriori error estimate $\|X^{(k)} - X^*\| \le \frac{q}{1-q} \, \|X^{(k)} - X^{(k-1)}\|$.

Proof: see [1]

Note: The norm $\| \, . \, \|$ in Theorem 1 can be *any* vector norm (however, the same in both the assumptions and the proposition).

Theorem 2 – the contraction property:

Let $D \subset R^n$, D be convex, $G : D \to R^n$ has continuous partial derivatives $\frac{\partial f_i}{\partial x_j}$ in D. Assume there exists $q < 1$ such that the matrix norm of the Jacobian $\|G'(X)\| < q$, $\forall X \in D$. Then G is a contraction in D and satisfies (1).

Proof: see [1]

Note: The norm $\| \, . \, \|$ in Theorem 2 can be *any* matrix norm consistent with the vector norm in (1).

Example continued – the Jacobi matrix

$$G'(X) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} \\ \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 0.2\,\cos(x_1 - 2x_2) & -0.4\,\cos(x_1 - 2x_2) \\ \\ \frac{1}{2\sqrt{x_1+x_2+4}} & \frac{1}{2\sqrt{x_1+x_2+4}} \end{bmatrix}$$

is continuous on $\Omega = \{X \in R^2 : x_1 > -\frac{1}{2}, x_2 > -\frac{1}{2}\}$ and Ω is convex.
Let us try the row norm first, because the computation seems to be the simplest:
$\|G'(X)\|_\infty = \max(|0.2\,\cos(x_1-2x_2)|+|-0.4\,\cos(x_1-2x_2)|,\ 2\,|\frac{1}{2\sqrt{x_1+x_2+4}}|) \leq \max(0.6, \frac{1}{\sqrt{3}}) =$
$= 0.6\,.$ Assumptions of Theorem 2 hold, so G is a contraction on Ω.
Choose $D = \{X \in R^2 : x_1 \geq 0, x_2 \geq 0\} \subset \Omega$. Then D is closed, $G(D) \subset D$ and according to Theorem 1, there exists a unique solution in D and FPI converge for any $X^{(0)} \in D$.
Starting at the origin, we have (rounded to 4 decimal places)

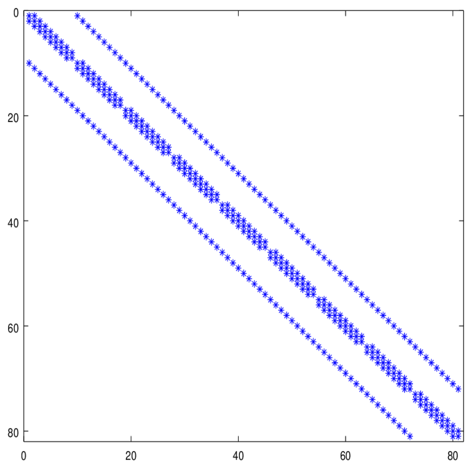
$$X^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ X^{(1)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ X^{(2)} = \begin{bmatrix} 0.9718 \\ 2.6458 \end{bmatrix},\ \ldots\ X^{(9)} = \begin{bmatrix} 1.1943 \\ 2.8333 \end{bmatrix} = X^{(10)} = X^{(11)},$$

so we stop and verify the result: $\quad X^{(9)} - G(X^{(9)}) = \begin{bmatrix} 3.7 \cdot 10^{-5} \\ -1.9 \cdot 10^{-6} \end{bmatrix}$.

Fixed point iterations for linear systems

Motivation

Typical matrix resulting from discretization of differential equations is *sparse*.
Example – discretization of Poisson equation in 2D square domain using finite differences (11×11 grid, zero Dirichlet boundary condition):



– a *banded* matrix 81×81, the *bandwidth* $h = 10$.
Consider $n \times n$ matrix with bandwidth $h = c \cdot n$ ($c \approx 0.12$ in our example).
For $n = 10^6$: 5 nonzero diagonals represent approximately $5n = 5 \cdot 10^6$ nonzeros
Gauss elimination fills in the whole band – approx. $h \cdot n = c \cdot n^2 = c \cdot 10^{12} \approx 10^{11}$ nonzeros
– about $2 \cdot 10^4$-times more computer memory is needed!

Fixed point iterations for $X = UX + V$

Assume $G(X) = UX + V$, where U is a $n \times n$ matrix, $V \in R^n$, so the fixed point equation now represents a system of linear equations $X = UX + V$:

$$\begin{aligned} x_1 &= u_{11}x_1 + u_{12}x_2 + \cdots + u_{1n}x_n + v_1 \\ x_2 &= u_{21}x_1 + u_{22}x_2 + \cdots + u_{2n}x_n + v_2 \\ \ldots & \\ x_n &= u_{n1}x_1 + u_{n2}x_2 + \cdots + u_{nn}x_n + v_n \ . \end{aligned}$$

Under which assumptions the convergence of the fixed point iterations

$$X^{(k+1)} = U\,X^{(k)} + V$$

is guaranteed on R^n ?

From properties of any norm on R^n and its consistent matrix norm,

$\|G(X) - G(Y)\| = \|UX + V - (UY + V)\| = \|UX - UY\| = \|U(X - Y)\| \le \|U\|\,\|X - Y\|$

holds $\forall\ X, Y \in R^n$, so the theorem follows:

Theorem 3 – sufficient condition for convergence of FPI:

If there exists a matrix norm such that $\|U\| < 1$,
then the mapping $G(X) = UX + V$ is a contraction on R^n.
Proof: above

Now from Theorem 1 it follows that fixed point iterations $X^{(k+1)} = U\,X^{(k)} + V$ converge to the (unique) fixed point X^* for any choice of $X^{(0)}$. Moreover, both a-priori and a-posterirori error estimates hold with choice of $q = \|U\| < 1$ (provided the matrix norm and the vector norms are consistent).

What can be said about convergence of FPI, if there is no norm found for which $\|U\| < 1$?

Analysis of an error

Let $e^{(k)} = X^{(k)} - X^*$ be an error in k-th iteration of FPI. Then

$e^{(k)} = X^{(k)} - X^* = (U\,X^{(k-1)} + V) - (U\,X^* + V) = U\,(X^{(k-1)} - X^*) = Ue^{(k-1)}$

$e^{(k)} = Ue^{(k-1)} = U^2e^{(k-2)} = \cdots = U^ke^{(0)}$

Theorem 4 – necessary and sufficient condition for convergence of FPI:

The iteration process $X^{(k+1)} = U\,X^{(k)} + V$ converges to the fixed point X^* for *any* choice of $X^{(0)}$, if and only if $\rho(U) < 1$.

Proof: follows from the analysis of an error above and from the property

$U^k \to 0 \iff \rho(U) < 1.$

(see [2], Th. 1.10)

Methods for solving $AX = B$ based on fixed point iterations

The idea: transform $AX = B$ to $X = UX + V$ and use the fixed point iterations.

Richardson method – the most straightforward method:

$$\begin{aligned} AX &= B \\ 0 &= B - AX \\ 0 &= \alpha\,(B - AX), \quad \alpha \neq 0 \\ X &= X + \alpha\,(B - AX) \\ X &= (I - \alpha\,A)\,X + \alpha\,B \end{aligned}$$

FPI: $X^{(k+1)} = U\,X^{(k)} + V$, where $U = I - \alpha\,A$ and $V = \alpha\,B$

Sufficient conditions (on matrix A) for convergence:

- Let A be symmetric positive definite (or sym. negative definite).
  Then there exists $\alpha \in R$ such that Richardson method converges.

  Proof: Let λ_i be the eigenvalues of A, assume real, $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.

  Then $\mu_i = 1 - \alpha\,\lambda_i$ are eigenvalues of U and

  $$\rho(U) < 1 \;\Leftrightarrow\; -1 < 1 - \alpha\,\lambda_i < 1 \;\Leftrightarrow\; \alpha\,\lambda_i < 2 \text{ and } 0 < \alpha\,\lambda_i.$$

  The first inequality can always be satisfied for some α,
  the second inequality can be satisfied only if all λ_i's have the same sign.

Jacobi and Gauss-Seidel methods

Decompose given matrix A as $A = L + D + R$, where D is a diagonal matrix, L is a lower triangular and R is an upper triangular matrix. Assume that A has no zero elements on diagonal, so that inverse D^{-1} of D exists and also $(L + D)^{-1}$ exists.

Jacobi method

$$\begin{aligned} AX &= B \\ (L + D + R)\,X &= B \\ DX &= -(L + R)\,X + B \\ X &= -D^{-1}(L + R)\,X + D^{-1}B \end{aligned}$$

FPI: $X^{(k+1)} = U_J\,X^{(k)} + V_J$, where $U_J = -D^{-1}(L + R)$ and $V_J = D^{-1}B$

Gauss-Seidel method

$$\begin{aligned} AX &= B \\ (L + D + R)\,X &= B \\ (L + D)\,X &= -R\,X + B \\ X &= -(L + D)^{-1}R\,X + (L + D)^{-1}B \end{aligned}$$

FPI: $X^{(k+1)} = U_G\,X^{(k)} + V_G$, where $U_G = -(L + D)^{-1}R$ and $V_G = (L + D)^{-1}B$

.

Sufficient conditions (on matrix A) for convergence of GS and Jac. methods:

- Let A be strictly diagonally dominant.
  Then both Jacobi and Gauss-Seidel methods converge for any $X^{(0)}$. ([2] Th. 4.9)
- Let A be symmetric positive definite.
  Then Gauss-Seidel method converges for any $X^{(0)}$.
  (see [2] Th. 4.10 – G-S is a special case of SOR for $\omega = 1$)

References

[1] Tobias von Petersdorff: Contraction mapping theorem
`http://terpconnect.umd.edu/~petersd/666/fixedpoint.pdf`

[2] Y. Saad: Iterative methods for sparse linear systems
`http://www-users.cs.umn.edu/~saad/IterMethBook_2ndEd.pdf`