FPI with Systems

Example: Here's an example before we get started. Suppose we have the following system:

$$\begin{aligned} 3u - v &= 5 \\ -u + 2v &= 4 \end{aligned}$$

The idea behind what is called Jacobi's method will be to solve the first equation for the first variable, and so on:

$$\begin{aligned} u &= 5/3 + (1/3)v \\ v &= 2 + (1/2)u \end{aligned} \qquad \Rightarrow \qquad \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 & 1/3 \\ 1/2 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} 5/3 \\ 2 \end{bmatrix}$$

We see that u, v form the fixed point for the iteration:

$$\boldsymbol{x}_{n+1} = T\boldsymbol{x}_n + \boldsymbol{c}$$

Here is some Matlab code for this example:

```
T=[0 1/3;1/2 0]; c=[5/3;2];
x=rand(2,1);

%Actual solution:
soln=inv(eye(2)-T)*c;

for j=1:10
    x=T*x+c;
    err(j)=norm(x-soln);
end
```

Your answer will be slightly different (we start with a random $\boldsymbol{x}$), but after 10 iterations, we should get pretty close to the solution.

In Chapter 2.5, we look at this technique for solving systems- Of course, we have very fast methods for small or moderate sized systems of equations (solve directly), so these techniques are typically used on very large systems.

First, we'll look at the algorithm, then some theory. Let's see if we can put the Jacobi method in matrix form. Consider splitting the matrix A as $A = D + L + U$, where:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & & & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} \qquad D = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

$$L = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ a_{21} & 0 & 0 & \cdots & 0 \\ a_{31} & a_{32} & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ a_{n1} & a_{n2} & & \cdots & 0 \end{bmatrix} \qquad U = \begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ \vdots & & & & \vdots \\ 0 & 0 & & \cdots & a_{n-1,n} \\ 0 & 0 & & \cdots & 0 \end{bmatrix}$$

Now,

$$A\boldsymbol{x} = \boldsymbol{b} \quad \Rightarrow \quad (D + L + U)\boldsymbol{x} = \boldsymbol{b} \quad \Rightarrow \quad D\boldsymbol{x} = -(L + U)\boldsymbol{x} + \boldsymbol{b}$$

Therefore, $A\boldsymbol{x} = \boldsymbol{b}$ can be re-written as:

$$\boldsymbol{x} = -D^{-1}(L + U)\boldsymbol{x} + D^{-1}\boldsymbol{b}$$

which is the fixed point in the form: $\boldsymbol{x}_{n+1} = T\boldsymbol{x} + \boldsymbol{c}$. This works extremely well, because D^{-1} is so easy to compute.

Some Theory

Earlier, we talked about FPI in one dimension, and now we want to understand iterating a system. For notation, let $\boldsymbol{x} \in \mathbb{R}^n$, T be an $n \times n$ matrix and $\boldsymbol{c}$ be a constant vector in $\mathbb{R}^n$. We iterate on an initial vector $\boldsymbol{x}_0$ and produce a sequence:

$$\boldsymbol{x}_{n+1} = T\boldsymbol{x}_n + \boldsymbol{c}$$

We define the *spectral radius* of the matrix T:

$$\rho(T) = \max\{|\lambda| \mid \lambda \text{ is an eigenvalue of } T\}$$

We also need the idea of a convergent matrix: A matrix A is convergent if $\|A^n\| \to 0$ as $n \to \infty$. This implies that every element of A^n goes to zero with n.

- Find the fixed point:

$$(I - T)\boldsymbol{r} = \boldsymbol{c} \quad \Rightarrow \quad \boldsymbol{r} = (I - T)^{-1}\boldsymbol{c}$$

- **Theorem:** If $\rho(T) < 1$, then $(I - T)^{-1}$ exists, and

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

  Proof:

$$T\boldsymbol{x} = \lambda\boldsymbol{x} \text{ iff } (I - T)\boldsymbol{x} = (1 - \lambda)\boldsymbol{x}$$

  Therefore, λ is an eigenvalue of T iff $1 - \lambda$ is an eigenvalue of $I - T$.

  Suppose $I - T$ is not invertible. Then $1 - \lambda = 0$ would be one of its eigenvalues. However, this implies that $\lambda = 1$ is an eigenvalue, but $\rho(T) < 1$. Therefore, $I - T$ is invertible.

  Secondly, let $S_m = I + T + T^2 + \cdots + T^m$. Then

$$(I - T)S_m = I - T^{m+1}$$

  Since $\rho(T) < 1$, the matrix T is convergent. Therefore,

$$\lim_{m\to\infty}(I - T)S_m = \lim_{m\to\infty} I - T^{m+1} = I$$

Therefore,

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

NOTE: This looks just like a geo series!

$$\frac{1}{1 - x} = 1 + x + x^2 + x^3 + \cdots$$

- **Theorem:** For any $\boldsymbol{x}_0 \in \mathbb{R}^n$, the sequence

$$\boldsymbol{x}_{n+1} = T\boldsymbol{x}_n + \boldsymbol{c}$$

converges to the fixed point $\boldsymbol{r}$ if and only if $\rho(T) < 1$.

Proof: We'll prove only one direction. First, assume that $\rho(T) < 1$. By an exercise,

$$\boldsymbol{x}_n = T^n \boldsymbol{x}_0 + \left(T^{n-1} + T^{n-2} + \cdots + T^2 + T + I\right) \boldsymbol{c}$$

and take the limit of both sides to see that

$$\lim_{n\to\infty} \boldsymbol{x}_n = \lim_{n\to\infty} T^n \boldsymbol{x}_0 + \lim_{n\to\infty} \left(T^{n-1} + T^{n-2} + \cdots + T^2 + T + I\right) \boldsymbol{c} =$$

$$\boldsymbol{0} + (I - T)^{-1}\boldsymbol{c}$$

Exercises

1. Let $\boldsymbol{x}_{n+1} = T\boldsymbol{x} + \boldsymbol{c}$. Show that

$$\boldsymbol{x}_n = T^n \boldsymbol{x}_0 + \left(T^{n-1} + T^{n-2} + \cdots + T^2 + T + I\right) \boldsymbol{c}$$

2. Show that, if $\boldsymbol{x}_{n+1} = T\boldsymbol{x}_n + \boldsymbol{c}$ and $\boldsymbol{r} = T\boldsymbol{r} + \boldsymbol{c}$, then using an induced matrix norm,

$$\|\boldsymbol{x}_k - \boldsymbol{r}\| \leq \|T\|^k \|\boldsymbol{x}_0 - \boldsymbol{r}\|$$

(Notice that this implies that, if $\|T\| < 1$, then the sequence $\boldsymbol{x}_n$ converges to the fixed point).

3. Exercise 1(a), 1(c) (only Jacobi). Also see the code online that will split the matrix A for you.

Appendix

The Spectral Radius and the Matrix Norm

The following theorems are presented here without proof. They are presented here to give you some sense of the relationship between the spectrum (or more specifically, the spectral radius), and matrix norms. The matrix A is $n \times n$, and the norm (if not stated) is an induced norm.

1. $\|A\|_2 = \sqrt{\rho(A^T A)}$
2. $\rho(A) \leq \|A\|$ for any induced norm $\|\cdot\|$.
3. The following statements are equivalent (therefore, if any of them are true, all are true):
    - A is convergent
    - $\rho(A) < 1$
    - $\lim_{n \to \infty} A^n \boldsymbol{x} = \mathbf{0}$ for every $\boldsymbol{x}$.