APPM 4650 — Fixed point for non-linear systems

Consider the problem of find $\begin{bmatrix}\alpha\\ \beta\end{bmatrix}$ such that

$$\begin{bmatrix}\alpha\\ \beta\end{bmatrix} = \begin{bmatrix} f(\alpha,\beta)\\ g(\alpha,\beta)\end{bmatrix}.$$

As in one dimension, we can create a fixed point iteration

$$\begin{bmatrix}x_{n+1}\\ y_{n+1}\end{bmatrix} = \begin{bmatrix} f(x_n,y_n)\\ g(x_n,y_n)\end{bmatrix} \quad \text{for } n=0,1,2,\dots$$

How do we know when this will converge?

Let's try to mimic what we did for one dimensional problems. Assume f and g are analytic. (This is more than sufficient.)

Let $\begin{bmatrix}x_n\\ y_n\end{bmatrix} = \begin{bmatrix}\alpha+\Delta x_n\\ \beta+\Delta y_n\end{bmatrix}$

Then the fixed point iteration becomes

$$\begin{bmatrix}x_{n+1}\\ y_{n+1}\end{bmatrix} = \begin{bmatrix} f(x_n,y_n)\\ g(x_n,y_n)\end{bmatrix}$$
$$\begin{bmatrix}\alpha+\Delta x_{n+1}\\ \beta+\Delta y_{n+1}\end{bmatrix} = \begin{bmatrix} f(\alpha+\Delta x_n,\beta+\Delta y_n)\\ g(\alpha+\Delta x_n,\beta+\Delta y_n)\end{bmatrix}$$

Now we can Taylor expand the right hand side centering at $\begin{bmatrix}\alpha\\ \beta\end{bmatrix}$. This results in

$$\begin{bmatrix}\alpha\\ \beta\end{bmatrix} + \begin{bmatrix}\Delta x_{n+1}\\ \Delta y_{n+1}\end{bmatrix} = \begin{bmatrix} f(\alpha,\beta)\\ g(\alpha,\beta)\end{bmatrix} + \underbrace{\begin{bmatrix} f_x & f_y\\ g_x & g_y\end{bmatrix}|_{(\alpha,\beta)}}_{\mathsf{G}} \begin{bmatrix}\Delta x_n\\ \Delta y_n\end{bmatrix} + \text{ higher order terms}$$

where f_x denotes the partial derivative of f with respect to x, etc.

Since $\begin{bmatrix}\alpha\\ \beta\end{bmatrix}$ is the fixed point $\begin{bmatrix} f(\alpha,\beta)\\ g(\alpha,\beta)\end{bmatrix} = \begin{bmatrix}\alpha\\ \beta\end{bmatrix}$. The matrix G is called the *Jacobian.*

$$\begin{bmatrix}\alpha\\ \beta\end{bmatrix} + \begin{bmatrix}\Delta x_{n+1}\\ \Delta y_{n+1}\end{bmatrix} = \begin{bmatrix}\alpha\\ \beta\end{bmatrix} + \underbrace{\begin{bmatrix} f_x & f_y\\ g_x & g_y\end{bmatrix}|_{(\alpha,\beta)}}_{\mathsf{G}} \begin{bmatrix}\Delta x_n\\ \Delta y_n\end{bmatrix} + \text{ higher order terms}$$

Dropping the higher order terms, we find

$$\begin{bmatrix}\Delta x_{n+1}\\ \Delta y_{n+1}\end{bmatrix} = \mathsf{G}\begin{bmatrix}\Delta x_n\\ \Delta y_n\end{bmatrix} = \mathsf{G}^n\begin{bmatrix}\Delta x_0\\ \Delta y_0\end{bmatrix}$$

So the convergence of the fixed point iteration depends on what happens to the powers of G.

To get an idea of what we need for convergence, let's recall some linear algebra.

Without loss of generality, let's assume that G is full rank. Let $\mathsf{G} \in \mathbb{R}^{m\times m}$. (In what we have done so far $m = 2$ but everything extends to other m.) Since G is full rank, it's eigenvectors form a basis for $\mathbb{R}^m$. Let $\{(\lambda_j, \boldsymbol{v}_j\}_{j=1}^m$ denote the eigenpairs. Then for all $\boldsymbol{x} \in \mathbb{R}^m$, there exists constants $\{\alpha_j\}_{j=1}^m$ such that $\boldsymbol{x} = \sum_{j=1}^m \alpha_j \boldsymbol{v}_j$.
This means that

$$\begin{aligned}
\mathsf{G}\boldsymbol{x} &= \mathsf{G}\sum_{j=1}^m \alpha_j \boldsymbol{v}_j \\
&= \sum_{j=1}^m \alpha_j \mathsf{G}\boldsymbol{v}_j \\
&= \sum_{j=1}^m \alpha_j \lambda_j \boldsymbol{v}_j
\end{aligned}$$

Thus

$$\mathsf{G}^n\boldsymbol{x} = \sum_{j=1}^m \alpha_j \lambda_j^n \boldsymbol{v}_j$$

The sum converges when $|\lambda_j| \leq 1$.
Really we want $\mathsf{G}^n \begin{bmatrix} \Delta x_0 \\ \Delta y_0 \end{bmatrix} \to 0$. This means that we need $|\lambda_j| < 1$. We need a stricter condition because we need

$$\begin{bmatrix} \Delta x_{n+1} \\ \Delta y_{n+1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In practice, we can look at other norms too.

$$\left\| \begin{bmatrix} \Delta x_{n+1} \\ \Delta y_{n+1} \end{bmatrix} \right\| = \left\| \mathsf{G}^n \begin{bmatrix} \Delta x_0 \\ \Delta y_0 \end{bmatrix} \right\| \leq \|\mathsf{G}\|^n \left\| \begin{bmatrix} \Delta x_0 \\ \Delta y_0 \end{bmatrix} \right\|$$

It is sufficent for $\|\mathsf{G}\| < 1$. In practice, most people use the l_2 norm.
Theorem 10.6 in the text has a more formal statement of the following.

Theorem 0.1. $\|\mathsf{G}\| \geq \rho(\mathsf{G})$ *where* $\rho(\mathsf{G})$ *denotes the spectral radius defined by* $\rho(\mathsf{G}) = \max_j |\lambda_j|$.