Roots and fixed points of vector-valued functions

A function $\mathbf{g}(\mathbf{x})$ from $\mathbb{R}^n$ to $\mathbb{R}^n$ has a fixed point at $\mathbf{p}$ if $\mathbf{g}(\mathbf{p}) = \mathbf{p}$.

Here is a fixed point theorem for vector-valued functions.

Theorem Let D be a closed, convex region in $\mathbb{R}^n$. Suppose $\mathbf{g}(\mathbf{x})$ is a continuous function that maps D into D. Then $\mathbf{g}(\mathbf{x})$ has a fixed point $\mathbf{p}$ in D. Further, suppose that all the component functions of $\mathbf{g}(\mathbf{x})$ have continuous first partial derivatives in all variables and there is a constant $K < 1$ such that

$$\left|\frac{\partial g_i(\mathbf{x})}{\partial x_j}\right| < \frac{K}{n}$$

for all i and j and all $\mathbf{x}$ in D. Then any sequence of iterates

$$\mathbf{x}^{(k+1)} = \mathbf{g}(\mathbf{x}^{(k)})$$

starting from any $\mathbf{x}^{(0)}$ in D converges to a unique fixed point $\mathbf{p}$ in D and

$$\|\mathbf{x}^{(k)} - \mathbf{p}\|_\infty \leq \frac{K^k}{1 - K} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|_\infty$$

Newton's Method

Let

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(x_1,x_2,\ldots,x_n) \\ f_2(x_1,x_2,\ldots,x_n) \\ \vdots \\ f_n(x_1,x_2,\ldots,x_n) \end{bmatrix}$$

be a function from $\mathbb{R}^n$ to $\mathbb{R}^n$. We say that $\mathbf{f}(\mathbf{x})$ has a root at $\mathbf{x}$ if $\mathbf{f}(\mathbf{x}) = \mathbf{0}$.

We are going to try to construct a root finding algorithm by constructing an argument that is analagous to the argument in chapter 2 that led to Newton's method.

Here is an analog of a theorem we saw in chapter 2.

Theorem Let $\mathbf{p}$ be a solution of $\mathbf{g}(\mathbf{x}) = \mathbf{x}$. Suppose that a number $\delta > 0$ exists with

1. $\dfrac{\partial g_i}{\partial x_j}$ is continuous on $N_\delta = \{\mathbf{x} | \|\mathbf{x} - \mathbf{p}\| < \delta\}$ for each $1 \leq i \leq n$ and $1 \leq j \leq n$.
2. $\dfrac{\partial^2 g_i(\mathbf{x})}{\partial x_j \partial x_k}$ is continuous, and $\left|\dfrac{\partial^2 g_i(\mathbf{x})}{\partial x_j \partial x_k}\right| \leq M$ for some constant M, whenever $\mathbf{x} \in N_\delta$ for each $1 \leq i \leq n$, $1 \leq j \leq n$, and $1 \leq k \leq n$.
3. $\dfrac{\partial g_i(\mathbf{p})}{\partial x_k} = 0$ for each $1 \leq i \leq n$ and $1 \leq k \leq n$.

Then a number $\hat{\delta} \leq \delta$ exists such that the sequence generated by $\mathbf{x}^{(k)} = \mathbf{g}(\mathbf{x}^{(k-1)})$ converges quadratically to $\mathbf{p}$ for any choice of $\mathbf{x}^{(0)}$, provided that $\|\mathbf{x}^{(0)} - \mathbf{p}\| < \delta$. Moreover,

$$\|\mathbf{x}^{(k)} - \mathbf{p}\|_\infty \le \frac{n^2 M}{2} \|\mathbf{x}^{(k-1)} - \mathbf{p}\|_\infty$$

for each $k \ge 1$.

We can use this theorem to construct Newton's method on $\mathbb{R}^n$ by seeking an n by n matrix function $\varphi(\mathbf{x})$ such that

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} - \varphi(\mathbf{x})\, \mathbf{f}(\mathbf{x})$$

satisfies the conditions of the theorem. We take partial derivatives of the coordinate functions of $\mathbf{g}(\mathbf{x})$ and see that

$$\frac{\partial g_i(\mathbf{x})}{\partial x_k} = 1 - \sum_{j=1}^{n} \left(\varphi_{i,j}(\mathbf{x}) \frac{\partial f_j(\mathbf{x})}{\partial x_k} + \frac{\partial \varphi_{i,j}(\mathbf{x})}{\partial x_k} f_j(\mathbf{x}) \right) \text{ for } i = k$$

$$\frac{\partial g_i(\mathbf{x})}{\partial x_k} = \sum_{j=1}^{n} \left(\varphi_{i,j}(\mathbf{x}) \frac{\partial f_j(\mathbf{x})}{\partial x_k} + \frac{\partial \varphi_{i,j}(\mathbf{x})}{\partial x_k} f_j(\mathbf{x}) \right) \text{ for } i \ne k$$

When $\mathbf{p}$ is a root of $\mathbf{f}(\mathbf{x})$, $\mathbf{p}$ is a fixed point of $\mathbf{g}(\mathbf{x})$ and we would like to have

$$0 = \frac{\partial g_i(\mathbf{p})}{\partial x_k} = 1 - \sum_{j=1}^{n} \varphi_{i,j}(\mathbf{p}) \frac{\partial f_j(\mathbf{p})}{\partial x_k} \text{ for } i = k$$

$$0 = \frac{\partial g_i(\mathbf{p})}{\partial x_k} = \sum_{j=1}^{n} \varphi_{i,j}(\mathbf{p}) \frac{\partial f_j(\mathbf{p})}{\partial x_k} \text{ for } i \ne k$$

If we look closely at these conditions, they tell us that the matrix $\varphi(\mathbf{p})$ is the inverse of the matrix of partial derivatives

$$J(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(\mathbf{x})}{\partial x_1} & \frac{\partial f_n(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_n(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

evaluated at $\mathbf{x} = \mathbf{p}$. This matrix is the *Jacobian matrix* for the function $\mathbf{f}(\mathbf{x})$ whose root we are trying to find.

The Newton iteration formula is

$$\mathbf{x}^{(k)} = \mathbf{x}^{(k-1)} - J^{-1}(\mathbf{x}^{(k-1)})\, \mathbf{f}(\mathbf{x}^{(k-1)})$$