Variation of Parameters Formula and Adapted Norm

Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be a continuously differentiable map and $\bar{q}$ be a nonsingular fixed point, i.e., the linearization $Df(\bar{q})$ is invertible. If $Df(\bar{q})$ is similar to a block diagonal matrix

$$Df(\bar{q}) \sim \operatorname{diag}(A, B)$$

then we can assume that after translating the fixed point to the origin 0 there is a coordinate system $(x, y) \in \mathbb{R}^d$ so that the mapping $(\bar{x}, \bar{y}) = f(x, y)$ can be written as

$$\begin{cases} \bar{x} = f_1(x, y) = Ax + h(x, y) \\ \bar{y} = f_2(x, y) = By + \tilde{h}(x, y). \end{cases} \tag{1}$$

By the Global Inverse Function Theorem we know that for $H(p) = f(p) - Df(\bar{q})p$ if $\|H\|_1 < \delta$ and δ sufficiently small, then f is globally invertible with inverse $f^{-1} = Df(\bar{q})^{-1} + G$ and $\|G\|_1 < \epsilon$ and $\epsilon = O(\delta)$ (i.e., $\lim_{\delta \to 0} \epsilon = 0$). So for $(x, y) = f^{-1}(\bar{x}, \bar{y})$ we can write

$$\begin{cases} x = [f^{-1}]_1(\bar{x}, \bar{y}) = A^{-1}\bar{x} + \tilde{g}(\bar{x}, \bar{y}) \\ y = [f^{-1}]_2(\bar{x}, \bar{y}) = B^{-1}\bar{y} + g(\bar{x}, \bar{y}). \end{cases} \tag{2}$$

Lemma 1 (Variation of Parameters Formula for Map). *Let $\bar{q}$ be a nonsingular fixed point of a continuously differentiable map f in $\mathbb{R}^d$. If $Df(\bar{q}) \sim \operatorname{diag}(A, B)$ for some matrixes A, B, then for sufficiently small $\delta > 0$, $\|f - Df(\bar{q})\|_1 < \delta$ implies there is a coordinate system so that $\bar{p} = f(p)$ is equivalent to*

$$\begin{cases} \bar{x} = Ax + h(x, y) \\ y = B^{-1}\bar{y} + g(\bar{x}, \bar{y}), \end{cases} \tag{3}$$

with the properties that $h(0,0) = 0$, $g(0,0) = 0$, $Dh(0,0) = 0$, $Dg(0,0) = 0$, and $\|(h, g)\|_1 \to 0$ as $\delta \to 0$. Moreover, the following variation of parameter formula holds for any orbit $(x_{n+1}, y_{n+1}) = f(x_n, y_n)$, $n \in \mathbb{Z}$

$$\begin{cases} x_n = A^{n-\ell}x_\ell + \sum_{i=\ell+1}^{n} A^{n-i}h(x_{i-1}, y_{i-1}) \\ y_n = B^{n-m}y_m + \sum_{i=n+1}^{m} B^{n+1-i}g(x_i, y_i) \end{cases} \tag{4}$$

for any $\ell \le n \le m$.

Proof. The discussion preceding the lemma shows that for sufficiently small δ, $\bar{p} = f(p)$ is equivalent to Eq.(1) and Eq.(2), together they imply Eq.(3). Also, it follows from the discussion that $h(0,0) = 0$, $g(0,0) = 0$, $Dh(0,0) = 0$, $Dg(0,0) = 0$, and $\|(h, g)\|_1 \to 0$ as $\delta \to 0$.

Conversely, assume a pair of points (x, y), $(\bar{x}, \bar{y})$ satisfy Eq.(3). Recall from the Global Inverse Function Theorem that for $H = f - J$ with $J = Df(\bar{q})$, the inverse of $f = J + H$ can be written as $f^{-1} = J^{-1} + G$ with $G = -J^{-1} \circ H \circ f^{-1}$.

In the coordinate system (x, y) for which $J = \mathrm{diag}(A, B)$, $H = (h, \tilde{h})$, $G = (\tilde{g}, g)$, we have

$$g = -B^{-1} \circ \tilde{h} \circ f^{-1}.$$

Hence, the second equation of Eq.(3) can be written as

$$\bar{y} = By - Bg(\bar{x}, \bar{y}) = By + \tilde{h} \circ f^{-1}(\bar{x}, \bar{y}).$$

Pairing it with the first equation of Eq.(3) we have

$$\begin{cases} \bar{x} = Ax + h(x, y) \\ \bar{y} = By + \tilde{h} \circ (J^{-1} + G)(\bar{x}, \bar{y}). \end{cases} \tag{5}$$

Treating it as a fixed point for the mapping defined by the right side of equation, $\bar{p} = S(\bar{p}, p)$, $S : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, we can conclude that if

$$\|H\|_1 (\|J\| + \|G\|_1) < 1 \tag{6}$$

then $S(\cdot, p)$ is a uniform contraction. Hence, for every $p = (x, y)$, there is a unique fixed point $(\bar{x}, \bar{y}) = f^*(x, y)$ parameterized by (x, y). Because $(\bar{x}, \bar{y}) = f(x, y)$ obviously satisfies Eq.(5), it is a fixed point of $S(\cdot, p)$. Therefore, by the uniqueness of the fixed point we must have $(\bar{x}, \bar{y}) = f(x, y) = f^*(x, y)$, proving the equivalence of Eq.(3) to Eq.(1). Because the assumption $\|f - Df(\bar{q})\|_1 < \delta$ implies (6) for small δ, the equivalence indeed holds.

As a consequence to Eq.(3), for any orbit $\gamma = \{p_n\}_{n=-\infty}^{\infty}$, with $(x_{n+1}, y_{n+1}) = f(x_n, y_n)$, we use the first equation of Eq.(3) to write

$$x_n = Ax_{n-1} + h(x_{n-1}, y_{n-1})$$

and then recursively apply it to itself to obtain the first equation of Eq.(4). Similarly, we use the second equation of Eq.(3) to write

$$y_n = B^{-1} y_{n+1} + g(x_{n+1}, y_{n+1})$$

and then recursively apply it to itself to obtain the second equation of Eq.(4). $\quad\square$

Denote $J = Df(\bar{q})$, $\sigma(J)$ the set of eigenvalues of the linearization, counting multiplicity. Denote $\sigma^s(J) = \sigma(J) \cap \{|z| < 1\}, \sigma^c(J) = \sigma(J) \cap \{|z| = 1\}, \sigma^u(J) = \sigma(J) \cap \{|z| > 1\}$ the set of eigenvalues inside, on, outside the unit circle, respectively. Denote $\mathbb{E}^s, \mathbb{E}^c, \mathbb{E}^u$ the corresponding generalized eigenspaces for eigenvalues of $\sigma^s, \sigma^c, \sigma^u$, respectively. Then $\mathbb{R}^d \cong \mathbb{E}^s \oplus \mathbb{E}^c \oplus \mathbb{E}^u$. In fact, the phase space can be split or combined in other different ways. Two splits we will need later are $\mathbb{R}^d \cong \mathbb{E}^{cs} \oplus \mathbb{E}^u$, $\mathbb{R}^d \cong \mathbb{E}^s \oplus \mathbb{E}^{cu}$, with $\mathbb{E}^{cs} = \mathbb{E}^s \oplus \mathbb{E}^c$ and $\mathbb{E}^{cu} = \mathbb{E}^c \oplus \mathbb{E}^u$, corresponding to $\sigma^{cs} = \sigma^s \cup \sigma^c, \sigma^{cu} = \sigma^c \cup \sigma^u$, etc. Let $d_i = \#(\sigma^i) = \dim(\mathbb{E}^i)$. Then $d = d_{cs} + d_u$, $d_{cs} = d_s + d_c$, etc. Also, $\mathbb{E}^{d_i} \cong \mathbb{R}^{d_i}$, for $i = s, c, u, sc, su$.

Depending on applications, a coordinate system (x, y) can be chosen so that $Df(\bar{q}) = \mathrm{diag}(A, B)$ with eigenvalues of A the set σ^{cs} and those of B the set

σ^u. Or A's eigenvalues are from σ^s, and B's eigenvalues are from σ^{cu}. Or in the case of hyperbolic fixed points, A's eigenvalues are from σ^s, and B's eigenvalues are from σ^u, because $\sigma^c = \varnothing$. Or a coordinate system (x, y, z) so that $Df(\bar{q}) = \mathrm{diag}(A, C, B)$ with $\sigma(A) = \sigma^s$, $\sigma(C) = \sigma^c$, and $\sigma(B) = \sigma^u$. In any case, for $\|f - Df(\bar{q})\|_1 < \delta$ with sufficiently small δ, the Variation of Parameters Formula (Lemma 1) applies. For the cases of two-matrixes splits for the linearization, functions h, g are all C^1 satisfying

$$h(0,0) = 0, \ Dh(0,0) = 0, \ g(0,0) = 0, \ Dg(0,0) = 0 \tag{7}$$

and they are globally Lipschitz with Lipschitz constants satisfying

$$L = \max\{\mathrm{Lip}(h), \mathrm{Lip}(g)\} \to 0 \quad \text{as} \ \|f - Df(\bar{q})\|_1 \to 0. \tag{8}$$

The same conditions also hold for $\tilde{h}, \tilde{g}$ but we do not need them usually. The Variation of Parameters Formula can also be generalized to three-matrixes split cases for the linearization.

We can further assume the coordinate is chosen so that the matrixes A, B, etc., are in their Jordan canonical forms. Specifically, for example, matrix A is a block-diagonal matrix, $A = \mathrm{diag}(D_1, \ldots, D_k)$ with each D_i being one of the following forms:

$$\begin{bmatrix} D & N & 0 & \ldots & 0 \\ 0 & D & N & \ldots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \ldots & D & N \\ 0 & 0 & \ldots & 0 & D \end{bmatrix}$$

where either $D = \lambda, N = 0$, or $D = \lambda, N = \epsilon$, or $D = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, $N = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, or $D = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, $N = \begin{bmatrix} \epsilon & 0 \\ 0 & \epsilon \end{bmatrix}$, with λ or $a + ib$ the eigenvalues of D_i, and ϵ being an arbitrarily small but nonzero number. Similar forms for B or for any other splitting also hold. Take a hyperbolic case as an example for which $\sigma^c = \varnothing$, $\sigma(A) = \sigma^s$, $\sigma(B) = \sigma^u$. Then for any fixed but arbitrary constants α, β satisfying

$$\max\{\sigma(A)\} < \alpha < 1 < \beta < \min\{\sigma(B)\}$$

we can choose a sufficiently small ϵ *a priori* and then a coordinate system (x, y) so that with respect the Euclidean norm for (x, y) the matrix norms for A and B satisfy

$$\|A\| < \alpha < 1 < \beta < \|B\| \ \text{ and } \|B^{-1}\| < 1/\beta. \tag{9}$$

Such a norm is referred to as an *adapted norm* for the linearization. Take a non-hyperbolic case for which $\sigma(A) = \sigma^{cs}$ and $\sigma(B) = \sigma^u$. Then for any constants α, β

$$\max\{\sigma(A)\} = 1 < \beta < \frac{1}{\alpha} < \min\{\sigma(B)\}$$

we can choose again a sufficiently small ϵ *a priori* and then a coordinate system (x, y) so that the matrix norms for A and B satisfy

$$\|A\| < \beta \ \text{ and } \|B^{-1}\| < \alpha < 1. \tag{10}$$

The last case as an example is for the splitting $\sigma(A) = \sigma^s$ and $\sigma(B) = \sigma^{cu}$, for which for any constants α, β

$$\max\{\sigma(A)\} < \alpha < 1 \ \text{ and } \ \max\{\sigma(B^{-1})\} = 1 < \beta < \frac{1}{\alpha}$$

we can choose an adapted coordinate (x, y) so that

$$\|A\| < \alpha < 1 \ \text{ and } \|B^{-1}\| < \beta. \tag{11}$$

Such coordinates will prove to be convenient in analyses of invariant manifolds of diffeomorphic maps.