Math 346 Lecture #10
7.4 The Implicit and Inverse Function Theorems

7.4.1 Implicit Function Theorem

For $A, B \in M_n(\mathbb{F})$, $c \in \mathbb{F}^n$, variables $\mathrm{y}, \mathrm{x} \in \mathbb{F}^n$, and $F(\mathrm{x}, \mathrm{y}) = A\mathrm{y} + B\mathrm{x}$, what condition is required to solve

$$c = F(\mathrm{x}, \mathrm{y})$$

for y as a function of x? It is the invertibility of $D_2F(\mathrm{x}, \mathrm{y}) = A$ from which we get

$$\mathrm{y} = A^{-1}(c - B\mathrm{x}),$$

i.e., y is a function of x on the level set $\mathrm{c} = F(\mathrm{x}, \mathrm{y})$.

For general $F : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}^n$, solving $c = F(\mathrm{x}, \mathrm{y})$ for y as a function of x, requires the invertibility of

$$D_2F(\mathrm{x}_0, \mathrm{y}_0)$$

at some point $(\mathrm{x}_0, \mathrm{y}_0)$ at which $c = F(\mathrm{x}_0, \mathrm{y}_0)$, to get that y is locally a function of x in a neighbourhood of x_0.

Theorem 7.4.2 (Implicit Function Theorem). For Banach spaces $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, and $(Z, \|\cdot\|_Z)$, let U be an open neighbourhood of $\mathrm{x}_0 \in X$, V an open neighbourhood of y_0 in Y, and $F \in C^k(U \times V, Z)$ for some $k \geq 1$. For $\mathrm{z}_0 = F(\mathrm{x}_0, \mathrm{y}_0)$, if $D_2F(\mathrm{x}_0, \mathrm{y}_0) \in \mathscr{B}(Y, Z)$ has bounded inverse, then there exists an open neighbourhood $U_0 \times V_0 \subset U \times V$ with $(\mathrm{x}_0, \mathrm{y}_0) \in U_0 \times V_0$, and a unique $f \in C^k(U_0, V_0)$ such that $f(\mathrm{x}_0) = \mathrm{y}_0$,

$$\{(\mathrm{x}, \mathrm{y}) \in U_0 \times V_0 : F(\mathrm{x}, \mathrm{y}) = \mathrm{z}_0\} = \{(\mathrm{x}, f(\mathrm{x})) : \mathrm{x} \in U_0\},$$

and for all $\mathrm{x} \in U_0$ there holds

$$Df(\mathrm{x}) = -D_2F(\mathrm{x}, f(\mathrm{x}))^{-1} D_1F(\mathrm{x}, f(\mathrm{x})).$$

Proof. By replacing $F(\mathrm{x}, \mathrm{y})$ with $F(\mathrm{x}, \mathrm{y}) - \mathrm{z}_0$ (which doesn't change derivatives), we assume WLOG that $F(\mathrm{x}_0, \mathrm{y}_0) = 0$.

Applying the quasi-Newton method for each fixed $\mathrm{x} \in U$ we define a function $G \in C^k(U \times V, Y)$ given by

$$G(\mathrm{x}, \mathrm{y}) = \mathrm{y} - D_2F(\mathrm{x}_0, \mathrm{y}_0)^{-1} F(\mathrm{x}, \mathrm{y}).$$

The bounded linear map $D_2F(\mathrm{x}_0, \mathrm{y}_0)^{-1}$ is invertible by hypothesis.

This implies that for each fixed $\mathrm{x} \in U$, we have $G(\mathrm{x}, \mathrm{y}) = \mathrm{y}$ if and only if $F(\mathrm{x}, \mathrm{y}) = 0$.

The derivative of G with respect to y evaluated at $(\mathrm{x}_0, \mathrm{y}_0)$ is

$$D_2G(\mathrm{x}_0, \mathrm{y}_0) = I - D_2F(\mathrm{x}_0, \mathrm{y}_0)^{-1} D_2F(\mathrm{x}_0, \mathrm{y}_0) = 0.$$

Since G is C^k on $U \times V$, there exist open sets $U_1 \subset U$ and $V_0 \subset V$ such that $\mathrm{x}_0 \in U_1$, $\mathrm{y}_0 \in V_0$, and for all $(\mathrm{x}, \mathrm{y}) \in U_1 \times \overline{V}_0$ there holds

$$\|D_2G(\mathrm{x}, \mathrm{y})\| < \frac{1}{2}.$$

WLOG we may assume that $V_0 = B(\mathrm{y}_0, \delta)$ for some $\delta > 0$.

Since F is C^k and vanishes at $(\mathrm{x}_0, \mathrm{y}_0)$, i.e., $F(\mathrm{x}_0, \mathrm{y}_0) = 0$, there exists an open $U_0 \subset U_1$ such that $\mathrm{x}_0 \in U_0$ and for all $\mathrm{x} \in U_0$ there holds

$$\|D_2F(\mathrm{x}_0, \mathrm{y}_0)^{-1}\| \, \|F(\mathrm{x}, \mathrm{y}_0\| < \frac{\delta}{2}.$$

By the triangle inequality, Integral Mean Value Theorem, and the definition of G we have

$$\begin{aligned}
\|G(\mathrm{x}, \mathrm{y}) - \mathrm{y}_0\| &\leq \|G(\mathrm{x}, \mathrm{y}) - G(\mathrm{x}, \mathrm{y}_0)\| + \|G(\mathrm{x}, \mathrm{y}_0) - \mathrm{y}_0\| \\
&\leq \sup_{\mathrm{c} \in \ell(\mathrm{y}_0, \mathrm{y})} \|D_2G(\mathrm{x}, \mathrm{c})\| \, \|\mathrm{y} - \mathrm{y}_0\| + \|D_2F(\mathrm{x}_0, \mathrm{y}_0)^{-1}F(\mathrm{x}, \mathrm{y}_0)\| \\
&\leq \frac{1}{2}\|\mathrm{y} - \mathrm{y}_0\| + \|D_2F(\mathrm{x}_0, \mathrm{y}_0)^{-1}\| \, \|F(\mathrm{x}, \mathrm{y}_0)\| \\
&< \frac{\delta}{2} + \frac{\delta}{2} \\
&= \delta
\end{aligned}$$

whenever $(\mathrm{x}, \mathrm{y}) \in U_0 \times \overline{V}_0$.

This implies that $G(\mathrm{x}, \mathrm{y}) \in \overline{B(\mathrm{y}_0, \delta)}$, so that G maps $U_0 \times \overline{V}_0$ to $\overline{V}_0$.

Applying the Integral Mean Value Theorem, for $\mathrm{x} \in U_0$ and $\mathrm{y}_1, \mathrm{y}_2 \in \overline{V}_0$ we have

$$\begin{aligned}
\|G(\mathrm{x}, \mathrm{y}_1) - G(\mathrm{x}, \mathrm{y}_2)\| &\leq \sup_{\lambda \in [0,1]} \|D_2G(\mathrm{x}, \lambda\mathrm{y}_1 + (1 - \lambda)\mathrm{y}_2)\| \, \|\mathrm{y}_1 - \mathrm{y}_2\| \\
&\leq \frac{1}{2}\|\mathrm{y}_1 - \mathrm{y}_2\|.
\end{aligned}$$

This implies that $G : U_0 \times \overline{V}_0 \to \overline{V}_0$ is a uniform contraction.

By the Uniform Contraction Mapping Principle (Theorem 7.2.4) there exists a unique $f \in C^k(U_0, \overline{V}_0)$ such that for all $\mathrm{x} \in U_0$ there holds

$$G(\mathrm{x}, f(\mathrm{x})) = f(\mathrm{x}).$$

By the equivalence of $G(\mathrm{x}, \mathrm{y}) = \mathrm{y}$ if and only if $F(\mathrm{x}, \mathrm{y}) = 0$, we obtain

$$F(\mathrm{x}, f(\mathrm{x})) = 0 \text{ for all } \mathrm{x} \in U_0.$$

Since $\|G(\mathrm{x}, \mathrm{y}) - \mathrm{y}_0\| < \delta$ on $U_0 \times \overline{V}_0$ we can restrict the codomain of f to V_0.

Differentiating the C^k function $F(\mathrm{x}, f(\mathrm{x})) = 0$ on U_0 gives

$$D_1F(\mathrm{x}, f(\mathrm{x})) + D_2F(\mathrm{x}, f(\mathrm{x}))Df(\mathrm{x}) = 0.$$

Since $D_2F(\mathrm{x}_0, \mathrm{y}_0)$ has bounded inverse, there exists a possible smaller choice of $U_0 \times V_0$ on which $D_2F(\mathrm{x}, \mathrm{y})$ has bounded inverse (see Lemma 7.3.11).

Thus we obtain $Df(\mathrm{x}) = -D_2F(\mathrm{x}, f(\mathrm{x}))D_1F(\mathrm{x}, f(\mathrm{x}))$ for all $\mathrm{x} \in U_0$. $\square$

Note. The condition of $D_2F(\mathrm{x}_0, \mathrm{y}_0) \in \mathscr{B}(Y, Z)$ having bounded inverse not only implies that $D_2F(\mathrm{x}_0, \mathrm{y}_0)$ is an isomorphism from Y to Z, but that it is a Banach space isomorphism, i.e., $D_2F(\mathrm{x}_0, \mathrm{y}_0)$ is also a homeomorphism from Y to Z. This means that Y and Z must be topologically equivalent as topological spaces in order to apply the Implicit Function Theorem. In finite dimensions this requires that Y and Z must have the same dimension. In most applications of the Implicit Function Theorem you will notice that $Y = Z$ so that $D_2F(\mathrm{x}_0, \mathrm{y}_0) \in \mathscr{B}(Y)$ and the condition that $D_2F(\mathrm{x}_0, \mathrm{y}_0)$ have bounded inverse means that $D_2F(\mathrm{x}_0, \mathrm{y}_0) \in \mathrm{GL}(Y)$. When Y is finite dimensional this means that $D_2F(\mathrm{x}_0, \mathrm{y}_0)$, when represented in coordinates, is an invertible matrix.

Example (in lieu of 7.4.5). Consider the surface in $\mathbb{R}^3$ that is the 0-level set of the function

$$F(x, y, z) = x^3 + 3y^2 + 8xz^2 - 3yz^3 - 9.$$

Find the points (x_0, y_0, z_0) near which the surface $F = 0$ is locally the graph of a function $z = f(x, y)$.

Since it is z we want to write as a function of (x, y) we want those points (x_0, y_0, z_0) on $F = 0$ where $D_3F(x_0, y_0, z_0)$ has bounded inverse.

Here $D_3F(x_0, y_0, z_0) \in \mathscr{L}(\mathbb{R})$, i.e., a 1×1 matrix, which has bounded inverse if and only if the derivative is nonzero.

We compute

$$D_3F(x, y, z) = 16xz - 9yz^2 = z(16x - 9yz).$$

The surface is locally the graph of a function $z = f(x, y)$ at those points (x_0, y_0, z_0) for which

$$D_3F(x_0, y_0, z_0) = z_0(16x_0 - 9y_0z_0) \neq 0.$$

The point $(1, 1, 1)$ lies on the surface $F = 0$ because $F(1, 1, 1) = 1 + 3 + 8 - 3 - 9 = 0$.

Since $D_3F(1, 1, 1) = 16 - 9 = 5 \neq 0$, the surface is locally the graph of a function $z = f(x, y)$ near $(1, 1, 1)$.

Example (in lieu of 7.4.6). Can we uniquely solve

$$\begin{aligned} xu + yvu^2 &= 2, \\ xu^3 + y^2v^4 &= 2, \end{aligned}$$

for (u, v) as a function of (x, y) near the point $(x, y, u, v) = (1, 1, 1, 1)$?

Check that the equations are satisfied at the given point $(1, 1, 1, 1)$. ✓

We have here a C^∞ function $F : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ whose components are

$$F_1(x, y, u, v) = xu + yvu^2 - 2, \quad F_2(x, y, u, v) = xu^3 + y^2v^4 - 2.$$

Since we want to express (u, v) as a function of (x, y), we want to compute the derivative

$$D_2F = \begin{bmatrix} \dfrac{\partial F_1}{\partial u} & \dfrac{\partial F_1}{\partial v} \\ \dfrac{\partial F_2}{\partial u} & \dfrac{\partial F_2}{\partial v} \end{bmatrix} = \begin{bmatrix} x + 2yuv & yu^2 \\ 3xu^2 & 4uy^2v^3 \end{bmatrix}.$$

Since

$$D_2F(1,1,1,1) = \begin{bmatrix} 3 & 1 \\ 3 & 4 \end{bmatrix}$$

has determinant $12 - 3 = 9 \neq 0$, the linear operator $D_2F(1,1,1,1)$ has bounded inverse.

Thus we can uniquely solve for (u,v) as functions of (x,y) near $(1,1,1,1)$.

This means that there are C^∞ functions $u = f(x,y)$ and $v = g(x,y)$ defined on an open neighbourhood of $(1,1)$ that satisfy $1 = f(1,1)$, $1 = g(1,1)$, and

$$xf(x,y) + yg(x,y)[f(x,y)]^2 = 2, \quad x[f(x,y)]^3 + y^2[g(x,y)]^4 = 2.$$

Taking the partial derivatives of these equations with respect to x gives

$$\begin{aligned} &f(x,y) + x\frac{\partial f}{\partial x} + y\frac{\partial g}{\partial x}[f(x,y)]^2 + 2yg(x,y)f(x,y)\frac{\partial f}{\partial x} = 0, \\ &[f(x,y)]^3 + 3x[f(x,y)]^2\frac{\partial f}{\partial x} + 4y^2[g(x,y)]^3\frac{\partial g}{\partial x} = 0. \end{aligned}$$

Evaluating these equations at $x = 1, y = 1$ gives

$$\begin{aligned} 3\frac{\partial f}{\partial x}(1,1) + \frac{\partial g}{\partial x}(1,1) &= -1, \\ 3\frac{\partial f}{\partial x}(1,1) + 4\frac{\partial g}{\partial x}(1,1) &= -1. \end{aligned}$$

This is a system of linear equations whose coefficient matrix is invertible, hence the system can be solved to give

$$\frac{\partial f}{\partial x}(1,1) = -\frac{1}{3}, \quad \frac{\partial g}{\partial x}(1,1) = 0.$$

Similarly we can compute

$$\frac{\partial f}{\partial y}(1,1) = -\frac{2}{9}, \quad \frac{\partial g}{\partial x}(1,1) = -\frac{1}{3}.$$

7.4.2 Inverse Function Theorem

Recall from single-variable Calculus that if $f : (a,b) \to \mathbb{R}$ is C^1 and at $c \in (a,b)$ we have $f'(c) \neq 0$, then f' is of one sign on an interval containing c, so that f is monotone, hence f has a local inverse near c. The conclusion of local invertibility extends to the general Banach space setting as a consequence of the derivative having bounded inverse.

Theorem 7.4.8 (The Inverse Function Theorem). For Banach spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$, let U be an open neighbourhood of $\mathrm{x}_0 \in X$, V an open neighbourhood of $\mathrm{y}_0 \in Y$, and $f \in C^k(U,V)$ for some $k \geq 1$ satisfying $f(\mathrm{x}_0) = \mathrm{y}_0$. If $Df(\mathrm{x}_0) \in \mathscr{B}(X,Y)$ has bounded inverse, then there exist open neighbourhoods $U_0 \subset U$ of x_0 and $V_0 \subset V$ of y_0, and a unique $g \in C^k(V_0, U_0)$ such that $g(f(\mathrm{x})) = \mathrm{x}$ for all $x \in U_0$, $f(g(\mathrm{y})) = \mathrm{y}$ for all $\mathrm{y} \in V_0$, and for all $\mathrm{y} \in V_0$ there holds

$$Dg(\mathrm{y}) = Df(g(\mathrm{y}))^{-1}.$$

Proof. Define the C^k function $F : U \times V \to V$ by $F(\mathrm{x}, \mathrm{y}) = f(\mathrm{x}) - \mathrm{y}$.

Since $D_1 F(\mathrm{x}_0, \mathrm{y}_0) = Df(\mathrm{x}_0)$ has bounded inverse, the Implicit Function Theorem gives the existence of an open neighbourhood $U_1 \times V_0 \subset U \times V$ of the point $(\mathrm{x}_0, \mathrm{y}_0)$ and $g \in C^k(V_0, U_1)$ such that $F(g(\mathrm{y}), \mathrm{y}) = 0$ for all $\mathrm{y} \in V_0$.

That is, we have $f(g(\mathrm{y})) = \mathrm{y}$ for all $\mathrm{y} \in V_0$.

This implies that g is injective (if $g(\mathrm{y}_1) = g(\mathrm{y}_2)$, then $\mathrm{y}_1 = f(g(\mathrm{y}_1)) = f(g(\mathrm{y}_2)) = \mathrm{y}_2$).

By restricting the codomain of g to $U_0 = g(V_0) \subset U_1$, we obtain a bijective function g.

This implies that $f : U_0 \to V_0$ is bijective: surjectivity, for $\mathrm{y} \in V_0$ there exists a unique $\mathrm{x} \in U_0$ such that $g(\mathrm{y}) = \mathrm{x}$ by the bijectivity of g, so that $f(g(\mathrm{y})) = \mathrm{y}$; injectivity, for $f(\mathrm{x}_1) = f(\mathrm{x}_2)$ there are unique $\mathrm{y}_1, \mathrm{y}_2 \in V_0$ such that $g(\mathrm{y}_1) = \mathrm{x}_1$ and $g(\mathrm{y}_2) = \mathrm{x}_2$ by the bijectivity of g, so that $\mathrm{y}_1 = f(g(\mathrm{y}_1)) = f(g(\mathrm{y}_2)) = \mathrm{y}_2$ which implies that $\mathrm{x}_1 = \mathrm{x}_2$.

Thus f and g are inverses of each other, so that $g(f(\mathrm{x})) = \mathrm{x}$ for all $\mathrm{x} \in U_0$.

That U_0 is open follows because $U_0 = g(V_0) = f^{-1}(V_0)$ where V_0 is open and f is continuous.

Differentiation of $f(g(\mathrm{y})) = \mathrm{y}$ gives $Df(g(\mathrm{y}))Dg(\mathrm{y}) = I$, so that with $Df(\mathrm{x}_0)$ having bounded inverse, there exists (by Lemma 7.3.11) possibly smaller open neighbourhoods U_0 and V_0 such that $Dg(\mathrm{y}) = Df(g(\mathrm{y}))^{-1}$ for all $\mathrm{y} \in V_0$. □

Note. The condition that $Df(\mathrm{x}_0) \in \mathscr{B}(X, Y)$ have a bounded inverse implies that X and Y are not only isomorphic but also homeomorphic as Banach spaces. In most applications of the Inverse Function Theorem you will notice that $X = Y$. The Inverse Function Theorem implies that $f \in C^k(U, V)$ is a local diffeomorphism near each $\mathrm{x}_0 \in U$ for which $Df(\mathrm{x}_0) \in \mathrm{GL}(X)$.

Example. For a Banach space $(X, \|\cdot\|)$ consider the function $f : \mathscr{B}(X) \to \mathscr{B}(X)$ given by

$$f(A) = I + A^2.$$

That $f(A) \in \mathscr{B}(X)$ when $A \in \mathscr{B}(X)$ follows because

$$\|(I + A^2)\mathrm{x}\| \leq \|I\mathrm{x}\| + \|A^2\mathrm{x}\| \leq \|\mathrm{x}\| + \|A\|^2\|\mathrm{x}\| = (1 + \|A\|^2)\|\mathrm{x}\|$$

implies that

$$\|I + A^2\| = \sup\left\{\frac{\|(I + A^2)\mathrm{x}\|}{\|\mathrm{x}\|} : \mathrm{x} \in X, \mathrm{x} \neq 0\right\} \leq 1 + \|A\|^2.$$

A candidate for the derivative of f at $A \in \mathscr{B}(X)$ is the function $L_A \in \mathscr{L}(\mathscr{B}(X), \mathscr{B}(X))$ defined by

$$L_A(H) = AH + HA.$$

That $L_A(H)$ belongs to $\mathscr{B}(X)$ for each $H \in \mathscr{B}(X)$ follows because

$$\|(AH + HA)\mathrm{x}\| \leq \|AH\mathrm{x}\| + \|HA\mathrm{x}\| \leq \|A\|\,\|H\|\,\|\mathrm{x}\| + \|H\|\,\|A\|\,\|\mathrm{x}\| = 2\|A\|\,\|H\|\,\|\mathrm{x}\|$$

implies that

$$\|L_A(H)\| = \sup\left\{\frac{\|L_A(H)\mathrm{x}\|}{\|\mathrm{x}\|} : \mathrm{x} \in X, \mathrm{x} \neq 0\right\} \leq 2\|A\|\,\|H\|.$$

For L_A to be the derivative of f at A requires that L_A be bounded, but this follows because

$$\|L_A(H)\| \leq \|AH\| + \|HA\| \leq \|A\|\,\|H\| + \|H\|\,\|A\| = 2\|A\|\,\|H\|$$

implies that

$$\|L_A\| = \sup\left\{\frac{\|L_A(H)\|}{\|H\|} : H \in \mathscr{B}(X), H \neq 0\right\} \leq 2\|A\|.$$

Thus $L_A \in \mathscr{B}(\mathscr{B}(X), \mathscr{B}(X))$ and is the derivative of f at A because

$$\begin{aligned} f(A+H) - f(A) - L_A(H) &= I + (A+H)^2 - I - A^2 - AH - HA \\ &= A^2 + AH + HA + H^2 - A^2 - AH - HA \\ &= H^2 \end{aligned}$$

implies that

$$\lim_{H\to 0} \frac{\|f(A+H) - f(A) - L_A(H)\|}{\|H\|} = \lim_{H\to 0} \frac{\|H^2\|}{\|H\|} \leq \lim_{H\to 0} \frac{\|H\|^2}{\|H\|} = \lim_{H\to 0} \|H\| = 0.$$

Thus f is differentiable on $\mathscr{B}(X)$ with derivative $Df(A) = L_A$.

We cannot apply the Inverse Function Theorem to f at $A = 0$ because $Df(A)H = AH + HA = 0$ for all H when $A = 0$, i.e., $Df(A)$ is not invertible so it cannot have a bounded inverse.

We can apply the Inverse Function Theorem to f at $A = 3I$ because $Df(A)H = AH + HA = 6H$ when $A = 3I$, which has the bounded inverse $H \to (1/6)H$.

Thus f is a local diffeomorphism near $3I$, i.e., there are open neighbourhoods U_0 of $3I$ and V_0 of $f(3I) = 10I$ such that $f : U \to V$ is a diffeomorphism.

Note that f is not invertible on $\mathscr{B}(X)$ because it is not injective, i.e., $f(3I) = f(-3I)$.

Theorem 7.4.12. The inverse and implicit function theorems are equivalent.

Proof. We have already derived the inverse function theorem from the implicit function theorem.

So it remains to show that the implicit function theorem can be derived from the inverse function theorem.

To this end for $F \in C^k(U \times V, Z)$ with $F(\mathrm{x}_0, \mathrm{y}_0) = 0$ and $D_2F(\mathrm{x}_0, \mathrm{y}_0) \in \mathscr{B}(Y, Z)$ having bounded inverse, let $G : U \times V \to X \times Z$ be the C^k function given by $G(\mathrm{x}, \mathrm{y}) = (\mathrm{x}, F(\mathrm{x}, \mathrm{y}))$.

Then

$$DG(\mathrm{x}_0, \mathrm{y}_0) = \begin{bmatrix} I & 0 \\ D_1F(\mathrm{x}_0, \mathrm{y}_0) & D_2F(\mathrm{x}_0, \mathrm{y}_0) \end{bmatrix} \in \mathscr{B}(X \times Y, X \times Z)$$

has the inverse

$$DG(\mathrm{x}_0,\mathrm{y}_0)^{-1} = \begin{bmatrix} I & 0 \\ -D_2F(\mathrm{x}_0,\mathrm{y}_0)^{-1}D_1F(\mathrm{x}_0,\mathrm{y}_0) & D_2F(\mathrm{x}_0,\mathrm{y}_0)^{-1} \end{bmatrix} \in \mathscr{B}(X \times Z, X \times Y).$$

The boundedness of $DG(\mathrm{x}_0,\mathrm{y}_0)^{-1}$ follows from the boundedness of $D_2F(\mathrm{x}_0,\mathrm{y}_0)^{-1}$ and $D_1F(\mathrm{x}_0,\mathrm{y}_0)$.

By the Inverse Function Theorem, the function G is a local diffeomorphism on a open neighbourhood $U_0 \times V_0$ of $(\mathrm{x}_0,\mathrm{y}_0)$.

Since $G(\mathrm{x}_0,\mathrm{y}_0) = (\mathrm{x}_0, 0)$ and $G(\mathrm{x},\mathrm{y}) = (\mathrm{x}, F(\mathrm{x},\mathrm{y}))$ on $U_0 \times V_0$, the C^k function G^{-1} has the form $G^{-1}(\mathrm{x},0) = (\mathrm{x}, f(\mathrm{x}))$ for a C^k function f that satisfies $f(\mathrm{x}_0) = \mathrm{y}_0$.

Hence $(\mathrm{x},0) = G(G^{-1}(\mathrm{x},0)) = G(\mathrm{x}, f(\mathrm{x})) = (\mathrm{x}, F(\mathrm{x}, f(\mathrm{x})))$ so that $F(\mathrm{x}, f(\mathrm{x})) = 0$ for all $\mathrm{x} \in U_0$. $\square$