Lecture 10: Appendix B: The Inverse and Implicit Function Theorems.

Contractions.
A map \( T : W \to W \) is called a contraction, if for \( x, y \in W \):

\[
\|T(x) - T(y)\| \leq K \|x - y\|, \quad K < 1
\]

(B1)

A point \( x \in W \) is called a fixed point if \( T(x) = x \). We have:

Lemma 2. Let \( T : W_0 \to W_0 \) be a contraction of a complete normed space \( W_0 \). Then \( T \) has a unique fixed point \( x \in W_0 \). In fact for any \( x_0 \in W_0 \), \( x_k = T^k(x_0) = T \circ \cdots \circ T(x_0) \) (\( k \) times) converges to \( x \); \( \|x - x_k\| \to 0 \), as \( k \to \infty \).

Proof. Using (B1) repeatedly we get

\[
\|x_{k+1} - x_k\| = \|T(x_k) - T(x_{k-1})\| \leq K \|x_k - x_{k-1}\| \leq \cdots \leq K^k \|x_1 - x_0\|
\]

(B2)

Here \( \|x_1 - x_0\| = \|T(x_0) - x_0\| = C \) is a fixed constant. For \( m > k \) we write \( x_m - x_k = (x_m - x_{m-1}) + (x_{m-1} - x_{m-2}) + \cdots + (x_{k+1} - x_k) \) and estimate the norm of each term by (B2):

\[
\|x_m - x_k\| \leq \|x_m - x_{m-1}\| + \cdots + \|x_{k+1} - x_k\| \leq (K^{m-1} + \cdots + K^{k-1})C
\]

(B3)

This is a geometric sum and since \( K < 1 \) the infinite sum converges; \( \sum_{\ell=k-1}^{m-1} K^\ell \leq \sum_{\ell=k-1}^{\infty} K^\ell = K^{k-1} \sum_{n=0}^{\infty} K^n = K^{k-1} / (1 - K) \). Hence

\[
\|x_m - x_k\| \leq \epsilon(N) = \frac{CK^{N-1}}{1 - K}, \quad \text{if } m, k \geq N,
\]

where \( \epsilon(N) \to 0 \) as \( N \to \infty \), i.e. \( x_k \) is a Cauchy sequence.

The uniqueness follows from (B1); if \( T(x) = x \) and \( T(y) = y \) then \( \|x - y\| = \|T(x) - T(y)\| \leq K\|x - y\| \) and since \( K < 1 \) it follows that \( \|x - y\| = 0 \) so \( x = y \). \( \square \)

Theorem 1. Suppose that \( F : \mathbb{R}^n \to \mathbb{R}^n \) is \( C^1 \). Let \( F(x_0) = y_0 \) and suppose that

\[
dF_{x_0} = \frac{\partial F}{\partial x}(x_0)
\]

(B9)

is invertible. Then for \( y \) close to \( y_0 \) there is an unique \( x \) close to \( x_0 \) such that

\[
F(x) = y
\]

Furthermre \( x = x(y) \) is a \( C^1 \) function of \( y \) close to \( y_0 \).

By Taylor’s formula, if \( F \in C^2 \),

\[
y - y_0 = F(x) - F(x_0) = dF_{x_0}(x - x_0) + O(|x - x_0|^2)
\]

(B10)

where the derivative \( dF_{x_0} : \mathbb{R}^n \to \mathbb{R}^n \) is the linear map that best approximates the function close to \( x_0 \) and \( O(|x - x_0|^2) \) means terms that are bounded by a constant times \( |x - x_0|^2 \) and hence much smaller than \( |x - x_0| \), when \( |x - x_0| \) is small. Therefore, to get a first approximation we must be able to invert the linear linear map, and we get that \( x - x_0 = (dF_{x_0})^{-1}(y - y_0) + O(|y - y_0|^2) \).

The proof of Theorem 1 uses the contraction mapping theorem. First by a translation replacing \( F(x) \) by \( F(x + x_0) - y_0 \) we can reduce to the case when \( x_0 = y_0 = 0 \). Furthermore by multiplying
both sides of (B10) by the matrix \((dF_0)^{-1}\) and making a change of variables replacing \(y\) by \((dF_0)^{-1}y\) we may assume that the equation (B10) takes the form

\[
y = x + \phi(x)
\]

where \(\phi(x)\) is small; \(\phi(0) = 0\) and \(d\phi_0 = 0\). We seek a solution in the form

\[
x = y + \psi(y)
\]

Then for \(\phi(y)\) we obtain the equation \(\psi(y) = -\phi(y + \psi(y))\) Consequently, the function \(\psi\) being sought is a fixed point of the mapping \(T\) defined by the formula

\[
(T\psi)(y) = -\phi(y + \psi(y))
\]

**Problem 1**: Show that \(T\) is a contraction in some norm for \(y\) sufficiently small. You have to use that since \(\phi\) is continuously differentiable and \(d\phi_0 = 0\) there is a neighborhood \(\delta > 0\) such that \(\|d\phi_2\| = \sup_{|z| \leq 1} |d\phi_2(x)|/|x| < 1/2\), when \(|z| < \delta\). Let \(W = \{\psi \in C^1([|y| \leq \delta/2]); |\psi(y)| \leq |y|\}\). With \(z(t) = y + \psi_1(y) + t(\psi_2(y) - \psi_1(y)), 0 \leq t \leq 1\), the line segment between \(y + \psi_1(y)\) and \(y + \psi_2(y)\):

\[
T(\psi_1)(y) - T(\psi_2)(y) = (\psi(y + \psi_2(y)) - \psi(y + \psi_1(y))) = \int_0^1 \frac{d}{dt} \phi(z(t))) dt = \int_0^t d\phi(z(t)) (\psi_2(y) - \psi_1(y)) dt.
\]

If \(\psi_1, \psi_2 \in W\) and \(|y| \leq \delta/2\) then \(|z(t)| \leq \delta, \) for \(0 \leq t \leq 1\) and hence

\[
|T(\psi_1)(y) - T(\psi_2)(y)| \leq \sup_{0 \leq t \leq 1} \|d\phi_2(t)\| \|\psi_2(y) - \psi_1(y)\| \leq \frac{1}{2} |\psi_2(y) - \psi_1(y)|, \quad \text{if } |y| < \delta/2.
\]

In particular if we take \(\psi_2(y) = -y\) we obtain

\[
|T(\psi_1)(y)| \leq \frac{1}{2} |\psi_1(y) + y| \leq |y|, \quad \text{if } |y| < \delta/2,
\]

i.e. \(T(\psi) \in W\) if \(\psi \in W\). If \(|y| < \delta/2\) and we set \(\psi_0(y) = 0\) and \(\psi_{n+1}(y) = T(\psi_n)(y)\), for \(n \geq 0\), then by the contraction lemma \(\psi_n(y) \rightarrow \psi(y)\) in \(W\), where \(T(\psi)(y) = \psi(y)\).

**Theorem 2**: Suppose that \(G : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^n\) is \(C^1\). Let \(G(x_0, y_0) = c_0\) and suppose that

\[
\frac{\partial G}{\partial y}(x_0, y_0)
\]

is invertible. Then for \(x\) close to \(x_0\) there is a unique \(y = g(x)\) close to \(y_0\) such that

\[
G(x, g(x)) = 0
\]

Furthermore \(y = g(x)\) is a \(C^1\) function of \(x\) close to \(y_0\).

**Problem 2**: Show that Theorem 2 follows from Theorem 1, by considering \(F(x, y) = (x, G(x, y))\).

**Problem 3**: Suppose that \(G(x_0, y_0, z_0) = 0\), and \(\text{grad } G(x_0, y_0, z_0) \neq 0\). Use Theorem 2 to deduce that close to \((x_0, y_0, z_0)\) the equation \(G(x, y, z) = c_0\) is a surface, i.e. show that one of the variables say \(z\) (if \(\partial G/\partial z \neq 0\) can by written as a graph \(z = g(x, y)\) so that \(G(x, y, g(x, y)) = 0\).