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

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

(B1) $$\|T(x) - T(y)\| \le K\|x - y\|, \qquad K < 1$$

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

(B2) $$\|x_{k+1} - x_k\| = \|T(x_k) - T(x_{k-1})\| \le K\|x_k - x_{k-1}\| \le \cdots \le K^k\|x_1 - x_0\|$$

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

(B3) $$\|x_m - x_k\| \le \|x_m - x_{m-1}\| + \cdots + \|x_{k+1} - x_k\| \le (K^{m-1} + \cdots + K^{k-1})C$$

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

(B4) $$\|x_m - x_k\| \le \varepsilon(N) = \frac{CK^{N-1}}{1-K}, \qquad \text{if} \quad m, k \ge N,$$

where $\varepsilon(N) \to 0$ 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 $\|x - y\| = \|T(x) - T(y)\| \le K\|x - y\|$ and since $K < 1$ it follows that $\|x - y\| = 0$ so $x = y$. □

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

(B9) $$dF_{x_0} = \frac{\partial F}{\partial x}(x_0)$$

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

(B10) $$F(x) = y$$

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

By Taylor's formula, if $F \in C^2$,

(B11) $$y - y_0 = F(x) - F(x_0) = dF_{x_0}(x - x_0) + O\big(|x - x_0|^2\big)$$

where the derivative $dF_{x_0} : \mathbf{R}^n \to \mathbf{R}^n$ is the linear map that best approximates the function close to x_0 and $O\big(|x - x_0|^2\big)$ 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

(B12) $$y = x + \phi(x)$$

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

(B13) $$x = y + \psi(y)$$

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

(B14) $$(T\psi)(y) = -\phi\big(y + \psi(y)\big)$$

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

$$T(\psi_1)(y)-T(\psi_2)(y) = \phi\big(y+\psi_2(y)\big)-\phi\big(y+\psi_1(y)\big) = \int_0^1 \frac{d}{dt}\,\phi\big(z(t)\big)\big)dt = \int_0^t d\phi_{z(t)}(\psi_2(y)-\psi_1(y))dt.$$

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

$$|T(\psi_1)(y) - T(\psi_2)(y)| \le \sup_{0\le t\le 1} \|d\phi_{z(t)}\|\, |\psi_2(y) - \psi_1(y)| \le \frac{1}{2}|\psi_2(y) - \psi_1(y)|, \quad \text{if} \quad |y| < \delta/2.$$

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

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

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 $n \ge 0$, then by the contraction lemma $\psi_n(y) \to \psi(y)$ in W, where $T(\psi)(y) = \psi(y)$.

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

(B15) $$\frac{\partial G}{\partial y}(x_0, y_0)$$

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

(B16) $$G(x, g(x)) = 0$$

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

Problem 2: Show that Theorem 2 follows from Theorem 1, by considering $F(x, y) = \big(x, G(x, y)\big)$.
Problem 3: Suppose that $G(x_0, y_0, z_0) = 0$, and **grad** $G(x_0, y_0, z_0) \ne 0$. 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 $\partial G/\partial z \ne 0$ can by written as a graph $z = g(x, y)$ so that $G(x, y, g(x, y)) = 0$.