Repetition.

Implicite Functions Theorems.

Task: solving a system of equations

$$F_1(\mathbf{x}, y_1, \ldots, y_n) = 0,$$
$$\ldots \quad \ldots \quad \ldots$$
$$F_n(\mathbf{x}, y_1, \ldots, y_n) = 0$$

in terms of y_i as well determined functions $f_i(\mathbf{x})$ (where $\mathbf{x} = (x_1, \ldots, x_n)$).

A. One equation: $F(\mathbf{x}, y) = 0$:

In a nbhood of $(\mathbf{x}^0, y_0)$ *assumed on* F*:*
cont. p. d. up to order $k \geq 1$*,*

$$F(\mathbf{x}^0, y_0) = 0 \quad \textit{and} \quad \left|\frac{\partial F(\mathbf{x}^0, y_0)}{\partial y}\right| \neq 0.$$

Then one has in some

$$\{\mathbf{x} \,|\, ||\mathbf{x} - \mathbf{x}^0|| < \delta\} \times (y_0 - \delta, y_0 + \delta)$$

unique solution $(\mathbf{x}, f(\mathbf{x}))$*, an the resulting* f *has cont.p.d. up to order* k*.*

Proved with one variable x.

The properties of f may be more transparent if we indicate the shifts $h_1 = h$ and $h_2 = f(x+h) - f(x)$ in Lagrange theorem by red:

$$0 = F(t+h, f(t+h)) - F(t, f(t)) =$$
$$= F(t+h, f(t) + (f(t+h) - f(t))) - F(t, f(t)) =$$
$$= \frac{\partial F(t+\theta h, f(t) + \theta(f(t+h) - f(t)))}{\partial x} h$$
$$+ \frac{\partial F(t+\theta h, f(t) + \theta(f(t+h) - f(t)))}{\partial y} (f(t+h) - f(t))$$

hence

$$f(t+h) - f(t) = -h \cdot \frac{\dfrac{\partial F(t+\theta h, f(t) + \theta(f(t+h) - f(t)))}{\partial x}}{\dfrac{\partial F(t+\theta h f(t) + \theta(f(t+h) - f(t)))}{\partial y}} \qquad (*)$$

for some θ between 0 and 1.Thus,

$$|f(t+h) - f(t)| \leq |h| \cdot \left|\frac{K}{a}\right|$$

Hence f is continuous, and from $(*)$ further

$$\lim_{h \to 0} \frac{f(t+h) - f(t)}{h} = -\frac{\dfrac{\partial F(t, f(t))}{\partial x}}{\dfrac{\partial F(t, f(t))}{\partial y}}$$

If we have partial derivatives of F in (x_0, y_0) we can compute from

$$f'(t) = -\frac{\dfrac{\partial F(t, f(t))}{\partial x}}{\dfrac{\partial F(t, f(t))}{\partial y}}$$

derivatives

$$f'(x_0), f''(x_0), f'''(x_0), ...$$

and hence

Taylor polynomials.

B. Two equations:

$$F_1(\mathbf{x}, y_1, y_2) = 0,$$
$$F_2(\mathbf{x}, y_1, y_2) = 0.$$

For F_1, F_2 with cont. p. d. up to order $k \geq 1$, in a nbh. of $(\mathbf{x}^0, y_1^0, y_2^0)$, with $F_i(\mathbf{x}^0, y_1^0, y_2^0) = 0$ we obtain in some

$\{\mathbf{x} \,|\, ||\mathbf{x} - \mathbf{x}^0|| < \delta\} \times (y_1^0 - \delta, y_1^0 + \delta) \times (y_2^0 - \delta, y_2^0 + \delta)$

solutions $(\mathbf{x}, f_1(\mathbf{x}), f_2(\mathbf{x}))$, f_i again with cont.p.d. up to order k.

Instead of $\left|\frac{\partial F(\mathbf{x}^0, y_0)}{\partial y}\right| \neq 0$ assumed

$$\begin{vmatrix} \frac{\partial F_1}{\partial y_1}, & \frac{\partial F_1}{\partial y_2} \\ \frac{\partial F_2}{\partial y_1}, & \frac{\partial F_2}{\partial y_2} \end{vmatrix} = \det \left(\frac{\partial F_i}{\partial y_j} \right)_{i,j} \neq 0.$$

Jacobi determinant.

For a sequence of functions

$$\mathbf{F}(\mathbf{x},\mathbf{y}) = (F_1(\mathbf{x}, y_1, \ldots, y_m), \ldots, F_m(\mathbf{x}, y_1, \ldots, y_m)).$$

and $\mathbf{y} = (y_1, \ldots, y_m)$ define the *Jacobi determinant* (briefly, the *Jacobian*)

$$\frac{\mathsf{D}(\mathbf{F})}{\mathsf{D}(\mathbf{y})} = \det\left(\frac{\partial F_i}{\partial y_j}\right)_{i,j=1,\ldots,m}$$

In a way it is an extension of a partial derivative of one function F by one y: we have

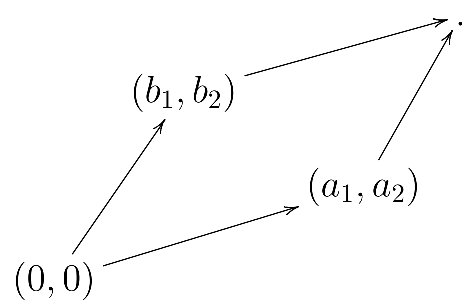
$$\frac{\mathsf{D}(F)}{\mathsf{D}(y)} = \frac{\partial F}{\partial y}.$$

hence the following theorem will come quite as an extension of the solution of one equation.

Aside. Hopefully the students know from linear algebra that (the absolute value) of the determinant

$$\begin{vmatrix} a_{11}, a_{12}, \ldots, a_{1n} \\ \ldots, \ldots, \ldots \\ a_{n1}, a_{n2}, \ldots, a_{nn} \end{vmatrix} \neq 0.$$

is the volume of the parallelepiped determined by the vectors $(a_{11}, a_{12}, \ldots, a_{1n})$,..., $(a_{11}, a_{12}, \ldots, a_{1n})$.
(As a simple exercise prove that the area of the parallelogram



is $a_1b_2 - a_2b_1 = \begin{vmatrix} a_1, a_2 \\ b_1, b_2 \end{vmatrix}$.)

Thus, like a function f transforming an interval (a, b) to $(f(a), f(b))$ stretches or compresses the lengths of small pieces of the interval around x in the rate of the (absolute) value of $\frac{\mathrm{d}f}{\mathrm{d}x}$ in x, a vector function $\mathbf{f} = (f_1, \ldots, f_n)$ transforming a domain $U \subseteq \mathbb{E}_n$ to $\mathbf{f}[U]$ stetches or compresses the volumes of small pieces of U around $\mathbf{x}$ in the rate of the (absolute) value of $\frac{\mathsf{D}(\mathbf{f})}{\mathsf{D}(\mathbf{x})}$.

Theorem. *Let* $F_i(\mathbf{x}, y_1, \dots, y_m)$*,* $i = 1, \dots, m$*, be functions of* $n + m$ *variables with continuous partial derivatives up to an order* $k \geq 1$*. Let*

$$\mathbf{F}(\mathbf{x}^0, \mathbf{y}^0) = \mathbf{o}$$

and let

$$\frac{\mathsf{D}(\mathbf{F})}{\mathsf{D}(\mathbf{y})}(\mathbf{x}^0, \mathbf{y}^0) \neq 0.$$

Then there exist $\delta > 0$ *and* $\Delta > 0$ *such that for every*

$$\mathbf{x} \in (x_1^0 - \delta, x_1^0 + \delta) \times \cdots \times (x_n^0 - \delta, x_n^0 + \delta)$$

there exists precisely one

$$\mathbf{y} \in (y_1^0 - \Delta, y_1^0 + \Delta) \times \cdots \times (y_m^0 - \Delta, x_m^0 + \Delta)$$

such that

$$\mathbf{F}(\mathbf{x}, \mathbf{y}) = 0.$$

(That is,

$$\begin{aligned} &F_1(\mathbf{x}, y_1, \dots, y_n) = 0, \\ &\dots \quad \dots \quad \dots \\ &F_n(\mathbf{x}, y_1, \dots, y_n) = 0. \quad) \end{aligned}$$

Furthermore, if we write this $\mathbf{y}$ *as a vector function* $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$*, then the functions* f_i *have continuous partial derivatives up to the order* k*.*

An application: extremes with constraints.

Local extremes of a function f in one variable.

f was defined, say, on an interval, and had a derivative in the interior. Then one considered the points in which the derivative was 0, and in addition the boundary points of the interval. Not much harder for more complex situations.

For functions of several variables, searching for candidates for local extremes *in the interiors of the domain* is equally easy (and for the same reason): at the points of local extreme $\mathbf{a}$, we must have

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) = 0, \quad i = 1, \ldots, n. \qquad (*)$$

But the exceptional points on the boundary are now typically infinitely many.

Example.

Find local extremes of

$$f(x, y) = x + 2y$$

on the disc

$$B = \{(x, y) \,|\, x^2 + y^2 \leq 1\}.$$

B is compact, and hence the function f attains a minimum and a maximum on B.

None of them is in the interior, though: we have, constantly, $\frac{\partial f}{\partial x} = 1$ and $\frac{\partial f}{\partial y} = 2$; thus, the extremes must be located somewhere in the infinte set $\{(x, y) \,|\, x^2 + y^2 = 1\}$, and the rule $(*)$ is of no use.

The approach: try to find local extremes of a function $f(x_1, \ldots, x_n)$ *subject to constraints* $g_i(x_1, \ldots, x_n) = 0$, $i = 1, \ldots, k$.

Theorem. *Let $f, g_1, \ldots, g_k$ be real functions defined in an open set $D \subseteq \mathbb{E}_n$, and let them have continuous partial derivatives. Suppose that the rank of the matrix*

$$M = \begin{pmatrix} \dfrac{\partial g_1}{\partial x_1}, & \ldots, & \dfrac{\partial g_1}{\partial x_n} \\ \ldots, & \ldots, & \ldots \\ \dfrac{\partial g_k}{\partial x_1}, & \ldots, & \dfrac{\partial g_k}{\partial x_n} \end{pmatrix}$$

is the largest possible, that is k, everywhere in D.

If the function f achieves at a point $\mathbf{a} = (a_1, \ldots, a_n)$ a local extreme subject to the constraints

$$g_i(x_1, \ldots, x_n) = 0, \quad i = 1, \ldots, k$$

then there exist numbers $\lambda_1, \ldots, \lambda_k$ such that for each $i = 1, \ldots, n$ we have

$$\frac{\partial f(\mathbf{a})}{\partial x_i} + \sum_{j=1}^{k} \lambda_j \cdot \frac{\partial g_j(\mathbf{a})}{\partial x_i} = 0.$$

Back to the example: How it helps.

We have $\frac{\partial f}{\partial x} = 1$ and $\frac{\partial f}{\partial y} = 2$, $g(x, y) = x^2 + y^2 - 1$ and hence $\frac{\partial g}{\partial x} = 2x$ and $\frac{\partial g}{\partial y} = 2y$. There is *one* λ that satisfies *two* equations

$$1 + \lambda \cdot 2x = 0 \quad \text{and} \quad 2 + \lambda \cdot 2y = 0.$$

This is possible only if $y = 2x$. Thus, as $x^2 + y^2 = 1$ we obtain $5x^2 = 1$ and hence $x = \pm\frac{1}{\sqrt{5}}$; this localizes the extremes to $(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$ and $(\frac{-1}{\sqrt{5}}\frac{-2}{\sqrt{5}})$.

Notes.

1. The functions f, g_i were assumed to be defined in an open D so that we can take derivatives whenever we need them. In typical applications one works with functions that can be extended to an open set containing the area in question.

2. **The force of the statement is in asserting the existence of**

$$\lambda_1, \dots, \lambda_k$$

that satisfy *more than* k equations, as we have seen in the solution of the task from the example.

3. The numbers λ_i are known as *Lagrange multipliers*.

Sketch of proof of Theorem. A matrix M has rank k iff at least one of the $k \times k$ submatrices of M is regular (and hence has a non-zero determinant). Let us have, say,

$$\begin{vmatrix} \frac{\partial g_1}{\partial x_1}, & \dots, & \frac{\partial g_1}{\partial x_k} \\ \dots, & \dots, & \dots \\ \frac{\partial g_k}{\partial x_1}, & \dots, & \frac{\partial g_k}{\partial x_k} \end{vmatrix} \neq 0. \tag{1}$$

Then by the Implicite Function Thm we have in a nbh of $\mathbf{a}$ functions $\phi_i(x_{k+1}, \dots, x_n)$ with cont. p. derivatives such that (write $\widetilde{\mathbf{x}}$ for $(x_{k+1}, \dots, x_n)$)

$$g_i(\phi_1(\widetilde{\mathbf{x}}), \dots, \phi_k(\widetilde{\mathbf{x}}), \widetilde{\mathbf{x}}) = 0 \quad \text{for} \quad i = 1, \dots, k.$$

Thus, a local maximum or a local minimum of $f(\mathbf{x})$ at $\mathbf{a}$ subject to the given constraints implies l. maximum or minimum (without constraints) of

$$F(\widetilde{\mathbf{x}}) = f(\phi_1(\widetilde{\mathbf{x}}), \ldots, \phi_k(\widetilde{\mathbf{x}}), \widetilde{\mathbf{x}}),$$

at $\widetilde{\mathbf{a}}$, and hence

$$\frac{\partial F(\widetilde{\mathbf{a}})}{\partial x_i} = 0 \quad \text{for} \quad i = k+1, \ldots, n,$$

that is, by the Chain Rule,

$$\sum_{r=1}^{k} \frac{\partial f(\mathbf{a})}{\partial x_r} \frac{\partial \phi_r(\widetilde{\mathbf{a}})}{\partial x_i} + \frac{\partial f(\mathbf{a})}{\partial x_i} \quad \text{for} \quad i = k+1, \ldots, n. \quad (2)$$

Taking derivatives of the constant functions $g_i(\phi_1(\widetilde{\mathbf{x}}), \ldots, \phi(\widetilde{\mathbf{x}}), \widetilde{\mathbf{x}}) = 0$ we obtain for $j = 1, \ldots, k$,

$$\sum_{r=1}^{k} \frac{\partial g_j(\mathbf{a})}{\partial x_r} \frac{\partial \phi_r(\widetilde{\mathbf{a}})}{\partial x_i} + \frac{\partial g_j(\mathbf{a})}{\partial x_i} \quad \text{for} \quad i = k+1, \ldots, n. \quad (3)$$

Use (1) (non-zero determinant) again. Because of the rank of the matrix, the system of linear equations

$$\frac{\partial f(\mathbf{a})}{\partial x_i} + \sum_{j=1}^{n} \lambda_j \cdot \frac{\partial g_j(\mathbf{a})}{\partial x_i} = 0, \quad i = 1, \ldots, k,$$

has a unique solution $\lambda_1, \ldots, \lambda_k$. These are the equalities from the statement for $i \leq k$ only. It remains to be shown that the same equalities hold also for $i > k$. By (2) and (3), for $i > k$

$$\frac{\partial f(\mathbf{a})}{\partial x_i} + \sum_{j=1}^{n} \lambda_j \cdot \frac{\partial g_j(\mathbf{a})}{\partial x_i} =$$

$$= -\sum_{r=1}^{k} \frac{\partial f(\mathbf{a})}{\partial x_r} \frac{\partial \phi_r(\widetilde{\mathbf{a}})}{\partial x_i} - \sum_{j=1}^{k} \lambda_j \sum_{r=1}^{k} \frac{\partial g_j(\mathbf{a})}{\partial x_r} \frac{\partial \phi_r(\widetilde{\mathbf{a}})}{\partial x_i} =$$

$$= -\sum_{r=1}^{n} \left(\frac{\partial f(\mathbf{a})}{\partial x_i} + \sum_{j=1}^{n} \lambda_j \cdot \frac{\partial g_j(\mathbf{a})}{\partial x_i} \right) \frac{\partial \phi_r(\widetilde{\mathbf{a}})}{\partial x_i} =$$

$$= -\sum_{r=1}^{n} 0 \cdot \frac{\partial \phi_r(\widetilde{\mathbf{a}})}{\partial x_i} = 0.$$

Another use of IFT: Regular maps.

Let $U \subseteq \mathbb{E}_n$ be open. Let

$$f_i, \quad i = 1, \ldots, n,$$

have continuous partial derivatives.

The resulting mapping

$$\mathbf{f} = (f_1, \ldots, f_n) : U \to \mathbb{E}_n$$

is *regular* if

$$\frac{\mathsf{D}(\mathbf{f})}{\mathsf{D}(\mathbf{x})}(\mathbf{x}) \neq 0$$

for all $\mathbf{x} \in U$.

Proposition. *If* $\mathbf{f} : U \to \mathbb{E}_n$ *is regular then the image* $\mathbf{f}[V]$ *of every open* $V \subseteq U$ *is open.*

Comment before proof: Images added to necessary preimages. Similarity with images of closed subsets in the compact case.

Proof. Let $f(\mathbf{x}^0) = \mathbf{y}^0$. Define $\mathbf{F} : V \times \mathbb{E}_n \to \mathbb{E}_n$ by setting

$$F_i(\mathbf{x}, \mathbf{y}) = f_i(\mathbf{x}) - y_i. \qquad (*)$$

then $\mathbf{F}(\mathbf{x}^0, \mathbf{y}^0) = \mathbf{o}$ and $\frac{\mathsf{D}(\mathbf{F})}{\mathsf{D}(\mathbf{x})} \neq 0$, and hence we can apply IFT to obtain $\delta > 0$ and $\Delta > 0$ such that for every $\mathbf{y}$ with $\|\mathbf{y} - \mathbf{y}^0\| < \delta$, there exists a $\mathbf{x}$ such that $\|\mathbf{x} - \mathbf{x}^0\| < \Delta$ and $F_i(\mathbf{x}, \mathbf{y}) = f_i(\mathbf{x}) - y_i = 0$. This means that we have $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ (note that y_i are here the variables, x_j are the wanted functions), and

$$\Omega(\mathbf{y}^0, \delta) = \{\mathbf{y} \mid \|\mathbf{y} - \mathbf{y}^0\| < \delta\} \subseteq \mathbf{f}[V].$$

Proposition. *Let* $\mathbf{f} : U \to \mathbb{E}_n$ *be a regular mapping. Then for each* $\mathbf{x}^0 \in U$ *there exists an open neighborhood* V *such that the restriction* $\mathbf{f}|V$ *is one-to-one. Moreover, the mapping* $\mathbf{g} : f[V] \to \mathbb{E}_n$ *inverse to* $\mathbf{f}|V$ *is regular.*

Proof. We will use again the mapping $\mathbf{F} = (F_1, \dots, F_n)$ with $F_i(\mathbf{x}, \mathbf{y}) = f_i(\mathbf{x}) - y_i$ as before. For a sufficiently small $\Delta > 0$ we have precisely one $\mathbf{x} = \mathbf{g}(\mathbf{y})$ such that $\mathbf{F}(\mathbf{x}, \mathbf{y}) = 0$ and $\|\mathbf{x} - \mathbf{x}^0\| < \Delta$. This $\mathbf{g}$ has, furthermore, continuous partial derivatives. We have

$$D(\mathrm{id}) = D(\mathbf{f} \circ \mathbf{g}) = D(\mathbf{f}) \cdot D(\mathbf{g}).$$

By the Chain Rule (and the theorem on product of determinants)

$$\frac{\mathsf{D}(\mathbf{f})}{\mathsf{D}(\mathbf{x})} \cdot \frac{D(\mathbf{g})}{D(\mathbf{y})} = \det D(\mathbf{f}) \cdot \det D(\mathbf{g}) = 1$$

and hence for each $\mathbf{y} \in \mathbf{f}[V]$, $\frac{\mathsf{D}(\mathbf{g})}{\mathsf{D}(\mathbf{y})}(\mathbf{y}) \neq 0$.

Corollary. *A one-to-one regular mapping* $\mathbf{f} : U \to \mathbb{E}_n$ *has a regular inverse* $\mathbf{g} : \mathbf{f}[U] \to \mathbb{E}_n$.

Details.

Text: Chapter XV, Sections 4, 6 and 5