The Implicit Function Theorem.

Theorem: $U \subset \mathbb{R}^m$ open, $V \subset \mathbb{R}^p$ open $\left(\Rightarrow U \times V \subset \mathbb{R}^{m+p} \text{ open}\right)$

$$F : U \times V \longrightarrow \mathbb{R}^p \text{ of class } C^1$$
$$(x, y) \longmapsto F(x,y) \qquad \text{ie } x \in \mathbb{R}^m, y \in \mathbb{R}^p$$

Let $(x_0, y_0) \in U \times V$.

If $D_y F(x_0, y_0)$ is invertible (ie $\det(D_y F(x_0,y_0)) \neq 0$)

then $\exists r, s > 0$ s.t. $B(x_0, r) \subset U$, $B(y_0, s) \subset V$ and $\varphi : B(x_0,r) \to B(y_0,s)$ of class C^1

such that $\forall (x,y) \in B(x_0,r) \times B(y_0,s)$, $F(x,y) = F(x_0,y_0) \Leftrightarrow y = \varphi(x)$.

Remark 0: $F(x,y) = F(x_0,y_0)$ defines implicitely a function $y = \varphi(x)$ around (x_0,y_0)

Remark 1: $D_y F(x_0,y_0)$ is the Jacobian matrix of $\begin{array}{l} V \xrightarrow{\subset \mathbb{R}^p} \mathbb{R}^p \\ y \mapsto F(x_0,y) \end{array}$

$$\text{ie } D_y F(x_0,y_0) = \begin{pmatrix} \frac{\partial F_1}{\partial y_1}(x_0,y_0) & \text{---} & \frac{\partial F_1}{\partial y_p}(x_0,y_0) \\ | & \diagdown & | \\ \frac{\partial F_p}{\partial y_1}(x_0,y_0) & \text{---} & \frac{\partial F_p}{\partial y_p}(x_0,y_0) \end{pmatrix} \in M_{p,p}(\mathbb{R})$$

Similarly we define

$$D_x F(x_0,y_0) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(x_0,y_0) & \text{---} & \frac{\partial F_1}{\partial x_m}(x_0,y_0) \\ | & \diagdown & | \\ \frac{\partial F_p}{\partial x_1}(x_0,y_0) & \text{---} & \frac{\partial F_p}{\partial x_m}(x_0,y_0) \end{pmatrix} \in M_{p,m}(\mathbb{R})$$

$$\underset{\in M_{p,m+p}(\mathbb{R})}{DF(x_0,y_0)} = \left(\begin{array}{ccc|ccc} \frac{\partial F_1}{\partial x_1}(x_0,y_0) & \cdots & \frac{\partial F_1}{\partial x_m}(x_0,y_0) & \frac{\partial F_1}{\partial y_1}(x_0,y_0) & \cdots & \frac{\partial F_1}{\partial y_p}(x_0,y_0) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_p}{\partial x_1}(x_0,y_0) & \cdots & \frac{\partial F_p}{\partial x_m}(x_0,y_0) & \frac{\partial F_p}{\partial y_1}(x_0,y_0) & \cdots & \frac{\partial F_p}{\partial y_p}(x_0,y_0) \end{array}\right)$$

$$\underbrace{\qquad\qquad\qquad}_{D_x F(x_0,y_0)} \quad \underbrace{\qquad\qquad\qquad}_{D_y F(x_0,y_0)}$$

Remark: $F(x_0, y_0) = F(x_0, y_0)$ so $y_0 = \varphi(x_0)$ by $(*)$

Remark: $F(x, \varphi(x)) = F(x_0, y_0) \quad \forall x \in B(x_0, r)$

$$\Rightarrow DF(x_0, \underbrace{\varphi(x_0)}_{y_0}) \begin{pmatrix} I_{m,m} \\ D\varphi(x_0) \end{pmatrix} = 0$$

↳ the RHS is constant

↳ by the chain rule applied to $F \circ G(x)$ where $G(x) = (x, \varphi(x))$

$$\Rightarrow \begin{pmatrix} D_x F(x_0, y_0) & D_y F(x_0, y_0) \end{pmatrix} \begin{pmatrix} I_{m,m} \\ D\varphi(x_0) \end{pmatrix} = 0$$

$$\Rightarrow D_x F(x_0, y_0) + D_y F(x_0, y_0) D\varphi(x_0) = 0$$

$$\Rightarrow D\varphi(x_0) = - [D_y F(x_0, y_0)]^{-1} D_x F(x_0, y_0)$$

recall that $D_y F(x_0, y_0)$ is invertible

Ccl: We know how to compute $D\varphi(x_0)$ in terms of F

$$\boxed{D\varphi(x_0) = - [D_y F(x_0, y_0)]^{-1} D_x F(x_0, y_0)}$$

You should know this formula (or better: be able to quickly recover it)

Special case of the IFT when $p = 1$

Theorem: $U \subset \mathbb{R}^m$ open, $I = (a, b)$, $F : \begin{matrix} U \times I \longrightarrow \mathbb{R} \\ (x_1, \dots, x_m, y) \mapsto F(x_1, \dots, x_m, y) \end{matrix}$ C^1

If $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$ then there exist $r, s > 0$ with $B(x_0, r) \subset U$ and $(y_0 - s, y_0 + s) \subset I$ and $\varphi : B(x_0, r) \to (y_0 - s, y_0 + s)$ C^1 s.t.

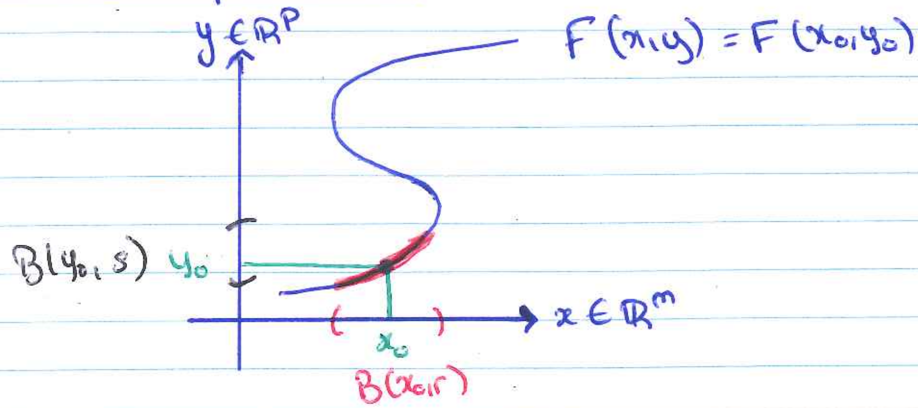
$$\forall (x, y) \in B(x_0, r) \times (y_0 - s, y_0 + s), \quad F(x, y) = F(x_0, y_0) \Leftrightarrow y = \varphi(x)$$

Remark: by computing the $\frac{\partial}{\partial x_i}$'s derivative at x_0 of $F(x, \varphi(x)) = F(x_0, y_0)$

we get:

$$\boxed{\frac{\partial \varphi}{\partial x_i}(x_0, y_0) = - \frac{\frac{\partial F}{\partial x_i}(x_0, y_0)}{\frac{\partial F}{\partial y}(x_0, y_0)}}$$

Geometric interpretation:

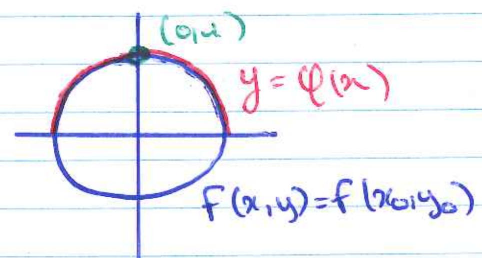


Under the assumptions of the IFT, the level set $F(x,y) = F(x_0,y_0)$ defines locally around (x_0,y_0) a function $y = \varphi(x)$ of class C^1

Example:

$$F(x,y) = x^2+y^2, \quad (x_0,y_0) = (0,1), \quad \frac{\partial F}{\partial y}(0,1) = 2 \neq 0$$

$$F(x,y) = F(x_0,y_0) \Leftrightarrow x^2+y^2 = 1$$



$$\varphi : \begin{array}{ccc} (-1,1) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sqrt{1-x^2} \end{array}$$

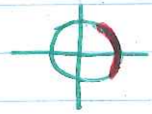
$$F(x,\varphi(x)) = 1 \Rightarrow x^2 + \varphi(x)^2 = 1 \Rightarrow 2x + 2\varphi(x)\varphi'(x) = 0$$

$$\Rightarrow 2\varphi(0)\varphi'(0) = 0$$

$$\Rightarrow \varphi'(0) = 0$$

Remark: at $(1,0)$ $\frac{\partial F}{\partial y}(1,0) = 0$ but $\frac{\partial F}{\partial x}(1,0) = 2 \neq 0$

so we can express $F(x,y) = 1$ as a function $x = \varphi(y)$



Homework: Questions from 3.1

Heuristic behind the IFT (it's not a proof!)

We want to solve $F(x,y) = F(x_0,y_0)$ around (x_0,y_0)

where the unknown is y (ie we want $F(x,\varphi(x)) = F(x_0,y_0)$)

↳ y in terms of x

By Taylor's theorem,

I am forgetting these terms: that's where I cheat...

$$F(x,y) = F(x_0,y_0) + D_xF(x_0,y_0)(x-x_0) + D_yF(x_0,y_0)(y-y_0) + \cdots$$

Hence $F(x_0,y_0) = F(x,y)$ becomes

$$\cancel{F(x_0,y_0)} = \cancel{F(x_0,y_0)} + D_xF(x_0,y_0)(x-x_0) + D_yF(x_0,y_0)(y-y_0) + \cdots$$

$$\Rightarrow 0 = D_xF(x_0,y_0)(x-x_0) + D_yF(x_0,y_0)\,y - D_yF(x_0,y_0)\,y_0 + \cdots$$

$$\Rightarrow D_yF(x_0,y_0)\,y = D_yF(x_0,y_0)\,y_0 + D_xF(x_0,y_0)\,x_0 - D_xF(x_0,y_0)\,x + \cdots$$

mult by $(D_yF)^{-1}$

$$\Rightarrow y = y_0 + [D_yF(x_0,y_0)]^{-1} D_xF(x_0,y_0)\,x_0 - [D_yF(x_0,y_0)]^{-1} D_xF(x_0,y_0)\,x + \cdots$$

↳ since $D_yF(x_0,y_0)$ is invertible

so we expressed y in terms of x (modulo some small errors in the ...)

and the linear part gives the differential of $y = \varphi(x)$:

$$D\varphi(x_0) = -[D_yF(x_0,y_0)]^{-1} D_xF(x_0,y_0)$$

Ex: Prove that $x+y+z+\sin(xyz)=0$ define z as a function of x and y in a neighborhood of $(0,0,0)$

what are $\frac{\partial z}{\partial x}(0,0)$ and $\frac{\partial z}{\partial y}(0,0)$?

Δ $f(x,y,z) = x+y+z+\sin(xyz)$ is C^1

$$\frac{\partial f}{\partial z}(0,0,0) = 1 \neq 0$$

Hence according to the IFT, $\exists\, r,\delta > 0$ and $g : \overbrace{B(0,0;r)}^{B_1} \to \overbrace{B(0;\delta)}^{B_2}$

s.t. $\forall ((x,y),(z)) \in B_1 \times B_2$, $g(x,y) = z \Leftrightarrow f(x,y,z) = 0$

By the formula given in class:

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

$$\frac{\partial g}{\partial y}(0,0) = -\frac{\frac{\partial f}{\partial y}(0,0)}{\frac{\partial f}{\partial z}(0,0)} = -1 \qquad \square$$

<u>Proof of Lagrange multipliers theorem:</u>

$f, g_1, \dots, g_p : \mathcal{U} \longrightarrow \mathbb{R}$ C^1, $\mathcal{U} \subset \mathbb{R}^m$ open

$X = \bigcap_{i=1}^{p} g_i^{-1}(0)$

If a is a local extremum of $f_{|X}$ and $\nabla g_1(a), \dots, \nabla g_p(a)$ are linearly independant

Then $\exists \lambda_1, \dots, \lambda_p \in \mathbb{R}$ s.t. $\nabla f(a) = \sum_{i=1}^{m} \lambda_i \nabla g_i(a)$

Δ We may assume that $p < m$. (if $p = m$ then $\nabla g_i(a)$ is a basis of $\mathbb{R}^m$ and $\nabla f(a) \in \mathbb{R}^m$, so there is nothing to prove, for $p > m$, $\nabla g_i(a)$ can't be lin indep)

Define $g : \mathcal{U} \longrightarrow \mathbb{R}^p$ by $g(x) = (g_1(x), \dots, g_p(x))$

$$Dg(a) = \begin{pmatrix} \frac{\partial g_1(a)}{\partial x_1} & \cdots & \frac{\partial g_1(a)}{\partial x_{m-p}} & \frac{\partial g_1(a)}{\partial x_{m-p+1}} & \cdots & \frac{\partial g_1(a)}{\partial x_m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_p}{\partial x_1}(a) & \cdots & \frac{\partial g_p(a)}{\partial x_{m-p}} & \frac{\partial g_p(a)}{\partial x_{m-p+1}} & \cdots & \frac{\partial g_p(a)}{\partial x_m} \end{pmatrix}$$

invertible (up to reording the variables) by the assumption that $\nabla g_1(a), \dots, \nabla g_p(a)$ lin indep

By the IFT, $\exists \varphi : \overbrace{B(a_1, \dots, a_{m-p}; r)}^{B_1} \longrightarrow \overbrace{B(a_{m-p+1}, \dots, a_m; \delta)}^{B_2}$

s.t. $g(u,v) = 0 \Leftrightarrow v = \varphi(u)$

hence $X \cap (B_1 \times B_2) = \{ (u,v) \in B_1 \times B_2 : v = \varphi(u) \}$ $(*)$

Define $h : B_1 \to \mathbb{R}$ by $h(x_1, \dots, x_{m-p}) = f(\underbrace{x_1, \dots, x_{m-p}, \varphi(x_1, \dots, x_{m-p})}_{\in X \text{ by } (*)})$

then $\alpha = (a_1, \dots, a_{m-p})$ is a local extremum of h and by the first derivative test and the chain rule

$\forall i = 1, \dots, m-p$

$$0 = \frac{\partial h}{\partial x_i}(\alpha) = \frac{\partial f}{\partial x_i}(a) + \sum_{j=1}^{p} \frac{\partial f}{\partial x_{m-p+1}}(a) \frac{\partial \varphi_j}{\partial x_i}(\alpha) \qquad (A)$$

From $g(x_1, \dots, x_{m-p}, \varphi(x_1, \dots, x_{m-p})) = 0$, we obtain

$\forall i = 1, \dots, m-p$

$\forall s = 1, \dots, p$

$$0 = \frac{\partial g_s}{\partial x_i}(a) + \sum_{j=1}^{p} \frac{\partial g_s}{\partial x_{m-p+1}}(a) \frac{\partial \varphi_j}{\partial x_i}(\alpha) \qquad (B)$$

Since the relations (A) and (B) are similar, the following matrix is of rank $\leq p$

$$\begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \cdots & \frac{\partial f}{\partial x_{m-p}}(a) & \frac{\partial f}{\partial x_{m-p+1}}(a) & \cdots & \frac{\partial f}{\partial x_m}(a) \\ \frac{\partial g_1}{\partial x_1}(a) & \cdots & \frac{\partial g_1}{\partial x_{m-p}}(a) & \frac{\partial g_1}{\partial x_{m-p+1}}(a) & \cdots & \frac{\partial g_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ \frac{\partial g_p}{\partial x_1}(a) & \cdots & \frac{\partial g_p}{\partial x_{m-p}}(a) & \frac{\partial g_p}{\partial x_{m-p+1}}(a) & \cdots & \frac{\partial g_p}{\partial x_m}(a) \end{pmatrix} \in M_{p+1,m}(\mathbb{R})$$

hence the rows are linearly dependant: $\exists \lambda_1, \dots, \lambda_p, \mu$ st.

$$\sum \lambda_i \nabla g_i(a) + \mu \nabla f(a) = 0$$

and $\mu \neq 0$ since the family $(\nabla g_i(a))$ is linearly independant $\square$