1 Recap

1. Bump functions
2. Einstein summation
3. Inverse function theorem

2 Implicit function theorem

Theorem 1 (Implicit function theorem (ImFT)). *Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set consisting of (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$. Let $F : U \to \mathbb{R}^m$ be a C^k ($1 \leq k \leq \infty$) function. Suppose $F(a, b) = 0$ and the matrix $\frac{\partial F^i}{\partial y^j}(a, b)$ is invertible. Then there exist neighbourhoods $a \in V_a, b \in W_b$ such that $V_a \times W_b \subset U$ and whenever $(x, y) \in V_a \times W_b$, $F(x, y) = 0 \Leftrightarrow y = g(x)$ where $g : V_a \to W_b$ is a C^k function. Moreover, $(D_x g)_{(a,b)} = -(D_y F)^{-1}_{(a,b)}(D_x F)_{(a,b)}$.*

Proof: Consider $G(x, y) = (x, F(x, y))$. Then G is C^k. Moreover, $\det(DG_{(a,b)}) = \det((D_y F)_{(a,b)}) \neq 0$. Thus by the IFT, G is a local C^k-diffeo from $V_a \times W_b$ to $G(V_a \times W_b)$. $G^{-1}(x, c) = (x, y)$ iff $c = F(x, y)$. Thus, $0 = F(x, y)$ iff $y = \pi_2 \circ G^{-1}(x, 0)$, i.e., y is a C^k function of x locally. The derivative formula follows from the chain rule. □

The proof shows that in fact, locally, $F(x, y) = c$ iff $y = G(x, c)$ where G is a C^k function. (So in a sense, we can "change coordinates" in a C^k manner from (x, y) to (x, c). In these new "coordinates", the level set $F(x, y) = a$, looks like $c = a$, i.e., we "flatten" the level sets.)

3 An application of ImFT - Lagrange's multipliers

Find the maximum value of $f(x, y, z) = x + y + z$ subject to $x^2 + y^2 + z^2 = 1$. Geometrically, it is easy to see that the answer is $\sqrt{3}$. How does one solve this using calculus alone? One possibility is to eliminate a variable using the constraint, i.e., $z(x, y) = \pm\sqrt{1 - x^2 - y^2}$ where $x^2 + y^2 \leq 1$. Thus $f(x, y, z) = f(x, y, z(x, y)) = x + y \pm \sqrt{1 - x^2 - y^2}$ over $x^2 + y^2 \leq 1$. As in one variable calculus, one can look at the interior and the boundary, and so on. This is too painful. More importantly, what if the constraint was $g(x, y, z) = x^{yz} + y^{\sin(xz^2)} + z^2 - 1 = 0$? We cannot always *explicitly* solve for one variable in terms of the others.

In general, suppose $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is a smooth function, and $a \in U$ is a point of local extremum of f subject to $g = 0$ where $g : U \to \mathbb{R}$ is a smooth function. Assume that $Dg_a \neq 0$.

Theorem 2 (Lagrange's multipliers). *In this case, there exists $\lambda \in \mathbb{R}$ such that $Df_a = \lambda Dg_a$.*

Proof: WLog assume that $\frac{\partial g}{\partial x^n}(a) \neq 0$. By the ImFT, locally, $x^n = h(x^1, \ldots, x^{n-1})$ where h is smooth. Thus, $k(x^1, \ldots, x^{n-1}) = f(x^1, \ldots, x^{n-1}, h(x^1, \ldots))$ achieves a local extremum at $(a^1, \ldots, a^{n-1})$. Thus, $Dk_{(a^1,\ldots,a^{n-1})} = 0$. Hence, $\frac{\partial f}{\partial x^i}(a) + \frac{\partial f}{\partial x^n}(a)\frac{\partial h}{\partial x^i}(a) = 0$

for all i. Since $g(x^1, \ldots, x^{n-1}, h) = 0$, we see that $\frac{\partial g}{\partial x^i}(a) + \frac{\partial g}{\partial x^n}(a)\frac{\partial h}{\partial x^i}(a) = 0$ for all i. Thus $Df_a = \lambda Dg_a$. □

In the problem above, $g = 0$ is a compact closed set. Thus f does attain a global maximum at some point a lying on $g = 0$. This point is a local extremum too. Indeed, since $Dg_a \neq (2a^1, 2a^2, 2a^2) \neq (0,0,0)$ (why?) using ImFT, locally, we can solve for one variable in a smooth manner in terms of the other variables. Since a is an interior point of the domain, the function k in the proof above attains a local maximum. Therefore, using Lagrange's theorem $Df_a = \lambda Dg_a$. Hence $(1,1,1) = \lambda(2a^1, 2a^2, 2a^2)$. Thus $a^1 = a^2 = a^3 = \frac{1}{\sqrt{3}}$. At this point, the value of f is $\sqrt{3}$.

The example of Lagrange's multipliers showed that it is easiest to optimise a "nice" function over a "nice" constraint, i.e., the constraining set is a "surface-like" object. We want to generalise this optimisation problem. To this end, we need to define "surface-like" objects (possibly with "boundary"). Our "surface-like" objects must hopefully be metric spaces at the least. Moreover, to use things like the first derivative test, we need to parametrise them locally using "coordinates". So these objects must locally look like $\mathbb{R}^n$, be Hausdorff and taking cue from Urysohn, second-countability is a natural requirement. In fact, compact Hausdorff metric spaces are second countable. (There is a better theorem called the Nagata-Smirnov theorem that needs a weaker condition called paracompactness (which is equivalent to second-countability if the manifold is connected).)

4 Topological manifolds

A topological space M is said to be a topological manifold of dimension n if M is

1. Hausdorff.
2. second countable.
3. locally euclidean of dimension n, i.e., every point of M has a neighbourhood that is homeomorphic to an open subset of $\mathbb{R}^n$.

Hausdorffness and second countability are inherited by subspaces and products. If M is *connected*, then automatically the dimension is *constant*. Topological manifolds are metrizable using the Urysohn metrization theorem.

The explicit local homeomorphism $\phi : U \subset M \to \hat{U} \subset \mathbb{R}^n$ is called a *coordinate chart*. If $\phi(p) = 0$, then (ϕ, U) is said to be a coordinate chart centred at p. The component functions $\phi(q) = (x^1(q), x^2(q), \ldots)$ are simply called "local coordinates". If $\hat{U}$ is a ball, then U is called a coordinate ball. It is fairly common to simply say "consider coordinates in a neighbourhood around p". Without loss of generality, one can assume that $\hat{U}$ is all of $\mathbb{R}^n$ itself.

Examples, non-examples:

- $\mathbb{R}^n$ with the usual topology is a topological n-fold.
- If M is a topological n-fold, then so is any *open* subset V of M: Indeed V is Hausdorff and second-countable. The restrictions of the coordinate charts give coordinate charts for V.