Lecture 17. Implicit function theorem; Tangent space; Lagrange multiplier

Let $g: \mathbb{R}^{n+m} \to \mathbb{R}^m$ continuous differentiable. $(x, y) = (x_1, \dots x_n, y_1, \dots y_m)$

Our goal is to construct $\varphi: \mathbb{R}^n \to \mathbb{R}^m$. s.t. the graph of φ

$\{(x, \varphi(x))\}$ is precisely the set $\{(x, y): g(x, y) = 0\}$

not always possible (e.g. $S^1 = \{(x, y): x^2 + y^2 = 1\}$). Fix (a, b). s.t.

$g(a, b) = 0$. then ask for a φ that works near (a, b).

Implicit function theorem: Let $g: \mathbb{R}^{n+m} \to \mathbb{R}^m$. $\in C^1$. $g(a, b) = 0$

If $D_y g(a, b)$, i.e. the Jacobian matrix $\left(\frac{\partial g_i}{\partial y_j}\right)_{1 \le i, j \le m}$ is invertible at (a, b).

Then $\exists$ open $a \in U \subseteq \mathbb{R}^n$. open $b \in V \subseteq \mathbb{R}^m$, and $\varphi: U \to V \in C^1$.

s.t. $\varphi(a) = b$. and $(x, y) \in U \times V$. $g(x, y) = 0$ iff $x \in U$. $y = \varphi(x)$.

Moreover, $D\varphi(x) = -\left(D_y g(x, y)\right)^{-1}_{m \times m} D_x g(x, y)_{m \times 1}$. $\forall x \in U$.

A simple example: $g: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$. $(x, y) \to y$.

Then $U = \mathbb{R}^n$. $V = \mathbb{R}^m$. $\varphi: \mathbb{R}^n \to \mathbb{R}^m$. $x \to 0$.

Submanifold version of implicit function theorem: let $M = \{(x, y): g(x, y) = 0\}$

If the derivative of g is "nondegenerate". then $\forall (x, y) \in M$. locally

M is the graph of $y = \varphi(x)$ for some $\varphi \in C^1$. So M is a submanifold.

In fact, let $g: \Omega \to \mathbb{R}^m$ is continuous differentiable. $\Omega \subseteq \mathbb{R}^{n+m}$ open.

$\forall c \in \mathbb{R}^m$. the inverse image of c $\{x \in \mathbb{R}^{n+m} : g(x) = c\}$ is a submanifold of codim $= m$ if $\forall x \in g^{-1}(c)$. rank $Dg(x) = m$.

x is called regular point if rank $Dg(x) = m$. and critical point otherwise.

Compared with linear algebra: $\{x \in \mathbb{R}^n : Ax = 0\}$ is a linear subspace of dim $n-m$ iff rank $A = m$.

we need $\nabla f^T(x^* - x) = 0$.

We now develop differential calculus on submanifolds. Recall that df is a linear operator to approximate f at a point. so we need linear space.

Tangent space: M is a differentiable submanifold. $p \in M \subseteq \mathbb{R}^n$.

$\gamma \in C^1: (-\varepsilon, \varepsilon) \to \mathbb{R}^n$. $t \to \gamma(t)$. where $\gamma(0) = p$.

tangent vector of γ at p is defined as $\gamma'(0) = \frac{d}{dt}\gamma\big|_{t=0}$.

tangent space of M at p: $T_pM \triangleq \{\gamma'(0) \in \mathbb{R}^n \mid \gamma((-\varepsilon, \varepsilon)) \subseteq M, \gamma(0) = p\}$

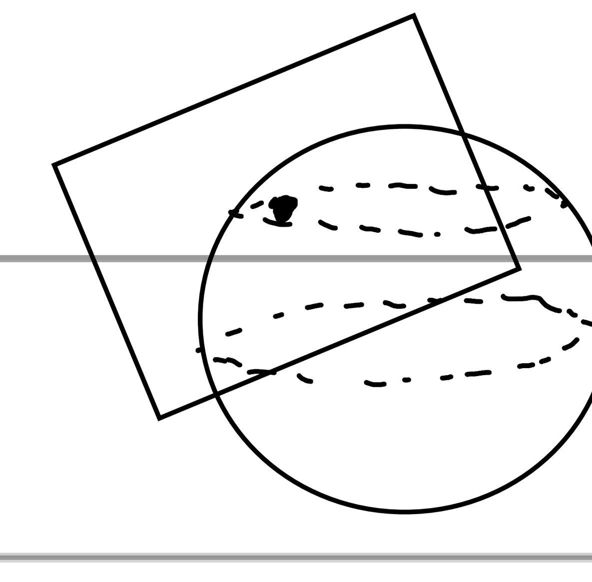
Example: $M = \mathbb{R}^d \times \{0\}^{n-d} = \{(x_1, \dots, x_d, 0, 0, \dots, 0) : x_i \in \mathbb{R}\}$.

$\forall p \in M$. $T_pM = \mathbb{R}^d \times \{0\}^{n-d}$. let $\gamma_i : t \to p + (0, \dots, 0, t, 0, \dots, 0)$.

on the other hand. $\forall \gamma : t \to (\gamma_1(t), \dots \gamma_d(t), 0, \dots, 0)$. $\gamma'(0) \in \mathbb{R}^d \times \{0\}^{n-d}$.

Remark: T_pM not depending on p is a coincidence.

Consider $M = S^2$. 2D-sphere.



T_pM is the tangent plane at p.

Our goal is to define df. linear submanifold. to approximate N

given $f: M \to N$. $df(p): T_pM \to T_{f(p)}N$. T_pM linear subspace?

Characterization of T_pM: (to show that T_pM is a d-dim linear space).

① $\exists$ open $U \subseteq \mathbb{R}^n$. open $V \subseteq \mathbb{R}^n$. diffeomorphism $\Phi: U \to V$. s.t. $p \in U$. and

$\Phi(U \cap M) = V \cap (\mathbb{R}^d \times \{0\}^{n-d})$. then. $T_pM = (D\Phi)^{-1}\big|_{\Phi(p)} (\mathbb{R}^d \times \{0\}^{n-d})$.

② assume M defined by zeros of $n-d$ $f_1, \dots f_{n-d} \in C^1$ namely. $\exists$ open U.

s.t. $p \in U$. and $M = \{x \in U : f_1(x) = \cdots = f_{n-d}(x) = 0\}$. and $\forall x \in M$.

$\nabla f_1(x), \dots \nabla f_{n-d}(x)$ are linearly independent. (x is a regular point).

then $T_pM = \bigcap_{k=1}^{n-d} \ker Df_i\big|_{x=p} = \ker Df\big|_{x=p}$. $f = (f_1, \dots, f_{n-d})$.

Remark: in some materials. ② is used to define tangent spaces. see [CZ].

Proof sketch: ① $\mathbb{R}^d \times \{0\}^{n-d} = T_{\Phi(p)}\Phi(M) = D\Phi\big|_{x=p}(T_pM)$.

② $\forall v \in T_pM$. $\exists \gamma: (-\varepsilon, \varepsilon) \to M$. $\gamma'(0) = v$. $\Rightarrow f_i(\gamma(t)) \equiv 0$. $\forall i$.

$\Rightarrow \nabla f_i(p)^T(\gamma'(0)) = 0 \Rightarrow T_pM \subseteq \ker Df_i$. $\dim \bigcap_{k=1}^{n-d} \ker Df_i \geq d$. $\Rightarrow$

$\cap \ker Df_i = T_pM$.

Example: tangent space and normal vector.

Suppose $M \subseteq \mathbb{R}^n$ is a (super)sphere. defined by $\{x \in \mathbb{R}^n : f(x) = 0\}$.

$f: \mathbb{R}^n \to \mathbb{R}^m \in C^1$. $\forall p = (x_1, \dots x_n) \in M$. $T_pM = \ker Df(p)$.

$T_pM = \{v \in \mathbb{R}^n : \langle v, \nabla f(p) \rangle = 0\}$. $\nabla f(p)$ is the normal vector of T_pM.

Example: cylinder $\{(x, y, z) : f(x, y, z) = x^2 + y^2 - 1 = 0\}$.

$\nabla f(p) = (2x, 2y, 0)$. $\nabla f(p) \perp \begin{matrix} (-x, y, 0) \\ (0, 0, 1) \end{matrix}$. so $T_pM = \mathrm{span}\{(-x, y, 0), (0, 0, 1)\}$

Mapping and differential on submanifolds:

Smooth function (or C' function): $M \subseteq \mathbb{R}^n$. $N \subseteq \mathbb{R}^m$. $f: M \to N$.

① $\forall p \in M$. $\exists \underset{\text{open}}{U \subseteq \mathbb{R}^n}$. $p \in U$. $\exists C^\infty$ (or C') function $F: U \to \mathbb{R}^m$.

s.t. $f|_{U \cap M} = F|_{U \cap M}$ (f is locally F restricted to M).

② $\forall p \in M$. $\exists \underset{\text{open}}{U \subseteq \mathbb{R}^n}$. $p \in U$. $\underset{\text{open}}{V \subseteq \mathbb{R}^n}$ and diffeomorphism $\Phi: U \to V$

s.t. $\Phi(U \cap M) = V \cap (\mathbb{R}^d \times \{0\}^{n-d})$. and $f_\Phi = f \circ \Phi \in C^\infty (C')$.

$f_\Phi : V \cap (\mathbb{R}^d \times \{0\}^{n \times d}) \longrightarrow \mathbb{R}^m$. then $f \in C^\infty$(or C') (M, N).

Remark: $M \subseteq \mathbb{R}^n$. $F: \mathbb{R}^n \to \mathbb{R}^m$ if $F(M) \subseteq N \subseteq \mathbb{R}^m$. then.

F restricted to M: $f = F|_M : M \to N$ is a smooth mapping (→ or C')

Differential: $f \in C'(M, N)$. $\forall p \in M$. $\exists \gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$. $\gamma(0) = p$. $\gamma'(0) = v$.

tangent vector $df(p)(v) = \frac{d}{dt}\Big|_{t=0} f(\gamma(t)) \in T_{f(p)}N$. $df(p): T_pM \to T_{f(p)}N$

Example: If $f = F|_M$. then $df(p)$ is $DF(p): \mathbb{R}^n \to \mathbb{R}^m$ restricted to T_pM.

Proposition: $f \in C^\infty$ (or C^1) (M, N). $df(p): T_pM \to T_{f(p)}N$ is linear operator.

The key lemma (first order condition for optimality on submanifolds).

Suppose $M \subseteq \mathbb{R}^n$ is a submanifold. $f \in C^1(M)$. p is a local extreme point of f. $p \in M$. Then $df(p) = 0$.

Proof. By definition. it suffices to show that $\forall v \in T_pM$. $\nabla_v f(p) = 0$.

Here we define $\nabla_v f(p) = \frac{d}{dt}\Big|_{t=0} f(\gamma(t))$ where $\gamma(0) = p$ and $\gamma'(0) = v$.

Now we consider $\gamma: (-\varepsilon, \varepsilon) \to M$. $t \to \gamma(t)$. $\gamma(0) = p$ and $\gamma'(0) = v$.

Since p is a local mininum point of $f(\gamma(t))$. $f'(\gamma(t)) \geq 0$.

$\Rightarrow \nabla_v f(p) \geq 0$. T_pM is a linear subspace $\Rightarrow -v \in T_pM$.

$\Rightarrow -\nabla_v f(p) \geq 0 \Rightarrow \nabla_v f(p) = 0, \quad \forall v \in T_pM$. $\square$

Lagrange multiplier: Let $f: \mathbb{R}^n \to \mathbb{R}$. $g: \mathbb{R}^n \to \mathbb{R}^m$. opt. $\begin{array}{l} \min f(x) \\ \text{s.t. } g(x) = 0 \end{array}$.

Suppose f. $g \in C^1$. and x^* is an optimal solution. If x^* is regular (i.e. rank $Dg(x^*) = \underset{<n}{m}$). Then $\exists$ unique λ^*. s.t. $\nabla L(x^*, \lambda^*) = 0$. Lagrangian.

Proof. $\forall v$. $df(x^*)(v) = \nabla_v f(x^*) = 0 \Rightarrow T_pM \subseteq \ker df(x^*)$. $T_pM = \ker dg$.

$\Rightarrow \ker df(x^*) \supseteq \ker dg(x^*)$. i.e. $\exists \lambda^*$. $df(x^*) = \lambda_1^* dg_1(x^*) + \cdots + \lambda_m^* dg_m(x^*)$.

$\square$