Analysis 2 Lecture Notes – Upendra Kulkarni

Soham Chatterjee

Abstract

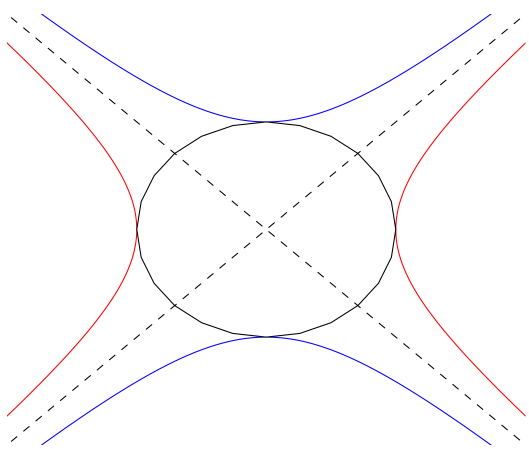
This is the lecture notes scribed by me. If you find any mistakes in the notes please email me at sohamc@cmi.ac.in.

The whole course is taken by Prof. Upendra Kulkarni, online. If you want the lectures then you can find them in this link. Sir mainly followed Prof. Pramath Sastry's Notes (https://www.cmi.ac.in/~pramath/teaching.html#ANA2). You can find all the assignments problems in the following drive link. Through out the course the books we followed is Principles of Mathematical Analysis by Walter Rudin.

Chapter 1

Constrained Optimizations and Lagrange Multipliers

Example: Optimize $f(x, y)y^2 - x^2$ subject to the constraint $h(x, y) = x^2 + y^2 = 1$



In other words we want to find extrema of $f|_M$ where M is the level curve for h at level 1 i.e. $M = h^{-1}(1)$

Form the way level sets of f interact with M, here we see that we have maximum at $(0, \pm 1)$ and minimum at $(\pm 1, 0)$

It also appears that the constrained graph and the level curve of the objective function f are $\underset{\overset{\downarrow}{\text{What it means}}}{\text{tangential}}$ to each other

We will define Tangent Space to a level set of a C^1 function at a point p on M

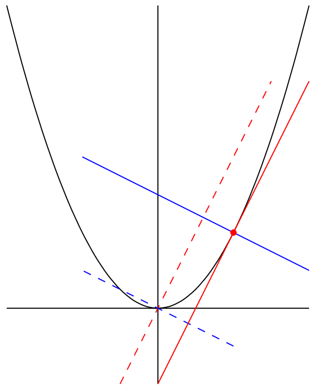
1.1 Tangent Space

Definition 1.1.1: Tangent Space

Tangent Space to a hypersurface $M = f^{-1}(c)$ in $\mathbb{R}^n$ where $f : (\text{Open } U \subset \mathbb{R}^n) \to \mathbb{R}$ is a C^1 function and $c \in \mathbb{R}$ at a point $p \in M$ is a subspace of $\mathbb{R}^n$ defined to be

$$T_pM = \ker(f'(p)) = \{v \in \mathbb{R}^n \mid f'(p)(v) = 0\} = \{v \in \mathbb{R}^n \mid \nabla f(p) \cdot v = 0\}$$

Geometric tangent space considering to our mental image $= T_pM + p =$ Shift T_pM by vector p. Likewise define Normal Space to be the set of vectors orthogonal to T_pM i.e. $T_pM^\perp$



Eg. $f(x,y) = y - x^2$, $M = f^{-1}(0)$. $p = (3,9) \in M$.
Here $f'(p) = \begin{bmatrix}-2x & 1\end{bmatrix}_{(3,9)} = \begin{bmatrix}-6 & 1\end{bmatrix}$

$$\begin{bmatrix}x\\y\end{bmatrix} \mapsto -6x + y$$

$T_pM = \ker(f'(p)) = \{(x,y) \mid y = 6x\} =$ red line.
$N_pM =$ line $y = -\frac{1}{6}x =$ blue line
Geometric tangent space $= T_pM + p$ and geometric normal space $= N_pM + p$

1.2 Lagrange Multiplier

Let U be open in $\mathbb{R}^n$. $f : U \to \mathbb{R}$ objective function and $h : U \to \mathbb{R}$ constraint function. Want to find extrema of f restricted to the level set $M = \{x \in U \mid h(x) = c\} = h^{-1}(c)$ for $c \in \mathbb{R}$

$f|_M$ has local maxima at $p \in M$ means for some $W \subset U$, $f(p) \geq f(x)\ \forall\ x \in W \cap M$

Definition 1.2.1: C^1 Path and Velocity Vector

A C^1 path centered at $p \in U$ in $U \subset \mathbb{R}^n$ is a C^1 map $\gamma : (-\varepsilon, \varepsilon) \to U$ where $0 \mapsto p$. We call $\gamma'(0) =$ velocity of γ at 0

Theorem 1.2.1 Lagrange Multiplier

Let U be open in $\mathbb{R}^n$. $f : U \to \mathbb{R}$, $h : U \to \mathbb{R}$. Let f, h are C^1 functions. Let $M = h^{-1}(c)$. If $h'(p) \neq 0$ and $f|_M$ has a local extrema at $p \in M$ then $\exists! \lambda \in \mathbb{R}$ such that

$$\nabla f(p) = \lambda \nabla h(p)$$

Proof. Consider paths on level set $M = h^{-1}(c)$ i.e.

$$\begin{tikzcd} \gamma : (-\varepsilon,\varepsilon) \arrow[r] \arrow[dr] & M \arrow[d, phantom, "\cap"] & = h^{-1}(c) \\ & U \arrow[r, "h"] & \mathbb{R} \end{tikzcd}$$

Then $h = \gamma(t) = c\ \forall\ t \in (-\varepsilon, \varepsilon)$. Hence by Chain Rule

$$h'(p)\gamma'(0) = \nabla h(p) \cdot \gamma'(0) = 0$$

i.e. $\{$velocity vectors of all paths γ on M centered at $p\} \subset T_pM$
Key Fact: When $h'(p) \neq 0$ we have equality! Proof of this fact uses Implicit Function Theorem

Now let's recall the objective function f and recall that p s assured to be a local max/min. If gm is a C^1 curve on M then in particular $f|_{\text{image}(\gamma)}$ also has a max/min at p. Therefore

$$0 = (f \circ \gamma)'(o) = \nabla f(p) \cdot \gamma'(0)$$

i.e. ∇f is orthogonal to velocity vectors to all curves centered at p.

By claim $\nabla f(p) \perp T_pM$, we already say $\nabla h(p) \perp T_pM$. We know $\nabla h(p) \neq 0$ by assumption. Hence $\exists! \lambda$ such that $\nabla f(p) = \lambda \nabla h(p)$ $\square$

1.3 Some Examples for Applications

(i) $f(x,y) = y^2 - h^2$ and $h(x,y) = x^2 + y^2$, $c = 1$. Therefore $M = h^{-1}(1) =$ Unit Circle

Suppose $p = \begin{bmatrix} a \\ b \end{bmatrix}$ is an extremum of $f|_M$

$$\nabla f(p) = \begin{bmatrix} -2x \\ 2y \end{bmatrix}_{(a,b)} = \begin{bmatrix} -2a \\ 2b \end{bmatrix} \qquad \nabla h(p) = \begin{bmatrix} 2x \\ 2y \end{bmatrix}_{(a,b)} = \begin{bmatrix} 2a \\ 2b \end{bmatrix}$$

We know that $\exists! \lambda \in \mathbb{R}$ such that

$$\begin{bmatrix} -2a \\ 2b \end{bmatrix} = \lambda \begin{bmatrix} 2a \\ 2b \end{bmatrix}$$

This is not possible unless one of a, b os 0. Therefore

$$\begin{aligned} a = 0 &\implies b = \pm 1 \text{ and } \lambda = 1 \\ b = 0 &\implies a = \pm 1 \text{ and } \lambda = -1 \end{aligned}$$

(ii) $f(x, y) = y$ is subject to constraint $h(x, y) = y - g(x) = 0$ where $g : \mathbb{R} \to \mathbb{R}$ is some C^1 function. This is equivalent to finding extrema of $y = g(x)$ as in school

Suppose $p = \begin{bmatrix} a \\ b \end{bmatrix}$ gives an extremum

$$\nabla f(p) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \lambda \nabla h(p) = \lambda \begin{bmatrix} -g'(a) \\ 1 \end{bmatrix}$$

i.e. $1 = \lambda \implies 0 = -\lambda g'(a) \implies g'(a) = 0$ as expected

(iii) $f(x, y) = x^2$ subject to $h(x, y) = y = 0$

$$\nabla f = \begin{bmatrix} 2x \\ 0 \end{bmatrix} = \lambda \nabla h = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies \lambda = 0, x = 0$$

Note:-

If we instead take $h(x,) = y^2$, then we get $(, xy) = (0, 0)$ but λ arbitrary

(iv) $f(x, y) = xy$ subject to $h(x, y) = \frac{x^2}{9} + \frac{y^2}{4} = 1$

$$\nabla f = \begin{bmatrix} y \\ x \end{bmatrix} = \lambda \nabla h = \lambda \begin{bmatrix} \frac{2x}{9} \\ \frac{y}{2} \end{bmatrix}$$

Therefore

$$y = \frac{2x}{9}\lambda, \quad x = \frac{y}{2}\lambda, \quad \frac{x^2}{9} + \frac{y^2}{4} = 1$$

$\lambda = \pm 3$. Find extrema. As constraint = ellipse, a compact set, evaluating f as candidates is enough to find max and min.

(v) Find the points on the sphere $x^2 + y^2 + z^2 = 9$ closest/furthest from $(a, b, c) \to$ arbitrary point in $\mathbb{R}^3$

$f(x, y, z) = (x - a)^2 + (y - b)^2 + (z - c)^2$ and $h(x, y, z) = x^2 + y^2 + z^2 = 9$. Complete this and see that geometrically obvious solution emerge

Next we will prove Inverse Function Theorem and Implicit Function Theorem and come back to justify the claim. In fact we will then be able to prove the general version of Lagrange Multiplier Method i.e. with multiple constraints

1.4 Generalized Lagrange Multiplier

Theorem 1.4.1 Generalized Lagrange Multiplier

U open $\subset \mathbb{R}^n = \mathbb{R}^{d+m}$ want to find extrema of objective function $f : U \to \mathbb{R}$ subject to constraint $h = c$

for a C^1 function: $U \to \mathbb{R}^m$ where $c \in \mathbb{R}^m$ i.e. we want to find extreme of $f|_{M=h^1(c)}$

Key Assumption: $\forall\ x \in M$, $h'(x)$ is surjective i.e. $h'(x) : \mathbb{R}^{d+m} \to \mathbb{R}^m$. (So $\ker(h'(x))$ has $\dim d$. Recall we called $\ker(h'(x)) = T_x M$)
Suppose $f|_M$ has a local extremum at $p \in M$ Then $\exists!$ real numbers $\lambda_1, \lambda_2, \ldots, \lambda_m$ such that

$$\nabla f(p) = \lambda_1 \nabla h_1(p) + \cdots + \lambda_n \nabla h_n(p)$$

where $h(p) = \begin{bmatrix} h_1(p) & \cdots & h_m(p) \end{bmatrix}^T \in \mathbb{R}^m$

Proof. we will show that

1. $\nabla f(p) \perp T_p M$
2. Any vector $\perp T_p M$ is a linear combination of $\nabla h_i(p)$

These are the steps.

1. Let $v \in T_p M = \ker(f'(p))$. By HW4 Problem v can be represented by some curve based at $p$ni.e. we can find a C^1 curve $\gamma : (-\varepsilon, \varepsilon) \to M \subset U$ where $0 \mapsto p$ such that $\gamma'(0) = v$.

   As we have an extremum of $f|M$ at p it is also an extreme point for $(-\varepsilon, \varepsilon) \xrightarrow{\gamma} M \xrightarrow{f} \mathbb{R}$. So by 1-Variable Calculus $(f \circ \gamma)'(0) = 0$ i.e. $f'(p)\gamma'(0) = 0$ i.e. $\nabla f(p) \cdot v = 0$

2. For every curve γ as above $h \circ \gamma =$ constant. Therefore $h'(p)\gamma'(0) = 0$ i.e. $\nabla h'(p) \cdot v = 0$

   $$h'(p) = \begin{bmatrix} \frac{\partial h_1}{\partial x_1}(p) & \cdots & \frac{\partial h_1}{\partial x_n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1}(p) & \cdots & \frac{\partial h_m}{\partial x_n}(p) \end{bmatrix} = \begin{bmatrix} \nabla h_1(p)^T \\ \vdots \\ \nabla h_m(p)^T \end{bmatrix} = \begin{bmatrix} \nabla h_1(p) & \cdots & \nabla h_m(p) \end{bmatrix}^T$$

   So

   $$\nabla h_1(p) \cdot v = 0, \ldots, \nabla h_m(p) \cdot v = 0$$

   Therefore $\underbrace{\nabla h_i(p)}_{\substack{m \text{ linearly} \\ \text{independent} \\ \text{vectors}}} \perp \underbrace{T_p M}_{\substack{\dim n-m \\ =d}}$. Everything is in $\mathbb{R}^n = \mathbb{R}m + d$. $\therefore (T_p M)^\perp$ has $\nabla h_1(p), \ldots, \nabla h_m(p)$ as a basis. i.e. 2 is proved

$\square$

Chapter 2

Implicit Function Theorem

Notation: For $n > m$ let $n = m + d$. Write points of $\mathbb{R}^n = \mathbb{R}^d \times \mathbb{R}^m$ as (x, y) where $x \in \mathbb{R}^d$, $y \in \mathbb{R}^m$

Theorem 2.1 Implicit Function Theorem

Let U open in $\mathbb{R}^{d+m}$. $\Phi : U \to \mathbb{R}^m$ is a C^1 map such that $\Phi'(p)$ is surjective (which means columns of the $m \times (d+m)$ matrix of $\Phi'(p)$ span $\mathbb{R}^m$). WLOG suppose the last m columns of $\Phi'(p)$ are linearly independent and hence span $\mathbb{R}^m$ i.e. the $m \times m$ matrix "$\left.\frac{\partial \Phi}{\partial y}\right|_p = \begin{bmatrix} D_{d+1}\Phi(p) & \cdots & D_{d+m}\Phi(p) \end{bmatrix}$" is invertible. Then

1. $\exists$ a neighorhood W of a in $\mathbb{R}^d$ and a unique C^1 map $W \xrightarrow{f} \mathbb{R}^m$ such that $\left.\begin{array}{l} f(a) = b, (x, f(x)) \in U \ \forall\ x \in W \text{ and} \\ \Phi(x, f(x)) = c \ \forall\ x \in W \end{array}\right\}$ i.e. f is an implicit solution to the equation $\Phi(x, y) = c$
2. One can calculate $f'(x)$ by "Implicit Differentiation"

To understand this, first examine two cases:

- When Φ is a linear map given by a matrix A. Here we are solving the equation $A \begin{bmatrix} x \\ y \end{bmatrix} = c$
- $d = m = 1$ i.e. $n = 2$ $\Phi(x, y) = x^2 + y^2 - 1$, solving $\Phi(x, y) = 0 = c$. When $\left.\frac{\partial |phi}{\partial y}\right|_{p=(a,b)} \neq 0$ we can locally solve for y in terms of x near p.
$$D\Phi = \left.\begin{bmatrix} \frac{\partial \Phi}{\partial x} & \frac{\partial \Phi}{\partial y} \end{bmatrix}\right|_{(a,b)} = \begin{bmatrix} 2a & 2b \end{bmatrix}$$
$2b = 0$ at $(\pm 1, 0)$

Proof. We will choose W later. Define

$$U \xrightarrow{\quad\psi\quad} \mathbb{R}^{d+m}$$

$$(x, y) \longmapsto (x, \Phi(x, y))$$

Note ψ' has the matrix $\begin{bmatrix} I & O \\ \frac{\partial \Phi}{\partial x} & \frac{\partial \Phi}{\partial y} \end{bmatrix}$. This is nonsingular in a neighborhood of p. So by Inverse Function Theorem ψ is invertible with C^1 inverse in a neighborhood V of p

$$\begin{array}{rcl} V & \longleftrightarrow & \psi(V) \\ (a, b) & \longmapsto & (a, c) \\ (x, y) & \longmapsto & (x, \Phi(x, y)) \\ (u, \alpha(u, v)) & \longleftarrow\!\!\!\shortmid & (u, v) \end{array}$$

Definition of $\alpha(u,v)$ defined on $\psi(V)$. This tells us $\alpha(a,c)=b$. Whenever $\Phi(x,y)=c$ i.e.

$$(x,y)\xrightarrow{\Phi}(x,c)\xrightarrow{\psi^{-1}}(x,\alpha(x,c))=(x,y)$$

i.e. $y=\alpha(x,c)$ and $\Phi(x,\alpha(x,c))=c$

So we are forced to define $f(x)=\alpha(x,c)$. But what should be the domain of this function f i.e. what should we take W to be.

$$(a,c)\in\psi(V) \text{ is open} \supset \begin{pmatrix}\text{open ball } W\\ \text{around } a \text{ in } \mathbb{R}^d\end{pmatrix}\times\{c\}$$

Now for any $x\in W$ we know $(x,c)\in\psi(V)$ i.e. $(x,\alpha(x,c))\in V$ so we define $f:W\to\mathbb{R}^m$ where $f(x)=\alpha(x,c)$ and we have derived the function. Now ϕ^{-1} is C^1 and α is component of ϕ^{-1} so all components of ϕ^{-1} is also C^1. hence f is C^1

Uniqueness of f is not true in general for arbitrary W. $\Phi(x,y)=x^2+y^2,\ c=1$. In $W=W_1\sqcup W_2$

$$f(x)=\begin{cases}\sqrt{1-x^2} & x\in W_1 \qquad \text{[is forced]}\\ \sqrt{1-x^2} \text{ or } -\sqrt{1-x^2} & x\in W_2\end{cases}$$

. We have choice for f on W_2.

If W is connected, f will be unique. Eg. take W to be a ball. Suppose g is another solution to $\Phi(x,y)=c$ i.e. $\Phi(x,g(x))=c$ for $x\in W$ and $g(a)=b$. Then consider the set $S=\{x\in W\mid f(x)=g(x)\}$. Show that this set is both closed (easy $S=(f-g)^{-1}(0)$) and open.

Calculate derivative of f using the fact that $\psi\circ\psi^{-1}=$ Identity and Chain Rule. □

Example 2.0.1 (Application of Implicit Function Theorem)

(i) Linear map $\Phi:\mathbb{R}^{d+m}\to\mathbb{R}^m$ given by matrix A. Given $A\begin{bmatrix}a\\b\end{bmatrix}=c$. Want to solve $A\begin{bmatrix}x\\y\end{bmatrix}=c$

$A=[P\mid Q]$ where P is $m\times d$ and Q is $m\times m$ and Q is invertible. i.e.

$$[P\mid Q]\begin{bmatrix}x\\y\end{bmatrix}=c \iff [Q^{-1}P\mid I]\begin{bmatrix}x\\y\end{bmatrix}=Q^{-1}c \iff Q^{-1}Px+y=Q^{-1}c$$
$$\iff y=Q^{-1}c-Q^{-1}Px$$

(ii) We can solve for y in terms of x near any (a,b) on the unit circle when $\left.\frac{\partial\Phi}{\partial y}\right|_{(a,b)}\neq 0$. [This is mate when $b\neq 0$ i.e. at all points except $(\pm1,0)$].

$$D\Phi|_{(a,b)}=\begin{bmatrix}2a & 2b\end{bmatrix}$$

We can see directly

$$\left.\begin{matrix}\text{when } b>0 & y=\sqrt{1-x^2}\\ \text{when } b<0 & y=-\sqrt{1-x^2}\end{matrix}\right\}\text{near (a,b) in fact } \forall\ x\in(-1,1)$$

Similarly we can solve for x in terms of y when $\left.\frac{\partial\Phi}{\partial x}\right|_{(a,b)}=2a\neq 0$ This is true when $a\neq 0$

Remark: Implicit Function Theorem gives a sufficient condition to be able to locally solve a system of linear equations

$$\left.\begin{aligned}\Phi_1(x_1,\dots,x_d,y_1,\dots,y_m) &= c_1\\ \Phi_2(x_1,\dots,x_d,y_1,\dots,y_m) &= c_1\\ \vdots \qquad\quad \vdots \qquad\quad &\vdots\\ \Phi_m(x_1,\dots,x_d,y_1,\dots,y_m) &= c_1\end{aligned}\right\} \begin{array}{l}\text{for } y_i\text{'s in terms of } x_i\text{'s}\\ \text{locally near a given solution}\\ y=b \text{ and } x=a\end{array}$$

Note:-

The condition of invertibility of submatrix of Φ is not necessary. Eg. $\Phi(x,y) = y - x^3$ near $(0,0)$

$$D\Phi\big|_{(0,0)} = \begin{bmatrix}-3x^2 & 1\end{bmatrix}\Big|_{(0,0)} = \begin{bmatrix}0, 1\end{bmatrix} \qquad \left.\frac{\partial \Phi}{\partial x}\right|(0,0) = 0$$

but still we can solve for x in terms of y: $x = \sqrt[3]{y}$