MATH 2860 RC 1

Presented by Yang Mo

Dear all.

It is my honor to see you still in the honors math course! It is the last honor math course and I think the MATH 2860 is quite hard. This course concerns mostly about differential equations, linear algebra, complex analysis and series expansions. For me, I think math is not learned until I can apply it. You'll find that, there are always differential equations when you learn about electrodynamics, there will always be linear algebra when you want to do robotics, there will always be Fourier and Laplace transformation when you are studying signals. Last but not least, when everything is put on a complex foundation, things are extremely ridiculous but somehow interesting. Don't think you're bad at math if you find this course hard. Hilbert once said, "You won't master a math course until you studied it for five times." You'll meet all your friends in math in signal processing, electrodynamics, control theories, robotics, ML theory, operating science, financial engineering, fluid dynamics and etc. all the time. Wish you enjoy this course and cherish your time in the last fundamental math course.

Yours sincerely,

Yang Mo

Implicit Theorem

2.8. Implicit Function Theorem. Let $\Omega \subset \mathbb{R}^n_x \times \mathbb{R}^m_y$ be open and $F \in C^1(\Omega, \mathbb{R}^m)$ such that $F(x_0, y_0) = 0$ for some $(x_0, y_0) \in \Omega$. Assume further that

$$\det DF(x_0, \cdot)|_{y_0} \neq 0. \tag{2.5}$$

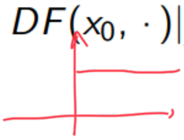
Then there exists an open ball $B_\varepsilon(x_0)$ and a C^1-map $g\colon B_\varepsilon(x_0) \to \mathbb{R}^m_y$ such that

$$F(x, g(x)) = 0 \qquad \text{for all } x \in B_\varepsilon(x_0).$$

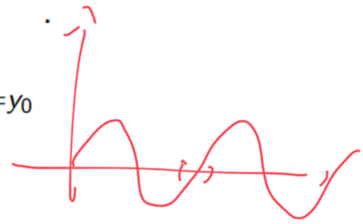
$y = g(x)$ Ex 1. $f(x,y) = x^2 + y$. $\Rightarrow y = -x^2$

We note that

Ex 2. $f(x,y) = \cos(x+y)\sin x$



$$DF(x_0, \cdot)|_{y_0} = \left.\begin{pmatrix} \frac{\partial F_1(x_0,y)}{\partial y_1} & \cdots & \frac{\partial F_1(x_0,y)}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m(x_0,y)}{\partial y_1} & \cdots & \frac{\partial F_m(x_0,y)}{\partial y_m} \end{pmatrix}\right|_{y=y_0}$$



Properties of the Implicitly Defined Function

Let us suppose again that $F\colon \mathbb{R}^2 \to \mathbb{R}$ and $F(x, y) = 0$, defines a curve. Suppose and that in some x-interval we can find a function g such that i.e.,

$$F(x, g(x)) = 0. \tag{2.3}$$

We might try to find the derivative of g as follows:

$$0 = \frac{d}{dx} F(x, g(x)) = \partial_1 F(x, g(x)) + \partial_2 F(x, g(x)) \cdot g'(x)$$

so

$$\boxed{g'(x) = -\frac{\partial_1 F(x, g(x))}{\partial_2 F(x, g(x))}} \tag{2.4}$$

if F and g are differentiable and if $\partial_2 F(x, y) \neq 0$. This last condition is actually the same as (2.1).

We consider the two-dimension case. For $F(x, y) = 0$, suppose we want to study whether the function can be write out explicitly at point (x_0, y_0). If F satisfies:

- $F(x, y) \in C^1(D)$
- $\exists (x_0, y_0) \in D \; s.t. \; F(x_0, y_0) = 0$
- $(F'_y(x_0, y_0))^2 + (F'_x(x_0, y_0))^2 \neq 0$

Then $F(x, y) = 0$ has a explicit form in the neighborhood of point (x_0, y_0) of $y = f(x)$. Besides, $f(x)$ posses a continuous derivative as $

$$\boxed{\frac{df}{dx} = -\frac{F'_x(x, y)}{F'_y(x, y)}} \tag{1}$$

Note: There is a minus sign!

Exercise

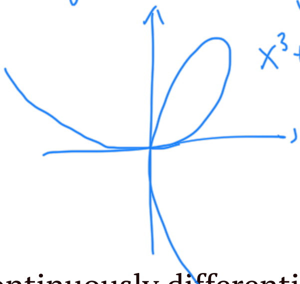
1. Find the derivative of $F(x, y) = x^3 + y^3 - 3axy$ at point $(x_0, y_0) = (0, 0)$.

$F(x,y) = x^3 + y^3 - 3axy.$

$F'_x = 3x^2 - 3ay. \quad F'_x(0,0) = 0$

$F'_y = 3y^2 - 3ax \quad F'_y(0,0) = 0$

No implicit function.



$x^3 + y^3 - 3axy = 0.$

(leaf-like curve)

2. Prove: For a continuously differentiable function F, if $F(x, y, z) = 0$, then

$$\frac{\partial x}{\partial y} \cdot \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial x} = -1 \tag{2}$$

$\frac{\partial x}{\partial y} = -\frac{F'_y}{F'_x}$ $\qquad$ $\frac{\partial x}{\partial y} \cdot \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial x} = -1.$

$\frac{\partial y}{\partial z} = -\frac{F'_z}{F'_y}$

$\frac{\partial z}{\partial x} = -\frac{F'_x}{F'_y}$

Inverse Function Theorem

Inverse Function Theorem

While all of the results in this section hold if X and Y are complete, infinite-dimensional vector spaces, we will from here on out assume that X and Y are finite-dimensional.

1.7. Inverse Function Theorem. Let X, Y be finite-dimensional, normed vector spaces, $\Omega \subset X$ an open set, $x_0 \in \Omega$ and $f: \Omega \to Y$ continuously differentiable.

Assume further that $Df|_{x_0}$ is invertible. Then f is locally C^1-invertible at x_0 and the locally defined inverse function f^{-1} has derivative

$$Df^{-1}|_{f(x)} = (Df|_x)^{-1} \in \mathscr{L}(Y, X).$$

You can understand that with the help of single-dimension case. Find in what case you can have a inverse function for $y = f(x)$. $Df^{-1}(x)$ $(Df(x))^{-1}$

Also, I want to provide another point of view here to understand the inverse function theorem. For example, for the function $x = x(u, v)$ and $y = y(u, v)$, we can see them as an implicit function: $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x(u,v) \\ y(u,v) \end{pmatrix}$, $\mathbb{R}^2 \to \mathbb{R}^2$.

$F(x, y, u, v) = x - x(u, v)$ and $G(x, y, u, v) = y - y(u, v) = 0$ and then try to solve $u = u(x, y)$ and $v = v(x, y)$.

$$\begin{cases} F(x,y,u,v) = x - x(u,v) = 0. \\ G(x,y,u,v) = y - y(u,v) = 0. \end{cases}$$

Find. whether can get $\begin{cases} u = u(x,y) \\ v = v(x,y) \end{cases}$?

Suppose we have inverse theorem.

$$\begin{cases} F(x, y, u(x,y), v(x,y)) = 0. \\ G(x, y, u(x,y), v(x,y)) = 0 \end{cases} \Rightarrow \begin{cases} F_x' + F_x' u_x' + F_x' v_x' = 0 \\ G_x' + G_x' u_x' + G_x' v_x' = 0, \\ F_y' + F_y' u_y' + F_y' v_y' = 0, \\ G_y' + G_y' u_y' + G_y' v_y' = 0 \end{cases}$$

$$\frac{\partial u}{\partial x} = \frac{\frac{\partial(F,G)}{\partial(x,v)}}{\frac{\partial(F,G)}{\partial(u,v)}} = \frac{y_v'}{x_u' y_v' - x_v' y_u'}$$

$\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, $\frac{\partial v}{\partial y}$.

Exercise

1. Find the derivative of the inverse map of $f(x, y) = (e^x cos(y), e^x sin(y))$.

$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e^x \cos y \\ e^y \sin y \end{pmatrix}$ $\quad Df|_{(x,y)} = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ -e^x \sin y & e^x \cos y \end{pmatrix}$

$\det(Df|_{(x,y)}) = e^{2x}$

$D^{-1}f|_{(u,v)} = \frac{1}{e^{2x}} \begin{pmatrix} e^x \cos y & e^x \sin y \\ -e^x \sin y & e^x \cos y \end{pmatrix}.$

$u^2+v^2 = (e^x)^2 \quad e^x = -\sqrt{u^2+v^2}. \quad RHS = \frac{1}{\sqrt{u^2+v^2}} \begin{pmatrix} u & v \\ -v & u \end{pmatrix}.$

2. Find the derivative of the inverse map of $f(x, y) = (x^2, y/x)$.

$f(x,y) = \begin{pmatrix} x^2 \\ \frac{y}{x} \end{pmatrix} \quad Df|_{(x,y)} = \begin{pmatrix} \frac{df_1}{dx} & \frac{df_1}{dy} \\ \frac{df_2}{dx} & \frac{df_2}{dy} \end{pmatrix} = \begin{pmatrix} 2x & 0 \\ -\frac{y}{x^2}, & \frac{1}{x} \end{pmatrix} \quad \det = 2.$

$(Df|_{(x,y)})^{-1} = \frac{1}{2} \begin{pmatrix} \frac{1}{x} & 0 \\ \frac{y}{x^2} & 2x \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \frac{1}{\sqrt{u}} & 0 \\ \frac{v}{\sqrt{u}} & 2\sqrt{u} \end{pmatrix}.$

$f^{-1}(u,v). \quad u = x^2 \quad v = \frac{y}{x} \quad x = \sqrt{u} \quad y = vx = v\sqrt{u}$

Lagrange Multipliers for Constrained Extrema

Finding Free Extrema

Let $\Omega \subset \mathbb{R}^n$ and $f : \Omega \to \mathbb{R}$.

1. f is said to have a *(global) maximum* at $\xi \in \Omega$ if
   $x \in \Omega \quad \Rightarrow \quad f(x) \le f(\xi).$
2. f is said to have a *strict (global) maximum* at $\xi \in \Omega$ if
   $x \in \Omega \setminus \{\xi\} \quad \Rightarrow \quad f(x) < f(\xi).$
3. f is said to have a *local maximum* at $\xi \in \Omega$ if there exists a $\delta > 0$ such that
   $x \in \Omega \cap B_\delta(\xi) \quad \Rightarrow \quad f(x) \le f(\xi).$
4. f is said to have a *strict local maximum* at $\xi \in \Omega$ if there exists a $\delta > 0$ such that
   $x \in \Omega \cap B_\delta(\xi) \setminus \{\xi\} \quad \Rightarrow \quad f(x) < f(\xi).$

The function f is said to have a (strict) *global/local minimum* at ξ if $-f$ has a (strict) global/local maximum at ξ.

Essential Condition. Let $\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$, and $\xi \in \operatorname{int} \Omega$. Assume that all partial derivatives of f exist at ξ and that f has a local extremum (maximum or minimum) in ξ. Then
$\nabla f(\xi) = 0$. If f is differentiable at ξ, this implies $Df|_\xi = 0$.

Hessian & Extrema. Let $\Omega \subset \mathbb{R}^n$ be open, $f \in C^2(\Omega)$, and $\xi \in \Omega$. Let $\nabla f(\xi) = 0$ (i.e., $Df|_\xi = 0$).

$$\text{Hess} = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}.$$

1. If Hess $f|_\xi$ is positive definite, f has a strict local minimum at ξ.
2. If Hess $f|_\xi$ is negative definite, f has a strict local maximum at ξ.
3. If Hess $f|_\xi$ is indefinite, f has no extremum at ξ.

Note: A critical point is not necessarily an extremum. $(0, 0)$ is a critical point for $y = x^3$. However, it is merely a *saddle point*.

The Normal Form of Question

In general, constrained optimization problems can be classified into equality constraints and inequality constraints. Here, we will focus on equality constraints. The general form of such problems is to find the extremum of $f(x)$ subject to the constraints $g_i(x) = 0$, where $i = 1, 2, 3, \ldots, n$. We can write the problem as:

Eg. $f(x,y) = x^2 + y$

$$\begin{gathered} \min(\max) f(x) \\ s.t.\ g_i(x) = 0, i = 1, 2, 3, \ldots, n \end{gathered} \tag{3}$$

min $f(x)$ for $x+y=1$

Lagrange Multipliers

Let $\Omega \subset \mathbb{R}^n$ be open and $f \in C^1(\Omega, \mathbb{R})$, $g \in C^1(\Omega, \mathbb{R}^m)$, $m < n$. Assume that f has a local extremum on the set $E = \{x \in \mathbb{R}^n : g(x) = 0\}$ at $\xi \in E$. Assume further that in

$$Dg|_\xi = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}, \tag{4}$$

there exists a submatrix consisting of m columns whose determinant does not vanish. Then there exist m numbers (called Lagrange multipliers)

$$\lambda_1, \dots, \lambda_m \in \mathbb{R}$$

$$Df|_\xi + \sum_{i=1}^{m} \lambda_i Dg_i|_\xi = 0. \tag{5}$$

In order to find the constrained extrema, we need to solve the $m + n$ equations

These equations are equivalent to the following:

$$\begin{aligned} &\frac{\partial f}{\partial x_i} + \lambda_1 \frac{\partial g_1}{\partial x_1} + \cdots + \lambda_m \frac{\partial g_m}{\partial x_i} = 0, \qquad i = 1, \dots, n, \\ &g_j(x) = 0, \qquad j = 1, \dots, m. \end{aligned} \tag{6}$$

$$F(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) := f(x) + \lambda_1 g_1(x) + \cdots + \lambda_m g_m(x). \tag{7}$$

Then at an extremal point, all partial derivatives of F will vanish.

Exercise

1. Calculate the maximum and minimum of function $u = x - 2y + 2z$ s.t. $x^2 + y^2 + z^2 = 1$.

$u = x - 2y + 2z.$

$g = x^2 + y^2 + z^2 - 1.$

$u + \lambda g = x - 2y + 2z + \lambda(x^2 + y^2 + z^2 - 1)$

$$\begin{cases} \frac{\partial F}{\partial x} = 1 + 2\lambda x \\ \frac{\partial F}{\partial y} = -2 + 2\lambda y. \\ \frac{\partial F}{\partial z} = 2 + 2\lambda z \\ \frac{\partial F}{\partial \lambda} = x^2 + y^2 + z^2 - 1 \end{cases}$$

$$\begin{cases} x = -\frac{1}{2\lambda} \\ y = \frac{1}{\lambda} \\ z = -\frac{1}{\lambda} \end{cases}$$

$\Rightarrow \frac{1}{4\lambda^2} + \frac{1}{\lambda^2} + \frac{1}{\lambda^2} = 1.$

$x^2 + y^2 + z^2 = 1 \Rightarrow \frac{1}{\lambda^2} \cdot \frac{3}{4} = 1.$

$\lambda^2 = \frac{3}{4}$

$\lambda = \pm \frac{\sqrt{3}}{2}.$

(x, y, z)

$= \left(-\frac{1}{\sqrt{3}}, \frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}\right)$

or $\left(\frac{1}{\sqrt{3}}, -\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}\right)$

2. If $x+y+z=3$, $x, y, z>0$, find the maximum of $x+xy+xyz$.

$y+z=3-x.$

$$
\begin{aligned}
& x+xy(1+z) \\
\le\ & x+x\left(\frac{1+y+z}{2}\right)^2 \\
=\ & x+x\left(\frac{4-x}{2}\right)^2 \\
=\ & \frac{1}{4}x^3-2x^2+5x
\end{aligned}
$$

$F'=0 \Rightarrow x_1=2.\quad x_2=\frac{10}{3}\ (\times).\quad >3.$

$x_1=2\ (\checkmark)\quad y+z=1.$

$\max = 4.$

Eigenvalue Problem

Eigenvalue Calculation

Definition: for a linear map $L \in L(V, V)$, if there exists a non-zero vector $v \in V$ and a number λ so that $Lv=\lambda v$, we can call λ to be an eigenvalue of L, and v an eigenvector of L. For a given λ, the set $V\lambda := v \in V | Lv = \lambda v$ is called the eigen-space for the eigenvalue λ. $dimV$ is called the geometric multiplicity of λ. **Independence of Eigenvector**: The eigenvector in different eigen-space are linearly independent.

Now we only discuss L as a matrix. suppose$L$ is an $n \times n$ matrix.

L — λ_1, λ_2, ..., λ_n; v_1, v_2, v_3, v_m, v_{m+1}

$Lv=\lambda v.$

calculation: since $Lv=\lambda v$, we have

$$\boxed{(L-\lambda I)v=0.} \tag{8}$$

In order for $v \neq 0$, $det(L-\lambda I)=0$. Here, $p(\lambda) := det(L-\lambda I)$ is called the **characteristic polynomial**. The degree is n. By the fundamental theorem of algebra, there are n eigenvalue of L if complex number and multiple root are taken into consideration. For each eigenvalue, its multiplicity of the root is called **algebraic multiplicity**.

$(L-\lambda I)v_1=0 \qquad v_1 \neq v_2.$

$v_2=0.$

Matrix Diagonalization

Theorem about algebraic multiplicity and geometric multiplicity: In matrix A, for a given eigenvalue, the algebraic multiplicity is greater or equal to geometric multiplicity. If for all the eigenvalues, the algebraic multiplicity is equal to geometric multiplicity, then all the eigenvectors will form a basis of R^n . In this kind of situation, we say the matrix A is diagonalizable.

$\boxed{(L-\lambda I)v=0.}$

① $D=\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & - & - & \lambda_n \end{pmatrix}.$

② $U=\begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}$

If A is diagonalizable, set $U = (v_1, v_2, \ldots, v_n)$to be the corresponding eigenvector of the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ (for multiple roots, the corresponding eigenvectors have more geometric multiplicity). We have $\boxed{D = U^{-1}AU}$ And $D = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$. We say D is the diagonal form of A.

Notice that $\boxed{A = UDU^{-1}}$, so $A^k = UD^kU^{-1}$. $D^k = diag(\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k)$. So as long as A is diagonalizable, we can calculate the close form of A k easily.

③ $A = UDU^{-1}$

Exercise

1. Diagonalize the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}$$

① $\det(A - \lambda I) = \lambda^2(14-\lambda) = 0.$

$\lambda_1 = \lambda_2 = 0. \quad \lambda_3 = 14$

② Discuss all eigenvalues.

1° $\lambda = 0. \quad AV = 0 \quad V = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$

$\Rightarrow x_1 + 2x_2 + 3x_3 = 0. \quad t_1 \underbrace{\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}}_{v_1} + t_2 \underbrace{\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}}_{v_2}$

2° $\lambda = 14.$

$\begin{pmatrix} -13 & 2 & 3 \\ 2 & -10 & 6 \\ 3 & 6 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$

$\Rightarrow t \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = 0.$

$U = \begin{pmatrix} -2 & -3 & \frac{1}{3} \\ 1 & 0 & \frac{2}{3} \\ 0 & 1 & 1 \end{pmatrix}$

③ Diagnolize,

$A = U \Lambda U^{-1}$

Spectral Theorem

In quantum mechanics, also called Hermitian matrix.

The adjoint of a linear map L is defined as L^* that satisfy

$$< x, Ly > = < L^* x, y > \tag{9}$$

If L is a real matrix, L^* is just L^T . If L is a complex matrix, L^* is $\bar{L}^T$. If a matrix A is Self-Adjoint, i.e.,$A = A^*$, then there are some good property regarding its eigenvalue and eigen-space, which is called spectral theorem.

$\langle x, Ly \rangle = \langle L^T, x \rangle$

If an $n \times n$ matrix A is Self-Adjoint, i.e., $A = A^*$, then we have:

- Eigenvalues of A are real.
- A is diagonalizable.
- The eigenvalues of A can form an orthonormal base of R^n .

Exercise

1. Q is a 12×12 orthogonal matrix $(Q^T Q = I)$. $x = (1, 1, \ldots, 1)^T$, $y = (1, -1, 1, -1, \ldots, -1)^T$. $x, y \in R^{12}$. Calculate $||Qx||$ and $< Qx, Qy >$

$||Qx||^2 = \langle Qx, Qx \rangle = \langle Q^T Q x, x \rangle$. $\quad Q^T = Q^{-1}$

$= \langle x, x \rangle = 12$. $\quad ||Qx|| = 2\sqrt{3}$.

$\langle Qx, Qy \rangle \ \langle Q^T Q x, y \rangle = \langle x, y \rangle = 0$.

References

- VV286, slide. Horst Hohberger
- RC, slides. TA-Wu Shuyu

For further questions, you can contact me through WeChat.

Fragments

芬兰

扫一扫上面的二维码图案，加我为朋友。


POMPO
THE CINÉPHILE
PETERSEN