Lecture 2: Review of Prerequisites

Disclaimer. As we have a textbook, this lecture note is for guidance and supplement only. It should not be relied on when preparing for exams.

In this lecture we review the prerequisites: Multivariable calculus, linear algebra, and basics of differential equations. We will not be able to cover everything that will be needed through the semester. We will only cover the basics, and leave some more sophisticated topics such as inverse/implicit function theorem for later lectures to explain.

Warning. In this lecture note we omit many technical assumptions such as differentiability/integrability of functions when stating formulas and theorems, because in 348 they are almost always satisfied. However please do not quote the statements here for your other courses.

Table of contents

1. Multivariable calculus

Multivariable calculus studies functions from $\mathbb{R}^m$ to $\mathbb{R}^n$ with either $m>1$ or $n>1$. In 348 we focus on functions from $\mathbb{R}$ to $\mathbb{R}^2/\mathbb{R}^3$ (curves) and from $\mathbb{R}^2$ to $\mathbb{R}^3$ (surfaces). In the following we review concepts, theorems, and techniques that are important to us.

1.1. The Euclidean space $\mathbb{R}^n$

- $\mathbb{R}^n$ is the mathematical representation of a n-dimensional flat space through setting up n orthogonal "axes". Along each axes, we define a unit vector: $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Then each point in the space is identified by an n-tuple $(x_1, \ldots, x_n)$.

  It is important to realize that $\mathbb{R}^n$ can be interpreted in two ways, in the context of describing a moving particle (along a curve, on a surface, etc.).

  - $\mathbb{R}^n$ as possible locations. At each moment, the location of the particle is identified with a point in $\mathbb{R}^n$: At time t, the particle is at $(x_1, \ldots, x_n)$.
  - $\mathbb{R}^n$ as possible velocities. To completely describe the motion of this particle, we also need to know its velocity. The velocity is also described by an n-tuple $(v_1, \ldots, v_n)$. Such n-tuples are called "vectors".

  Thus at each point x in the "location" space $\mathbb{R}^n$, we "overlay" a "velocity" space whose mathematical representation is also $\mathbb{R}^n$. This overlay will later be called the "tangent space" of the location space at x. Such distinction seems a bit silly now, but will be very useful later when we "curve" the location space. The key idea of differential geometry is that when the location space is curved, the velocity space stays flat. To visualize, think of a 2-dimensional surface as the curved locations space, then the velocity space at a location is the tangent plane of the surface there.

1.2. Operations on $\mathbb{R}^n$

- When $\mathbb{R}^n$ is interpreted as the overlaying "velocity" space, the most important concepts are

  - Norm. When $v = (v_1, \ldots, v_n)$ describes a velocity vector, its speed is given by its norm:
  $$\|v\| := \sqrt{v_1^2 + \cdots + v_n^2}. \tag{1}$$
  - Inner product. The inner product of two vectors $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ is defined as
  $$u \cdot v := u_1 v_1 + \cdots + u_n v_n. \tag{2}$$
  The importance of inner product comes from the following property:

  Proposition 1. *Let θ be the angle between two vectors $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$, then*
  $$\cos\theta = \frac{u \cdot v}{\|u\| \, \|v\|}. \tag{3}$$

Exercise 1. Obviously θ is not uniquely defined. What does this not matter?

In particular, $u \cdot v = 0 \Longleftrightarrow u \perp v$, that is u is perpendicular(orthogonal) to v.

- CROSS PRODUCT/VECTOR PRODUCT. Let $u = (u_1, u_2, u_3), v = (v_1, v_2, v_3)$ be two vectors in $\mathbb{R}^3$. Their cross product (also called "vector product") is defined as

$$u \times v := (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1). \tag{4}$$

Exercise 2. Prove the following.

i. $u \times v = -v \times u$;

ii. Let $a, b \in \mathbb{R}$ be arbitrary. Then

$$(a\, u + b\, w) \times v = a\, u \times v + b\, w \times v. \tag{5}$$

iii. $u \times v = 0$ if and only if u and v are "parallel".

iv. $(u \times v) \cdot u = 0, (u \times v) \cdot v = 0.$

v. $(u \times v) \cdot w = (v \times w) \cdot u = (w \times u) \cdot v.$

Also note that

$$\|u \times v\| = \|u\| \, \|v\| \sin \theta \tag{6}$$

where θ is the angle between u, v.

Remark 2. It turns out that reasonable "cross-product"s can only be defined in $\mathbb{R}^3$ and $\mathbb{R}^7$. See W. S. Massey, *Cross products of vectors in higher dimensional Euclidean spaces*, The American Mathematical Monthly, 90(10): 697–701, 1983.[1]

- When $\mathbb{R}^n$ is interpreted as the underlying "location" space, the most important concepts are

  - DISTANCE. Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. For now, the only distance we know is the distance along straight lines, given by the Pythagorean theorem:

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}. \tag{7}$$

Remark 3. We see that $d(x, y) = \|x - y\|$. However this will cease to be true when the location space is curved. One important theme in 348 is to generalize (7) for curved spaces.

  - AREA. The foundation of the definition of area is the following agreement:

    i. The unit square has area 1.

    ii. The area of a union of disjoint regions is the sum of the areas of these regions.

1. The idea of the proof is that if a cross product can be defined on $\mathbb{R}^n$ such that i) $v \times w$ is bilinear in v and w, ii) $v \times w$ is perpendicular to both v, w, iii) $|v \times w|^2 + (v \cdot w)^2 = \|v\|^2 \, \|w\|^2$, then $\mathbb{R}^{n+1}$ can be made into a normed division ring with identity. By Hurwitz's Theorem $n + 1 = 1, 2, 4, 8$ and the conclusion follows.

iii. Moving a region around "rigidly" (that is not changing the distance of any pair of two points in the region) does not change its area.

From these one could develop the whole theory of integration in two-variable calculus.

Example 4. Let $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$ in $\mathbb{R}^3$. Then the area of the triangle $(0,0,0) - (x_1, x_2, x_3) - (y_1, y_2, y_3) - (0,0,0)$ is given by $\frac{1}{2} \|x \times y\|$.

- VOLUME. The foundation of the definition of volume is the following agreement:

  i. The unit cube has volume 1.

  ii. The volume of a union of disjoint regions is the sum of the volumes of these regions.

  iii. Moving a region around "rigidly" (that is not changing the distance of any pair of two points in the region) does not change its volume.

  With these agreement one could develop the whole theory of integration in three-variable calculus.

  Example 5. Let $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$, $z = (z_1, z_2, z_3)$ in $\mathbb{R}^3$. Then the volume of the parallelopiped spanned by these three vectors is given by

$$V = |(x \times y) \cdot z|. \tag{8}$$

  Exercise 3. Check your old textbook and explain what the sign of $(x \times y) \cdot z$ means.

1.3. Functions from $\mathbb{R}^m$ to $\mathbb{R}^n$

Let $\Omega \subseteq \mathbb{R}^m$ be a region in $\mathbb{R}^m$. A function $f: \Omega \mapsto \mathbb{R}^n$ is a correspondence between points in Ω and points in $\mathbb{R}^n$ satisfying the following: Each $x \in \Omega$ corresponds to exactly one $y \in \mathbb{R}^n$. One classifies functions into two categories: linear and nonlinear. In 348, we will see that linear functions are often considered in the context of "velocity" spaces.

The most important functions are linear functions (linear transformations) and bilinear forms.

- LINEAR FUNCTIONS. A function $f: \Omega \mapsto \mathbb{R}^n$ is linear if and only if for every $a, b \in \mathbb{R}$ and every $x, y \in \Omega$, one has

$$f(a\,x + b\,y) = a\,f(x) + b\,f(y). \tag{9}$$

  Exercise 4. Prove that if f is linear then $f(0) = 0$.

  Exercise 5. Let $f: \mathbb{R} \mapsto \mathbb{R}$ be linear. Prove that there is a constant $c \in \mathbb{R}$ such that $f(x) = c\,x$.

  NOTATION. *We often denote a linear function by a capital letter such as T.*[2]

2. First letter of "transformation".

PROPOSITION 6. *Let $T: \Omega \subseteq \mathbb{R}^m \mapsto \mathbb{R}^n$ be linear. Then there are $m\,n$ numbers a_{11}, $a_{12}, \ldots, a_{1m}, a_{21}, \ldots, a_{n1}, \ldots, a_{nm} \in \mathbb{R}$, such that for every $v \in \Omega$, the following holds for $w = T(v)$:*

$$w_1 = a_{11} v_1 + \cdots + a_{1m} v_m \tag{10}$$

$$\vdots \quad \vdots \quad \vdots$$

$$w_n = a_{n1} v_1 + \cdots + a_{nm} v_m. \tag{11}$$

In other words, we have $T(v) = A\,v$ where A is the matrix $\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$.

Exercise 6. Prove Proposition 6.

- BILINEAR FORMS. A bilinear form is a special type of nonlinear function: $B: \mathbb{R}^{2n} \mapsto \mathbb{R}$, satisfying for every $a, b \in \mathbb{R}$ and every $u, v, w \in \mathbb{R}^n$,

$$B(a\,u + b\,w, v) = a\,B(u, v) + b\,B(w, v), \tag{12}$$

$$B(u, a\,w + b\,v) = a\,B(u, w) + b\,B(u, v). \tag{13}$$

Exercise 7. Prove that the inner product: $f(u, v) := u \cdot v$ is a bilinear form.

Exercise 8. Prove or disprove: The area $A(u, v) := \|u \times v\|$ is a bilinear form.

PROPOSITION 7. *Let B: $\mathbb{R}^{2n} \mapsto \mathbb{R}$ be a bilinear form. Then there are n^2 numbers $a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{n1}, \ldots, a_{nn}$ such that*

$$B(u, v) = \sum_{i=1}^n \sum_{j=1}^n u_i\, a_{ij}\, v_j. \tag{14}$$

Exercise 9. Let $A := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$. Prove that $B(u, v) = u \cdot A\,v = (A\,u) \cdot v$ where $\cdot$ denotes inner product.

1.4. Taking derivatives of functions from $\mathbb{R}^m$ to $\mathbb{R}^n$

- DIFFERENTIABILITY.

DEFINITION 8. *A function $f: \Omega \subseteq \mathbb{R}^m \mapsto \mathbb{R}^n$ is differentiable at $x \in \Omega$ if and only if the following holds: There is a linear function $T: \mathbb{R}^m \mapsto \mathbb{R}^n$ such that*

$$\lim_{v \longrightarrow 0} \frac{\|f(x+v) - f(x) - T(v)\|}{\|v\|} = 0. \tag{15}$$

The linear function T is called the "differential" of f at x and will be denoted $Df(x)$.

Remark 9. Note that f is defined on Ω which belongs to the "location" space $\mathbb{R}^m$, and f takes value in another "location" space $\mathbb{R}^n$; while T is defined on $\mathbb{R}^m$ which is the "velocity" overlay space on top of Ω, centered at x, and takes value in another "velocity space" centered at $f(x)$.

By Proposition 6 there is a matrix A such that $T(v) = A\, v$ for all $v \in \mathbb{R}^m$. The numbers a_{ij} in the matrix A is given by the partial derivatives of f at x:

DEFINITION 10. *Let $f\colon \Omega \subseteq \mathbb{R}^m \mapsto \mathbb{R}$. Its j-th partial derivative at $x = (x_1, \ldots, x_m) \in \Omega$ is defined as*

$$\frac{\partial f}{\partial x_j}(x) := \frac{\mathrm{d} f(x_1, \ldots, x_{j-1}, t, x_{j+1}, \ldots, x_m)}{\mathrm{d}t}\Big|_{t = x_j}. \tag{16}$$

Exercise 10. Prove that $a_{ij} = \frac{\partial f_i}{\partial x_j}(x)$.

Note. This matrix (representation of the differential of f) is called the **Jacobian matrix** of f at x.

Exercise 11. Let $f\colon \mathbb{R}^m \mapsto \mathbb{R}^n$ be differentiable everywhere. Prove that if $Df(x) = 0$ for all $x \in \mathbb{R}^m$, then f is a constant.

- CHAIN RULE. Let $f(x)\colon \mathbb{R}^m \mapsto \mathbb{R}^n$ and $g(y)\colon \mathbb{R}^n \mapsto \mathbb{R}^k$. Then the composite function $g \circ f\colon \mathbb{R}^m \mapsto \mathbb{R}^k$ is defined as

$$(g \circ f)(x) = g(f(x)). \tag{17}$$

The partial derivatives of $g \circ f$ can be calculated through

$$\frac{\partial (g \circ f)_i}{\partial x_j} = \sum_{k=1}^{n} \frac{\partial g_i}{\partial y_k}(f(x))\, \frac{\partial f_k(x)}{\partial x_j}. \tag{18}$$

Exercise 12. Find a calculus textbook and work on a few chain rule problems.

- HIGHER ORDER DERIVATIVES. We consider the simplest case $f\colon \Omega \subseteq \mathbb{R}^m \mapsto \mathbb{R}$. At every $x \in \Omega$, it has a differential which is a linear function $Df(x)\colon \mathbb{R}^m \mapsto \mathbb{R}$. As this linear function can be represented by a $1 \times m$ matrix, it can also be identified with an m-vector. Thus at each point $x \in \Omega$, we have a vector $g(x) \in \mathbb{R}^m$. Differentiating this function g we obtain the second order derivative of f at x, denoted $D^2 f(x)$.

Exercise 13. Prove or disprove: Let $x \in \Omega$ be arbitrary. The second order derivative of f can be seen as a bilinear form on $\mathbb{R}^{2m} = \mathbb{R}^m \times \mathbb{R}^m$, where each $\mathbb{R}^m$ is the "velocity" space at x.

Exercise 14. Prove that the matrix representing the bilinear form is given by

$$a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right). \tag{19}$$

As we can see here, for functions between $\mathbb{R}^m$ and $\mathbb{R}^n$ with m or $n > 1$, higher order derivatives get more and more complicated. However there is one exception.

Exercise 15. What happens if we take successive derivatives of a function $f\colon \mathbb{R} \mapsto \mathbb{R}^n$?

Note. The matrix representation of $D^2 f(x)$ is called the **Hessian matrix** of f at x.

Exercise 16. What are the Jacobian and Hessian matrices for $f\colon \mathbb{R} \mapsto \mathbb{R}$?

- TAYLOR EXPANSION.

Let $f : \Omega \subseteq \mathbb{R}^m \mapsto \mathbb{R}$. Then there holds

$$f(x+v) = f(x) + Df(x)(v) + \frac{1}{2} D^2 f(x)(v,v) + R \tag{20}$$

where $\lim_{v \to 0} \frac{\|R\|}{\|v\|^2} = 0$. (20) is the **Taylor expansion of f at x to second order**.

Exercise 17. Prove the following.

$$Df(x)(v) = \sum_{i=1}^m \frac{\partial f(x)}{\partial x_i} v_i, \qquad D^2 f(x)(v,v) = \sum_{i=1}^m \sum_{j=1}^m \frac{\partial^2 f(x)}{\partial x_i \partial x_j} v_i v_j. \tag{21}$$

- Examples.

Example 11. Let $f(x,y) := e^{xy}$. We calculate its Taylor expansion at $(0,0)$ to second order. We have

$$\begin{aligned}
f(0,0) &= e^0 = 1;\\
\frac{\partial f}{\partial x} = y\, e^{xy} \Longrightarrow \frac{\partial f}{\partial x}(0,0) &= 0;\\
\frac{\partial f}{\partial y} = x\, e^{xy} \Longrightarrow \frac{\partial f}{\partial y}(0,0) &= 0;\\
\frac{\partial^2 f}{\partial x^2} = y^2\, e^{xy} \Longrightarrow \frac{\partial^2 f}{\partial x^2}(0,0) &= 0;\\
\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = e^{xy} \Longrightarrow \frac{\partial^2 f}{\partial x \partial y}(0,0) &= 1;\\
\frac{\partial^2 f}{\partial y^2} = x^2\, e^{xy} \Longrightarrow \frac{\partial^2 f}{\partial y^2}(0,0) &= 0.
\end{aligned}$$

Therefore the expansion reads

$$\begin{aligned}
f(u,v) &= f(0,0) + Df(0,0)(u,v) + \frac{1}{2} D^2 f(0,0)(u,v) + R\\
&= 1 + u\,v + R.
\end{aligned} \tag{22}$$

Remark 12. (INTERPRETING f AND Df) Functions are often called "maps". For classical differential geometry this is pertinent. Consider, e.g., Google Map with GPS. This can be seen as a function f from $\mathbb{R}^2$ to the earth surface. Thus $f(x_0)$ is the "real world" location of a car that is at location x_0 on the map. More importantly, if the car symbol is moving with velocity v on the map, the velocity of the car itself in the real world is given by $[Df(x_0)](v)$ (remember that $Df(x_0)$ is a function!). It is woth mentioning that, although the car is located in a curved surface (earth surface), its possible velocities at a particular location (time) form a plane, the "tangent plane" of earth surface at location $f(x_0)$.

2. Linear algebra

The fundamental idea of calculus is to study functions through their first and second (or higher) derivatives. We have seen that such derivatives can be viewed as linear transformations between $\mathbb{R}^m$ and $\mathbb{R}^n$. Therefore to understand them we need linear algebra.

2.1. Operations on matrices

- DETERMINANT.
  - Determinant is defined for square matrices.
  - Determinant is uniquely defined. Let $f(v_1, \ldots, v_n)\colon \mathbb{R}^n \times \cdots \times \mathbb{R}^n \mapsto \mathbb{R}$ be such that

    i. f is linear in each of the v_i's. That is, e.g., f is a linear function of v_1 when $v_2, \ldots, v_n$ are fixed.

    ii. $f(e_1, \ldots, e_n) = 1$. Where e_i is the vector with one 1 at the ith entry and 0's elsewhere.

    iii. f changes sign when two columns are switched. e.g. $f(v_1, v_2, v_3, \ldots, v_n) = -f(v_2, v_1, v_3, \ldots, v_n)$.

    Then $f(v_1, \ldots, v_n) = \det V$ where V is the matrix with v_1 as its first column, v_2 as its second column, and so on.

    Exercise 18. Prove the above claim for $n = 2, 3$.

  - 2×2. Let $A := \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Then

    $$\det A := a_{11} a_{22} - a_{12} a_{21}. \tag{23}$$

    Exercise 19. Let $u, v, x, y \in \mathbb{R}^3$. Prove that

    $$(u \times v) \cdot (x \times y) = \det \begin{pmatrix} u \cdot x & v \cdot x \\ u \cdot y & v \cdot y \end{pmatrix}. \tag{24}$$

  - 3×3. Let $A := \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Then

    $$\det A := a_{11}\, a_{22}\, a_{33} + a_{12}\, a_{23}\, a_{31} + a_{21}\, a_{32}\, a_{13} - a_{13}\, a_{22}\, a_{31} - a_{12}\, a_{21}\, a_{33} - a_{23} a_{32} a_{11}. \tag{25}$$

    Exercise 20. Let $u, v, w \in \mathbb{R}^3$. Let $A := \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}$. Prove that $\det A = (u \times v) \cdot w$.

  - A square matrix A is invertible, that is there is a matrix B such that $A\, B = B\, A = I = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$ if and only if $\det A \neq 0$.

- EIGENVALUES AND EIGENVECTORS.
  - In most cases the matrix A under study is the mathematical representation of a linear function T.
  - We could try to understand an $n \times n$ matrix through its n^2 entries. However a more efficient way is through its eigenvalues/eigenvectors.
  - The eigenvalues of an $n \times n$ matrix A are the solutions to the equation

    $$\det(\lambda\, I - A) = 0. \tag{26}$$

The eigenvectors corresponding to an eigenvalue λ_0 are those vectors satisfying

$$(\lambda I - A)\, v = 0 \Longleftrightarrow A\, v = \lambda\, v. \tag{27}$$

- ○ In many situations, one could choose n linearly independent eigenvectors $\{v_1, \ldots, v_n\}$ to form a "coordinate system", that is every vector $u \in \mathbb{R}^n$ can be uniquely represented as $u = a_1\, v_1 + \cdots + a_n\, v_n$ where $a_1, \ldots, a_n \in \mathbb{R}$. Then the linear function T can be very easily understood:

$$T(u) = \lambda_1\, a_1\, v_1 + \cdots + \lambda_n\, a_n\, v_n. \tag{28}$$

- ○ The most useful situation for us is when A is symmetric: $a_{ij} = a_{ji}$ for all i, $j = 1, 2, \ldots, n$. In this case we can choose n eigenvectors $v_1, \ldots, v_n$ satisfying

  i. $v_i \perp v_j$, for all $i \neq j$;

  ii. $\|v_i\| = 1$ for all $i = 1, 2, \ldots, n$.

2.2. Linear dependence and independence

- Linear dependence/independence. Let $v_1, \ldots, v_k \in \mathbb{R}^n$. We say they are linearly dependent if there are $a_1, \ldots, a_k \in \mathbb{R}$, not all equal to 0, such that

$$a_1\, v_1 + \cdots + a_k\, v_k = 0. \tag{29}$$

  If no such $a_1, \ldots, a_k$ exist, we say the vectors are linearly independent.

  Exercise 21. Prove that, if $v_1, \ldots, v_k$ are linearly independent, then the following holds. If $a_1, \ldots, a_k \in \mathbb{R}$ satisfy $a_1\, v_1 + \cdots + a_k\, v_k = 0$, then $a_1 = a_2 = \cdots = a_k = 0$.

  Exercise 22. Prove that if $v_1, \ldots, v_k$ are linearly dependent, then at least one of them is "redundant" in the sense that it equals a linear combination of the other vectors.

- n vectors $v_1 = (v_{11}, \ldots, v_{1n}), \ldots, v_n = (v_{n1}, \ldots, v_{nn}) \in \mathbb{R}^n$ are linearly independent if and only if $\det A \neq 0$ where

$$A = \begin{pmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{n1} & \cdots & v_{nn} \end{pmatrix}. \tag{30}$$

- n vectors $v_1, \ldots, v_n \in \mathbb{R}^n$ are linearly independent if and only if they form a "base" of $\mathbb{R}^n$, that is every $u \in \mathbb{R}^n$ can be written uniquely as a linear combination of $v_1, \ldots, v_n$.

3. Differential equations

A differential equation is an equation involving the derivative (and/or higher order derivatives) of functions.

3.1. The simplest and second simplest differential equations

In the following $x(t)\colon \mathbb{R} \mapsto \mathbb{R}$ is a single variable function.

- The simplets ODE.

Consider

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(t). \tag{31}$$

Its solution, by definition, is the indefinite integral

$$x(t) = \int f(s)\,\mathrm{d}s \tag{32}$$

which by the fundamental theorem of calculus further equals

$$\int_0^t f(s)\,\mathrm{d}s + C \tag{33}$$

where C is an arbitrary constant.

If we further specify $x(t_0) = x_0$, then this constant is fixed.

Exercise 23. What is the solution to

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(t), \qquad x(t_0) = x_0? \tag{34}$$

Exercise 24. Solve

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x), \qquad x(t_0) = x_0. \tag{35}$$

- The next simplest ODE.

Consider

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} + c\,x(t) = f(t), \qquad x(0) = x_0. \tag{36}$$

where $c \in \mathbb{R}$. Multiplying both sides by e^{ct} and apply Leibniz rule of differentiation of product of functions, we reach

$$\frac{\mathrm{d}}{\mathrm{d}t}(e^{ct}\,x(t)) = e^{ct}\,f(t), \tag{37}$$

which is the simplest ODE now. The solution of (37) is

$$e^{ct}\,x(t) = x_0 + \int_0^t e^{cs}\,f(s)\,\mathrm{d}s \tag{38}$$

which now gives

$$x(t) = e^{-ct}\,x_0 + \int_0^t e^{c(s-t)}\,f(s)\,\mathrm{d}s. \tag{39}$$

Exercise 25. Solve

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} + c\,x(t) = f(t), \qquad x(t_0) = x_0. \tag{40}$$

In general, the equation

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} + c(t)\,x(t) = f(t) \tag{41}$$

could be solved by multiplying both sides by $e^{C(t)}$ where $C(t)$ satisfies $C'(t) = c(t)$.

Exercise 26. Solve

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} + c(t)\,x(t) = f(t), \qquad x(0) = x_0. \tag{42}$$

3.2. Existence and uniqueness of a general first order ODE

- General situation we face in 348.

  ◦ In 348 we often need to study systems of differential equations:

$$\begin{aligned} \frac{\mathrm{d}x_1(t)}{\mathrm{d}t} &= f_1(x_1(t), \ldots, x_n(t)), \qquad x_1(0) = x_{01} \qquad (43)\\ \vdots &\quad \vdots \quad \vdots \\ \frac{\mathrm{d}x_n(t)}{\mathrm{d}t} &= f_n(x_1(t), \ldots, x_n(t)). \qquad x_n(0) = x_{0n} \qquad (44) \end{aligned}$$

  ◦ Taking advantage of multivariable calculus, we can re-write the above to a more compact form:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(x(t)), \qquad x(0) = x_0 \tag{45}$$

  where $x\colon \mathbb{R} \mapsto \mathbb{R}^n$, $f\colon \mathbb{R}^n \mapsto \mathbb{R}^n$, and $x_0 = (x_{01}, \ldots, x_{0n})$.

- The one ODE system that we could really solve.

  ◦ For general f there is little hope explicitly solving (45). However there is a special, very useful, situation that we could solve.

  ◦ Consider the case where

$$f_i(x_1, \ldots, x_n) = a_{i1}\, x_1 + \cdots + a_{in}\, x_n, \qquad i = 1, 2, \ldots, n \tag{46}$$

  where $a_{ij} \in \mathbb{R}$ for all $i, j = 1, 2, \ldots, n$. In other words, we have

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = A\, x(t), \qquad x(0) = x_0. \tag{47}$$

  ◦ Now to further simplify the situation, we assume A has n linearly independent eigenvectors $v_1, \ldots, v_n$. Writing $x(t) = z_1(t)\ v_1 + \cdots + z_n(t)\ v_n$ and $x_0 = z_{01}\, v_1 + \cdots + z_{0n}\, v_n$ we obtain

$$\frac{\mathrm{d}z_i(t)}{\mathrm{d}t} = \lambda_i\, z_i(t), \qquad z_i(0) = z_{i0}. \tag{48}$$

  Exercise 27. Justify (48).

  Exercise 28. Finish the solution.

- Existence and uniqueness theorem.

  Although there is little hope explicitly solving (45), we have the following qualitative understanding.

Theorem 13. *Let $f\colon \mathbb{R}^n \mapsto \mathbb{R}^n$ be differentiable with continuous derivatives. Then there is $T > 0$ and a unique function $x(t)$ satisfying $x(0) = x_0$, and for all $0 < t < T$,*

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(x). \tag{49}$$

Exercise 29. Let $A(t)=\begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & \ddots & \vdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{pmatrix}$ be such that $a_{ij}(t)$ are continuous functions of t for every i, $j=1,2,\ldots,n$. Prove that the solution to

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=A(t)\,x(t) \tag{50}$$

exists and is unique.

Remark 14. We will see later that existence of curves with certain desirable properties would be equivalent to the existence of solution to an ODE of the form (45). Then Theorem 13 would guarantee us that this curve exists.