

中文课程名称

课程讲义模板

将封面图命名为 `cover.jpg` 并放入 `figures/` 目录后，会自动显示在这里。

教师： 授课教师
学期： 2026 春季
单位： 学校 / 机构名称
日期： 2026 年 7 月 6 日

目录

第一讲 课程导论	1
1.1 课程说明	1
1.2 表格示例	1
1.3 图片示例	1
1.4 例题与练习	2
第二讲 第二讲标题	3
2.1 知识点一	3
2.1.1 更小的层级	3
2.2 长表格示例	3
第三讲 大模型预训练与微调方法	4
3.1 大模型预训练	4
第四讲 深度学习中的连续动力学机制	5
4.1 数学基础	5
4.1.1 常微分方程的基本概念	5
4.1.2 常微分方程的求解：积分表示与数值方法	5
4.1.3 深度学习模型的常微分方程解释	9
4.2 神经常微分方程与流模型	9
4.2.1 神经常微分方程	9
4.2.2 伴随灵敏度法	9
4.2.3 基于流的生成模型	9
4.2.4 流匹配	9
4.3 视觉中的扩散模型	9
4.3.1 问题设定	9
4.3.2 随机视角：随机微分方程	9
4.3.3 确定性视角：概率流常微分方程	9
4.3.4 迈向更高效的生成	9
4.4 语言中的扩散模型	9
4.4.1 词元序列生成	9

4.4.2	连续空间中的扩散：嵌入空间扩散	9
4.4.3	词元空间中的扩散：离散扩散	9
4.4.4	关于离散扩散的进一步讨论	9
附录 A 附录		10
A.1	常用 LaTeX 片段	10
参考资料		11

第一讲 课程导论

学习目标

- 了解本课程的学习目标、评价方式与课堂要求。
- 熟悉讲义中的例题、练习、阅读材料与作业模块。
- 掌握如何在模板中插入章节、表格、图片和重点提示。

1.1 课程说明

这里填写本讲的正文内容。中文可直接输入，不需要额外转码。若要突出关键词，可使用 **关键词强调**，也可以使用普通的 **加粗**、*斜体* 等格式。

重点提示

这里可以写课堂重点、易错点、考试提示，或者需要学生特别记住的概念。

1.2 表格示例

表格建议使用 booktabs 提供的 `\toprule`、`\midrule`、`\bottomrule`，视觉会比普通竖线表格更清爽。

表 1.1: 课程安排示例

周次	主题	任务
第 1 周	课程导论与学习方法	阅读讲义
第 2 周	核心概念一	完成练习
第 3 周	核心概念二	小组讨论

1.3 图片示例

将图片放入 `figures/` 目录，然后用 `\includegraphics` 插入。下面的示例会在 `figures/example.png` 存在时显示图片，否则显示一个占位框。



图片占位：将图片保存为 `figures/example.png`

图 1.1: 图片插入示例

1.4 例题与练习

例题 1

这里放置例题内容。可以包含公式：

$$a^2 + b^2 = c^2$$

也可以写解题步骤或分析。

课堂练习

1. 请用自己的话概括本讲的三个重点。
2. 根据表 1.1，说明第 2 周需要完成什么任务。
3. 观察图 1.1，写出你能得到的信息。

第二讲 第二讲标题

2.1 知识点一

从这里开始写第二讲内容。新增章节时，复制下面结构即可：

```
\chapter{新的一讲}
\section{小节标题}
正文内容……
```

2.1.1 更小的层级

如果一讲内部内容较多，可以继续使用 `\subsection` 或 `\subsubsection` 组织层次。

2.2 长表格示例

如果表格跨页，可以使用 `longtable`。

表 2.1: 长表格结构示例

模块	内容	备注
课前	预习阅读材料	可附链接或二维码图片
课中	讲授、讨论、练习	建议加入例题与即时反馈
课后	作业、复盘、扩展阅读	可列出提交要求

第三讲 大模型预训练与微调方法

3.1 大模型预训练

第四讲 深度学习中的连续动力学机制

4.1 数学基础

常微分方程 (ODE) 这一概念源自更广义的微分方程概念。该学科研究变量的变化规律，有着悠久且丰硕的发展历程 [爱德华兹, 1994; 邓纳姆, 2005]。如今，微分方程已成为现代科学与工程领域的核心工具语言，应用场景十分广泛。本节将介绍微分方程的基础概念，重点围绕常微分方程展开讲解，并举例说明这类数学工具在深度学习中的若干应用。

4.1.1 常微分方程的基本概念

我们首先考虑一个简单的汽车运动问题，并在本节中始终以此为例来引出若干关键概念。假设我们观察到一辆汽车沿直线行驶。令 $t \in \mathbb{R}$ 表示时间，并将其作为自变量。记汽车在任意时刻 t 的位置为 $x(t)$ 。其中， $x(t)$ 是因变量，其取值依赖于自变量 t 。在某些情况下，为了简化记号，我们会省略函数的自变量，即用 x 代替 $x(t)$ 。

我们知道，汽车的速度是其位置关于时间的变化率。在微积分中，这一变化率记为导数 $\frac{dx(t)}{dt}$ 。其中， $dx(t)$ 和 dt 称为微分，分别表示因变量和自变量的无穷小变化。导数 $\frac{dx(t)}{dt}$ 描述了位置 $x(t)$ 关于时间 t 的瞬时变化率。因此，这一概念定义在某一时刻，而不是某个有限的时间区间上。从数学上看， $\frac{dx(t)}{dt}$ 可以写成差商的极限

$$\frac{dx(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{x(t + \Delta t) - x(t)}{\Delta t}. \quad (4.1)$$

如果汽车的速度由某种依赖于其当前位置和时间的规律所决定，例如道路摩擦力或风阻随位置和时间发生变化，那么我们可以使用函数 $f(\cdot)$ 来表示这一关系

$$\frac{dx(t)}{dt} = f(x(t), t). \quad (4.2)$$

这个方程被正式称为常微分方程 (ordinary differential equation, ODE)。“常”意味着因变量 $x(t)$ 是单个自变量 t 的函数。与之相对的是偏微分方程 (partial differential equation, PDE)，其中未知函数依赖于多个自变量。

4.1.2 常微分方程的求解：积分表示与数值方法

我们已经说明，常微分方程 (ODE) 刻画的是状态 $x(t)$ 的导数。一个自然的问题是：为什么我们要用导数来定义系统，而不是直接描述状态 $x(t)$ 本身？事实上，如

果整个轨迹 $x(t)$ 都已知，那么 ODE 的描述是多余的。然而，在许多实际应用中，我们无法得到 $x(t)$ 的解析表达式。我们通常只给定系统的动力学函数 $f(\cdot)$ 以及初始条件 $x(0)$ 。这类似于驾驶汽车：我们可以很容易知道当前时刻的速度，但要精确确定相对于起点的位置却更困难。在 ODE 语言中，速度对应 $f(\cdot)$ ，起点对应 $x(0)$ ，而我们的目标是重建整个轨迹 $x(t)$ 的“路径”。

在这种情况下，我们需要求解常微分方程，这通常被称为初值问题 (IVP)。即给定 ODE 以及初始状态 $x(0)$ ，希望从导数信息中恢复任意时刻 t 的状态 $x(t)$ 。

根据微积分基本定理，时刻 t 的状态可以表示为初始状态与在区间 $[0, t]$ 上所有瞬时变化的累积：

$$x(t) = x(0) + \int_0^t \frac{dx(\tau)}{d\tau} d\tau$$

由于 $\frac{dx(\tau)}{d\tau} = f(x(\tau), \tau)$ ，因此可以写为：

$$x(t) = x(0) + \int_0^t f(x(\tau), \tau) d\tau \quad (8)$$

如图 2 所示，该公式具有非常直观的几何解释：积分项

$$\int_0^t f(x(\tau), \tau) d\tau$$

表示函数 $f(x(\tau), \tau)$ 在区间 $[0, t]$ 上曲线下的带符号面积。

该面积刻画了状态在时间区间 $[0, t]$ 内的累计变化量。因此，该等式说明：某一时刻的状态等于初始状态加上累计变化。

计算挑战与数值求解

然而，这种积分形式在计算上存在困难。尽管表达式看起来简单，但由于被积函数 $f(x(\tau), \tau)$ 依赖于未知的轨迹 $x(\tau)$ 本身，因此无法直接计算。

此外，即使轨迹是显式的，由于动力系统通常具有非线性（例如由深度神经网络建模），解析积分也往往不可行。

因此，我们必须使用数值方法来近似该积分，这一过程被称为“离散化 (discretization)”。

其核心思想是：不再连续地追踪系统在每一个无穷小时间点的变化，而是在一系列离散时间点上近似轨迹：

$$\{t_0, t_1, \dots, t_N\}$$

由于在有限个点上评估函数 $f(\cdot)$ 计算成本较低，因此这种方法在实践中非常高效。

Riemann 积分视角

从黎曼积分的角度来看，可以通过将积分区间划分为多个小区间，并用矩形面积之和来近似连续积分。

具体而言，在区间 $[0, t]$ 上构造时间网格：

$$0 = t_0 < t_1 < \cdots < t_N = t$$

在每个小区间上，用函数值乘以时间步长来近似积分贡献，从而得到整体近似。

4.1.3 深度学习模型的常微分方程解释

4.2 神经常微分方程与流模型

4.2.1 神经常微分方程

4.2.2 伴随灵敏度法

4.2.3 基于流的生成模型

向量场、轨迹与流映射

连续归一化流

4.2.4 流匹配

4.3 视觉中的扩散模型

4.3.1 问题设定

生成建模作为概率输运

前向过程与反向过程的对偶性

训练与推理

4.3.2 随机视角：随机微分方程

前向与反向随机微分方程

分数函数的学习

方差爆炸与方差保持随机微分方程

数值求解器

4.3.3 确定性视角：概率流常微分方程

从随机微分方程到常微分方程

流匹配

4.3.4 迈向更高效的生成

高阶与专用求解器

轨迹校直

一致性模型

附录 A 附录

A.1 常用 LaTeX 片段

% 插入普通图片

```
\begin{figure}[htbp]
  \centering
  \includegraphics[width=0.8\linewidth]{figures/your-image.png}
  \caption{图片标题}
\end{figure}
```

% 插入三线表

```
\begin{table}[htbp]
  \centering
  \caption{表格标题}
  \begin{tabular}{lll}
    \toprule
    A & B & C \\
    \midrule
    内容 & 内容 & 内容 \\
    \bottomrule
  \end{tabular}
\end{table}
```

参考资料

这里可以列出教材、论文、网站或其他阅读材料。

参考文献

- [1] 作者. 书名. 出版社, 年份.
- [2] 网站名称. 文章标题. <https://example.com>