



# 机器学习沉思录

作者：张润丰

时间：2023/09/05

版本：0.1

# 目录

|  |    |
|--|----|
| 第一章 机器学习基本概念                                       | 1  |
| 第二章 理论机器学习基础                                       | 2  |
| 第三章 机器学习中的优化方法                                     | 3  |
| 3.1 梯度下降方法 (Gradient Descent Method)               | 3  |
| 3.2 随机梯度下降方法 (Stochastic Gradient Descent Method)  | 3  |
| 3.3 Momentum                                       | 3  |
| 3.4 Nesterov                                       | 4  |
| 3.5 Adagrad  | 4  |
| 3.6 RMSprop  | 4  |
| 3.7 Adam   | 5  |
| 3.8 AdamW  | 5  |
| 第四章 机器学习中的损失函数                                     | 6  |
| 4.1 回归问题的损失函数                                      | 6  |
| 4.2 分类问题的损失函数                                      | 7  |
| 4.2.1 二分类问题的损失函数                                   | 7  |
| 4.2.2 多分类问题的损失函数                                   | 7  |
| 第五章 机器学习模型的评估指标                                    | 8  |
| 5.1 分类模型的评估指标                                      | 8  |
| 5.1.1 二分类模型  | 8  |
| 5.1.2 多分类模型  | 8  |
| 5.2 回归问题的评估指标                                      | 8  |
| 第六章 树模型  | 10 |
| 6.1 分类回归树 (CART)                                   | 10 |
| 6.1.1 分类回归树中的分类问题                                  | 10 |
| 6.1.2 分类回归树中的回归问题                                  | 10 |
| 第七章 提升方法 <b>Boosting</b>                           | 11 |
| 7.1 Forward Stagewise Additive Modeling            | 11 |
| 7.2 Gradient Boosting                              | 11 |
| 7.3 XGBoost  | 11 |
| 第八章 推荐系统   | 12 |
| 8.1 召回模型   | 12 |
| 8.1.1 协同过滤 (Collaborative Filtering)               | 12 |
| 8.1.2 Embedding 召回方法                               | 12 |
| 8.1.2.1 KD-Tree 算法                                 | 12 |
| 8.1.2.2 局部敏感哈希 (Local Sensitive Hash)              | 12 |
| 8.1.2.3 Hierarchical Navigable Small Worlds (HNSW) | 12 |
| 8.2 精排模型   | 13 |

|  |           |
|--|-----------|
| 8.2.1 因子分解机 (Factorization Machine)                  | 13        |
| 8.2.2 DeepFM   | 13        |
| 8.3 线上实验评估方法   | 13        |
| 8.3.1 A/B Test                                       | 13        |
| 8.3.2 Interleaving                                   | 13        |
| <b>第九章 Learning to Rank</b>                          | <b>14</b> |
| <b>第十章 自然语言处理</b>                                    | <b>15</b> |
| 10.1 注意力机制 (Attention Mechanism)                     | 15        |
| 10.1.1 为什么要用多头注意力机制?                                 | 15        |
| 10.2 Transformer                                     | 15        |
| 10.2.1 Batch Normalization 还是 Layer Normalization?   | 15        |
| 10.3 基于 BERT 的检索模型                                   | 15        |
| 10.3.1 Bi-Encoder                                    | 15        |
| 10.3.1.1 负样本采样                                       | 15        |
| 10.3.1.2 对比学习  | 17        |
| 10.3.2 Cross-Encoder                                 | 17        |
| <b>第十一章 Elegant<math>\text{\LaTeX}</math> 系列模板介绍</b> | <b>18</b> |
| 11.1 模板安装与更新   | 18        |
| 11.1.1 在线使用模板  | 18        |
| 11.1.2 本地免安装使用                                       | 18        |
| 11.1.3 发行版安装与更新                                      | 18        |
| 11.1.4 其他发行版本  | 19        |
| <b>第十二章 ElegantBook 设置说明</b>                         | <b>20</b> |
| 12.1 语言模式  | 20        |
| 12.2 设备选项  | 20        |
| 12.3 颜色主题  | 20        |
| 12.4 封面  | 21        |
| 12.4.1 封面个性化   | 21        |
| 12.4.2 封面图   | 21        |
| 12.4.3 徽标  | 22        |
| 12.4.4 自定义封面   | 22        |
| 12.5 章标标题  | 22        |
| 12.6 数学环境简介  | 22        |
| 12.6.1 定理类环境的使用                                      | 23        |
| 12.6.1.1 fancy 模式                                    | 23        |
| 12.6.1.2 simple 模式                                   | 24        |
| 12.6.2 修改计数器   | 25        |
| 12.6.3 自定义定理类环境                                      | 25        |
| 12.6.4 其他环境的使用                                       | 26        |
| 12.7 列表环境  | 27        |
| 12.8 参考文献  | 27        |
| 12.8.1 打印文献  | 27        |

|                              |           |
|------------------------------|-----------|
| 12.8.2 修改文献格式 . . . . .      | 27        |
| 12.9 添加序章 . . . . .          | 28        |
| 12.10 目录选项与深度 . . . . .      | 28        |
| 12.11 章节摘要 . . . . .         | 28        |
| 12.12 章后习题 . . . . .         | 29        |
| 第十二章 练习 . . . . .            | 29        |
| 12.13 旁注 . . . . .           | 29        |
| <b>第十三章 字体选项</b>             | <b>31</b> |
| 13.1 数学字体选项 . . . . .        | 31        |
| 13.2 使用 newtx 系列字体 . . . . . | 31        |
| 13.2.1 连字符 . . . . .         | 31        |
| 13.2.2 宏包冲突 . . . . .        | 31        |
| 13.3 中文字体选项 . . . . .        | 32        |
| 13.3.1 方正字体选项 . . . . .      | 32        |
| 13.3.2 其他中文字体 . . . . .      | 32        |
| <b>第十四章 ElegantBook 写作示例</b> | <b>34</b> |
| 14.1 Lebesgue 积分 . . . . .   | 34        |
| 14.1.1 积分的定义 . . . . .       | 34        |
| 第十四章 练习 . . . . .            | 36        |
| <b>第十五章 常见问题集</b>            | <b>37</b> |
| <b>第十六章 版本更新历史</b>           | <b>38</b> |
| <b>参考文献</b>                  | <b>41</b> |
| <b>附录 A 基本数学工具</b>           | <b>42</b> |
| A.1 求和算子与描述统计量 . . . . .     | 42        |

# 第一章 机器学习基本概念

机器学习由几个要素所组成：数据，模型和优化方法。数据提供模型优化的目标。在有监督学习中，模型总是通过优化模型实现最小化经验风险，并期望借助于模型的泛化能力实现在未观测数据上进行预测。机器学习模型包括模型的结构，以及损失函数 (**Loss Function**)。经验风险即为训练数据对于当前损失函数的样本均值。优化方法则提供了这样一种工具：针对当前损失函数，求得满足条件的参数配置。

## 第二章 理论机器学习基础

### 定理 2.1 (PAC 理论)

对于某个固定的假设检验  $h$ ，当抽样样本量  $N$  足够大时，抽样样本上的样本均值  $E_{in}(h)$  将 probably close to 样本分布上期望  $E_{out}(h)$ 。其中，

$$E_{in}(h) = \frac{1}{N} \sum_{n=1}^N \| h(\mathbf{x}_n) \neq y_n \|$$

$$E_{out}(h) = E_{\mathbf{x} \sim q} [\| h(\mathbf{x}) \neq y \|].$$

这里，probably close 是指

$$\mathbb{P}(|E_{in}(h) - E_{out}(h)| > \epsilon) \leq 2 \exp(-2\epsilon^2 N) \quad (2.1)$$



## 第三章 机器学习中的优化方法

### 内容提要

- |  |                                  |
|--|----------------------------------|
| <input type="checkbox"/> Gradient Descent            | <input type="checkbox"/> Adagrad |
| <input type="checkbox"/> Stochastic Gradient Descent | <input type="checkbox"/> RMSprop |
| <input type="checkbox"/> Momentum                    | <input type="checkbox"/> Adam    |
| <input type="checkbox"/> Nesterov                    | <input type="checkbox"/> AdamW   |

我们可以把优化的问题简单的理解成一个“求最值”的问题。对于简单的情形，比如求  $x^2 + ax + b$  的最小值，我们很容易能够得到这个问题的解析解  $b - \frac{a^2}{4}$ 。但对于大多数机器学习问题，由于模型结构复杂，参数众多，甚至带有复杂的约束条件，我们很难得到解析解。此时，我们会采用一阶或者二阶优化方法，逐步逼近最优解。尤其在深度学习时代，采用更加便捷的一阶优化方法成为主流。当然这也是由于深度模型过于大和复杂导致的迫不得已的选择。

### 3.1 梯度下降方法 (Gradient Descent Method)

梯度下降算法是一个一阶优化算法，我们针对损失函数  $J(\theta)$ ，会首先将参数初始化为  $\theta_0$ ，之后采用迭代的方法不断更新参数  $\theta$ 。

$$\theta_{t+1} \leftarrow \theta_t - \nabla_{\theta} J(\theta) \quad (3.1)$$

 **练习 3.1** 请证明优化目标函数  $J(\theta)$  在当前参数  $\theta$  下的最优参数更新方向为梯度反方向  $-\nabla_{\theta} J(\theta)$ 。

**证明** 即  $D(x)$  在  $[0, 1]$  上是 Lebesgue 可积的并且积分值为零。但  $D(x)$  在  $[0, 1]$  上不是 Riemann 可积的。

### 3.2 随机梯度下降方法 (Stochastic Gradient Descent Method)

梯度下降方法在每一步迭代过程中，会计算优化目标函数关于所有样本  $\mathbf{x}_i$  的梯度，这会随着数据集中样本数量  $n$  的增长而线性增加，即计算复杂度为  $O(n)$ 。这导致的问题是，参数的一次迭代更新会非常耗时，且消耗大量计算资源。随机梯度下降方法则在参数迭代过程中只针对当前一个样本计算梯度  $\nabla f(\mathbf{x}_i)$ ，并以此更新模型参数。在样本是随机同分布的情况下，随机梯度的期望显然等同于在全部数据集上的。

### 3.3 Momentum

在梯度下降的基础上，引入“速度”的概念。在梯度下降时，不以当前的梯度  $\mathbf{x}_i$  进行参数更新，而是以速度  $\mathbf{m}$  进行更新：

$$\begin{aligned} \mathbf{m} &\leftarrow \beta \mathbf{m} - \eta \nabla_{\theta} J(\theta), \\ \theta &\leftarrow \theta + \mathbf{m}. \end{aligned} \quad (3.2)$$

书中经常举的例子是，对于一个狭长的半椭球体 (elongated half-ellipsoid) 损失函数，例如  $f(\mathbf{x}) = 0.1x_1^2 + 2x_2^2$ ，采用 Momentum 优化方法会使得在较小梯度方向  $\frac{\partial f(\mathbf{x})}{\partial x_1}$  不断加速。而在较大梯度方向  $\frac{\partial f(\mathbf{x})}{\partial x_2}$  上，通过梯度间的不断抵消，避免了在  $x_2$  上的不必要的大梯度，且减弱了梯度震荡的影响。Momentum 方法引入了一个新的参数  $\beta$ ，通常令  $\beta = 0.9$  是一个不错的选择。



### 3.4 Nesterov

Nesterov 优化是基于 Momentum 优化的改进：对于计算梯度的选点，由当前参数  $\theta$  更改为  $\theta + \beta m$ 。这是因为使用 Momentum 方法更新参数时，当前速度  $\beta m$  也会对参数进行更新，进行上述梯度修正，往往会得到一个更好的参数更新方向。

$$\begin{aligned} m &\leftarrow \beta m - \eta \nabla_{\theta} J(\theta + \beta m), \\ \theta &\leftarrow \theta + m. \end{aligned} \quad (3.3)$$

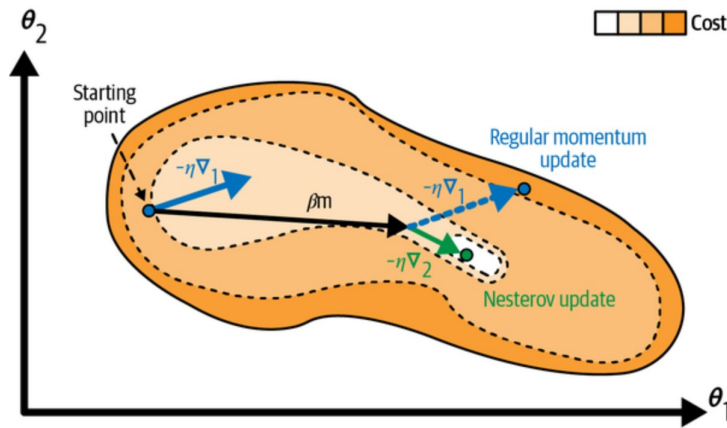


图 3.1: Nesterov 优化方法相较 Momentum 方法具有更好的参数更新方向

图 3.1 形象化说明了 Nesterov 比 Momentum 在计算参数更新方向上的优势 [4]<sup>1</sup>。

### 3.5 Adagrad

Adagrad 是在梯度下降的基础上的改进：对于参数在不同的维度上施加不同的 decay rate。对于某些更加陡峭的维度，decay rate 会相对更大，对于较为平缓的维度，decay rate 会相对更小。

$$\begin{aligned} s &\leftarrow s + \nabla_{\theta} J(\theta) \odot \nabla_{\theta} J(\theta), \\ \theta &\leftarrow \theta - \eta \nabla_{\theta} J(\theta) \oslash \sqrt{s + \epsilon}. \end{aligned} \quad (3.4)$$

由3.4可以看出，Adagrad 中用于 decay 梯度的  $s$  会随着迭代不断累积，这导致一个问题就是 AdaGrad 往往会 early stop 到一个不是很好的解。在实际应用中，尤其是深度学习场景，不推荐使用该优化方法。

### 3.6 RMSprop

RMSprop 是基于 AdaGrad 的改进。在 AdaGrad 中，计算用于平衡不同维度间下降速率的 decay weight 的  $s$  是不断累积的，因而会 blew up 进而使得模型训练过早地停止。RMSprop 在  $s$  的更新中引入了 exponential decay，从而避免产生过大的 decay weight，进而缓解了 stop too early 的问题。

<sup>1</sup> 图片来源于 [4]



$$\begin{aligned} s &\leftarrow \rho s + (1 - \rho) \nabla_{\theta} J(\theta) \odot \nabla_{\theta} J(\theta), \\ \theta &\leftarrow \theta - \eta \nabla_{\theta} J(\theta) \odot \sqrt{s + \epsilon}. \end{aligned} \quad (3.5)$$

RMSprop 相较之前的优化方法，多了一个参数  $\rho$ 。这个参数不需要过多调整，我们通常将其设置为 0.9。

### 3.7 Adam

Adam 优化方法可以理解为 RMSprop 和 Momentum 的“合体”。RMSprop 在更新参数中使用原始梯度，而 Adam 则使用了 Momentum 方法中的 Velocity 对参数更新：

$$\begin{aligned} m &\leftarrow \beta_1 m + (1 - \beta_1) \nabla_{\theta} J(\theta), \\ s &\leftarrow \beta_2 s + (1 - \beta_2) \nabla_{\theta} J(\theta) \odot \nabla_{\theta} J(\theta), \\ \hat{m} &\leftarrow \frac{m}{1 - \beta_1^t}, \\ \hat{s} &\leftarrow \frac{s}{1 - \beta_2^t}, \\ \theta &\leftarrow \theta - \eta \hat{m} \odot \sqrt{\hat{s} + \epsilon}. \end{aligned} \quad (3.6)$$

一般我们也不用去调其中  $\beta_1$  和  $\beta_2$  这两个参数。经验来看，我们会设定  $\beta_1$  为 0.9， $\beta_2$  为 0.99。我们注意到 Adam 里会对  $m$  和  $s$  进行修正，这种修正在迭代初期发挥比较大的作用。比如说，当  $t = 1$  时， $m = (1 - \beta_1) \nabla_{\theta} J(\theta)$ ，这相当于我们对梯度有一个放大的效果，因此可以通过除  $1 - \beta_1$  来进行修正。简单来说明一下分母是  $1 - \beta_1^t$  的原因：假设  $t = 1, \dots, T$  时，所有的梯度都相等且为  $\nabla$ ，那么：

$$\begin{aligned} m &= -(1 - \beta_1) \beta_1^{T-1} \nabla - (1 - \beta_1) \beta_1^{T-2} \nabla + \dots + (1 - \beta_1) \nabla \\ &= (1 - \beta_1^T) \nabla \end{aligned} \quad (3.7)$$

可以看出，我们期望的梯度  $\nabla$  和实际经过 exponential decay 后的速率  $m$  正好是  $1 - \beta_1^T$  的比例关系。

### 3.8 AdamW

AdamW 是 Adam 和 Weight Decay 的结合。所谓 Weight Decay 是将模型的参数在每次迭代之后乘一个 decay factor，比如 0.99。显然，当我们用最原始的 SGD 时，weight decay 相当于 l2 regularization。但是如果优化方法是 Adam 时，Adam + l2-reg 与 Adam + Weight Decay 并不等价。Adam + l2-reg 在实操中泛化能力不够理想，而 Adam + Weight Decay 一般效果更佳。作者本人在对 BERT 等 NLP 模型进行 FineTuning 时，一般默认选择 AdamW 优化方法。

## 第四章 机器学习中的损失函数

### 内容提要

❑ regression loss

❑ multi-class classification loss

❑ binary classification loss

### 4.1 回归问题的损失函数

常见的回归问题的损失函数包括：

- 平方损失函数 (square loss 或 l2-loss):  $L(\hat{y}, y) = (\hat{y} - y)^2$
- 绝对值损失函数 (absolute loss 或者 l1-loss):  $L(\hat{y}, y) = |\hat{y} - y|$
- Huber 损失函数: 当  $|\hat{y} - y| \leq \delta$ ,  $L(\hat{y}, y)$  为  $(\hat{y} - y)^2$ ; 当  $|\hat{y} - y| > \delta$ ,  $L(\hat{y}, y)$  为  $|\hat{y} - y|$

平方损失函数 (square loss) 在处理回归问题时最常见, 且其具有光滑可导、凸性等优点。但平方损失函数的缺点在于, 它对异常点 (outliers) 敏感。这里, 敏感指的是由于平方损失函数的形式, 异常点会最终“贡献”较大的损失值, 从而导致模型在训练过程中过于受异常点的影响。我们可以从最大似然的角度解释平方损失函数: 对于样本点  $\mathbf{x}_i$ , 样本标签  $y_i$  的噪声  $\epsilon^{(i)}$  独立同分布, 且服从于均值为 0, 标准差为  $\delta$  的高斯分布  $\mathcal{N}(0, \delta)$ 。那么在整个训练集上的似然为:

$$L(\theta) = \prod_{i=1}^n p(y^{(i)} | \mathbf{x}^{(i)}; \theta) \quad (4.1)$$

绝对值损失函数 (absolute loss) 相对平方损失函数对异常点较不敏感。但该损失函数在  $y = 0$  时不可导, 且当  $\hat{y}$  非常接近  $y$  时,  $|\frac{\partial L}{\partial \hat{y}}| = 1$ , 导致优化过程中当参数接近极值点时可能发生震荡, 此时的优化收敛性相较于使用平方损失函数时差。绝对值损失函数也可以从最大似然角度解释, 只不过样本标签的噪声满足 Laplace 分布:

#### 定义 4.1 (Laplace 分布)


一元维度上的 Laplace 分布密度函数为:

$$Lap(x | \mu, b) = \frac{1}{2b} \exp\left(-\frac{|x - \mu|}{b}\right) \quad (4.2)$$

基于此定义的 Laplace 分布的均值为  $\mu$ , 方差为  $2b^2$ 。

♣

Huber 损失函数可以看作平方损失函数和绝对值损失函数的结合, 对异常点没有平方损失函数那样敏感, 损失函数光滑可导, 并且在靠近极值点时具有平方损失函数的性质。

 **练习 4.1** 是否能够将线性回归用于二分类任务? 即假定二分类任务的 label 为 +1, -1, 利用 Linear Regression 去拟合样本点 label。

**解** 做法可行, 往往效果也不会太差。我们由

$$L(f, y) = (f - y)^2 \quad (4.3)$$

$$= (yf - 1)^2 \quad (4.4)$$

可以看出, 当 margin  $yf$  过大时, 损失函数的值也会很大。但实际如果 margin 是一个较大的正值, 恰恰说明模型能够以高置信度将样本判为正确的 label。也就是说, 这样的损失函数限制了模型的学习能力 (限制了一部分本可以很好完成当前二分类任务的 hypothesis)。

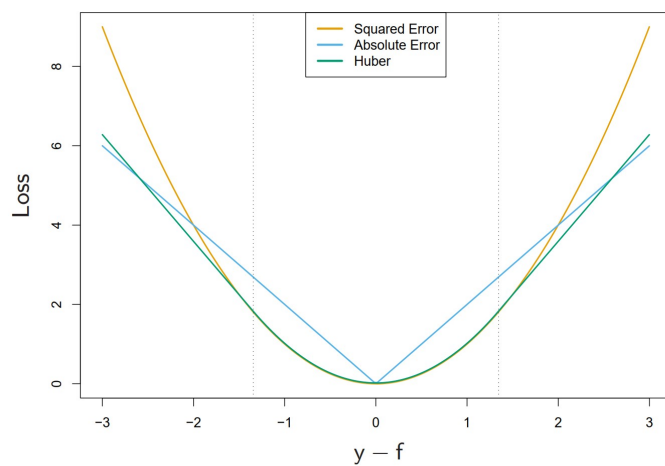


图 4.1: 平方损失函数 / 绝对值损失函数 / Huber 损失函数对比

## 4.2 分类问题的损失函数

### 4.2.1 二分类问题的损失函数

最符合人类思维的二分类损失函数即为 0-1 loss:

$$L(\hat{y}, y) = \mathbb{1} [\hat{y} \neq y] \quad (4.5)$$

### 4.2.2 多分类问题的损失函数

## 第五章 机器学习模型的评估指标

### 5.1 分类模型的评估指标

#### 5.1.1 二分类模型

对于分类最为简单直接的评估指标就是准确率 (accuracy)，即分类正确的样本个数在样本总体中的比例。准确率是很符合人类直观感受的一个评估指标，但并非适合很多场景，例如样本不均衡的情况。推荐系统中的转化率预估模型是一个二分类模型，在训练样本中大部分样本为负样本，如果只考虑准确率则甚至可能将所有预测样本预测为负样本会得到很高的准确率。显然，准确率无法满足我们衡量模型好坏的需求。

考虑到二分类任务中样本 label 只有正样本 (Positive) 和负样本 (Negative) 两种，预测 label 亦是如此。那么对于一个样本则可能有四种情况：

- True Positive (TP): label 为 +1 且预测结果为 +1
- False Negative (FN): label 为 +1 但预测结果为 -1
- False Positive (FP): label 为 -1 但预测结果为 +1
- True Negative (TN): label 为 -1 且预测结果为 -1

藉此，我们可以定义精确率 (Precision) 和召回率 (Recall) 两个指标：

$$Precision = \frac{TP}{TP + FP} \quad (5.1)$$

$$Recall = \frac{TP}{TP + FN} \quad (5.2)$$

直观理解，Precision 表示判断为正样本的集合中确实是正样本的比例，而 Recall 则表示预测为正样本的集合覆盖所有正样本的比例。对于二分类问题，通常最后的模型输出是样本为正样本的概率  $p(y = 1|\mathbf{x})$ 。我们会比较  $p(y = 1|\mathbf{x})$  和阈值  $\alpha$  的大小关系来决定最终模型预测的结果  $y_{pred}$ ：

$$y_{pred} = \begin{cases} 1, & p(y = 1|\mathbf{x}) \geq \alpha \\ 0, & p(y = 1|\mathbf{x}) < \alpha \end{cases} \quad (5.3)$$

对于同一个模型，当我们设置较大的阈值  $\alpha$  时，结果 Precision 会增大，但同时由于相当于我们施加了更严的标准，Recall 会相应降低。在实际应用中，我们往往希望模型达到一定的召回率，或者满足一定的精确度。我们在比较两个二分类模型的性能时，可以观察两个模型在相同召回下的精确率，亦或者比较相同精确率下的召回比例。如果我们希望更加全面地刻画模型的性能，我们可以不断增大阈值（从 0 到 1），以 Precision 为纵轴，Recall 为横轴，绘出如图 5.1 的 PR 曲线。我们希望 Precision 和 Recall 同时取得较大值，因此 PR 曲线越偏向右上方则越理想。

#### 5.1.2 多分类模型

### 5.2 回归问题的评估指标

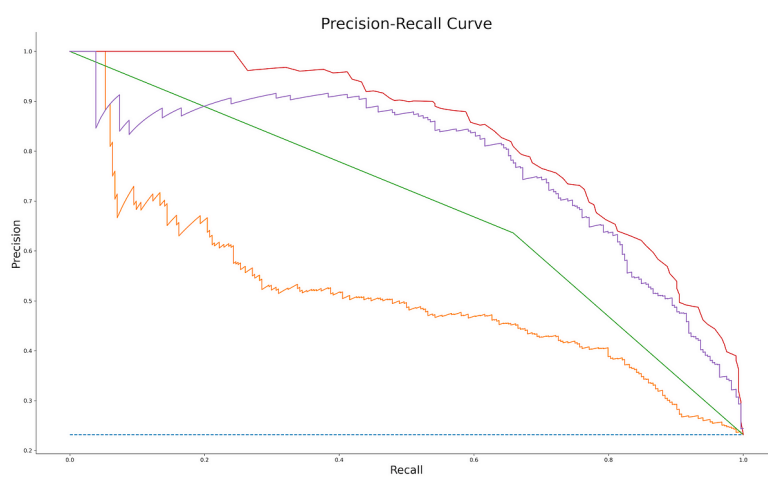


图 5.1: PR 曲线

## 第六章 树模型

### 内容提要

❑ ID3

❑ C4.5

❑ CART

### 6.1 分类回归树 (CART)

#### 6.1.1 分类回归树中的分类问题

#### 6.1.2 分类回归树中的回归问题

## 第七章 提升方法 Boosting

### 7.1 Forward Stagewise Additive Modeling

我们可以将 Additive Modeling 理解成为一组基函数  $b(\mathbf{x}; \gamma_m)$  的线性组合 (basis function expansion):

$$f(\mathbf{x}) = \sum_{m=1}^M \beta_m b(\mathbf{x}; \gamma_m) \quad (7.1)$$

其中,  $\beta_m \in \mathcal{R}$  是线性组合的系数,  $\{b(\mathbf{x}; \gamma_m)\}$  为 hypothesis space 中的基函数,  $\gamma_m$  为刻画基函数的参数。这里, 线性组合  $f(\mathbf{x})$  可以用于回归任务, 亦可通过拟合 logits 来用于分类任务 (后面将结合具体的例子理解)。

基于上述的 Additive Modeling 的形式, 机器学习的优化目标为:

$$\min_{\{\beta_m, \gamma_m\}_1^M} \sum_{i=1}^N L \left( y_i, \sum_{m=1}^M \beta_m b(\mathbf{x}_i; \gamma_m) \right) \quad (7.2)$$

但是遗憾的是该问题是 NP-Hard 问题, 我们很难直接针对该 additive model 进行优化。这里就引入了 Forward Stagewise 的概念, 即假设当前已拟合的函数为  $f_{m-1}(\mathbf{x})$ , 那么我们通过优化

$$\min_{\{\beta, \gamma\}} \sum_{i=1}^N L(y_i, f_{m-1}(\mathbf{x}_i) + \beta b(\mathbf{x}_i; \gamma)) \quad (7.3)$$

来进一步提升模型的 performance。

### 7.2 Gradient Boosting

### 7.3 XGBoost



## 第八章 推荐系统

### 内容提要

- ❑ 协同过滤
- ❑ 矩阵分解
- ❑ 图卷积神经网络
- ❑ BPR
- ❑ 因子分解机
- ❑ DeepFM
- ❑ A/B Test
- ❑ Interleaving

### 8.1 召回模型

#### 8.1.1 协同过滤 (Collaborative Filtering)

#### 8.1.2 Embedding 召回方法

##### 8.1.2.1 KD-Tree 算法

**KD-Tree 构建：**选择高维向量中的某一维（一般采用轮转的方式），将当前结点上的数据点按照这些数据点在该维度上的中位数进行划分，重复上述过程直到树的结点不包含数据点为止。

**KD-Tree 检索：**1) 首先由树的根结点寻到目标点所在的叶子结点，并以当前叶子结点作为最近邻的候选点；2) 递归地判断当前结点的兄弟结点是否与以目标点为中心，以当前最短距离为半径的超球体相交；3) 如果相交则进入兄弟结点进行搜索，搜索过程对最短距离进行实时更新；如果不相交则往父结点回退直到回退到根结点终止。

 **笔记** 读者朋友如果希望通过例子来更加直观理解 KD-Tree 的检索过程，可以参考[KD 树算法之详细篇](#)。KD-Tree 检索 K 近邻与检索最近邻类似，只需在检索过程中维护一个存储当前 K 近邻的 list 或者大顶堆即可。


##### 8.1.2.2 局部敏感哈希 (Local Sensitive Hash)

局部敏感哈希的基本思路是：两个高维空间上距离近的点，映射到低维空间中的两个投影点也相对靠近；两个高维空间上距离较远的点，映射到低维空间中的投影点则可能较远也可能较近。那么，如果我们随机构建若干投影映射，如果两个点在多个投影映射下的投影点都有相近的距离，那么大概率这两个点在原始空间中也距离较近。

我们可以假设有  $N$  个投影映射  $\{(p_1, b_1), (p_2, b_2), \dots, (p_N, b_N)\}$ ，且对应的分桶宽度为  $\{w_1, w_2, \dots, w_N\}$ ，那么对于高维空间中的向量  $x$ ，落在映射  $i$  的分桶为：

$$h_i(x) = \lfloor \frac{p_i^T x + b_i}{w_i} \rfloor \quad (8.1)$$

其中， $h_i(x) \in \mathbb{Z}$  为分桶标号。对于高维空间中的两个点  $x_1, x_2$ ，如果各自的分桶标号  $(h_1(x_1), h_2(x_1), \dots, h_N(x_1))$  和  $(h_1(x_2), h_2(x_2), \dots, h_N(x_2))$  在大多数维度上重合，则  $x_1, x_2$  在原始空间中相近。

 **笔记** 关于局部敏感哈希的详细介绍，可以参见 Pinecore 的 [Locality Sensitive Hashing: The Illustrated Guide](#)。

##### 8.1.2.3 Hierarchical Navigable Small Worlds (HNSW)

HNSW 算法构建了一个分层的索引图，如图8.1所示<sup>1</sup>。位于上层的索引图较为“稀疏”，点与点之间有较大距离。越往下层，索引图越“稠密”，点与点之间的距离则更近。在搜索时，先从上层索引进行搜索，找到近似最近

<sup>1</sup> 图片来源于 Pinecone。

邻点，再跳转至下层对应的点，并重复搜索过程，直到在最底层的索引图上找到最近邻点。HNSW 算法构建索引图以及在索引图上搜索的详细过程请参见 [Hierarchical Navigable Small Worlds](#)。

HNSW 索引是工业场景中最受欢迎的索引方法之一，尤其适合索引规模特别庞大的场景（Billion 级别甚至更大），但构建该索引过程较为耗时，且会占用大量的存储空间。

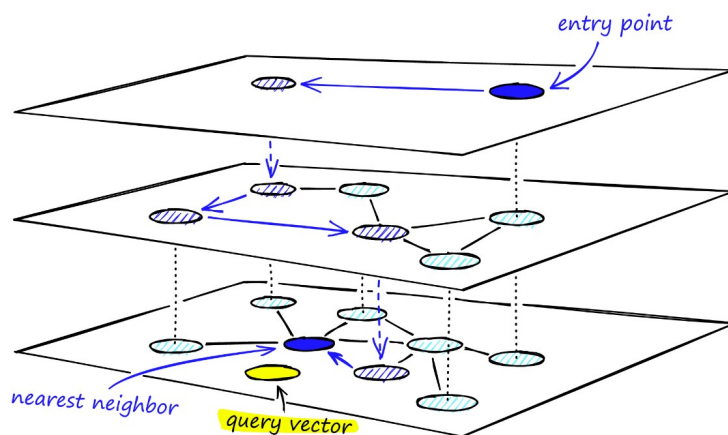


图 8.1: HNSW 检索最近邻过程

## 8.2 精排模型

召回模型从海量的数据中捞出一批候选集合，而精排模型则在召回的集合中“精挑细选”，将候选集合中的对象打分排序。相对与召回模型，精排模型要求更加精确，它往往会利用更多的特征或者采用更加复杂的模型结构，藉此来充分挖掘数据之间的关联性。简而言之，召回模型需要尽量做到不遗漏，精排模型力求更加精准。

### 8.2.1 因子分解机 (Factorization Machine)

### 8.2.2 DeepFM

## 8.3 线上实验评估方法

### 8.3.1 A/B Test

### 8.3.2 Interleaving

## 第九章 Learning to Rank

# 第十章 自然语言处理

## 内容提要

❑ Attention Mechanism

❑ Transformer

❑ BERT

❑ GPT

❑ Bi-Encoder

❑ Cross-Encoder

## 10.1 注意力机制 (Attention Mechanism)

### 10.1.1 为什么要用多头注意力机制?

通俗的解释是，多头注意力机制会学习不同的 Q、K、V 子空间，因而能够将这种不同维度的信息综合起来，使得模型有更好的表征能力。这里也可以参照 cs224n 中 [assignment5](#) 举的例子来解释：多头注意力机制对于 K 的长度变化更加不敏感，因而增强了模型的鲁棒性和泛化能力。

## 10.2 Transformer

### 10.2.1 Batch Normalization 还是 Layer Normalization?

## 10.3 基于 BERT 的检索模型

本章节将介绍两种业界常见的基于 BERT 的信息检索 (Information Retrieval) 模型：Bi-Encoder 和 Cross-Encoder。其中 Bi-Encoder 是双塔结构，主要用于信息的召回阶段；Cross-Encoder 为单塔结构，主要用于召回后的精排或者粗排。

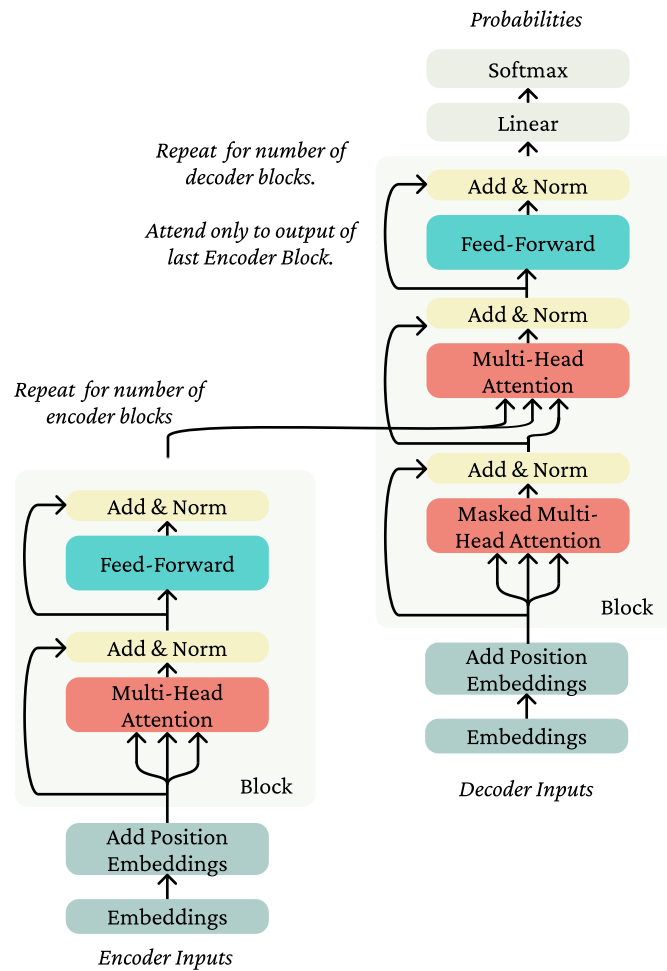
### 10.3.1 Bi-Encoder

双塔结构广泛应用于推荐和搜索场景，由于该模型结构能够解耦合查询的 query 和待查询的 document pool 或者用户与待推荐的商品集合。我们下面以搜索场景介绍：当模型训练完成后，我们对 document pool 中的所有 document 进行模型推理。得到 embedding 后构建近似搜索的索引。当新的 query 产生，双塔模型中的 query tower 将进行 inference 生成 embedding。我们即可用此 embedding 在 document embedding index 中检索。搜索模型中双塔模型想要达到比较好的效果，在以下几个方面需要有所考量：

#### 10.3.1.1 负样本采样

负样本采样的“招式”很多，既有 in-batch negative sampling，又有 offline hard negative mining。其中，in-batch 的负样本采样指的是对于某个正例样本组  $(q, d_+)$ ，我们可以将在模型训练过程中同一个 batch 内的其它样本组的 document 取出作为负样本。可以取一部分，例如按照  $q$  与负样本的 score 来排序截取。也可以全部拿来作为负样本使用。

Offline hard negative mining 相对复杂一些：根据当前模型从 document 集合中召回样本，剔除其中包含的正例样本。这也是其中“negative”的由来，因为按照此法得到的 document 实为与  $q$  很相关但又没有被用户所点击。这里涉及到一个 False Negative 的问题，



Transformer Encoder-Decoder

图 10.1: Transformer 架构

### **10.3.1.2 对比学习**

### **10.3.2 Cross-Encoder**

# 第十一章 $\text{Elegant}\text{L}\text{A}\text{T}\text{E}\text{X}$ 系列模板介绍

本模板自 2023 年 1 月 1 日开始，不再维护，不建议使用本系列模板！为了保证之前版本的用户仍然能查到说明文档，本说明文档仍然保留过去的信息。

$\text{Elegant}\text{L}\text{A}\text{T}\text{E}\text{X}$  项目组致力于打造一系列美观、优雅、简便的模板方便用户使用。目前由  $\text{ElegantNote}$ ,  $\text{ElegantBook}$ ,  $\text{ElegantPaper}$  组成，分别用于排版笔记、书籍和工作论文。大版本改动较大，请关注版本信息，在未开始使用模板前，建议直接选择最新正式版本！

本文将介绍本模板的一些设置内容以及基本使用方法。如果您有其他问题，建议或者意见，欢迎在 GitHub 上给我们提交 `issues` 或者邮件联系我们。我们的联系方式如下，建议加入用户 QQ 群提问，这样能更快获得准确的反馈，加群时请备注  $\text{L}\text{A}\text{T}\text{E}\text{X}$  或者  $\text{Elegant}\text{L}\text{A}\text{T}\text{E}\text{X}$  相关内容。

- 官网: <https://elegantlatex.org/>
- GitHub 地址: <https://github.com/ElegantLaTeX/>
- CTAN 地址: <https://ctan.org/pkg/elegantbook>
- 下载地址: 正式发行版, 最新版
- Bilibili:  $\text{ElegantLaTeX}$
- 用户 QQ 群: 692108391

## 11.1 模板安装与更新

你可以通过免安装的方式使用本模板，包括在线使用和本地（文件夹内）使用两种方式，也可以通过  $\text{T}\text{E}\text{X}$  发行版安装使用。

### 11.1.1 在线使用模板

我们把三套模板全部上传到  $\text{Overleaf}$  上了，网络便利的用户可以直接通过  $\text{Overleaf}$  在线使用我们的模板。使用  $\text{Overleaf}$  的好处是无需安装  $\text{T}\text{E}\text{X}$  Live，可以随时随地访问自己的文件。查找模板，请在  $\text{Overleaf}$  模板库里面搜索 `elegantlatex` 即可，你也可以直接访问[搜索结果](#)。选择适当的模板之后，将其 `Open as Template`，即可把模板存到自己账户下，然后可以自由编辑以及与他人一起协作。更多关于  $\text{Overleaf}$  的介绍和使用，请参考  $\text{Overleaf}$  的[官方文档](#)。

### 11.1.2 本地免安装使用

免安装使用方法如下：从 GitHub 或者 CTAN 下载最新版，严格意义上只需要类文件 `elegantbook.cls`。然后将模板文件放在你的工作目录下即可使用。这样使用的好处是，无需安装，简便；缺点是，当模板更新之后，你需要手动替换 `cls` 文件。

### 11.1.3 发行版安装与更新

本模板测试环境为

1. Win10 +  $\text{T}\text{E}\text{X}$  Live 2022;
2. Ubuntu 20.04 +  $\text{T}\text{E}\text{X}$  Live 2022;
3. macOS Monterey +  $\text{MacT}\text{E}\text{X}$  2022.

$\text{T}\text{E}\text{X}$ Live/ $\text{MacT}\text{E}\text{X}$  的安装请参考[嗝行的一份简短的关于安装  \$\text{L}\text{A}\text{T}\text{E}\text{X}\$  安装的介绍](#)。



安装  $\text{T}_{\text{E}}\text{X}$  Live 之后, 安装后建议升级全部宏包, 升级方法: 使用 `cmd` 或 `terminal` 运行 `tlmgr update --all`, 如果 `tlmgr` 需要更新, 请使用 `cmd` 运行 `tlmgr update --self`, 如果更新过程中出现了中断, 请改用 `tlmgr update --self --all --reinstall-forcibly-removed` 更新, 也即

```
tlmgr update --self
tlmgr update --all
tlmgr update --self --all --reinstall-forcibly-removed
```

更多的内容请参考 [How do I update my  \$\text{T}\_{\text{E}}\text{X}\$  distribution?](#)

#### 11.1.4 其他发行版本

由于宏包版本问题, 本模板不支持  $\text{C}_{\text{T}}\text{E}_{\text{X}}$  套装, 请务必安装  $\text{TeX}$  Live/ $\text{MacT}_{\text{E}}\text{X}$ 。更多关于  $\text{T}_{\text{E}}\text{X}$  Live 的安装使用以及  $\text{C}_{\text{T}}\text{E}_{\text{X}}$  与  $\text{T}_{\text{E}}\text{X}$  Live 的兼容、系统路径问题, 请参考官方文档以及嘯行的一份简短的关于安装  $\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$  安装的介绍。

## 第十二章 ElegantBook 设置说明

本模板基于基础的 book 文类，所以 book 的选项对于本模板也是有效的（纸张无效，因为模板有设备选项）。默认编码为 UTF-8，推荐使用  $\text{\TeX}$  Live 编译。

### 12.1 语言模式

本模板内含两套基础语言环境 `lang=cn`、`lang=en`。改变语言环境会改变图表标题的引导词（图，表），文章结构词（比如目录，参考文献等），以及定理环境中的引导词（比如定理，引理等）。不同语言模式的启用如下：

```
\documentclass[cn]{elegantbook}
\documentclass[lang=cn]{elegantbook}
```

除模板自带的两套语言设定之外，由网友提供的其他语言环境设置如下：

- 由 VincentMVV 提供的意大利语翻译 `lang=it`，相关讨论见 [Italian translation](#)；
- 由 abfek66 提供的法语翻译 `lang=fr`，相关讨论见 [Italian translation](#)；
- 由 inktvis75 提供的荷兰语翻译 `lang=nl`，相关讨论见 [Dutch Translation](#)；
- 由 palkotamas 提供的匈牙利语翻译 `lang=hu`，相关讨论见 [Hungarian translation](#)；
- 由 Lisa 提供的德语翻译 `lang=de`，相关讨论见 [Deutsch translation](#)；
- 由 Gustavo A. Corradi 提供的西班牙语的翻译 `lang=es`，相关讨论见 [Spanish translation](#)；
- 由 Altantsooj 提供的蒙古语的翻译 `lang=mn`，相关讨论见 [Mongolian translation](#)；
- 由 inusturbo 提供的日语的翻译 `lang=jp`，相关讨论见 [Japanese Translation](#)。

注 以上各个语言的设定均为网友设定，我们未对上述翻译进行过校对，如果有问题，请在对应的 issue 下评论。并且，只有中文环境（`lang=cn`）才可以输入中文。

### 12.2 设备选项

最早我们在 ElegantNote 模板中加入了设备选项（`device`），后来，我们觉得这个设备选项的设置可以应用到 ElegantBook 中<sup>1</sup>，而且 Book 一般内容比较多，如果在 iPad 上看无需切边，放大，那用户的阅读体验将会得到巨大提升。你可以使用下面的选项将版面设置为 iPad 设备模式<sup>2</sup>

```
\documentclass[pad]{elegantbook} %or
\documentclass[device=pad]{elegantbook}
```

### 12.3 颜色主题

本模板内置 5 组颜色主题，分别为 `green`<sup>3</sup>、`cyan`、`blue`（默认）、`gray`、`black`。另外还有一个自定义的选项 `nocolor`。调用颜色主题 `green` 的方法为





















```
\documentclass[green]{elegantbook} %or
\documentclass[color=green]{elegantbook}
```

<sup>1</sup>不过因为 ElegantBook 模板封面图片的存在，在修改页面设计时，需要对图片进行裁剪。

<sup>2</sup>默认为 normal 模式，也即 A4 纸张大小。

<sup>3</sup>为原先默认主题。

表 12.1: ElegantBook 模板中的颜色主题

|           | green   | cyan  | blue  | gray  | black   | 主要使用的环境                     |
|-----------|---|---|---|---|---|-----------------------------|
| structure |  |  |  |  |  | chapter section subsection  |
| main      |  |  |  |  |  | definition exercise problem |
| second    |  |  |  |  |  | theorem lemma corollary     |
| third     |  |  |  |  |  | proposition                 |

如果需要自定义颜色的话请选择 `nocolor` 选项或者使用 `color=none`, 然后在导言区定义 `structurecolor`、`main`、`second`、`third` 颜色, 具体方法如下:

```
\definecolor{structurecolor}{RGB}{0,0,0}
\definecolor{main}{RGB}{70,70,70}
\definecolor{second}{RGB}{115,45,2}
\definecolor{third}{RGB}{0,80,80}
```

## 12.4 封面

### 12.4.1 封面个性化

从 3.10 版本开始, 封面更加弹性化, 用户可以自行选择输出的内容, 包括 `\title` 在内的所有封面元素都可为空。目前封面的元素有

表 12.2: 封面元素信息

| 信息 | 命令                      | 信息  | 命令                     | 信息 | 命令                    |
|----|-------------------------|-----|------------------------|----|-----------------------|
| 标题 | <code>\title</code>     | 副标题 | <code>\subtitle</code> | 作者 | <code>\author</code>  |
| 机构 | <code>\institute</code> | 日期  | <code>\date</code>     | 版本 | <code>\version</code> |
| 箴言 | <code>\extrainfo</code> | 封面图 | <code>\cover</code>    | 徽标 | <code>\logo</code>    |

另外, 额外增加一个 `\bioinfo` 命令, 有两个选项, 分别是信息标题以及信息内容。比如需要显示 User Name: 111520, 则可以使用

```
\bioinfo{User Name}{111520}
```

封面中间位置的色块的颜色可以使用下面命令进行修改:

```
\definecolor{customcolor}{RGB}{32,178,170}
\colorlet{coverlinecolor}{customcolor}
```

### 12.4.2 封面图

本模板使用的封面图片来源于 [pixabay.com](http://pixabay.com)<sup>4</sup>, 图片完全免费, 可用于任何场景。封面图片的尺寸为 1280×1024, 更换图片的时候请严格按照封面图片尺寸进行裁剪。推荐一个免费的在线图片裁剪网站 [fotor.com](http://fotor.com)。用户 QQ 群内有一些合适尺寸的封面, 欢迎取用。

<sup>4</sup>感谢 ChinaTeX 提供免费图源网站, 另外还推荐 [pexels.com](http://pexels.com)。

### 12.4.3 徽标

本文用到的 Logo 比例为 1:1，也即正方形图片，在更换图片的时候请选择合适的图片进行替换。

### 12.4.4 自定义封面

另外，如果使用自定义的封面，比如 Adobe illustrator 或者其他软件制作的 A4 PDF 文档，请把 `\maketitle` 注释掉，然后借助 `pdfpages` 宏包将自制封面插入即可。如果使用 `titlepage` 环境，也是类似。如果需要 2.x 版本的封面，请参考 `etitlepage`。

## 12.5 章标标题

本模板内置 2 套章标题显示风格，包含 `hang`（默认）与 `display` 两种风格，区别在于章标题单行显示（`hang`）与双行显示（`display`），本说明使用了 `hang`。调用方式为

```
\documentclass[hang]{elegantbook} %or
\documentclass[titlestyle=hang]{elegantbook}
```

在章标题内，章节编号默认是以数字显示，也即第 1 章，第 2 章等等，如果想要把数字改为中文，可以使用

```
\documentclass[chinese]{elegantbook} %or
\documentclass[scheme=chinese]{elegantbook}
```

## 12.6 数学环境简介

在我们这个模板中，我们定义了两种不同的定理模式 `mode`，包括简单模式（`simple`）和炫彩模式（`fancy`），默认为 `fancy` 模式，不同模式的选择为

```
\documentclass[simple]{elegantbook} %or
\documentclass[mode=simple]{elegantbook}
```

本模板定义了四大类环境

- 定理类环境，包含标题和内容两部分，全部定理类环境的编号均以章节编号。根据格式的不同分为 3 种
  - `definition` 环境，颜色为 `main`；
  - `theorem`、`lemma`、`corollary`、`axiom`、`postulate` 环境，颜色为 `second`；
  - `proposition` 环境，颜色为 `third`。
- 示例类环境，有 `example`、`problem`、`exercise` 环境（对应于例、例题、练习），自动编号，编号以章节为单位，其中 `exercise` 有提示符。
- 提示类环境，有 `note` 环境，特点是：无编号，有引导符。
- 结论类环境，有 `conclusion`、`assumption`、`property`、`remark`、`solution` 环境<sup>5</sup>，三者均以粗体的引导词为开头，和普通段落格式一致。

其中，定理类环境均有带星号的版本：`definition*`、`theorem*`、`lemma*`、`corollary*`、`axiom*`、`postulate*`、`proposition*`，带星号的定理类环境不会编号。

<sup>5</sup>本模板还添加了一个 `result` 选项，用于隐藏 `solution` 和 `proof` 环境，默认为显示（`result=answer`），隐藏使用 `result=noanswer`。

## 12.6.1 定理类环境的使用

### 12.6.1.1 fancy 模式

在 fancy 模式下使用了 `tcolorbox` 宏包来定制定理类环境，所以和普通的定理环境的使用有些许区别，有编号定理的使用方法如下：

```
% 有名字，有标签
\begin{theorem}{theorem name}{label}
  这是一个有名字和标签的定理。
  用 \ref{thm:label} 来引用这个定理。
\end{theorem}

% 无名字，有标签
\begin{theorem}{}{label no name}
  这是一个没有名字但是有标签的定理。
  用 \ref{thm:label no name} 来引用这个定理。
\end{theorem}

% 有名字，无标签
\begin{theorem}{theorem name}{}
  这是一个有名字但是没有标签的定理。
  这个定理不能被引用。
  最后的 {} 可以省略不写。
\end{theorem}

% 无名字，无标签
\begin{theorem}{}{}
  这是一个没有名字也没有标签的定理。
  这个定理不能被引用。
  两个 {} 均可以省略不写。
\end{theorem}
```

第一个选项 `theorem name` 是定理的名字，如果定理没有名字请使用 {}<sup>6</sup>。第二个选项 `label` 是交叉引用时所用到的标签，如果定理没有标签，可以不写 `{label}` 或使用 {}。交叉引用的方法为 `\ref{thm:label}`。请注意，交叉引用时必须加上前缀 `thm:`。其他相同用法的定理类环境见表 12.3。

不编号的定理使用方法如下：

```
% 有名字
\begin{theorem*}{theorem name}
  这是一个不编号的有名字的定理。
  这个定理不能被引用。
\end{theorem*}

% 无名字
\begin{theorem*}{}
  这是一个不编号且没有名字的定理。
  最后的 {} 可以省略不写。
  这个定理不能被引用。
\end{theorem*}
```

其中的选项 `theorem name` 是定理的名字，如果没有名字可以不写 `{theorem name}` 或使用 {}

在用户多次反馈下，在 4.1 版本引入了原生定理的支持方式，使用方法如下：

<sup>6</sup>除非这个定理也没有标签，否则不能省略。

表 12.3: 定理类环境

| 环境名         | 标签名   | 前缀  | 交叉引用                         |
|-------------|-------|-----|------------------------------|
| definition  | label | def | <code>\ref{def:label}</code> |
| theorem     | label | thm | <code>\ref{thm:label}</code> |
| postulate   | label | pos | <code>\ref{pos:label}</code> |
| axiom       | label | axi | <code>\ref{axi:label}</code> |
| lemma       | label | lem | <code>\ref{lem:label}</code> |
| corollary   | label | cor | <code>\ref{cor:label}</code> |
| proposition | label | pro | <code>\ref{pro:label}</code> |

```
% 有名字，有标签
\begin{theorem}[theorem name]\label{thm:label}
  这是一个有名字和标签的定理。
  用 \ref{thm:label} 来引用这个定理。
\end{theorem}

% 无名字，有标签
\begin{theorem}\label{thm:label no name}
  这是一个没有名字但是有标签的定理。
  用 \ref{thm:label no name} 来引用这个定理。
\end{theorem}

% 有名字，无标签
\begin{theorem}[theorem name]
  这是一个有名字但是没有标签的定理。
  这个定理不能被引用。
\end{theorem}

% 无名字，无标签
\begin{theorem}
  这是一个没有名字也没有标签的定理。
  这个定理不能被引用。
\end{theorem}
```

这时引用不需要加前缀，`\ref` 中的内容与 `\label` 中的内容相同，换句话说，这时的使用方法与使用原生定理环境的时候几乎相同。

### 12.6.1.2 simple 模式

在 `simple` 模式下使用了 `amsthm` 宏包，定理类环境使用方法与原生一致，定理的使用方法如下：

```
% 有名字，有标签
\begin{theorem}[theorem name]\label{thm:label}
  这是一个有名字和标签的定理。
  用 \ref{thm:label} 来引用这个定理。
\end{theorem}

% 无名字，有标签
\begin{theorem}\label{thm:label no name}
  这是一个没有名字但是有标签的定理。
  用 \ref{thm:label no name} 来引用这个定理。
\end{theorem}

% 有名字，无标签
```

```
\begin{theorem}[theorem name]
  这是一个有名字但是没有标签的定理。
  这个定理不能被引用。
\end{theorem}
% 无名字，无标签
\begin{theorem}
  这是一个没有名字也没有标签的定理。
  这个定理不能被引用。
\end{theorem}
```

### 12.6.2 修改计数器

当前定理等环境计数器按章计数，如果想修改定理类环境按节计数，可以修改计数器选项 `thmcnt`，可用选项为 `chapter`（默认）与 `section`：

```
\documentclass[section]{elegantbook} %or
\documentclass[thmcnt=section]{elegantbook}
```

如果希望全局的定理类环境使用同一个计数器，可以使用文档类选项 `usesamecnt`：

```
\documentclass[usesamecnt]{elegantbook}
```

### 12.6.3 自定义定理类环境

4.4 版本新增了一个自定义定理类环境的命令：`\elegantnewtheorem`，它的参数含义如下：

```
% fancy 模式（默认）
\elegantnewtheorem{env}{title}{style}{prefix}[numbered-like]
% simple 模式
\elegantnewtheorem{env}{title}{style}[numbered-like]
```

该命令可以同时定义编号环境 `env` 和不编号环境 `env*`。

其中 `style` 支持的参数有：`defstyle`，`thmstyle`，`prostyle`，分别对应“定义”，“定理”，“命题”三种样式。

如果添加了可选参数 `numbered-like`，将会使该定理类环境与名为 `numbered-like` 的定理类环境使用同一计数器。**注意：**该参数在使用 `usesamecnt` 选项时不起作用，并且会在终端以及 `.log` 文件中输出一个警告，来提示用户该选项不起作用：

```
[numbered-like] won't make sence with option
`usesamecnt'.
```

- 在炫彩模式（`fancy`）下，需要 5 个参数来定义一个新的定理类环境，分别是：定理类环境名，定理类环境的标题，定理类环境的样式，该定理类环境的前缀，（可选）该定理类环境继承的定理类环境：

```
% 导言区
\elegantnewtheorem{examplefancy}{自定义定理类环境}{thmstyle}{exfancy}
% 正文
\begin{examplefancy}{定理名}{label}
  这里是自定义定理类环境 \ref{exfancy:label}
\end{examplefancy}
\begin{examplefancy*}{定理名}
  这里是无编号自定义定理类环境
\end{examplefancy*}
```



如果不给出第四个参数，或第四个参数置空的话，将会用定理类环境名来作为默认前缀，即

```
% 引言区
\elegantnewtheorem{test}{TEST}{thmstyle}
% 或
\elegantnewtheorem{test}{TEST}{thmstyle}{}
% 正文
\begin{test}{name}{label}
    默认前缀为 test。
    使用 \ref{test:label} 来引用这个定理类环境。
\end{test}
```

这时会在终端以及 .log 文件中输出一个警告信息来提示用户没有定义前缀：

```
Class elegantbook Warning: Because you didn't provide a prefix.
(elegantbook) We use test as the default prefix.
(elegantbook) You have to use
(elegantbook) \ref{test:label} to refer a
(elegantbook) \begin{test}{name}{label} environment.
(elegantbook) on input line 3.
```

- 在简单模式（simple）下，需要 4 个参数来定义一个新的定理类环境，分别是：定理类环境名，定理类环境的标题，定理类环境的样式，该定理类环境的前缀，（可选）该定理类环境继承的定理类环境：

```
% 引言区
\elegantnewtheorem{examplesimple}{自定义定理类环境}{thmstyle}
% 正文
\begin{examplesimple}[定理名]\label{exsimple:label}
    这里是自定义定理类环境 \ref{exsimple:label}
\end{examplesimple}
\begin{examplesimple*}[定理名]
    这里是无编号自定义定理类环境
\end{examplesimple*}
```

如果此时错误地给出了第四个参数，那么将会在终端以及 .log 文件中输出一个错误信息：

```
% elegantbook-cn.tex
\elegantnewtheorem{test}{TEST}{thmstyle}{}
% .log file
./elegantbook-cn.tex:3: Class elegantbook Error: You can't set a prefix in mode ``simple''.
(elegantbook) Just use
(elegantbook) \elegantnewtheorem {test}{TEST}{thmstyle} .
```

### 12.6.4 其他环境的使用

其他三种环境没有选项，可以直接使用，比如 example 环境的使用方法与效果：

```
\begin{example}
    This is the content of example environment.
\end{example}
```

这几个都是同一类环境，区别在于

- 示例环境 (example)、练习 (exercise) 与例题 (problem) 章节自动编号;
- 注意 (note)、练习 (exercise) 环境有提醒引导符;
- 结论 (conclusion) 等环境都是普通段落环境, 引导词加粗。

## 12.7 列表环境

本模板借助于 `tikz` 定制了 `itemize` 和 `enumerate` 环境, 其中 `itemize` 环境修改了 3 层嵌套, 而 `enumerate` 环境修改了 4 层嵌套 (仅改变颜色)。示例如下

- |                           |                             |
|---------------------------|-----------------------------|
| • first item of nesti;    | 1. first item of nesti;     |
| • second item of nesti;   | 2. second item of nesti;    |
| • first item of nestii;   | (a). first item of nestii;  |
| • second item of nestii;  | (b). second item of nestii; |
| • first item of nestiii;  | I. first item of nestiii;   |
| • second item of nestiii. | II. second item of nestiii. |

## 12.8 参考文献

### 12.8.1 打印文献

之前我们将文献调用的命令放在模板里面, 然后用户反馈 `\cite` 命令无法自动补全, 因此我们新版本将其拿到外面来, 新版本打印参考文献的命令的方法是, 在导言区 (也即 `\begin{document}` 之前), 加入:

```
\addbibresource[location=local]{reference.bib}
```

然后再需要打印文献的地方使用:

```
\printbibliography[heading=bibintoc, title=\ebibName]
```

其中 `reference.bib` 为参考文献存放的文件, 需要放在项目文件夹下。

### 12.8.2 修改文献格式

此外, 本模板调用了 `biblatex` 宏包, 并提供了 `biber` (默认) 和 `bibtex` 两个后端选项, 可以使用 `bibend` 进行修改:

```
\documentclass[bibtex]{elegantbook}
\documentclass[bibend=bibtex]{elegantbook}
```

关于文献条目 (bib item), 你可以在谷歌学术, Mendeley, Endnote 中取, 然后把它们添加到 `reference.bib` 中。在文中引用的时候, 引用它们的键值 (bib key) 即可。

为了方便文献样式修改, 模板引入了 `bibstyle` 和 `citestyle` 选项, 默认均为数字格式 (numeric), 参考文献示例: [1, 2, 5] 使用了中国一个大型的 P2P 平台 (人人贷) 的数据来检验男性投资者和女性投资者在投资表现上是否有显著差异。

如果需要设置为国标 GB7714-2015, 需要使用:

```
\documentclass[citestyle=gb7714-2015, bibstyle=gb7714-2015]{elegantbook}
```

在使用

```
\documentclass[citestyle=gb7714-2015, bibstyle=gb7714-2015]{elegantbook}
```

后，排序方式为按引用先后排序。如果不使用国标 GB7714-2015 并且需要添加排序方式，可以在导言区加入下面命令：

```
\ExecuteBibliographyOptions{sorting=<name>}
```

其中 <name> 可以是 nty, nyt, nyvt, anyt, anyvt, ynt, ydnt, none, count, debug 的其中之一，更多排序相关请参考 biblatex 宏包手册的 3.1.2.1 节。

## 12.9 添加序章

如果你想在第一章前面添序章，不改变原本章节序号，可以在第一章内容前面使用

```
\chapter*{Introduction}
\markboth{Introduction}{Introduction}
The content of introduction.
```

## 12.10 目录选项与深度

本模板添加了一个目录选项 `toc`，可以设置目录为单栏 (`onecol`) 和双栏 (`twocol`) 显示，比如双栏显示可以使用

```
\documentclass[twocol]{elegantbook}
\documentclass[toc=twocol]{elegantbook}
```

默认本模板目录深度为 1，你可以在导言区使用

```
\setcounter{tocdepth}{2}
```

将其修改为 2 级目录（章与节）显示。

## 12.11 章节摘要

模板新增了一个章节摘要环境 (`introduction`)，使用示例

```
\begin{introduction}
  \item Definition of Theorem
  \item Ask for help
  \item Optimization Problem
  \item Property of Cauchy Series
  \item Angle of Corner
\end{introduction}
```

效果如下：

### 内容提要

- |  |  |
|--|--|
| <input type="checkbox"/> Definition of Theorem | <input type="checkbox"/> Property of Cauchy Series |
| <input type="checkbox"/> Ask for help          | <input type="checkbox"/> Angle of Corner           |
| <input type="checkbox"/> Optimization Problem  |  |

环境的标题文字可以通过这个环境的可选参数进行修改，修改方法为：

```
\begin{introduction}[Brief Introduction]
...
\end{introduction}
```

## 12.12 章后习题

前面我们介绍了例题和练习两个环境，这里我们再加一个，章后习题（`problemset`）环境，用于在每一章结尾，显示本章的练习。使用方法如下

```
\begin{problemset}
  \item exercise 1
  \item exercise 2
  \item exercise 3
\end{problemset}
```

效果如下：

### 我的题目

1. exercise 1
2. exercise 2
3. exercise 3
4. 测试数学公式

$$a^2 + b^2 = c^2, (1, 2) [1, 23] \quad (12.1)$$

**注** 如果你想把 `problemset` 环境的标题改为其他文字，你可以类似于 `introduction` 环境修改 `problemset` 的可选参数。另外，目前这个环境会自动出现在目录中，但是不会出现在页眉页脚信息中（待解决）。

**解** 如果你想把 `problemset` 环境的标题改为其他文字，你可以类似于 `introduction` 环境修改 `problemset` 的可选参数。另外，目前这个环境会自动出现在目录中，但是不会出现在页眉页脚信息中（待解决）。

## 12.13 旁注

在 3.08 版本中，我们引入了旁注设置选项 `marginpar=margintrue` 以及测试命令 `\elegantpar`，但是由此带来一堆问题。我们决定在 3.09 版本中将其删除，并且，在旁注命令得到大幅度优化之前，不会将此命令再次引入书籍模板中。对此造成各位用户的不方便，非常抱歉！不过我们保留了 `marginpar` 这个选项，你可以使用 `marginpar=margintrue` 获得保留右侧旁注的版面设计。然后使用系统自带的 `\marginpar` 或者 `marginnote` 宏包的 `\marginnote` 命令。

**注** 在使用旁注的时候，需要注意的是，文本和公式可以直接在旁注中使用。

```
% text
\marginpar{margin paragraph text}

% equation
\marginpar{
  \begin{equation}
    a^2 + b^2 = c^2
  \end{equation}
}
```

但是浮动体（表格、图片）需要注意，不能用浮动体环境，需要使用直接插图命令或者表格命令环境。然后使用 `\captionof` 为其设置标题。为了得到居中的图表，可以使用 `\centerline` 命令或者 `center` 环境。更多详情请参考：Caption of Figure in Marginpar。

```
% graph with centerline command
\marginpar{
  \centerline{
    \includegraphics[width=0.2\textwidth]{logo.png}
  }
  \captionof{figure}{your figure caption}
}

% graph with center environment
\marginpar{
  \begin{center}
    \includegraphics[width=0.2\textwidth]{logo.png}
    \captionof{figure}{your figure caption}
  \end{center}
}
```

## 第十三章 字体选项

字体选项独立成章的原因是，我们希望本模板的用户关心模板使用的字体，知晓自己使用的字体以及遇到字体相关的问题能更加便捷地找到答案。

**重要提示：**从 3.10 版本更新之后，沿用至今的 `newtx` 系列字体被重新更改为 `cm` 字体。并且新增中文字体（`chinesefont`）选项。

### 13.1 数学字体选项

本模板定义了一个数学字体选项（`math`），可选项有三个：

1. `math=cm`（默认），使用 L<sup>A</sup>T<sub>E</sub>X 默认数学字体（推荐，无需声明）；
2. `math=newtx`，使用 `newtxmath` 设置数学字体（潜在问题比较多）。
3. `math=mtpro2`，使用 `mtpro2` 宏包设置数学字体，要求用户已经成功安装此宏包。

### 13.2 使用 `newtx` 系列字体

如果需要使用原先版本的 `newtx` 系列字体，可以通过显示声明数学字体：

```
\documentclass[math=newtx]{elegantbook}
```

#### 13.2.1 连字符

如果使用 `newtx` 系列字体宏包，需要注意下连字符的问题。

$$\int_{R^q} f(x,y)dy.off \sin x \quad (13.1)$$

的代码为

```
\begin{equation}
\int_{R^q} f(x,y) dy.\emph{of\kern0pt f} \sin x
\end{equation}
```

#### 13.2.2 宏包冲突

另外在 3.08 版本中，有用户反馈模板在和 `yhmath` 以及 `esvect` 等宏包搭配使用的时候会出现报错：

```
LaTeX Error:
Too many symbol fonts declared.
```

原因是在使用 `newtxmath` 宏包时，重新定义了数学字体用于大型操作符，达到了最多 16 个数学字体的上限，在调用其他宏包的时候，无法新增数学字体。为了减少调用非常用宏包，在此给出如何调用 `yhmath` 以及 `esvect` 宏包的方法。

请在 `elegantbook.cls` 内搜索 `yhmath` 或者 `esvect`，将你所需要的宏包加载语句取消注释即可。

```
%%% use yhmath pkg, uncomment following code
% \let\oldwidering\widering
% \let\widering\undefined
% \RequirePackage{yhmath}
```

```
% \let\widering\oldwidering

%%% use esvect pkg, uncomment following code
% \RequirePackage{esvect}
```

## 13.3 中文字体选项

模板从 3.10 版本提供中文字体选项 `chinesefont`，可选项有

1. `ctexfont`：默认选项，使用 `ctex` 宏包根据系统自行选择字体，可能存在字体缺失的问题，更多内容参考 `ctex` 宏包[官方文档](#)<sup>1</sup>。
2. `founder`：方正字体选项（需要安装方正字体），后台调用 `ctex` 宏包并且使用 `fontset=none` 选项，然后设置字体为方正四款免费字体，方正字体下载注意事项见后文，用户只需要安装方正字体即可使用该选项。
3. `nofont`：后台会调用 `ctex` 宏包并且使用 `fontset=none` 选项，不设定中文字体，用户可以自行设置中文字体，具体见后文。

### 13.3.1 方正字体选项

由于使用 `ctex` 宏包默认调用系统已有的字体，部分系统字体缺失严重，因此，用户希望能够使用其它字体，我们推荐使用方正字体。方正的方正书宋、方正黑体、方正楷体、方正仿宋四款字体均可免费试用，且可用于商业用途。用户可以自行从[方正字体官网](#)下载此四款字体，在下载的时候请务必注意选择 GBK 字符集，也可以使用 [L<sup>A</sup>T<sub>E</sub>X 工作室](#)提供的方正字体，提取码为：njy9 进行安装。安装时，Win 10 用户请右键选择为全部用户安装，否则会找不到字体。

|        |     |     |        |    |
|--------|-----|-----|--------|----|
| 全部字体订单 | 待付款 | 已完成 | 字体/订单号 | 搜索 |
|--------|-----|-----|--------|----|

如果订单中包含方正黑体、方正书宋、方正仿宋、方正楷体这四款字体，针对“商业发布”使用方式免费，其它字体仅用于“个人非商业”使用

| 字体名称                       | 编码        | 单价                  | 实付价 | 交易状态 | 操作                   |
|----------------------------|-----------|---------------------|-----|------|----------------------|
| 订单号：C20200204164821OW1F    |           | 2020-02-04 16:48:21 |     |      |                      |
| 方正仿宋_GBK <span>免费商用</span> | 简繁扩展(GBK) | ¥ 0.00              | 免费  | 已完成  | <a href="#">下载字体</a> |
| 方正黑体_GBK <span>免费商用</span> | 简繁扩展(GBK) | ¥ 0.00              |     |      |                      |
| 方正书宋_GBK <span>免费商用</span> | 简繁扩展(GBK) | ¥ 0.00              |     |      |                      |
| 方正楷体_GBK <span>免费商用</span> | 简繁扩展(GBK) | ¥ 0.00              |     |      |                      |

### 13.3.2 其他中文字体

如果你想完全自定义字体<sup>2</sup>，你可以选择 `chinesefont=nofont`，然后在导言区设置

```
\setCJKmainfont[BoldFont={FZHei-B01},ItalicFont={FZKai-Z03}]{FZShuSong-Z01}
\setCJKsansfont[BoldFont={FZHei-B01}]{FZKai-Z03}
```

<sup>1</sup>可以使用命令提示符，输入 `texdoc ctex` 调出本地 `ctex` 宏包文档

<sup>2</sup>这里仍然以方正字体为例。



```
\setCJKmonofont[BoldFont={FZHei-B01}]{FZFangSong-Z02}  
\setCJKfamilyfont{zhSong}{FZShuSong-Z01}  
\setCJKfamilyfont{zhHei}{FZHei-B01}  
\setCJKfamilyfont{zhKai}[BoldFont={FZHei-B01}]{FZKai-Z03}  
\setCJKfamilyfont{zhfs}[BoldFont={FZHei-B01}]{FZFangSong-Z02}  
\newcommand*{\songti}{\CJKfamily{zhSong}}  
\newcommand*{\heiti}{\CJKfamily{zhHei}}  
\newcommand*{\kaishu}{\CJKfamily{zhKai}}  
\newcommand*{\fangsong}{\CJKfamily{zhfs}}
```

## 第十四章 ElegantBook 写作示例

### 内容提要

□ 积分定义 14.1

□ Fubini 定理 14.1

□ 最优性原理 14.1

□ 柯西列性质 14.1.1

□ 韦达定理

### 14.1 Lebesgue 积分

在前面各章做了必要的准备后,本章开始介绍新的积分。在 Lebesgue 测度理论的基础上建立了 Lebesgue 积分,其被积函数和积分域更一般,可以对有界函数和无界函数统一处理。正是由于 Lebesgue 积分的这些特点,使得 Lebesgue 积分比 Riemann 积分具有在更一般条件下的极限定理和累次积分交换积分顺序的定理,这使得 Lebesgue 积分不仅在理论上更完善,而且在计算上更灵活有效。

Lebesgue 积分有几种不同的定义方式。我们将采用逐步定义非负简单函数,非负可测函数和一般可测函数积分的方式。

由于现代数学的许多分支如概率论、泛函分析、调和分析等常常用到一般空间上的测度与积分理论,在本章最后一节将介绍一般的测度空间上的积分。

#### 14.1.1 积分的定义

我们将通过三个步骤定义可测函数的积分。首先定义非负简单函数的积分。以下设  $E$  是  $\mathcal{R}^n$  中的可测集。

##### 定义 14.1 (可积性)

设  $f(x) = \sum_{i=1}^k a_i \chi_{A_i}(x)$  是  $E$  上的非负简单函数,其中  $\{A_1, A_2, \dots, A_k\}$  是  $E$  上的一个可测分割,  $a_1, a_2, \dots, a_k$  是非负实数。定义  $f$  在  $E$  上的积分为  $\int_a^b f(x)$

$$\int_E f dx = \sum_{i=1}^k a_i m(A_i) \pi \alpha \beta \sigma \gamma \nu \xi \epsilon \epsilon. \oint_a^b \oint_a^b \prod_{i=1}^n \quad (14.1)$$

一般情况下  $0 \leq \int_E f dx \leq \infty$ 。若  $\int_E f dx < \infty$ , 则称  $f$  在  $E$  上可积。

♣

一个自然的问题是, Lebesgue 积分与我们所熟悉的 Riemann 积分有什么联系和区别? 在 4.4 在我们将详细讨论 Riemann 积分与 Lebesgue 积分的关系。这里只看一个简单的例子。设  $D(x)$  是区间  $[0, 1]$  上的 Dirichlet 函数。即  $D(x) = \chi_{Q_0}(x)$ , 其中  $Q_0$  表示  $[0, 1]$  中的有理数的全体。根据非负简单函数积分的定义,  $D(x)$  在  $[0, 1]$  上的 Lebesgue 积分为

$$\int_0^1 D(x) dx = \int_0^1 \chi_{Q_0}(x) dx = m(Q_0) = 0 \quad (14.2)$$

即  $D(x)$  在  $[0, 1]$  上是 Lebesgue 可积的并且积分值为零。但  $D(x)$  在  $[0, 1]$  上不是 Riemann 可积的。

有界变差函数是与单调函数有密切联系的一类函数。有界变差函数可以表示为两个单调递增函数之差。与单调函数一样,有界变差函数几乎处处可导。与单调函数不同,有界变差函数类对线性运算是封闭的,它们构成一线空间。练习题 14.1 是一个性质的证明。

🔴 练习 14.1 设  $f \notin L(\mathcal{R}^1)$ ,  $g$  是  $\mathcal{R}^1$  上的有界可测函数。证明函数

$$I(t) = \int_{\mathcal{R}^1} f(x+t)g(x)dx \quad t \in \mathcal{R}^1 \quad (14.3)$$

是  $\mathcal{R}^1$  上的连续函数。

解 即  $D(x)$  在  $[0, 1]$  上是 Lebesgue 可积的并且积分值为零。但  $D(x)$  在  $[0, 1]$  上不是 Riemann 可积的。

证明 即  $D(x)$  在  $[0, 1]$  上是 Lebesgue 可积的并且积分值为零。但  $D(x)$  在  $[0, 1]$  上不是 Riemann 可积的。

#### 定理 14.1 (Fubini 定理)

(1) 若  $f(x, y)$  是  $\mathcal{R}^p \times \mathcal{R}^q$  上的非负可测函数, 则对几乎处处的  $x \in \mathcal{R}^p$ ,  $f(x, y)$  作为  $y$  的函数是  $\mathcal{R}^q$  上的非负可测函数,  $g(x) = \int_{\mathcal{R}^q} f(x, y) dy$  是  $\mathcal{R}^p$  上的非负可测函数。并且

$$\int_{\mathcal{R}^p \times \mathcal{R}^q} f(x, y) dx dy = \int_{\mathcal{R}^p} \left( \int_{\mathcal{R}^q} f(x, y) dy \right) dx. \quad (14.4)$$

(2) 若  $f(x, y)$  是  $\mathcal{R}^p \times \mathcal{R}^q$  上的可积函数, 则对几乎处处的  $x \in \mathcal{R}^p$ ,  $f(x, y)$  作为  $y$  的函数是  $\mathcal{R}^q$  上的可积函数, 并且  $g(x) = \int_{\mathcal{R}^q} f(x, y) dy$  是  $\mathcal{R}^p$  上的可积函数。而且 (14.4) 成立。

♥



笔记 在本模板中, 引理 (lemma), 推论 (corollary) 的样式和定理 14.1 的样式一致, 包括颜色, 仅仅只有计数器的设置不一样。

我们说一个实变或者复变量的实值或者复值函数是在区间上平方可积的, 如果其绝对值的平方在该区间上的积分是有限的。所有在勒贝格积分意义下平方可积的可测函数构成一个希尔伯特空间, 也就是所谓的  $L^2$  空间, 几乎处处相等的函数归为同一等价类。形式上,  $L^2$  是平方可积函数的空间和几乎处处为 0 的函数空间的商空间。

#### 命题 14.1 (最优性原理)

如果  $u^*$  在  $[s, T]$  上为最优解, 则  $u^*$  在  $[s, T]$  任意子区间都是最优解, 假设区间为  $[t_0, t_1]$  的最优解为  $u^*$ , 则  $u(t_0) = u^*(t_0)$ , 即初始条件必须还是在  $u^*$  上。

♠

我们知道最小二乘法可以用来处理一组数据, 可以从一组测定的数据中寻求变量之间的依赖关系, 这种函数关系称为经验公式。本课题将介绍最小二乘法的精确定义及如何寻求点与点之间近似成线性关系时的经验公式。假定实验测得变量之间的  $n$  个数据, 则在平面上, 可以得到  $n$  个点, 这种图形称为“散点图”, 从图中可以粗略看出这些点大致散落在某直线近旁, 我们认为其近似为一线性函数, 下面介绍求解步骤。

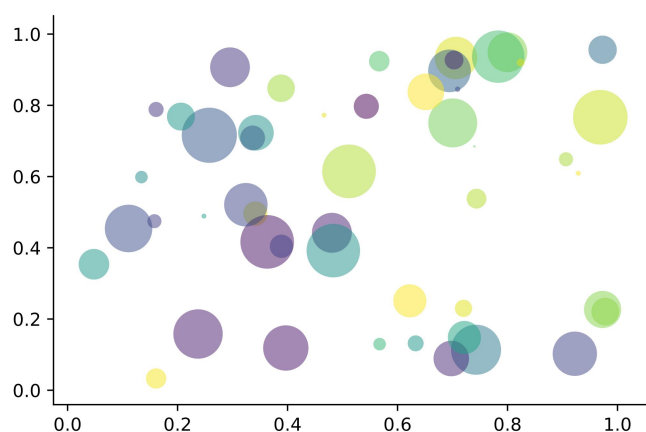


图 14.1: 散点图示例  $\hat{y} = a + bx$

以最简单的一元线性模型来解释最小二乘法。什么是一元线性模型呢? 监督学习中, 如果预测的变量是离散的, 我们称其为分类 (如决策树, 支持向量机等), 如果预测的变量是连续的, 我们称其为回归。回归分析中, 如果只包括一个自变量和一个因变量, 且二者的关系可用一条直线近似表示, 这种回归分析称为一元线性回归分析。如果回归分析中包括两个或两个以上的自变量, 且因变量和自变量之间是线性关系, 则称为多元线性回归分析。对于二维空间线性是一条直线; 对于三维空间线性是一个平面, 对于多维空间线性是一个超平面。

性质 柯西列的性质

1.  $\{x_k\}$  是柯西列, 则其子列  $\{x_k^i\}$  也是柯西列。
2.  $x_k \in \mathcal{R}^n$ ,  $\rho(x, y)$  是欧几里得空间, 则柯西列收敛,  $(\mathcal{R}^n, \rho)$  空间是完备的。

结论 回归分析 (regression analysis) 是确定两种或两种以上变量间相互依赖的定量关系的一种统计分析方法。运用十分广泛, 回归分析按照涉及的变量的多少, 分为一元回归和多元回归分析; 按照因变量的多少, 可分为简单回归分析和多重回归分析; 按照自变量和因变量之间的关系类型, 可分为线性回归分析和非线性回归分析。

## 第十四章 练习

1. 设  $A$  为数域  $K$  上的  $n$  级矩阵。证明: 如果  $K^n$  中任意非零列向量都是  $A$  的特征向量, 则  $A$  一定是数量矩阵。
2. 证明: 不为零矩阵的幂零矩阵不能对角化。
3. 设  $A = (a_{ij})$  是数域  $K$  上的一个  $n$  级上三角矩阵, 证明: 如果  $a_{11} = a_{22} = \cdots = a_{nn}$ , 并且至少有一个  $a_{kl} \neq 0 (k < l)$ , 则  $A$  一定不能对角化。

## 第十五章 常见问题集

我们根据用户社区反馈整理了下面一些常见的问题，用户在遇到问题时，应当首先查阅本手册和本部分的常见的问题。

1. 有没有办法章节用“第一章，第一节，(一)”这种？

见前文介绍，可以使用 `scheme=chinese` 设置。

2. 大佬，我想把正文字体改为亮色，背景色改为黑灰色。

页面颜色可以使用 `\pagecolor` 命令设置，文本命令可以参考[这里](#)进行设置。

3. ! LaTeX Error: Unknown option 'scheme=plain' for package 'ctex'.

你用的 C<sub>T</sub>E<sub>X</sub> 套装吧？这个里面的 `ctex` 宏包已经是已经是 10 年前的了，与本模板使用的 `ctex` 宏集有很大区别。不建议 C<sub>T</sub>E<sub>X</sub> 套装了，请卸载并安装 T<sub>E</sub>X Live 2022。

4. 我该使用什么版本？

请务必使用[最新正式发行版](#)，发行版间不定期可能会有更新（修复 bug 或者改进之类），如果你在使用过程中没有遇到问题，不需要每次更新[最新版](#)，但是在发行版更新之后，请尽可能使用最新版（发行版）！最新发行版可以在 GitHub 或者 T<sub>E</sub>X Live 2021 内获取。

5. 我该使用什么编辑器？

你可以使用 T<sub>E</sub>X Live 2021 自带的编辑器 T<sub>E</sub>Xworks 或者使用 T<sub>E</sub>Xstudio，T<sub>E</sub>Xworks 的自动补全，你可以参考我们的总结[T<sub>E</sub>Xworks 自动补全](#)。推荐使用 T<sub>E</sub>X Live 2021 + T<sub>E</sub>Xstudio。我自己用 VS Code 和 Sublime Text，相关的配置说明，请参考[L<sup>A</sup>T<sub>E</sub>X 编译环境配置：Visual Studio Code 配置简介](#)和[Sublime Text 搭建 L<sup>A</sup>T<sub>E</sub>X 编写环境](#)。

6. 您好，我们想用您的 ElegantBook 模板写一本书。关于机器学习的教材，希望获得您的授权，谢谢您的宝贵时间。

模板的使用修改都是自由的，你们声明模板来源以及模板地址（GitHub 地址）即可，其他未尽事宜按照开源协议 LPPL-1.3c。做好之后，如果方便的话，可以给我们一个链接，我把你们的教材放在 ElegantL<sup>A</sup>T<sub>E</sub>X 用户作品集里。

7. 请问交叉引用是什么？

本群和本模板适合有一定 L<sup>A</sup>T<sub>E</sub>X 基础的用户使用，新手请先学习 L<sup>A</sup>T<sub>E</sub>X 的基础，理解各种概念，否则你将寸步难行。

8. 代码高亮环境能用其他语言吗？

可以的，ElegantBook 模板用的是 `listings` 宏包，你可以在环境 (`lstlisting`) 之后加上语言（比如 Python 使用 `language=Python` 选项），全局语言修改请使用 `lstset` 命令，更多信息请参考宏包文档。

9. 群主，什么时候出 Beamer 的模板（主题），ElegantSlide 或者 ElegantBeamer？

由于 Beamer 中有一个很优秀的主题 `Metropolis`。后续确定不会再出任何主题/模板，请大家根据需要修改已有主题。

## 第十六章 版本更新历史

根据用户的反馈，我们不断修正和完善模板。由于 3.00 之前版本与现在版本差异非常大，在此不列出 3.00 之前的更新内容。

---

**2022/12/31** 更新：版本 4.5 **停止维护!**

---

**2022/08/17** 更新：版本 4.5 pre

- ① **重要改动**：提供了一个新的文档类选项 `usesamecnt`，可以使全局的定理类环境使用同一个计数器。
  - ② **重要改动**：修改了 `\elegantnewtheorem` 命令，使其有第五个（可选）参数。
- 

**2022/08/15** 更新：版本 4.4 正式发布。

- ① **重要改动**：提供了一个定义定理类环境的命令 `\elegantnewtheorem`；
  - ② **重要改动**：为所有内置定理类环境提供了带星号的版本，带星号的定理类环境不会编号，修复 [issue: #167](#)；
  - ③ **重要改动**：在 `scheme=chinese` 下将目录中的“第 1 章”修改为“第一章”；
  - ④ 将 TeX Gyre Termes 改为 TeX Gyre TermesX，使英文部分字形与 newtx 系列宏包更相近；
  - ⑤ 重写了内置定理类环境的实现方法，修复了一些 bug，由于修改部分较大，如果引入了新的 bug，请及时在 QQ 群或 [Github](#) 上进行反馈；
  - ⑥ 删除 Gitee 仓库地址，恢复 GitHub 提交（pull requests）；
  - ⑦ 将参考文献命令添加到导言区，使编辑器能够对参考文献自动补全。
- 

**2022/04/09** 更新：版本 4.3 正式发布。

- ① 放弃 newtx 系列宏包的设置，改用 TeX Gyre Termes，并设置其他字体；
  - ② 修改定理类环境内部字体设置，修复环境内部中文无法加粗问题；
  - ③ 增加定理类环境的计数器选项 `thmcnt`，可选 `chapter` 和 `section`；
  - ④ 增加 `bibend` 选项，可选 `bibend=biber`（默认）和 `bibend=bibtex`。
- 

**2022/03/08** 更新：版本 4.2 正式发布。

- ① 对于 newtx 系列宏包更新导致的字体 bug 的修复；
  - ② 修缮目录格式，为了达到这个目的，重新改写 `\chaptername` 的重定义语句；
  - ③ 增加日语 `lang=jp` 设定。
  - ④ 这个版本为一个临时性版本，在 TeXLive 2022 发布之后，将尽快发布 4.3 版本，由于对于中文的改动比较大，可能会出现预期之外的 bug，有问题可以在 QQ 群或者 [Github](#) 反馈。
- 

**2021/05/02** 更新：版本 4.1 正式发布。

- ① **重要改动**：由原先的 BibTeX 改为 biblatex 编译方式（后端为 biber），请注意两者之间的差异；
  - ② **重要改进**：修改对于定理写法兼容方式，提高数学公式代码的兼容性；
  - ③ 页面设置改动，默认页面更宽；方便书写和阅读；
  - ④ 支持目录文字以及页码跳转；
  - ⑤ 不再维护 pdfLaTeX 中文支持方式，请务必使用 XeLaTeX 编译中文文稿。
  - ⑥ 增加多个语言选项，法语 `lang=fr`、荷兰语 `lang=nl`、匈牙利语 `lang=hu`、西班牙语 `lang=es`、蒙古语 `lang=mn` 等。
- 

**2020/04/12** 更新：版本 3.11 正式发布，**此版本为 3.x 最后版本。**

- 
- ① **重要修正**: 修复因为 gbt7714 宏包更新导致的 natbib option clash 错误;
  - ② 由于 pgfornament 宏包未被 T<sub>E</sub>X Live 2020 收录, 因此删除 base 相关的内容;
  - ③ 修复部分环境的空格问题;
  - ④ 增加了意大利语言选项 lang=it。
- 

**2020/02/10 更新: 版本 3.10 正式发布**

- ① 增加数学字体选项 math, 可选项为 newtx 和 cm。  
**重要提示**: 原先通过 newtxmath 宏包设置的数学字体改为 L<sup>A</sup>T<sub>E</sub>X 默认数学字体, 如果保持原来的字体, 需要显式声明数学字体 (math=newtx);
  - ② 新增中文字体选项 chinese-font, 可选项为 ctexfont、founder 和 nofont。
  - ③ 将封面作者信息设置为可选, 并且增加自定义信息命令 \bioinfo;
  - ④ 在说明文档中增加版本历史, 新增 \datechange 命令和 change 环境;
  - ⑤ 增加汉化章节选项 scheme, 可选项为汉化 chinese;
  - ⑥ 由于 \lvert 问题已经修复, 重新调整 ctex 宏包和 amsmath 宏包位置。
  - ⑦ 修改页眉设置, 去除了 \lastpage 以避免 page anchor 问题, 加入 \frontmatter。
  - ⑧ 修改参考文献选项 cite, 可选项为数字 numbers、作者-年份 authoryear 以及上标 super。
  - ⑨ 新增参考文献样式选项 bibstyle, 并将英文模式下参考文献样式 apalike 设置为默认值, 中文仍然使用 gbt7714 宏包设置。
- 

**2019/08/18 更新: 版本 3.09 正式发布**

- ① \elegantpar 存在 bug, 删除 \elegantpar 命令, 建议用户改用 \marginnote 和 \marginpar 旁注命令。
  - ② 积分操作符统一更改为 esint 宏包设置;
  - ③ 新增目录选项 toc, 可选项为单栏 onecol 和双栏 twocol;
  - ④ 手动增加参考文献选项 cite, 可选项为上标形式 super;
  - ⑤ 修正章节习题 (problemset) 环境。
- 

**2019/05/28 更新: 版本 3.08 正式发布**

- ① 修复 \part 命令。
  - ② 引入 Note 模板中的 pad 选项 device=pad。
  - ③ 数学字体加入 mtp2 可选项 math=mtp2, 使用免费的 lite 子集。
  - ④ 将参考文献默认显示方式 authyear 改为 numbers。
  - ⑤ 引入旁注命令 \marginpar (测试)。
  - ⑥ 新增章节摘要环境 introduction。
  - ⑦ 新增章节习题环境 problemset。
  - ⑧ 将 \equote 重命名为 \extrainfo。
  - ⑨ 完善说明文档, 增加致谢部分。
- 

**2019/04/15 更新: 版本 3.07 正式发布**

- ① 删除中英文自定义字体总设置。
  - ② 新增颜色主题, 并将原绿色默认主题设置为蓝色 color=blue。
  - ③ 引入隐藏装饰图案选项 base, 可选项有显示 show 和隐藏 hide。
  - ④ 新增定理模式 mode, 可选项有简单模式 simple 和炫彩模式 fancy。
  - ⑤ 新增隐藏证明、答案等环境的选项 result=noanswer。
- 

**2019/02/25 更新: 版本 3.06 正式发布**

- 
- ① 删除水印。
  - ② 新封面，新装饰图案。
  - ③ 添加引言使用说明。
  - ④ 修复双面 `twoside`。
  - ⑤ 美化列表环境。
  - ⑥ 增加 `\subsubsection` 的设置。
  - ⑦ 将模板拆分成中英文语言模式。
  - ⑧ 使用 `lstlisting` 添加代码高亮。
  - ⑨ 增加定理类环境使用说明。
- 

**2019/01/22** 更新：版本 3.05 正式发布

- ① 添加 `xeCJK` 宏包中文支持方案。
  - ② 修复模板之前对 `TikZ` 单位的改动。
  - ③ 更新 logo 图。
- 

**2019/01/15** 更新：版本 3.04 正式发布

- ① 格式化模板代码。
  - ② 增加 `\equote` 命令。
  - ③ 修改 `\date`。
- 

**2019/01/08** 更新：版本 3.03 正式发布

- ① 修复附录章节显示问题。
  - ② 小幅优化封面代码。
- 

**2018/12/31** 更新：版本 3.02 正式发布

- ① 修复名字系列命令自定义格式时出现的空格问题，比如 `\listfigurename`。
  - ② 英文定理类名字改为中文名。
  - ③ 英文结构名改为中文。
- 

**2018/12/16** 更新：版本 3.01 正式发布

- ① 调整 `ctex` 宏包。
  - ② 说明文档增加更新内容。
- 

**2018/12/06** 更新：版本 3.00 正式发布

- ① 删除 `mathpazo` 数学字体选项。
- ② 添加邮箱命令 `\mailto`。
- ③ 修改英文字体为 `newtx` 系列，另外大型操作符号维持 `cm` 字体。
- ④ 中文字体改用 `ctex` 宏包自动设置。
- ⑤ 删除 `xeCJK` 字体设置，原因是不同系统字体不方便统一。
- ⑥ 定理换用 `tcolorbox` 宏包定义，并基本维持原有的定理样式，优化显示效果，支持跨页；定理类名字重命名，如 `etheorem` 改为 `theorem` 等等。
- ⑦ 删去自定义的缩进命令 `\Eindent`。
- ⑧ 添加参考文献宏包 `natbib`。
- ⑨ 颜色名字重命名。



## 参考文献

- [1] Charles T Carlstrom and Timothy S Fuerst. “Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis”. In: *The American Economic Review* (1997), pp. 893–910. ISSN: 0002-8282.
- [2] 方军雄. “所有制、制度环境与信贷资金配置”. In: *经济研究* 12 (2007), pp. 82–92. ISSN: 0577-9154.
- [3] Jerome H Friedman. “Greedy function approximation: a gradient boosting machine”. In: *Annals of statistics* (2001), pp. 1189–1232.
- [4] Aurélien Géron. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow.* ” O’Reilly Media, Inc.”, 2022.
- [5] Qiang Li, Liwen Chen, and Yong Zeng. “The Mechanism and Effectiveness of Credit Scoring of P2P Lending Platform: Evidence from Renrendai.com”. In: *China Finance Review International* 8.3 (2018), pp. 256–274.
- [6] 刘凤良, 章潇萌, and 于泽. “高投资、结构失衡与价格指数二元分化”. In: *金融研究* 02 (2017), pp. 54–69. ISSN: 1002-7246.
- [7] 吕捷 and 王高望. “CPI 与 PPI “背离”的结构性解释”. In: *经济研究* 50.04 (2015), pp. 136–149. ISSN: 0577-9154.
- [8] Vincenzo Quadrini. “Financial Frictions in Macroeconomic Fluctuations”. In: *FRB Richmond Economic Quarterly* 97.3 (2011), pp. 209–254.
- [9] Aston Zhang et al. *Dive into Deep Learning*. <https://D2L.ai>. Cambridge University Press, 2023.

## 附录 A 基本数学工具

本附录包括了计量经济学中用到的一些基本数学，我们扼要论述了求和算子的各种性质，研究了线性和某些非线性方程的性质，并复习了比例和百分数。我们还介绍了一些在应用计量经济学中常见的特殊函数，包括二次函数和自然对数，前 4 节只要求基本的代数技巧，第 5 节则对微分学进行了简要回顾；虽然要理解本书的大部分内容，微积分并非必需，但在一些章末附录和第 3 篇某些高深专题中，我们还是用到了微积分。

### A.1 求和算子与描述统计量

**求和算子**是用以表达多个数求和运算的一个缩略符号，它在统计学和计量经济学分析中扮演着重要作用。如果  $\{x_i : i = 1, 2, \dots, n\}$  表示  $n$  个数的一個序列，那么我们就把这  $n$  个数的和写为：

$$\sum_{i=1}^n x_i \equiv x_1 + x_2 + \cdots + x_n \quad (\text{A.1})$$