



链滴

# 统计学习基础 (ESL) 中文版

作者: [wizardforcel](#)

原文链接: <https://ld246.com/article/1543139716206>

来源网站: 链滴

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

译者: [szcf-weiya](#)

ESL 指的是 [The Elements of Statistical Learning](#)。因为 (译者) 自己也是统计学专业, 所以想研读本书, 同时实现书中的算法及其例子, 并尝试解决习题。

## 说明

1. 参考文献保留原书的写法, 如 “Efron and Tibshirani (1993)” 指的是 “Efron, B. and Tibshirani R. (1993). An Introduction to the Bootstrap, Chapman and Hall, London.” ; 并且为了查阅方便, 特别在每章的 “Bibliographic Notes” 中用脚注列出了相应的参考文献, 查阅过的文献, 还附上下载链接。
2. 该项目不仅仅是单纯地翻译原书, 也在实现书中的一些算法或者例子。这些代码都可以在 [这里](#) 找, 也可以在 Issues 中找到部分习题的解答。
3. 限于 (译者) 本人英语水平以及统计知识, 很多地方可能翻译不到位, 欢迎大家指出来, Issue 或评论均可, 错别字以及数学公式解析错误的也可以指出。

## 目录

- [主页](#)
- [序言](#)
  - [第二版序言](#)
  - [第一版序言](#)
- [1 简介](#)
  - [1.1 导言](#)
- [2 监督学习概要](#)
  - [2.1 导言](#)
  - [2.2 变量类型和术语](#)
  - [2.3 两种预测的简单方法](#)
  - [2.4 统计判别理论](#)
  - [2.5 高维问题的局部方法](#)
  - [2.6 统计模型, 监督学习和函数逼近](#)
  - [2.7 结构化的回归模型](#)
  - [2.8 限制性估计的种类](#)
  - [2.9 模型选择和偏差-方差的权衡](#)
  - [文献笔记](#)
- [3 回归的线性方法](#)
  - [3.1 导言](#)
  - [3.2 线性回归模型和最小二乘法](#)
  - [3.3 子集的选择](#)

- 3.4 收缩的方法
  - 3.5 运用派生输入方向的方法
  - 3.6 选择和收缩方法的比较
  - 3.7 多重输出的收缩和选择
  - 3.8 Lasso 和相关路径算法的补充
  - 3.9 计算上的考虑
  - 文献笔记
- 4 分类的线性方法
  - 4.1 引言
  - 4.2 指示矩阵的线性回归
  - 4.3 线性判别分析
  - 4.4 逻辑斯蒂回归
  - 4.5 分离超平面
  - 文献笔记
- 5 基展开和正规化
  - 5.1 引言
  - 5.2 分段多项式和样条
  - 5.3 滤波和特征提取
  - 5.4 光滑样条
  - 5.5 光滑参数的自动选择
  - 5.6 非参逻辑斯蒂回归
  - 5.7 多维样条
  - 5.8 正则化和再生核希尔伯特空间理论
  - 5.9 小波光滑
  - 文献笔记
  - 附录-B 样条的计算
- 6 核光滑方法
  - 6.0 引言
  - 6.1 一维核光滑器
  - 6.2 选择核的宽度
  - 6.3  $\mathbb{R}^p$ 中的局部回归
  - 6.4  $\mathbb{R}^p$ 中的结构化局部回归模型
  - 6.5 局部似然和其他模型
  - 6.6 核密度估计和分类
  - 6.7 径向基函数和核
  - 6.8 混合模型的密度估计和分类

- 6.9 计算上的考虑
  - 文献笔记
- 7 模型评估及选择
  - 7.1 引言
  - 7.2 偏差, 方差和模型复杂度
  - 7.3 偏差-方差分解
  - 7.4 测试误差率的 optimism
  - 7.5 样本内预测误差的估计
  - 7.6 参数的有效个数
  - 7.7 贝叶斯方法和 BIC
  - 7.8 最小描述长度
  - 7.9 VC 维
  - 7.10 交叉验证
  - 7.11 自助法
  - 7.12 条件测试误差或期望测试误差
  - 文献笔记
- 8 模型推断和平均
  - 8.1 引言
  - 8.2 自助法和最大似然法
  - 8.3 贝叶斯方法
  - 8.4 自助法和贝叶斯推断之间的关系
  - 8.5 EM 算法
  - 8.6 从后验分布采样的 MCMC
  - 8.7 袋装法
  - 8.8 模型平均和堆栈
  - 8.9 随机搜索
  - 文献笔记
- 9 增广模型, 树, 以及相关方法
  - 9.0 引言
  - 9.1 广义可加模型
  - 9.2 基于树的方法
  - 9.3 PRIM
  - 9.4 多变量自适应回归样条
  - 9.5 专家的分层混合
  - 9.6 缺失数据
  - 9.7 计算上的考虑

- 文献笔记
- 10 增强和可加树
  - 10.1 boosting 方法
  - 10.2 boosting 拟合可加模型
  - 10.3 向前逐步加性建模
  - 10.4 指数损失和 AdaBoost
  - 10.5 为什么是指数损失
  - 10.6 损失函数和鲁棒性
  - 10.7 数据挖掘的现货方法
  - 10.8 垃圾邮件的例子
  - 10.9 boosting 树
  - 10.10 Gradient Boosting 的数值优化
  - 10.11 大小合适的 boosting 树
  - 10.12 正则化
  - 文献笔记
- 11 神经网络
  - 11.1 导言
  - 11.2 投影寻踪回归
  - 11.3 神经网络
  - 11.4 拟合神经网络
  - 11.5 训练神经网络的一些问题
  - 11.6 模拟数据的例子
  - 11.7 邮编数字的例子
  - 文献笔记
- 12 支持向量机和灵活的判别方法
  - 12.1 导言
  - 12.2 支持向量分类器
  - 12.3 支持向量机和核
  - 12.4 广义线性判别分析
  - 12.5 FDA
  - 12.6 PDA
  - 12.7 混合判别分析
  - 计算上的考虑
  - 文献笔记
- 13 原型方法和最近邻

- 13.1 引言
  - 13.2 原型方法
  - 13.3 k 最近邻分类器
  - 13.4 自适应的最近邻方法
  - 13.5 计算上的考虑
  - 文献笔记
- 14 非监督学习
  - 14.1 引言
  - 14.2 关联规则
  - 14.3 聚类分析
  - 14.4 自组织图
  - 14.5 主成分, 主曲线以及主曲面
  - 14.6 非负矩阵分解
  - 14.7 独立成分分析和探索投影寻踪
  - 14.8 多维缩放
  - 14.9 非线性降维和局部多维缩放
  - 14.10 谷歌的 PageRank 算法
  - 文献笔记
- 15 随机森林
  - 15.1 引言
  - 15.2 随机森林的定义
  - 15.3 随机森林的细节
  - 15.4 随机森林的分析
  - 文献笔记
- 16 集成学习
  - 16.1 引言
  - 16.2 增强和正则路径
  - 16.3 学习集成
  - 文献笔记
- 17 无向图模型
  - 17.1 引言
  - 17.2 马尔科夫图及其性质
  - 17.3 连续变量的无向图模型
  - 17.4 离散变量的无向图模型
  - 文献笔记

- 18 高维问题
  - 18.1 当  $p$  大于  $N$
  - 18.2 对角线性判别分析和最近收缩重心
  - 18.3 二次正则的线性分类器
  - 18.4 一次正则的线性分类器
  - 18.5 当特征不可用时的分类
  - 18.6 有监督的主成分
  - 18.7 特征评估和多重检验问题
  - 文献笔记