

# 入门知识准备

唐珂

# 说明

- 本文档主要面向即将读研（硕士或博士）的同学，属于在正式开展具体科研/工程工作之前的前期准备与建议。
- 文档中覆盖的知识点或技能，若同学已经具备，则可跳过。
- 应在半年内完成本文档所述事宜，越早越好，但不能操之过急。

# 阶段1：数学基础知识

- 简要复习线性代数、高数/计算方法、概率论知识。**对以下知识点尽量深入理解：**
- 线性代数
  - 矩阵求逆
  - 特征值与特征向量、特征值分解
  - 矩阵的秩、迹
  - 正定、半正定矩阵
  - 线性相关、正交
- 高数/计算方法，必要时可参考一些数值优化的书（Numerical Optimization）
  - 导数、链式法则
  - 拉格朗日法
  - 梯度下降
  - 二次规划（问题）的求解算法（注意其与梯度下降、拉格朗日法的关联）

# 阶段1：数学基础知识

- 概率与统计
  - 贝叶斯定理
  - 先验、后验
  - 概率分布、概率密度函数（以正态分布、伯努利分布为例）
  - 最大似然估计
  - 大数定理
  - 中心极限定理
  - 关于期望/均值/方差的若干概念与定理
  - 统计检验（包括参数检验、非参检验及具体方法）
  - 线性回归

# 阶段1：计算机基础知识

- 最好能使用1-2门编程语言，C++/java/python等均可。
- 简要复习算法设计与分析、人工智能，**对以下知识点尽量深入理解：**
  - 参考书：《Introduction to Algorithms》、《Artificial Intelligence: A Modern Approach》
- 算法设计与分析
  - 时间复杂度的定义及其内涵
  - 贪心算法与动态规划的基本思想
  - NP-Completeness、NP-Hard问题及其代表性例子（可参考wiki）
  - NP-hard问题的精确与近似算法（可参考wiki）
- 人工智能
  - 树搜索、启发式搜索
  - NPC/NP-hard问题与启发式搜索的关系
  - 逻辑推理及其与搜索的关系

## 阶段2：领域知识储备-机器学习

- 阅读相关书籍，上公开课，建立对领域的通盘认识，应覆盖《智能数据分析》课程的主要知识点。
  - 换言之，如果已经上了这门课，学得也还不错，储备期就可以算结束了。
  - 如果还没有上过这门课，可以自己找组里师兄、师姐们了解一下课程内容，自己在网上上一下其他的公开课，也行。
- 储备结束后，应该能使用常用的机器学习工具（比如tensorflow、pytorch之类，解决一个具体问题）。

## 阶段2：领域知识储备-演化计算

- 可以找1-2本参考书读一读，例如《Evolutionary Computation – A unified approach》、《Introduction to Evolutionary Computing》
- 演化计算涉及到的知识面更宽，不能只是看书，还应自己找资料/论文了解以下知识：
  - 最优化、约束优化、多目标优化的概念
  - 模拟退火
  - 禁忌搜索 (Tabu Search)
  - 遗传算法
  - 粒子群优化、差分演化
  - 分布估计算法 (Estimation of Distribution Algorithm)
  - Evolution Strategies, Natural Evolution Strategies, CMA-ES
  - 贝叶斯优化
  - 复习经典启发式算法
  - 如果有余力，应该了解一些经典运筹学、凸优化的知识（商学院好像开过类似课程）。
- 思考上述几种方法之间的关联和演进

## 阶段3：实践训练

- 指导老师提供2-3篇经典论文，同学实现论文中的方法及所有实验结果，并以ppt形式汇报。