

建模模型概念

线性规划

目标函数及约束条件均为线性函数，故被称为线性规划问题。线性规划问题是在一组线性约束条件的限制下,求一线性目标函数最大或最小的问题。

案例：某机床厂生产甲、乙两种机床，每台销售后的利润分别为4千元与3千元。生产甲机床需用A、B机器加工，加工时间分别为每台2小时和1小时;生产乙机床需用A、B、C三种机器加工,加工时间为每台各一小时。若每天可用于加工的机器时数分别为A机器10小时、B机器8小时和C机器7小时，求该厂应生产甲、乙机床各几台，才能使总利润最大。

标准形式:

$$\begin{aligned} & \min c^T x \\ & x \\ & x.t. \begin{cases} Ax \leq b \\ Aeq \cdot x = beq \\ lb \leq x \leq ub \end{cases} \end{aligned}$$

整数规划

数学规划中的变量（部分或全部）限制为整数时，成为整数规划。若在线性规划模型中，变量限制为整数，则称为整数线性规划。目前所流行的求解整数规划的方法，往往只适用于整数线性规划。目前还没有一种方法 能有效地求解一切整数规划。

$$\begin{cases} \text{纯（完全）整数规划} \\ \text{混合整数规划} \end{cases}$$

特点：

原线性规划有最优解，当自变量限制为整数 后，其整数规划解出现下述情况

- 1. 原线性规划最优解全是整数，则整数规划最优 解与线性规划最优解一致
- 2. 整数规划无可行解
- 3. 有可行解（当然就存在最优解），但最优解值变差

整数规划最优解不能按照实数最优解（例如线性规划求解）简单取整而获得。

案例：

设用某型号的圆钢下零件，A2，.....，Am的毛胚。在一根圆钢上下料的方式有B1，B2，.....，Bn种。每种下料方式可以得到各种零件的毛胚数以及每种零件的需要量，如表所示。问怎样安排下料方式，使得即满足需要，所用的原材料又最少。

零件个数	B1.....Bn (方式)	零件毛胚数
(零件)		
A1	a11.....a1n	b1
.....
Am	am1.....amn	bm

一般形式：

$$\begin{aligned} \max(\min) z &= \sum_{j=1}^n c_j x_j \\ s.t. \quad &\begin{cases} \sum_{j=1}^n a_{ij} x_j \leq (=, \geq) b_i \quad (i = 1, 2, \dots, n) \\ x_j \geq 0, x_j \text{ 为整数} \quad (j = 1, 2, \dots, n) \end{cases} \end{aligned}$$

1. 纯整数规划：所有决策变量要求取非负整数（这时引进的松弛变量和剩余变量可以不要求取整数）。
松弛变量： $x_1 + x_2 \leq 10, x_1 + x_2 + x_3 = 10, x_3 \geq 0, x_3$ 为松弛变量
剩余变量： $x_1 + x_2 \geq 10, x_1 + x_2 - x_3 = 10, x_3 \leq 0, -x_3$ 为剩余变量
2. 全整数规划：除了所有决策变量要求取非负整数外，系数 a_{ij} 和常数 b_i 也要求取整数（这时引进的松弛变量和剩余变量也必须是整数）。
3. 混合整数规划：只有一部分的决策变量要求取非负整数，另一部分可以取非负实数。
4. 0-1整数规划：所有决策变量只能取0或1两个整数。（工作安排问题）

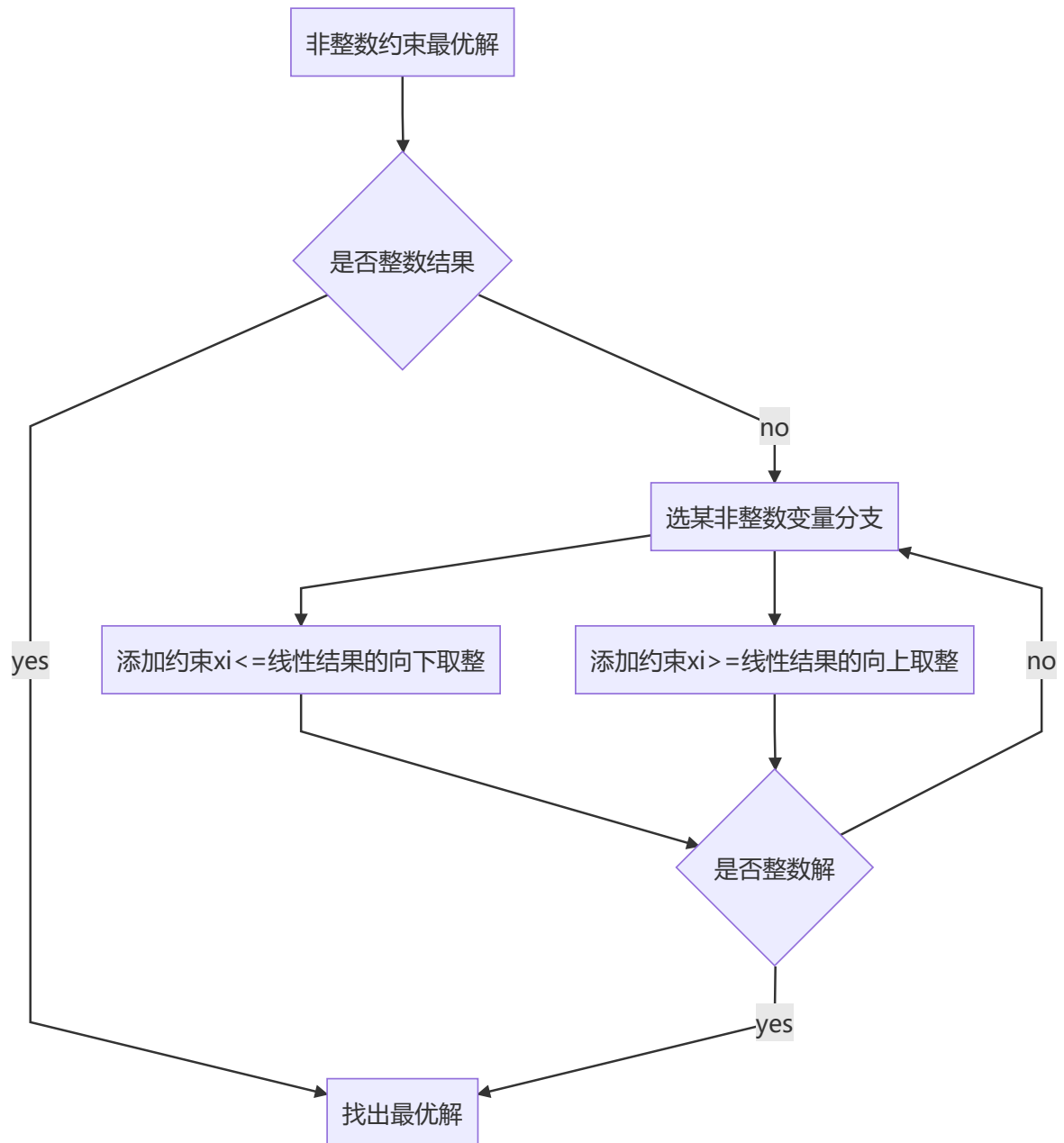
整数规划可行解是线性规划可行域中的整数格点

线性规划无可行解，则整数规划无可行解

ILP最优解小于或等于线性规划的最优值

线性规划最优解满足整数要求，则该最优解为整数规划的最优解

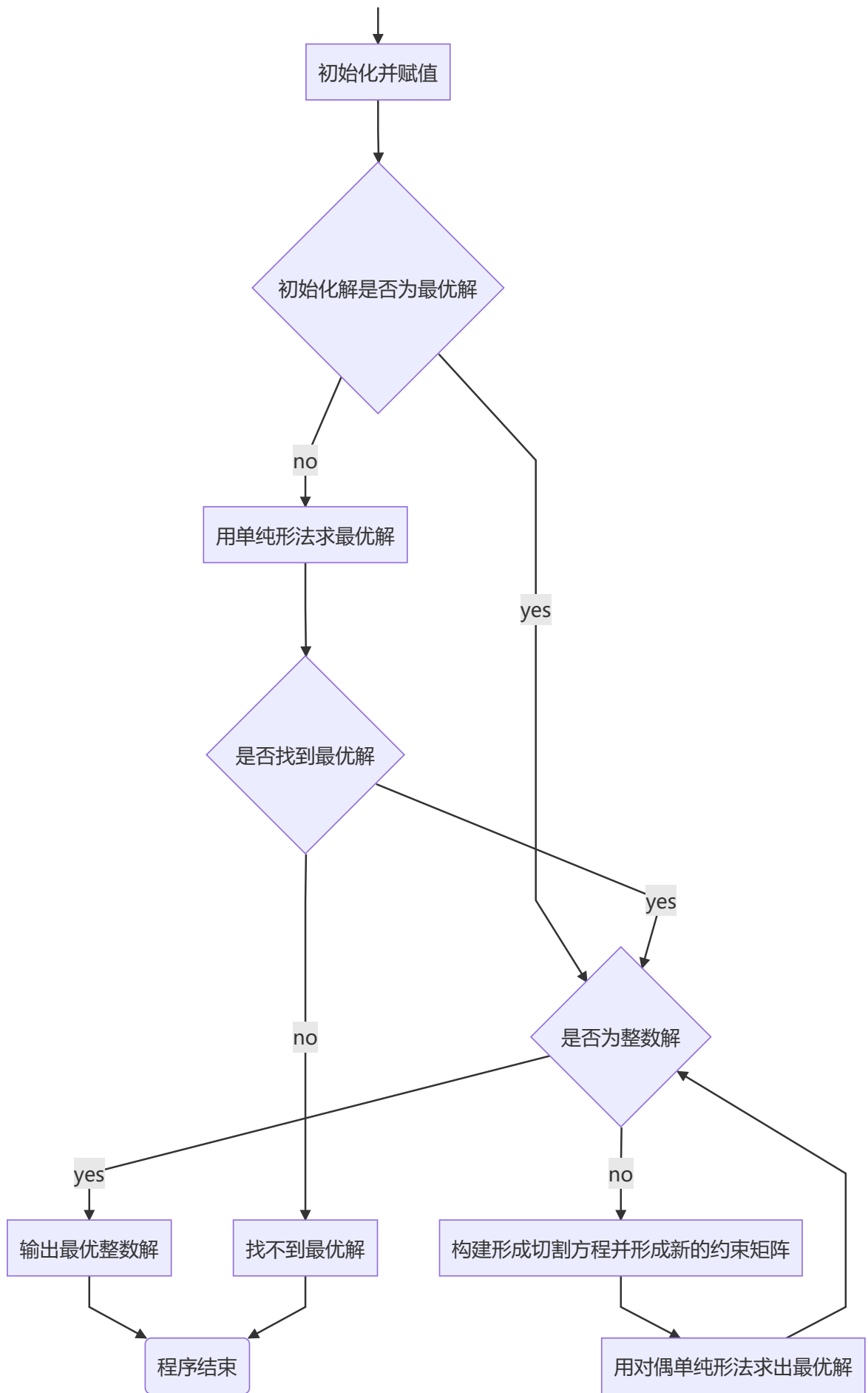
分支定界算法



割平面算法

即对 (P_0) 增加一个线性约束，将 (P_0) 的可行区域割掉一块，使得非整数解恰好在割掉的一块中，但是又没有割掉原问题 (P) 的可行解，得到问题 (P_1) ，重复上述的过程。

开始



匈牙利算法

0-1规划问题

典型：指派问题

所谓的指派问题就比如：甲乙丙三个人去做ABC三件事情。每个人做每件事情所花的时间可能不一样。每个人只能安排一件事情，问怎样安排才能使三个人所工作的时间之和最小？

扩展成 n 个人 n 件事也可以，但要求是：事情数和人数一样多，并且每人只能做一件事。

数学模型：

$$x_{ij} = \begin{cases} 1 & \text{第 } i \text{ 个人做第 } j \text{ 项工作} \\ 0 & \text{第 } i \text{ 个人不做第 } j \text{ 项工作} \end{cases} (i, j = 1, \dots, n)$$

$$X = (x_{ij})_{n \times n}$$

$$\min z = \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij}$$

$$s. t. \begin{cases} \sum_{j=1}^n x_{ij} = 1 (i = 1, \dots, n) \\ \sum_{i=1}^n x_{ij} = 1 (j = 1, \dots, n) \\ x_{ij} = 0 \text{ 或 } 1 (i, j = 1, \dots, n) \end{cases}$$

指派问题可行解中，每列每行有且仅有一个1

非标准形式的指派问题：

1. 最大指派问题

$$C = (c_{ij})_{n \times n} \text{ 中最大元素为 } m, \text{ 令 } B = (b_{ij})_{n \times n} = (m - c_{ij})_{n \times n}$$

2. 人数和工作数不等

人少工作多：添加虚拟的“人”，代价都为0

人多工作少：添加虚拟的工作，代价都为0

3. 一个人可做多件工作

该人可化为几个相同的“人”

4. 某工作一定不能由某人做

该人做该工作的相应代价取足够大 M

非线性规划

如果目标函数或约束条件中包含非线性函数，就称这种规划问题为非线性规划问题。一般来说，解非线性规划要比解线性规划问题困难很多。而且，也不像线性规划有单纯形法这一通用方法，非线性规划目前还没有适于各种问题的一般算法，各个方法都有自己特定的适用范围。

案例：投资决策问题

某企业有 n 个项目可供选择投资，并且至少要对其中一个项目投资。已知该企业拥有投资总资金 A 元，投资于第 $i, i=1, 2, \dots, n$ 个项目需要花资金 a_i 元，并预计可收益 b_i 元，试选择最佳的投资方案。

一般形式:

$$\begin{aligned} & \min f(x) \\ \text{s.t. } & \begin{cases} h_j(x) \leq 0, j = 1, 2, \dots, q \\ g_i(x) = 0, i = 1, 2, \dots, p \end{cases} \end{aligned}$$

在一组等式或不等式的约束下, 求一个函数的最大值(或最小值)问题, 其中至少有一个非线性函数

特殊情况

二次规划: 若某非线性规划的目标函数为自变量的二次函数, 约束条件又全是线性的, 就称这种规划为二次规划。

层次分析法

层次分析法是一种解决多目标的复杂问题的定性与定量相结合的决策分析方法。该方法将定量分析与定性结合起来, 用决策者的经验判断各衡量目标能否实现的标准之间的相对重要程度, 并合理地给出每个决策方案的每个标准的权数, 利用权数求出各方案的优劣次序, 比较有效第应用于那些难以用定量方法解决的课题。

糊的问题作出决策的简易方法, 它特别适用于那些难于完全定量分析的问题。常用于**相互关联、相互制约的众多因素构成的复杂而往往缺少定量数据的系统**。

案例:

1. 买钢笔, 一般要依据质量、颜色、实用性、价格等方面的因素来选择某一只钢笔。
2. 我们如果想要给一些问题的指标设定权重, 而又减弱主观因素在里面。

根据问题的性质和要达到的总目标, 将问题分解为不同的组成因素, 并按照因素间相互关联影响以及隶属关系将因数按不同层次的聚集组合, 形成一个多层次的分析结构模型, 从而最终使文通归结为最低层(供决策的方案、措施等)相对于最高层(总目标)的相对重要权值的确定或相对优劣次序的排定。

步骤:

1. 建立层次结构模型

- 将决策目标、考虑的因素(决策准则)和决策对象按它们之间的相互关系分为最高层、中间层和最底层, 绘出层次结构图。

最高层: 决策的目的、要解决的问题。

最底层: 决策时的备选方案。

中间层: 考虑的因素、决策的准则。

- 对于相邻的两层, 最高层为目标层, 低层为因素层。

2. 构造判断(成对比较)矩阵

1. 不把所有因素放在一起比较, 而是两两相互比较
 2. 对此时采用的相对尺度, 以尽可能减少性质不同的诸因素相互比较的困难, 以提高准确度。
- 判断矩阵式表示本层所有因素针对上一层某一个因素的相对重要性的比较。
 - 心理学家认为成对比较的因素不宜超过9个, 即每层不要超过9个因素。

3. 层次单排序及其一致性检验

对于判断矩阵最大特征根的特征向量, 经归一化(使向量中各元素之和等于1)后记为W。

W的元素为同一层次因素对于上一层次因素某因素相对重要性的排序权值, 这一过程称为层次单排序。

能否确认层次单排序，需要进行一致性检验，所谓一致性检验是指对A确定不一致的允许范围。

定理：n阶一致矩阵的唯一非零特征根为n

定理：n阶正互反阵A的最大特征根 $\lambda \geq n$,当且仅当 $\lambda = n$ 时A为一致矩阵

4. 层次总排序及其一致性检验

- 计算某一层次所有因素对于最高层（总目标）相对重要性的权值，称为层次总排序。
- 这一过程是从最高层次到最低层次依次进行的。

灰色系统理论

应用范畴：

1. 灰色关联分析。
2. 灰色预测：人口预测；灾变预测...等等。
3. 灰色决策。
4. 灰色预测控制。

灰色系统理论是人们认识客观系统改造客观系统的一个新型的理论工具。

概念：

1. 灰色系统、白色系统和黑色系统

- 白色系统是指一个系统的内部特征是完全已知的，即系统的信息是完全充分的。
- 黑色系统是指一个系统的内部信息对外界来说是一无所知的，只能通过它与外界的联系来加以观测研究。
- 灰色系统内的一部分信息是已知的，另外一部分信息是未知的，系统内各因素间有不确定的关系。

2. 灰色预测法

- 灰色预测法是一种对含有不确定因素的系统进行预测的方法
- 灰色预测是对既含有已知信息又含有不确定信息的系统进行预测，就是对在一定范围内变化的、与时间有关的灰色过程进行预测。

灰色预测通过鉴别系统因素之间发展趋势的相异程度，即进行关联分析，并可对原始数据进行生成处理来寻找系统变动的规律，生成有较强规律性的数据序列，然后建立相应的微分方程模型，从而预测事物未来发展趋势的状况。

灰色预测法用等时距观测到的反应预测对象特征的一系列数量值构造灰色预测模型，预测未来某一时刻的特征量，或达到某一特征量的时间。

3. 常见类型

- 灰色时间序列预测

即用观察到的反映预测对象特征的时间序列来构造灰色预测模型，预测未来某一时刻的特征量，或达到某一特征量的时间。

- 畸变预测

即通过灰色模型预测异常值出现的时刻，预测异常值什么时候出现在特定时区内。

- 系统预测

通过对系统行为特征指标建立一组相互关联的灰色预测模型，预测系统中众多变量间的相互协调关系的变化。

- 拓扑预测

将原始数据做曲线，在曲线上按定值寻找该定值变化发生的所有时点，并以该定值为框架构成长成时点数列，然后建立模型预测该定值所发生的时点。

灰色关联度与优势分析

大千世界里的客观事物往往现象复杂，因素繁多。我们经常要对系统进行因素分析，这些因素中哪些对系统来讲是主要的，哪些是次要的，哪些需要发展，哪些需要抑制，哪些是潜在的，哪些是明显的。一般来讲，这些都是我们极为关心的问题。事实上，因素间关联性如何、关联程度如何量化等问题是系统分析的关键。

灰色关联度一定是分析向量与向量之间以及矩阵与矩阵之间的关联度。既然计算关联度，一定是计算某一个待比较的数列与参照物之间的相关程度。而回归分析方法需要大量数据、计算量大以及可能出现反常情况等。

灰色预测

灰色生成数列

灰色系统理论认为，尽管客观表象复杂，但总是有整体功能的，因此必然蕴含某种内在规律。关键在于如何选择适当的方式去挖掘和利用它。灰色系统是通过原始数据的整理来寻求其变化规律的，这是一种就数据寻求数据的现实规律的途径，即为灰色序列的生成。一切灰色序列都能通过某种生成弱化其随机性，显现其规律性。数据生成的常用方式有累加生成、累减生成和加权累加生成。

1. 累加生成

把数列各项数据依次累加的过程称为累加生成过程，由累加生成过程所得的数列称为累加生成数列。

2. 累减生成

对于原始数据列依次做前后相邻的两个数据相减，与累加过程互为逆。

3. 加权邻值生成

设原始数列为 $x^0 = (x^0(1), x^0(2), \dots, x^0(n))$

对于常数 $\alpha \in [0, 1]$

令 $z^0(k) = \alpha x^0(k) + (1 - \alpha)x^0(k - 1), k = 2, 3, \dots, n$

得到数列 z^0 为数列 x^0 在权 α 下的邻值生成数，权 α 也称为生成系数

特别的当生成系数 $\alpha = 0.5$ 时，称 z^0 为均值生成数，也称等权邻值生成数

插值拟合

插值问题

例题：在一天24小时内，从零点开始每间隔2小时测得的环境温度数据为

12, 8, 8, 10, 18, 24, 28

27, 25, 20, 18, 15, 13

推测中午1点温度。

- 已知函数在某区间内若干点处的值，求函数在该区间内其它点处的值。

拉格朗日插值法的高次插值的Runge现象

- 在实际中不应使用七次以上插值。避免Runge现象的常用方法：将插值区间分成若干小区间，在小区间内用低次插值，即分段低次插值，如样条函数插值。

一维插值优先选用样条函数插值

拟合问题

例题：从1点到12点每隔一小时测量一次温度，测得的温度一次为

5, 8, 9, 15, 25, 29

31, 30, 22, 25, 27, 24

推测 $t=13.5$ 时的温度

拟合问题和插值问题的区别：

1. 插值函数过已知点，而拟合函数不一定过已知点
2. 插值主要用于求函数值，而拟合的主要目的求函数关系，从而进行预测等进一步的分析

模糊综合评价

- 模糊数学

现实世界中许多现象和关系具有不确定性。这些不确定性的表现形式多种多样，如随机性、灰色性、模糊性和粗糙性等。

模糊数学正是利用模糊集及其运算研究、处理模糊不确定现象和关系的数学分支学科。

许多数学建模问题包括模糊现象和关系，这类问题往往可以用模糊数学方法处理。

- 模糊集

现实中的许多现象及关系比较模糊。如高与矮，长与短，大与小，多与少，穷与富，好与差，年轻与年老等。

此类现象不满足“非此即彼”的排中律，而具有“亦此亦彼”的模糊性。

需要指出的是，模糊不确定不同于随机不确定。随机不确定是因果律破损造成的不确定，而模糊不确定是由于排中律破损造成的不确定。

定义 设给定论域 U ，所谓 U 上的一个模糊集 A 是指对于任意 x 属于 U ，都能确定一个正数 $u_A(x)$ 属于 $[0$

, 1]用其表示 x 属于 A 的程度。

映射

$$x \in U \rightarrow u_A(x) \in [0, 1]$$

称为 A 的隶属函数，函数值 $u_A(x)$ 称为 x 对 A 的隶属度。

每个元素都有隶属度的集合即为模糊集，确定模糊集的关键是构造隶属函数。

模糊集运算：

由于模糊集中没有元素和集合间的绝对隶属关系，所以模糊集的运算时通过隶属函数完成的。

隶属度函数的确定：

由于模糊集的概念可知，模糊数学的基本思想是隶属度，所以应用模糊数学方法建立数学模型的关键是建立符合实际的隶属关系。然而，如何确定一个模糊集的隶属函数至今还是尚未完全解决的问题。

- 确定隶属度的常用方法是模糊分布法

模糊分布法将隶属函数看成一种模糊分布，首先依据问题性质选取适当的模糊分布，然后再依据相关数据确定分布中的参数。

1. 偏小型

$$A(x) = s.t. \begin{cases} 1, x < a \\ \frac{b-x}{b-a}, a \leq x \leq b \\ 0, x > b \end{cases}$$

2. 偏大型

$$A(x) = s.t. \begin{cases} 0, x < a \\ \frac{x-a}{b-a}, a \leq x \leq b \\ 1, x > b \end{cases}$$

3. 中间型

$$A(x) = s.t. \begin{cases} 0, x < a \\ \frac{x-a}{b-a}, a \leq x < b \\ 1, b \leq x < c \\ \frac{d-x}{d-c}, c \leq x < d \\ 0, x \geq d \end{cases}$$

中间型一般适用于描述处于中间状态的模糊现象。

• 模糊综合评价

基本思想：

以模糊数学为基础，应用模糊关系合成原理，将一些边界不清、不易定量的因素定量化，从多个因素对评价事物隶属等级状况进行综合评价。

具体步骤：

首先确定被评价对象的因素集和评价集，然后再分别确定各因素的权重及它们的隶属度向量，获得模糊评价矩阵，最后将模糊评价矩阵与因素的权向量进行模糊运算并归一化，从而得到模糊评价综合结果。

案例：

某服装厂欲采用模糊综合评价法来了解顾客对某种服装的欢迎程度。

顾客是否喜欢某种服饰，通常与这种服装的花色、样式、价格、耐用度和舒适度等因素有关，故确定评价服装的因素集为 $U=\{\text{花色, 样式, 价格, 耐用度, 舒适度}\}$ 。

综合评价的目的是弄清楚顾客对衣服各方面的欢迎程度。因此，评价集应为 $V=\{\text{很欢迎, 欢迎, 一般, 不欢迎}\}$ 。

相关性分析

在统计学上，我们通常这样判断变量之间是否相关：如果一个变量的取值发生变化，另一个变量的取值也相应发生变化，则这两个变量有关。如果一个变量的变化不引起另一个变量的变化则二者无关。

变量关系强度的含义：指两个变量相关程度的高低。统计学中是以准实验的思想来分析相关变量的。通常从以下的角度分析：

1. 两变量是否相互独立。
2. 两变量是否有共变趋势。
3. 一变量的变化多大程度上能由另一变量的变化来解释。

四种基本变量：

- 1. 定类变量：根据定性的原则区分总体各个案类别的变量
案例：性别、民族、婚姻情况
- 2. 定序变量：区别同一类别个案中等级次序的变量
案例：文化程度、工厂规模、年龄大小
- 3. 定距变量：区别同一类别个案中等级次序及其距离的变量（没有零点的概念）
案例：摄氏温度、比率、智力水平
- 4. 定比变量：也是区别同一类别个案等级次序及其距离的变量
案例：收入、价格、市场占有率

Pearson相关系数

使用于定距、定比类型的变量。是运用最广的一种相关程度统计量。检验用t统计量：其中统计量t服从自由度（n-2）的分布。

使用条件：

- 1. 两变量均应由测量得到的连续变量。
- 2. 两变量所来自的总体都应是正态分布，或接近正太的单峰对称分布。
- 3. 变量必须是成对的数据。
- 4. 两变量间为线性关系。

Spearman等级相关系数

适用于度量定序变量与定序变量之间的关系。检验用t统计量。

Kendall tua-b等级相关系数

它也是用来度量定序变量间的线性相关关系。这里的度量原理是把所有的样本点配对，看每一对中x和y是否都增加来判断总体模式。

• 双变量关系强度测量的主要指标

	定类	定序	定距（比）
定类	卡方类检测	卡方类检测	Eta系数
定序		Spearman相关系数	Spearman相关系数
定距			Person相关系数

数据降维

多变量大样本资料无疑能给研究人员或决策者提供很多有价值的信息，但在分析处理多变量问题时，由于众变量之间往往存在一定的相关性，使得观测数据所反映的信息存在重叠现象。因此为了尽量避免信息重叠和减轻工作量，人们就往往希望能找出少数几个互不相关的综合变量来尽可能地反映原来数据所含有的绝大部分信息，处理数据的冗余，而主成分分析和因子分析正是为解决此类问题而产生的多元统计分析方法。

两种方法的出发点都是变量的相关系数矩阵，在损失较少信息的前提下，把多个变量（这些变量之间要求存在较强的相关性，以保证能从原始变量中提取主成分）综合成少数几个综合变量来研究总体各方面信息的多元统计方法，且这少数几个综合变量所代表的信息不能重叠，即变量间不相关。

主成分分析是通过变量变换把注意力集中在具有较大变差的那些主成分上，而舍弃那些变差小的主成分；**因子分析**是因子模型把注意力集中在少数不可观测的潜在变量（即公共因子）上，而舍弃特殊因子。

回归分析

回归分析是研究自变量与因变量之间数量变化关系的一种分析方法，它主要是通过因变量Y与影响它的自变量 X_i ($i=1,2,3,\dots$) 之间的回归模型，衡量自变量 X_i 对因变量Y的影响能力的，进而可以用来预测因变量Y的发展趋势。

相关分析与回归分析的联系与区别：

联系：

- 两者均为研究与测度两个或两个以上变量之间关系的方法。

在实际工作中，一般先进行相关分析，计算相关函数，然后建立回归模型，最后用回归模型进行推算或预测。

- 区别：

1. 相关分析研究的都是随机变量，并且不分自变量与因变量；
2. 回归分析研究的变量要定义出自变量与因变量，并且自变量是确定的普通变量，因变量是随机变量；
3. 相关分析描述的是两个变量之间相关关系的密切程度；
4. 回归分析不仅可以揭示变量X对变量Y的影响程度，还可以根据回归模型进行预测。

模糊综合和灰色关联分析评价

综合评价

综合评价就是运用多个指标对多个参评单位进行评价的方法，称为多变量综合评价方法，又称综合评价法，其基本思想是将多个指标转化为一个能够反映综合情况的指标来进行评价。

- 特点：

1. 评价过程不是逐个指标顺次完成的，而是通过一些特殊方法将多个指标的评价同时完成的。
2. 在综合评价过程中，一般要根据指标的重要性进行加权处理。
3. 评价结果不再是具有具体含义的统计指标，而是以指数或分值表示参评单位“综合状况”的排序。

- 目的：

综合评价一般表现为以下几类问题：

1. 分类——对所研究对象的全部个体进行分类；
2. 比较、排序（直接对全部评价单位排序，或在分类基础上对各小类按优劣排序）；
3. 考察某一综合目标的整体实现程度（对某一事物作出整体评价）。如小康目标的实现程度、现代化的实现程度。

构成综合评价问题的要素：

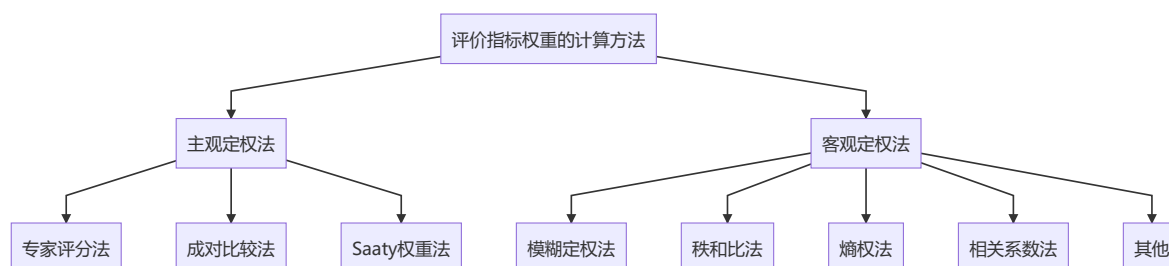
1. 被评价对象：被评价对象就是综合平均问题中所研究的对象，或称为系统。
2. 评价指标：评价指标是反映被评价对象的运行或发展状况的基本要素。通常的问题都是由多项指标构成，每一项指标都是从不同的侧面刻画系统所具有的某种特征大小的一个量度。一个综

合评价问题的指标一般可用一个向量表示，其中每一个分量就是从侧面反映系统的状态，即称为综合评价的指标体系。

3. 权重系数：任一综合评价的问题都有相应的评价项目，针对某种评价目的，各评价指标之间的相对重要性是不同的，评价指标之间的这种相对重要性的大小可以用权重系数来刻画。权重系数确定的合理与否，直接关系到综合评价结果的可信度，甚至影响到最后决策的正确性。
4. 综合评价模型：对于多指标（多因素）的综合类评价问题，就是要通过建立合适的综合评价数学模型将多个评价指标综合称为一个整体的综合评价指标，作为综合评价的依据，从而得到相应的评价结果。
5. 评价者

• 步骤：

1. 确定综合评价目的
2. 建立评价指标体系
3. 对指标数据作预处理
 - 使所有的指标都从同一角度说明总体，这就提出了如何使用指标一致化的问题
 - 所有的指标可以相加，这就提出了如何消除指标之间不同计量单位对指标数值大小的影响和不能加总的问题，即对指标进行无量纲化处理——计算单项评价价值
4. 确定各个评价指标的权重
5. 求综合评价值——将单项评价值综合而成



• 规范化处理

1. 指标的类型

定性：优，良，中，一般，差；很高，高，一般，低，很低。

定量：极大型，极小型，居中型。

2. 定性指标的量化（评分法）

等级	很低	低	一般	高	很高
分值	1	3	5	6	9

3. 定量指标类型的一致化

极大型指标、极小型指标、中间型指标、区间型指标

极小值指标转化为极大型指标

$$bij = M - aij, (M > 0)$$

或

$$bij = \frac{1}{aij}$$

$$(i = 1, 2, \dots, n)$$

剧中型指标转化为极大型指标

$$bik = \begin{cases} \frac{2(aik - m)}{M - m}, m \leq aik \leq \frac{m + M}{2} \\ \frac{2(M - aij)}{M - m}, \frac{m + M}{2} \leq aik \leq M \end{cases}$$

$$(i = 1, 2, \dots, n)$$

其中m和M分别为指标允许的下界和上界

4. 评价指标的无量纲化

防止“大数吃小数”，无量纲化处理又称为指标数据的标准化，或规范化处理。

常用方法：标准差方法、极值差方法和功效系数方法

- 常用综合评价方法

1. 线性加权综合法

使用条件：各评价指标之间相互独立

对不完全独立的情况，其结果将导致各指标间信息的重复，使评价结果不能客观地反映实际。

特点：

1. 各评价指标间作用得到线性补偿
2. 权重系数的对评价结果的影响补偿

2. 非线性加权综合法

使用条件：各指标间又较强关联性

特点：

1. 对数据要求较高，指标数值不能为0、负数
2. 乘除法容易拉开评价档次，对较小数值的变动更敏感

3. 逼近理想点（TOPSIS）方法

设定系统指标的一个理想点 $(x_1^*, x_2^*, \dots, x_m^*)$ ，将每一个被评价对象与理想点进行比较。

模糊综合评价

对方案、人才、成果的评价，人们的考虑的因素很多，而且有些描述很难给出确切的表达，这时采用模糊评价方法。它可对人、事、物进行比较全面而又定量化的评价。

- 模糊概念

秃子悖论：天下所有的人都是秃子

设头发根数n n=1显然

若n=k为秃子 n=k+1亦为秃子

模糊概念：从属于该概念到不属于该概念之间无明显分界线

灰色关联分析评价

- 灰色系统

是指部分信息已知，部分信息未知的“小样本，贫信息”不确定性系统，即信息不完全的系统。灰色系统理论主要通过对“部分”已知信息的挖掘、开发，提取有价值的信息，实现对系统运行行为、演化规律的正确描述和有效监控。

小样本分析

元胞自动机模型

元胞自动机模型是网络动力学模型中最常用的一种。网络动力学模型主要包括元胞自动机、布尔网络、神经网络以及L-系统等。元胞自动机，是一种时间、空间和状态都离散的动力模型。元胞自动机可以被视为由一个元胞空间和定义于该空间的变换函数组成。

元胞自动机不是由严格定义的物理方程或函数确定，而是用一系列模型构造的规则构成。凡是满足这些规则的模型都可以算作是元胞自动机模型。因此，元胞自动机是一类模型的总称，或者说是一个方法框架。

在CA模型中，散布在规则格网(Lattice Grid)中的每-元胞(Cell)取有限的离散状态，遵循同样的作用规则，依据确定的局部规则作同步更新。大量元胞通过简单的相互作用而构成动态系统的演化。

数学表达式为：

$$A=(L,d,S,N,f)A=(L,d,S,N,f)$$

其中：

A：一个元胞自动机系统

LdLd：一个规则划分的网格空间，L通常为一个一维或二维空间，但理论上可以是任意正整数维的规则空间

S：一个离散的有限集合，用来表示各个元胞的状态

N：元胞的邻居集合，对于任何元胞的邻居集合N属于L，设邻居集合内元素数目为n，那么N可以表示为一个所有邻域内元胞的组合（包括中心元胞）

ff：一个映射函数，根据tt时刻某个元胞的所有邻居的状态组合来确定 t+1t+1时刻该元胞的状态值。通常又被称为状态转换函数，或局部规则

图论

摘自CSDN

图 (graph) 是数据结构和算法学中最强大的框架之一（或许没有之一）。图几乎可以用来表现所有类型的结构或系统，从交通网络到通信网络，从下棋游戏到最优流程，从任务分配到人际交互网络，图都有广阔的用武之地。

图 (graph) 并不是指图形图像 (image) 或地图 (map)。通常来说，我们会把图视为一种由“顶点”组成的抽象网络，网络中的各顶点可以通过“边”实现彼此的连接，表示两顶点有关联。注意上面图定义中的两个关键字，由此得到我们最基础最基本的2个概念，顶点 (vertex) 和边 (edge)。

一、顶点 (vertex)

上图中黑色的带数字的点就是顶点，表示某个事物或对象。由于图的术语没有标准化，因此，称顶点为点、节点、结点、端点等都是可以的。叫什么无所谓，理解是什么才是关键。

二、边 (edge)

上图中顶点之间蓝色的线条就是边，表示事物与事物之间的关系。需要注意的是边表示的是顶点之间的逻辑关系，粗细长短都无所谓。包括上面的顶点也一样，表示逻辑事物或对象，画的时候大小形状都无所谓。

三、同构 (Isomorphism)

前面提到顶点和边指的是事物和事物的逻辑关系，不管顶点的位置在哪，边的粗细长短如何，只要不改变顶点代表的事物本身，不改变顶点之间的逻辑关系，那么就代表这些图拥有相同的信息，是同一个图。同构的图区别仅在于画法不同。

四、有向/无向图 (Directed Graph/ Undirected Graph)

最基本的图通常被定义为“无向图”，与之对应的则被称为“有向图”。两者唯一的区别在于，有向图中的边是有方向性的。

五、权重 (weight)

边的权重（或者称为权值、开销、长度等），也是一个非常核心的概念，即每条边都有与之对应的值。例如当顶点代表某些物理地点时，两个顶点间边的权重可以设置为路网中的开车距离。

在图上任取两顶点，分别作为起点 (start vertex) 和终点 (end vertex)，我们可以规划许多条由起点到终点的路线。不会来来回回绕圈子、不会重复经过同一个点和同一条边的路线，就是一条“路径”。两点之间存在路径，则称这2个顶点是连通的 (connected)。

路径也有权重。路径经过的每一条边，沿路加权重，权重总和就是路径的权重（通常只加边的权重，而不考虑顶点的权重）。在路网中，路径的权重，可以想象成路径的总长度。在有向图中，路径还必须跟随边的方向。

值得注意的是，一条路径包含了顶点和边，因此路径本身也构成了图结构，只不过是一种特殊的图结构。

七、环 (loop)

环，也成为环路，是一个与路径相似的概念。在路径的终点添加一条指向起点的边，就构成一条环路。通俗点说就是绕圈。

八、连通图/连通分量 (connected graph/connected component)

如果在图G中，任意2个顶点之间都存在路径，那么称G为连通图（注意是任意2顶点）。上面那张城市之间的图，每个城市之间都有路径，因此是连通图。而下面这张图中，顶点8和顶点2之间就不存在路径，因此下图不是一个连通图，当然该图中还有很多顶点之间不存在路径。连通分量有如下特点：

- 1) 是子图；
- 2) 子图是连通的；
- 3) 子图含有最大顶点数。

神经网络算法

人工神经网络 (Artificial Neural Networks, 简称为ANNs) 也简称为神经网络 (NNs) 或称作连接模型 (Connection Model)，它是一种模仿动物神经网络行为特征，进行分布式并行信息处理的算法数学模型。这种网络依靠系统的复杂程度，通过调整内部大量节点之间相互连接的关系，从而达到处理信息的目的。

多层向前神经网络由以下几个部分组成：

输入层 (input [layer](#))，隐藏层 (Hidden layer)，输出层 (output layer)

特点：

- 1、每层由单元 (units) 组成
- 2、输入层是有训练集的实例特征向量传入
- 3、经过连接接点的权重 (weight) 传入下一层，一层的输出是下一层的输入

- 4、隐藏层的个数可以是任意的，输入层有一层，输出层有一层
- 5、每个单元也可以称之为神经结点，根据生物学来源定义
- 6、以上成为两层的神经网络，输入层是不算在里面的
- 7、一层中加权求和，然后根据非线性方程转化输出
- 8、作为多层向前神经网络，理论上，如果有足够的隐藏层，和足够的训练集，可以模拟出任何方程

建立和应用神经网络的步骤：

1. 网络结构的确定

包含网络的拓扑结构和每个神经元相应函数的选取

2. 权值和阈值的确定

通过学习得到，为有指导的学习，也就是利用已知的一组正确的输入、输出数据，调整权和阈值使得网络输出与理想输出偏差尽量小

3. 工作阶段

用带有确定权重和阈值的神经网络解决实际问题的过程，也叫模拟(simulate)