

2020 年第五届“数维杯”大学生 数学建模竞赛论文

题 目 全面建成小康社会下的数据分析

摘 要

新时代的进程下，由于经济迅速地发展，全面建成小康社会的目标则是越来越近，通过对我国历史发展过程中的经济建设数据进行分析已成为一大重要手段，在这个问题上予以启示。本问题主要分为四个小问题，其中第一小问要求通过数据分析对建设过程进行经验总结以及可持续发展预测；第二小问则是通过综合指标对每个地区的富裕程度予以评价排序；第三小问则是选取合适目标地区，对可持续发展予以预测对比；第四小问则是提供一个城市间的合作方案。

对于问题 1，首先筛选合适的的数据，采用主成分分析（PCA）对数据进行分析筛选，获得对经济建设中影响比较大的因素为人口项因素；其次通过查找相关文献，从三大方面对结果进行归纳总结；最终则是通过数据与文献两方面，从五方面出发对发展进行规划。

对于问题 2，通过六大方面的筛选，构建 PSR 综合指标评价体系，采用层次分析法（AHP）对这些数据进行加权处理，得到我国 31 个省市在 2018 年的小康程度指数排名，其中江苏省最高（0.572），符合我国国情，并最终绘制出了指数分布地图（图 6）。

对于问题 3，选取了陕西省为研究目标，通过如问题 2 中的评价流程，计算可持续发展得分，采用时间序列分析对不同时间跨度进行预测，最终结果如图 7 所示。

对于问题 4，为确定必要的帮扶方案，建立了了 0-1 整数规划对问题进行分析求解，并通过 Lingo 编程获得如四川—西藏的前七组帮扶城市配对结果（表 3）。

最终对整体模型以及求解方法进行回顾，分析了优缺点，并对最终如何建设小康社会提出了相应的建议。

关键词：数据分析；主成分分析；层次分析；时间序列分析；0-1 整数规划

目 录

一、问题重述	1
1.1 问题背景	1
1.2 要解决的问题	1
二、问题分析	1
2.1 问题 1 的分析	1
2.2 问题 2 的分析	2
2.3 问题 3 的分析	2
2.4 问题 4 的分析	2
三、模型假设	2
四、符号说明	3
五、模型的建立与求解	3
5.1 主成分分析模型的建立与求解	3
5.2 PSR 评价模型的建立与求解	10
5.3 时间序列预测模型的建立与求解	16
5.4 0-1 整数规划模型的建立与求解	17
六、模型优缺点	19
6.1 模型优点	19
6.2 模型缺点	20
参考文献	20
附录 A 层次分析法	21
附录 B 0-1 整数规划	21

一、问题重述

1.1 问题背景

经过百年奋斗，我国人民在 2020 年迎来全面建成小康社会的里程碑。然而，新冠肺炎突然侵袭。从内部角度来看，国内从武汉开始向全国范围内扩散，虽然我国目前已经很好地控制住了疫情，但是疫情对我国经济发展还是造成了重大影响；从外部角度来看，全球范围内疫情依旧大肆传播、尚未有效控制，这使得我国也面临严峻的国际环境。我国经济发展面临内忧外患，如何对社会发展状况进行评价、预测与规划也成了一个重要问题。

1.2 要解决的问题

结合数学建模方法对如下问题进行合理的优化建模与规划：

1、总结归纳过去 20 年间我国全面建成小康社会历程中的成功经验，并结合全球气候变暖、生态环境与人口压力倍增等现状提出未来 20 年的我国可持续发展思路。

2、结合官方数据，总结归纳我国各省全面建成小康社会完成难度系数、完成比率及未来可持续发展能力等在内的多重指标，对各省全面建成小康社会总体情况进行合理的评价与排序。

3、针对世界处于大变局及新冠疫情全球性蔓延的大环境，充分考虑人口变化、老龄化趋势、城乡居民收入、资源禀赋及生态环境承载力等关键性因素，建立以某省为代表的短期及长远的经济社会可持续发展模型。

4、充分考虑各省份经济社会发展状况及资源禀赋，提供一个能够对双方产生积极促进作用、使总体效率提升的省际合作计划。

二、问题分析

2.1 问题 1 的分析

问题 1 要求总结我国过去 20 多年全面建成小康社会进程中的先进经验，并结合多个要素（全球气候变暖、生态环境与人口压力倍增等），对我国未来 20 多年的发展提出可持续发展建议思路。其中第二步——针对多个发展壁垒构建综合变量、进行数据分析是这一问的建模关键。

在建模方面，需要对全球气候变暖、人口压力倍增等状况进行量化，对变量（如，GDP）进行数据分析，本文主要采用主成分分析，将初始的多个变量重组成新的较少的

几个综合变量体系来评估建成小康社会进程中面临的障碍。在文字归纳方面，首先，需要对我国过去建设小康社会的经历进行总结归纳，提炼出成功经验；其次，需要结合前面变量分析给出未来我国可持续发展的建议思路。

2.2 问题 2 的分析

问题 2 主要要求建立模型来对我国各省全面建成小康社会的总体情况进行评价与排序。

首先，要对包括各省全面建成小康社会完成难度系数、完成比率及未来可持续发展能力等在内的多重指标进行量化，本文采用 PSR 指标构建方法，指标体系内部既有联系又有区别，共同反映小康水平。其次是采用 AHP 层次分析评价得到最终表达式，然后通过表达式对数据进行计算分析，进而完成各省总体情况的排序。

2.3 问题 3 的分析

问题 3 要求结合世界大变局以及疫情蔓延的发展环境，综合人口变化、老龄化趋势、城乡居民收入、资源禀赋及生态环境承载力等关键性因素，针对某一省提出短期和长期的经济可持续发展预测模型。显然，这一问的关键在于预测模型的构建。这里需要结合世界变局及疫情的大环境，摒除了季节因素等影响，文中采用时间序列分析。该问主要采用问题 1 的指标以及问题 2 的评价方法。

近年来关于小康社会的进程的研究更多偏向于定性研究、定量研究相对较少，使用模型进行数据测量多局限于某省范围。该题同样要求针对某一省提出预测模型。地区发展水平存在差异，全面建成小康社会的指标设定以及权重选择也相应的有所区别，因而此处针对一省提出经济可持续发展模型可以代表一类省市地区的预测。

2.4 问题 4 的分析

问题 4 要求结合各省份经济社会发展状况及资源禀赋提出一个能够促进各自发展、提升总体效率的省际合作计划。

这一问主要是针对问题 2 的结果进行模型优化、建立优化模型对城市之间的合作关系进行分析，最终得到城市的合作结果。

三、模型假设

- 假设我国国内生产总值只受全球气候变暖、生态环境和人口压力的影响；

- 全球气候变暖只与温度及二氧化碳排放量有关;
- 我国过去人口的变化主要由于二胎政策的影响, 人口结构的变化只与老龄化问题有关;
- 考虑到地域问题, 限制合作的选择双方必须为邻省, 避免地域问题消除选择;
- 发展平稳, 国内生产总值没有骤增骤降;
- 忽略除本文考虑因素外的其他因素的影响。

四、符号说明

符号	符号含义
GDP_i	第 i 年国内生产总值
T_i	第 i 年温度变化
P_i	第 i 年人口变化
O_i	第 i 年老龄人口变化
F_i	经过处理后的指标
$PGDP_j$	地区人均生产总值
C_j	地区居民收入
V_j	地区农民收入

注: 其余符号见正文部分

五、模型的建立与求解

5.1 主成分分析模型的建立与求解

通过对问题 1 的分析以及讨论, 问题 1 中主要包含三个小问题, 其一是对过去 20 年间我国全面建成小康社会历程中的成功经验的总结与概括; 其二则是通过多方面的指标, 对整个所存在的问题进行分析讨论; 其三则是对于我国未来 20 年的可持续发展规划的思路进行提出与研究。要解决这个问题, 应当结合必要的文献以及通过过去 20 年相应指标进行的数据分析这两大方面, 对三个问题予以解决。

整体的流程按照过去 20 年数据分析、通过分析结果以及经验进行概括评价以及对 20 年后的发展目标进行规划的流程进行实现。本问题的具体叙述逻辑也按照这三步骤进行描述分析。

Step 1: 数据分析

首先则是对于过去二十年的数据分析，从而从更加理性的角度对过去 20 年的经验进行总结以及归纳。这一方面目的是通过必要的数据分析，从而进一步挖掘出过去 20 年里经济发展对应的相关因子，根据这些因子的贡献程度大小，从而对整个问题的结果进行更加详细且贴切的分析。

第一步则是相关指标的选取，通过中国国家统计局^[1]所得到的相关数据，从全球气候变暖、生态环境与人口压力倍增这三个方面进行详细的数据分析。所选取的指标分别如下：

- 每一年国内生产总值 (GDP_i):

选取国内地区生产总值可以作为每一年所对应的经济发展指标，通过对于每一年生产总值的分析以及讨论，将其作为因变量，可以通过必要的数据分析方法进行适当的共性因子的挖掘。

- 每一年温度变化 (T_i):

选取每一年温度变化，可以体现出过去 20 年的气候变暖状况，通过这一指标侧面反映出二氧化碳的排放量，衡量过去发展过程中对环境所造成的影响，可以反映出发展的绿色性、环保性。

- 每一年二氧化碳变化 (CO_{2i}):

选取每一年二氧化碳变化，与温度变化相对，可以直接反映出该地区气候的变化状况，从而对可持续发展的重要性进行侧面相对应的映照以及分析评价。

- 每一年人口变化 (P_i):

每一年人口变化可以反应我国过去人口压力的变化程度，由于计划生育到二胎政策相对的改变，我国的人口压力也会出现较为不一样的改变以及分析变化，通过这一点可以反映出人口总压力对于结果的影响。

- 每一年老龄人口比例变化 (O_i):

每一年老龄人口的比例改变，则会从侧面反应出人口结构的改变。由于人口结构的改变 (如中国的人口老龄化问题)，使得市场上现有的劳动力发生了必要的改变，这一政策则会使得每一年的国家 GDP 发生必须的变化。

以上则是选取的各项指标，如果需要对这些指标进一步的分析，则需要做相应的归一化处理，考虑采取最大最小归一化的方法进行讨论分析，如式 1 所示：

$$F_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

其中 F_i 为经过归一化后之后的因子值， X_i 为原始的因子值。

通过指标选取，需要对原始数据过 20 年的变化情况进行进一步的分析以及讨论，首先则是对每一年数据指标变化情况的分析总结，各项指标的变化折线图如图 1 与 2 所示：

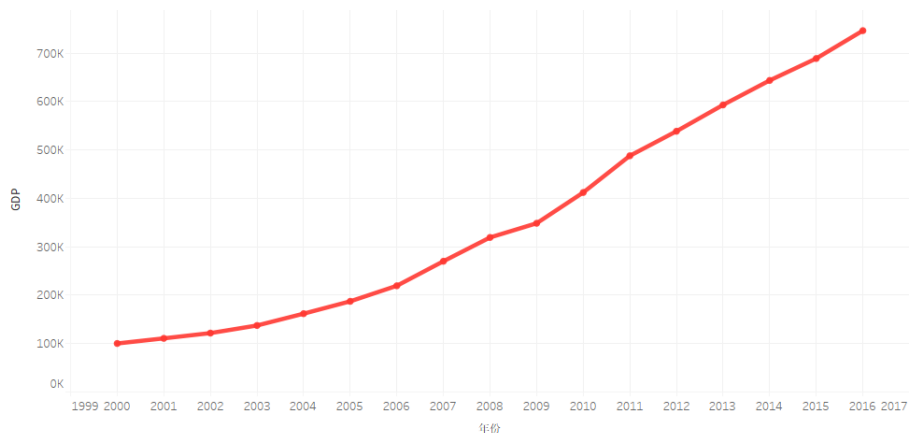


图 1 GDP 含量变化图

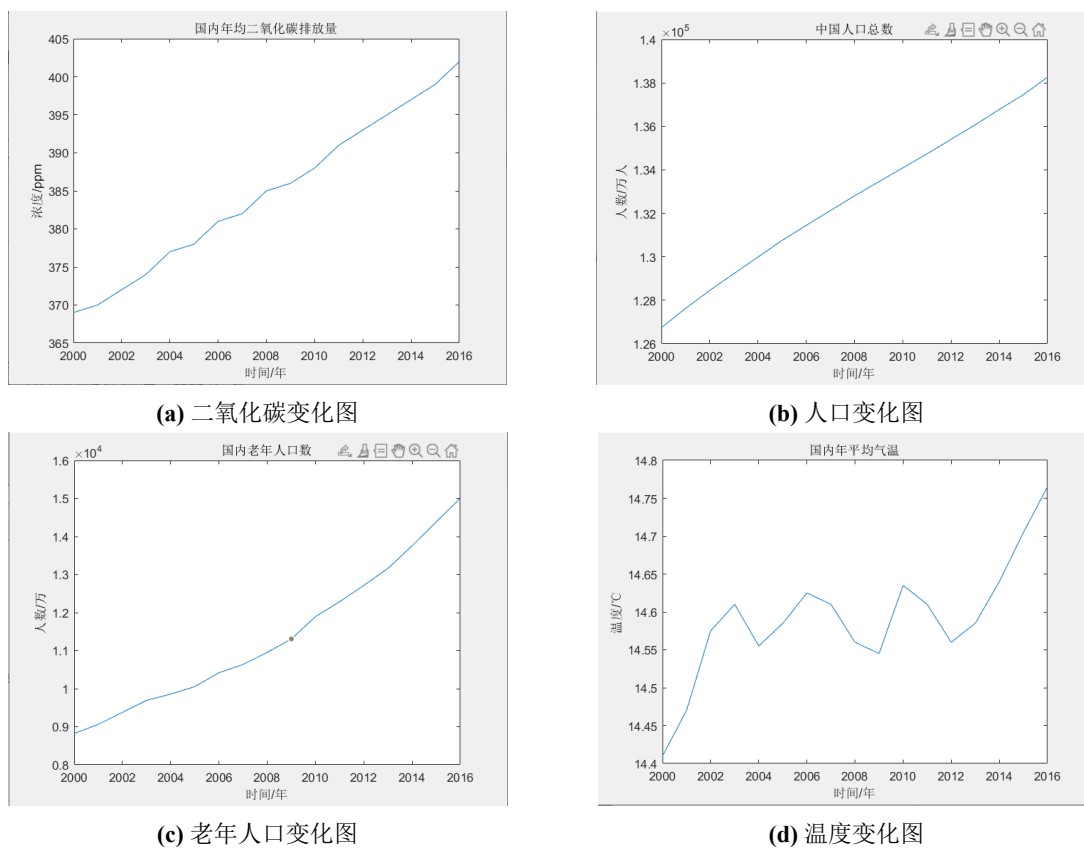


图 2 四大自变量时序变化图

通过对于图 1 与图 2 的对比分析，不难发现，对于衡量的 GDP 含量而言，会呈现一种每一年逐年增长的趋势，而需要衡量的所选取的四大自变量变化也是总体呈现一

种增长的趋势。这四点因素对最终的 GDP 含量均有一定的影响，但由于影响机理的错综复杂性，很难从原始数据中得到更加详细以及有用的信息。

经过原始数据分析后，不难发现，影响国家 GDP 的因素是错综复杂的，为需要进一步的分析讨论，则需要通过自变量组与因变量组之间数据的相关性进行分析讨论。结合文献 [2]，对这些因素进行主成分分析 (PCA)，结合结果进行更深层次的分析讨论。主成分分析原理如下：

主成分估计采用的方法是将原来的回归自变量变换到另另一组变量，即主成分，选择其中一部分重要的主成分作为新的自变量（此时丢弃了一部分，影响不大的自变量，这实际达到了降维的目的），然后用最小二乘法对选取主成分后的模型参数进行估计，最后再变换回原来的模型求出参数的估计。结合文献 [4]，主成分分析 PCA 的主要步骤如以下几个步骤所示：

设有 p 个回归（自）变量 x_1, x_2, x_3, x_n ，它在第 i 次试验中的取值为： $x_{1p}, x_{2p}, x_{3p}, x_{np}$ ，将它们写成矩阵形式：

$$X = (x_1, x_2, \dots, x_n) = \begin{pmatrix} x_1 & x_1 & \cdots & x_1 \\ x_1 & x_1 & \cdots & x_1 \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_1 & \cdots & x_1 \end{pmatrix} \quad (2)$$

考虑线性模型：

$$Y = |\beta_0 1 + X\beta + \varepsilon, \varepsilon \in N(0, \sigma^2 I)| \quad (3)$$

其中 Y 为 $n \times 1$ 向量，为未知参数， 1 为所有元素均为 1 的 n 维列向量，为 $p \times 1$ 未知参数向量，为 $n \times 1$ 误差向量。假定 X 已经标准化：

$$\beta_0 = \bar{Y} = \frac{1}{n} \sum Y_i \quad (4)$$

对于自变量的任意一个线性组合

$$z = \sum c_i p_i, \sum_{j=1}^p c_j^2 = 1 \quad (5)$$

将 z 视为一个新的变量。

记 $w = (c_1, c_2, \dots, c_p)^T$, 则

$$M_2 = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2 = \frac{1}{n} (Xw)^T (Xw) \quad (6)$$

对于新变量 z 来说, 如果在 n 次试验之下它的取值变化不大, 即是说 M_2 较小, 则这个新变量可以去掉。反之, M_2 较大, 那么这个新变量有较大的变化, 它的作用比较明显。注意到 i 与 z 的取值与 i 与 c 的选取有关。

现在回到线性模型, 将 x_p 变为 z_p 易有:

$$Z = (z_1, z_2, \dots, z_n) = \begin{pmatrix} z_1 & z_1 & \cdots & z_1 \\ z_1 & z_1 & \cdots & z_1 \\ \vdots & \vdots & \ddots & \vdots \\ z_1 & z_1 & \cdots & z_1 \end{pmatrix} \quad (7)$$

则最终得到的主成分分析形式为:

$$\hat{\beta} = (Q_1, Q_2) \begin{pmatrix} \alpha_1 \\ 0 \end{pmatrix} = Q_1 \alpha_1 \quad (8)$$

最终通过计算得, 主成分分析结果如图 3 所示:

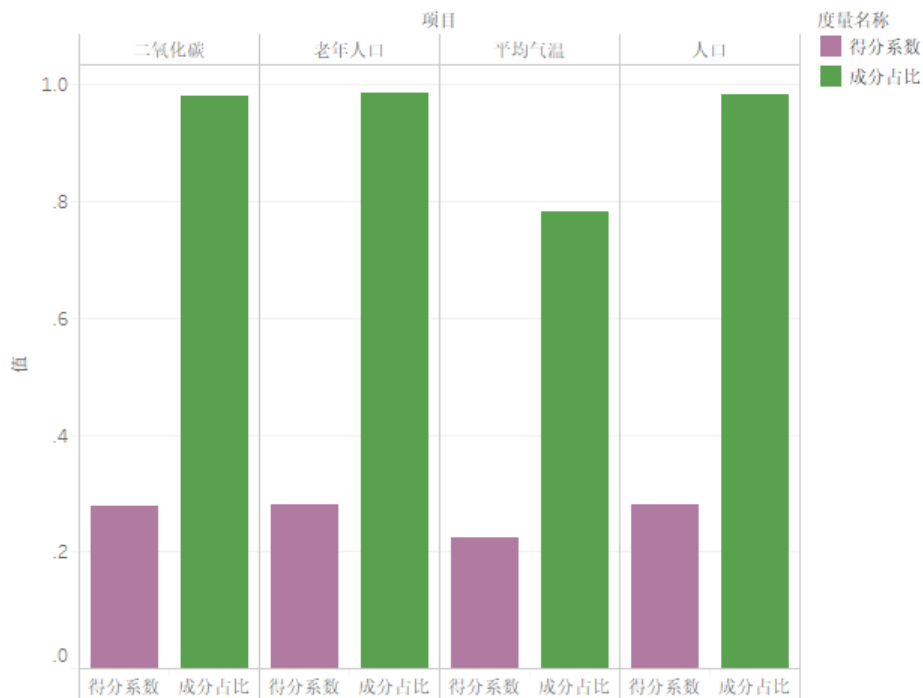


图 3 主成分分析结果图

通过对主成分结果的分析，对于这四点因素而言，其对与最终 GDP 变化的结果是差不多的，但相比之下，与人口有关的人口数量以及人口结构项目对最终结果的影响是要显著高于其余二者的。通过这一点，不难发现，我国在以前 20 年的发展历程中，GDP 增长往往通过劳动力的数量堆积来形成增长的动力。但由于我国资源有限，国家 GDP 的增长需要避免这种“粗放”型的出现，为实现可持续发展，经济类型的改革是势在必行的。

Step 2: 经验总结

《诗经·大雅·民劳》是最早使用“小康”一词的文字记载，文中提到：“民亦劳止，汔可小康；惠此中国，以绥四方。”当时将小康作为生活水平的一个标准。1979 年，邓小平同志首先将“小康”用于现代语境，当时用了“小康之家”的说法。1982 年，“小康”被首次在政府报告中出现。党的十二大报告中的“小康”包含的是经济发展内涵——“从 1981 年到 20 世纪末的 20 年，我国经济建设总的奋斗目标是，在不断提高经济效益的前提下，工农业的年总产值翻两番，城乡人民的收入将成倍增长，人民的物质文化生活可以达到小康水平。”在 1997 年时我国经济发展就已经实现了小康水平，生产力得到大幅提升，将小康仅仅定位于经济水平是不全面的。于是，2002 年，十六大正式提出了从政治、经济、文化方面全面建设小康社会，小康从经济领域的内涵拓展到了文化、政治；随着发展，人口问题、生态问题等跃入人们的视线，十七大则将格局提升到了政治、经济、文化、民生和生态的小康，使得小康社会的内涵更加全面，多方位考量人民生活发展的需求。2012 年，十八大将“全面建设小康社会”变更为“全面建成小康社会”，明确了要从经济持续健康发展，人民民主不断扩大，文化软实力显著增强，人民生活水平全面提高，资源节约型、环境友好型社会建设取得重大进展五个方面进行总体布局。这样的更新是有时代内涵的——我国当时的经济总量已经达到第二位，国际竞争力明显提升，为全面建成小康社会奠定了坚实的基础。2017 年，十九大提出“决胜全面建成小康社会”。这是立足于“我国社会主要矛盾转变为人民日益增长的美好生活需要和不平衡不充分的发展之间的矛盾”的发展现状所提出的发展目标。

1、小康的内涵从传统中脱胎而出，在现代社会主义内涵中不断丰富。

从经济内涵到政治、经济、文化、民生和生态五位一体，从全面建设小康社会到全面建成小康社会。前者是对发展目标的内涵丰富，后者是对发展程度的不断推进。社会前进是动态的，而这些都是根据社会演进程度、国情发展状态和人民发展需求有针对性地提出的。

2、确立战略思维，从国家大局宏观部署、长远安排；发现差异化，推进区域协调发展。

在小康社会的不断实践中，我们始终明确我国处于社会主义初级阶段，以此为宏观

背景，以发展生产力为核心要义来推进小康的进程。另一方面，我国建设小康社会始终立足于人民对于美好生活的向往与追求。从人民“能吃饱”的基本生存需求到政治、经济、文化、民生和生态五位一体全面需求，逐层推进，但人民这个核心是不变的。由于地理位置、自然资源等方面的影响，各地区天然地存在差异，而有针对性地提出区域发展规划、发展区域比较优势、消除地区劣势、推进区域协调发展也是我国发展中一直奉行的核心要义。从振兴东北老工业、京津冀协同发展、长三角一体化发展到雄安新区的建设、黄河流域生态保护，等等一系列规划都在有效地扬长避短、促进区域的全面协调发展、规避地区发展不平衡。

3、聚焦小康社会发展进程中的“后进生”，全面扫清建成小康社会进程中的障碍。

我国人口数量居于世界前列，社会人口压力相对较大，尤其面临老龄化问题。面对巨大的基数，缓解压力、全体推进小康进程成了摆在眼前的重要问题。习近平总书记曾多次强调“小康不小康，关键看老乡”，大的人口基数中，贫困人口是全面小康不容忽视的最突出短板。2020年，李克强总理提出我国依旧有6亿人月收入仅1000元。这些群众的脱贫直接牵动着全面建成小康社会的进程。2019年12月10日举行的中央经济工作会议上，将脱贫攻坚订立为“三大攻坚战”的首要战役，确定将政策、资金倾向“三区三州”等贫困地区。这是对小康社会的助力，同时也是完善与补充。

Step 3: 发展规划

1、引导产业转型，由规模至上转向质量至上。

为了改变一直以来的粗放型劳动密集产业、高能耗高污染产业，相关部门应当推出相应的产业升级鼓励政策与支持措施。在制度方面：明确限制高能耗、高污染的产业类型，同时明确鼓励产业升级、向集约型经济发展。在投资方面：首先，合理约束粗放型产业的政府投资规模，政府投资的相对减少会一定程度上敦促相关产业进行优化；同时，应予以适当的产业过渡支持。其次，加大对低能耗、低污染产业以投资支持和政策倾向，鼓励提升对生产要素的利用程度。

2、科技驱动发展，建立新的科技创新型经济。

首先，应建立科研创新基金、完善奖励机制，驱动更多人才、企业走向科技创新。其次，应当优化创新人才管理制度、引进大量先进科技人才。这不仅仅包括创新人才的合理开发，同时还有人才评估专家的培养，后者有利于对人才的合理利用以及对人才执行项目效果的评估。然后，还要建立科技创新项目管理办法，将科技效益纳入政府考核评估体系。

3、保护生态环境，构建环境友好型发展体系。

以人为本是可持续发展观的核心，需要达到既满足现代人的需求又不损害后代人满

足需求的能力的水平。环境与资源的永续利用是可持续发展的重要保障，应当将生态效益与经济效益有机结合。一方面，完善环境保护制度，提升生态保护政策在国家政策体系中的重要程度，尤其要加强对违规污染行为的处罚力度；另一方面，提升对环境友好型产业的激励程度。此外，还要关注对已污染生态的修复工作，创造并利用先进科技手段治理污染区域。

4、调整生育政策，健全现有劳动力开发体系。

目前，我国人口结构持续走向老龄化，是世界上老年人口最多的国家，人口的老龄化会带来劳动力结构的老龄化，青年劳动力的缺乏则必然会对社会的可持续发展带来一定阻碍，因而相关部门可以推出适当的生育鼓励政策，这一点也是我国正在努力践行的。对于现有老年人口，可以适当延迟退休年龄、完善养老保险制度与医疗卫生条件。从另一个角度来看，老年人口也是对于医药保健、娱乐服务等产业的助力，相关部门可以鼓励相关产业的优化创新。

5、培育智力资本，由人口红利转向人才红利。

随着科技与生产力的不断发展，人口单纯堆砌而成的红利越来越贫瘠。这种模式下人力资源的可替代性太强，而且劳动生产效率不高。我国是人口大国，在人力资源方面有着一定优势，关键在于如和高效利用这种关键的生产要素。首先，应当加大教育投资，优化教育资源的配置，提升教育水平，弘扬科教兴国战略。除了常规的教育结构，也应当加强专业教育、职业教育。其次，应当加强对于人才的合理利用，使得人尽其才。这一点的关键在于人才如何有效配置，可以设置鼓励人才流动、降低人才流动成本的政策，通过教育、服务、住房等优惠条件为人才流动提供绿色通道，同时也要充分发挥市场对于资源优化配置的作用。另外，还应当鼓励科技产业的发展、提供优质岗位。

5.2 PSR 评价模型的建立与求解

通过上述的问题分析，问题 2 主要要求从我国各省全面建成小康社会完成难度系数、完成比率及未来可持续发展能力等在内的多重指标出发，结合数据，进行综合指标评价体系的构建以及地区小康建设指数的计算。此问题为一个典型的评价问题，针对这个问题，结合文献^{[3][4]}，考虑采用 PSR 构建指标评价体系，运用层次分析法 (AHP) 进行加权的方法，对整个问题进行分析求解。

本问主要的流程如图 4 所示：

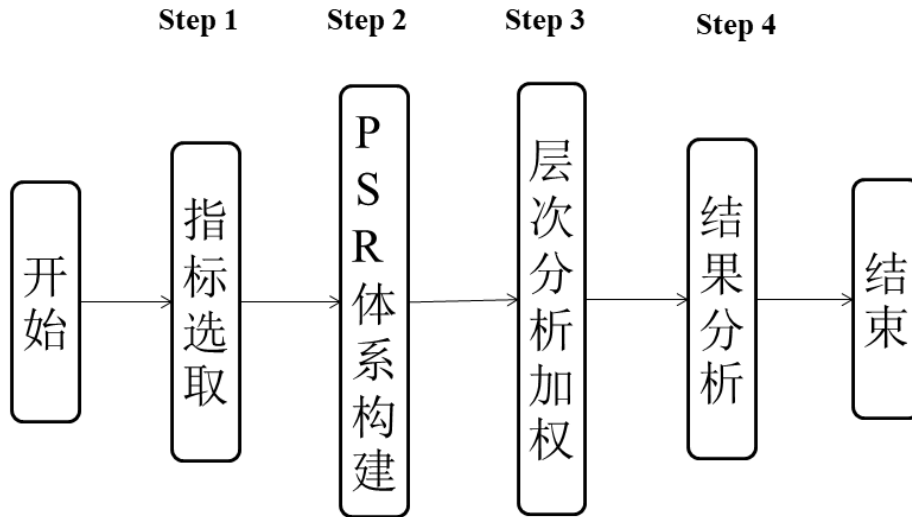


图 4 问题 2 求解流程图

如图 4 所示，本问的叙述主要分为四步按照流程图进行展开，分别为指标的选取部分 (六大方面指标)、构建 PSR 综合指标评价体系、运用层次分析法 (AHP) 进行加权处理以及最终的结果分析。通过最终得到的式子，对整个问题进行必要的阐述以及模型的建立与求解。

Step 1: 指标的选取

第一步是对于包括我国各省全面建成小康社会完成难度系数、完成比率及未来可持续发展能力等在内的多重指标的筛选分析，通过指标的筛选，从而确定相关的数据，最终才能进行进一步综合指标评价体系的构建。

结合文献 [3][5]，2008 年国家统计局出台了由经济发展、生活质量、社会和谐、民主法治、资源环境和文化教育 6 个子系统 23 项指标组成的《全面建成小康社会统计监测指标体系》，由于其全面性、具体性与可获得性，深受广大学者的认可。但这些指标存在重复性，考虑到数据的侧重点，筛选以下六个指标进行分析评价：

- 经济发展方面：

人均 GDP：人均 GDP 是最能反应一个小康社会的经济发展方面的因素。由于经济体系所存在的问题，地区发展不平衡，这一点条件可以看成每个地区对于小康社会的建设程度完成程度的分析，其具体表达如下：

$$PGDP_j = GDP_j / P_j \quad (9)$$

其中 $PGDP_j$ 为该地区人均 GDP， GDP_j 为该区 GDP 总量， P_j 为人口数量。

- 生活质量方面：

生活质量方面，可采用恩格尔系数作为衡量标志，恩格尔系数（Engel's Coefficient）是食品支出总额占个人消费支出总额的比重。19 世纪德国统计学家恩格尔根据统计资料，对消费结构的变化得出一个规律：一个家庭收入越少，家庭收入中（或总支出中）用来购买食物的支出所占的比例就越大，随着家庭收入的增加，家庭收入中（或总支出中）用来购买食物的支出比例则会下降。推而广之，一个国家越穷，每个国民的平均收入中（或平均支出中），用于购买食物的支出所占比例就越大，随着国家的富裕，这个比例呈下降趋势。

- 社会和谐方面：

社会和谐方面，可采用城乡居民收入比进行适当的衡量以及表示，当一个地区城乡之间存在具体的差异，城乡收入比值越大，越能反映地区发展的不平衡，其表达式为：

$$B_j = C_j/V_j \quad (10)$$

其中 C_j 为该地区城市人均平均收入， V_j 为该地区农村人均平均收入。

- 民主法治方面：

民主法治方面可采用公民自身民主权利满意度来衡量，由于这一点难以量化，所以采用表格方式进行模糊划分，其划分的标准如表 1 所示：

表 1 权利满意度打分表

水平	不满意	一般	满意
打分	0-0.3	0.3-0.6	0.6-1

- 资源环境方面：

资源环境方面可采用单位 GDP 能耗程度进行分析讨论，当能耗越大，会产生对环境的伤害越大，此项表达式如式？所示：

$$E_j = CO_{2j}/GDP_j \quad (11)$$

其中 CO_{2j} ，为该地区二氧化碳的排放量。

- 文化教育方面：

文化水平的高低程度是反映文化教育方面的一大因素，可以用平均受教育年限 \bar{Y}_j 来进行表示。

以上便是指标的选取以及分析，通过上述的分析以及量化，可在此基础上进行指标评价体系的构建。

Step 2:PSR 评价体系的构建

在指标量化之后，则需要构建相应的指标评价体系进行分析讨论。考虑到整个体系的正负反馈作用，通过文献^[4]，则需要通过 PSR(压力-状态-反映) 来构建整个指标的评价体系。

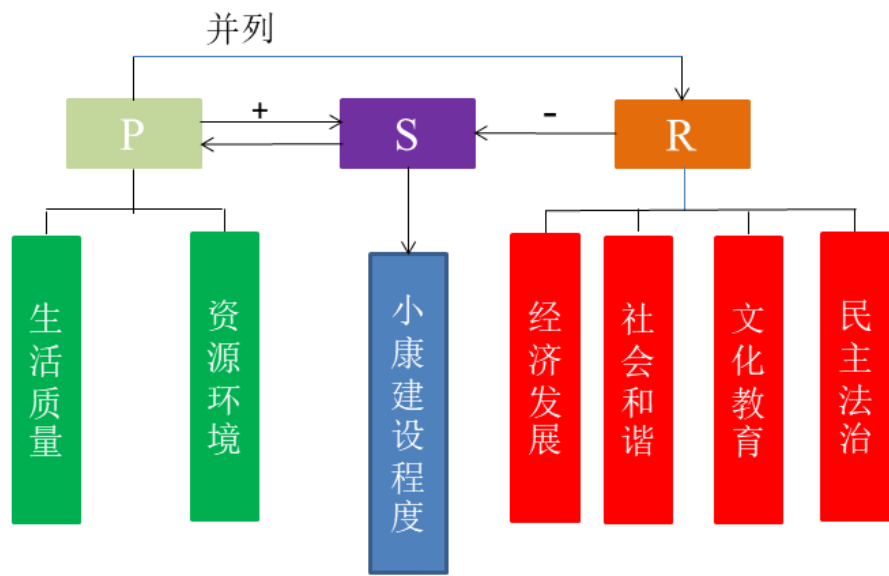


图 5 PSR 体系构建图

PSR 模型区分了 3 类指标，即压力指标、状态指标和响应指标。其中，压力指标表征对整个体系的负作用，具体可以理解为对整个疫情造成的破坏和扰动；状态指标表征特定时间阶段的疫情状态和流行性变化情况；响应指标指社会和个人如何行动来减轻、阻止、恢复和预防人类活动对疫情的负面影响，以及对已经发生的不利变化进行补救的措施。

对上述指标进行必要的逻辑分析，构建的体系如图 5 所示。如图 5 所示，最终根据第一步筛选的指标构建图如图所示。其中左侧的压力项目代表对于整个系统的压力大小，当系统所承受的压力越大，整个系统的最终的状态越差，其中压力系统项目包含了生活质量以及资源环境两大方面；而右侧的响应状态项则是起到一个负反馈的效应，因此若各省之间的经济发展、文化发展、社会和谐、民主法治项目所对应的指标数值越大，会缓解系统最终状态所承受的压力，会变得更好。

Step 3: 层次分析法加权

通过体系的构建后，则需要对象对应的因素进行加权分析，从而计算出每个地区的得分，从而对中国各个省市的评价结果进行对比。结合上述的分析，采用层次分析法 (AHP) 进行加权处理。

接下来是针对 PSR 模型对应的加权过程，对于 PSR 模型而言，其具体的加权过程则是分为 P 与 R 两个方面进行评价分析的。结合文献 [5]，考虑采用 AHP（层次分析法）对两部分分别进行加权处理。

判断矩阵 A 对应于最大特征值 λ_{\max} 的特征向量 W ，经归一化后即为同一层次相应因素对于上一层次某因素相对重要性的排序权值，这一过程称为层次单排序。上述构造成对比较判断矩阵的办法虽能减少其它因素的干扰，较客观地反映出一对因子影响力的差别。但综合全部比较结果时，其中难免包含一定程度的非一致性。如果比较结果是前后完全一致的，则矩阵 A 的元素还应当满足：

$$a_{ij}a_{jk} = a_{ik}, \forall i, j, k = 1, 2, \dots, n \quad (12)$$

需要检验构造出来的（正互反）判断矩阵 A 是否严重地非一致，以便确定是否接受。

$$A = \begin{bmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \dots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \dots & \frac{w_2}{w_n} \\ \dots & \dots & \dots & \dots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \dots & \frac{w_n}{w_n} \end{bmatrix} \quad (13)$$

n 阶正互反矩阵 A 为一致矩阵当且仅当其最大特征根 λ_{\max} ，且当正互反矩阵 A 非一致时，必有 $\lambda_{\max} > n$ 。

可以由根 λ_{\max} 是否等于 n 来检验判断矩阵 A 是否为一致矩阵。由于特征根连续地依赖于 a_{ij} ，故比 n 大得越多， A 的非一致性程度也就越严重，对应的标准化特征向量也就越不能真实地反映出 λ_{\max} 在对因素 $X = (x_1, x_2, \dots, x_n)$ 的影响中所占的比重。因此，对决策者提供的判断矩阵有必要作一次一致性检验，以决定是否能接受它。

对判断矩阵的一致性检验的步骤如下：

计算一致性指标 CI ：

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad (14)$$

查找相应的平均随机一致性指标 RI 。

对 $n = 1, 2, \dots, 9$ ，Saaty 给出了 RI 的值如表 2 所示：

表 2 一致性对照表

n	1	2	3	4	5	6
RI	0	0.1	0.58	0.90	1.12	1.24

用随机方法构造 500 个样本矩阵：随机地从 1 9 及其倒数中抽取数字构造正互反矩阵，求得最大特征根的平均值 λ_{\max} ，并定义：

$$RI = \frac{\lambda'_{\max} - n}{n - 1} \quad (15)$$

计算一致性比例 CR ：

$$CR = \frac{CI}{RI} \quad (16)$$

当 $CR < 0.10$ 时，认为判断矩阵的一致性是可以接受的，否则应对判断矩阵作适当修正。

最终的加权结果为：

$$SCORE = \sum_P w_P X_P - \sum_R w_R X_R \quad (17)$$

通过计算，六个方面最终所对应的权重分别为：生活 (0.324)、资源 (0.676)、经济 (0.6058)、社会 (0.1288)、文化 (0.2070) 与民主 (0.0584)。

Step 4: 结果分析

最终通过 Tableau 可视化化后结果如图 6 所示，通过对于图片的分析，不难发现，针对小康社会的满意度与否，东西部存在着较大差异，如东部沿海由于经济发达，教育水平较高，所以对应的结果则是十分可观，小康指数也是大幅升高。而对于西部而言，教育水平以及经济成为限制其最大的条件，所以全面小康程度是十分落后的。针对这个问题，实现东西部帮扶是一种较为可观的选择。



图 6 全国小康程度可视化图

5.3 时间序列预测模型的建立与求解

通过上述问题分析，问题 3 则是要求选取以某省为代表，从多方面因素出发，对经济可持续性进行相对应的分析以及评价预测。此问题可以看成是前两问的综合讨论，采取问题 1 的时间序列数据，在问题 2 的基础上从人口变化、老龄化趋势、城乡居民收入、资源禀赋及生态环境承载力等关键性因素角度出发，建立相关联的评价模型进行预测分析。

首先是对于指标的界定，结合文献^[6]，指标的选取应当基于相应的时间序列，选取代表性的指标，选取的指标如下：

1. 每一年人口的变化状况 (选自问题 1)。
2. 每一年人口结构变化状况 (选自问题 1)。
3. 每一年城乡居民收入比变化状况 (选自问题 2)。
4. 每一年 GDP 变化状况 (选自问题 1)。
5. 每一年单位 GDP 能耗变化状况 (选自问题 2)。

同样采用 AHP 加权，结合文献^[7]，最终通过时间序列分析对结果进行预测。

通过判断计算后的数据存在一定线性趋势，当线性趋势出现在时间序列中时，可以通过趋势移动平均法修改一阶指数平滑方法，并可以使用滞后偏差定律建立线性趋势模型，计算公式如下：

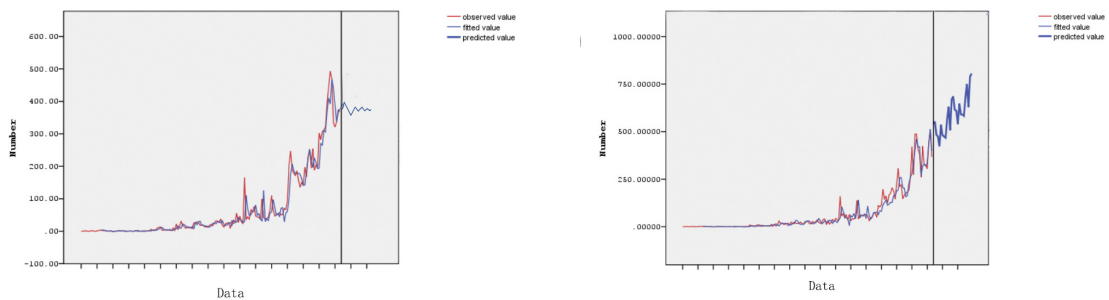
$$\begin{cases} S_t^{(1)} = \alpha y_t + (1 - \alpha)S_{t-1}^{(1)} \\ S_t^{(2)} = \alpha S_t^{(1)} + (1 - \alpha)S_{t-1}^{(2)} \\ at = 2S_t^{(1)} - S_t^{(2)} \\ bt = \frac{\alpha}{1-\alpha}(S_t^{(1)} - S_t^{(2)}) \end{cases} \quad (18)$$

在上面的公式中， $S_t^{(1)}$ 是主指数的平滑值； $S_t^{(2)}$ 是二次指数的平滑值。线性趋势模型的方程如下：

$$\hat{y}_t + T = at + btT, T = 1, 2, \dots \quad (19)$$

当时间序列 \hat{y}_t 在一定时期内具有线性趋势时，类似于趋势移动平均法，可以使用上述线性趋势模型来解决问题。

选取陕西省的相关数据，最终的预测结果如图 7 所示：



(a) 陕西省短期可持续预测

(b) 陕西省长期可持续预测

图 7 预测结果图

对于短时间实行可持续发展战略而言，总体的评分会变得比较平稳短时间内收效甚微，但是长时间而看，可持续发展则会起到较为长远的效果。

5.4 0-1 整数规划模型的建立与求解

问题 4 要求提供一个必要的合作计划，通过该合作计划，从而对双方的经济合作从而起到必要的促进作用。通过分析不难发现，这个问题实质上是在问题 2 的基础上进行进一步地讨论以及改进优化，这一点可以看做双方经济选择的变化。最终结果为给出省市相对应的合作方案。通过必要的问题分析不难发现，这一点实质上是对于选择问题的一种变式，可以通过建立 0-1 整数规划模型来构建相对应的合作网络，从而对其合作方案予以给出以及提升。

给定选取地区小康指数集合 L :

$$L = \{L_1, L_2, \dots, L_{31}\} \quad (20)$$

此集合表示合作前相对应的指数模式，其中所对应的角标 1 到 31 反映出了每个省市合作前的对应小康指数值。

针对选择而言，构建选择函数：

$$C_{ij} = \begin{cases} 1, Yes \\ 0, No \end{cases} \quad (21)$$

考虑到地域问题，可限制选择的条件为，合作的选择双方必须为邻省，否则会因为地域问题消除选择。

未体现互惠互利规则，选择后双方会产生一种共同进步的趋势，其中主要体现二者幸福差值越大所对应的提升越大，定义选择后为二者总提升为原有差值的百分之三十 [8]。

$$\Delta L_{ij} = |L_i - L_j| \times 30\% \quad (22)$$

通过上述定义，可考虑目标函数为选择后的总收益最大：

$$z = \max \sum_{i,j} \Delta L_{ij} C_{ij} \quad (23)$$

以下为约束条件的考虑：

- 对于每一个省所选择的合作省份不能超过一个：

$$\sum_{i=1}^{33} C_{ij} \leq 1 \quad (24)$$

- 每个省份的双向选择只计算一遍：

$$C_{ij} = C_{ji} \leq 1 \quad (25)$$

综上所述，该优化模型表示为：

$$z = \max \sum_{i,j} \Delta L_{ij} C_{ij}$$

$$s.t. \begin{cases} \sum_{i=1}^{33} C_{ij} \leq 1 \\ C_{ij} = C_{ji} \leq 1 \\ \Delta L_{ij} = |L_i - L_j| \times 30\% \\ C_{ij} = \begin{cases} 1, Yes \\ 0, No \end{cases} \end{cases} \quad (26)$$

为对问题进行简要说明，通过模型求解实现经济合作的省市地区前七个结果如表 3 所示：

表 3 城市合作互助表

序号	城市 A	城市 B
1	四川	西藏
2	甘肃	宁夏
3	北京	河北
4	广东	广西
5	陕西	宁夏
6	江苏	安徽
7	湖南	江西

通过对于表 3 的简要分析不难发现，所选取的帮扶城市往往是存在东西分界线或者南北分界线上。造成一点原因的不同便是因为东西部经济差距过大，造成发展不平衡。往往是东部的城市对于西部的邻省进行帮扶，从而缓解这一项发展存在的弊病。

六、模型优缺点

6.1 模型优点

- 1 充分考虑了每一个小问之间的关系，对于指标的选取较为透彻。
- 2 对于问题 1 所建立的模型，从数据以及文献综述两个角度出发，结果较有说服力。

- 3 对于问题 2 中所建立的评价模型,充分考虑了指标评价体系的存在,运用 PSR 反映了指标之间的构建情况。
- 4 对于问题 4,采用了 0-1 整数规划模型,比较全面且透彻的说明了相关的问题,数学语言描述精炼。

6.2 模型缺点

- 1 数据查找筛选方面,数据样本量由于网站的原因,比较少,使结果的说服力产生了影响。
- 2 问题 4 中所建立的模型,只考虑到了邻省的变化,并未进一步的分析介绍。
- 3 层次分析法加权过于主观性,由于工作量的不足,使得加权结果正确性无法得到保证。
- 4 问题 2 中所选取的 6 大指标过于片面,并未从更加详细的角度对结果进行探讨分析。

参考文献

- [1] <http://data.stats.gov.cn/>
- [2] 王利利. 甘肃省定西市全面建成小康社会实现度测评分析 [J]. 河北农业大学学报 (社会科学版),2020,22(01):101-107.
- [3] 高静, 武彤. 突发公共卫生事件的政府响应机制与效能诊断:PSR 模型的应用 [J]. 重庆社会科学,2020(05):6-18.
- [4] 谢再统. 广西全面建成小康社会进程预测及短板分析 [J]. 广西经济,2019(11):28-30.
- [5] 彭贞贞, 马骥. 湖南省贫困县域全面建成小康社会进程的评价与比较分析 [J]. 现代化农业,2019(10):52-55.
- [6] 张倩. 甘肃省全面建成小康社会进程研究 [D]. 兰州财经大学,2019.
- [7] 胡俊航, 常向丽, 史小俊. 河南省全面建成小康社会水平评价与区域差异分析 [J]. 黄冈职业技术学院学报,2016,18(02):82-85.
- [8] 负晓晴. 基于层次分析法全面建设小康水平的有效性评价 [J]. 电脑知识与技术,2010,6(18):5039-5040.

附录 A 层次分析法

```

%层次分析法 (AHP)
disp('请输入判断矩阵A (n阶) ');
A = input('A=');
[n,n] = size(A);
x = ones(n,100);
y = ones(n,100);
m = zeros(1,100);
m(1) = max(x(:,1));
y(:,1) = x(:,1);
x(:,2) = A*y(:,1);
m(2) = max(x(:,2));
y(:,2) = x(:,2)/m(2);
p=0.0001; i=2; k=abs(m(2)-m(1));
while k>p
i=i+1;
x(:,i) = A*y(:,i-1);
m(i) = max(x(:,i));
y(:,i) = x(:,i)/m(i);
k=abs(m(i)-m(i-1));
end
a = sum(y(:,i));
w = y(:,i)/a;
t = m(i);
disp(w);
%一致性检验
CI = (t-n)/(n-1);
RI = [0 0 0.52 0.89 1.12 1.36 1.41 1.46 1.49 1.52 1.54 1.56 1.58 1.59];
CR = CI/RI(n);
if CR<0.10
disp('此矩阵一致性可以接受! ');
disp('CI=');disp(CI);
disp('CR=');disp(CR);
end

```

附录 B 0-1 整数规划

```

model:
sets:
first/1..13/:u;

```

```
!second/1..168/:v;
link011(first,first):C011,D011;
endsets
min = @sum(link011(i,j)|i#ne#j:D011(i,j)*C011(i,j));
n = @size(first);
@sum(first(j)|j#gt#1:C011(1,j))>=1;
@for(first(i)|i#ne#1:@sum(first(j)|i#ne#j:C011(j,i))=1);
@for(link011(i,j)|i#ne#j:u(i)-u(j)+n*C011(i,j)<=n-1);
@for(link011(i,j):@bin(C011(i,j)));
data:
D011 = @file('D:\1AAA_programmer\lingo\2020xiaosai\D011_t.txt');
enddata
end
```