# 景气建模逻辑分析文档

## 初选指标

确定初始需要纳入分析的所有指标。

宏观经济指标：

投资情况指标：

政务大数据指标：

企业指标：

## 数据预处理

1. 数据类型转换。

将所有指标数据转换为np.float，日期转换为np.datetime。

1. 对时序频度转换。

确定统一的频度（年/月/天）；高频转为低频，确定是对数据进行聚合还是采样（时间窗的第一个值、最后一个值）；低频转为高频，缺失值填补：向前填充法、向后填充法、均值填充法。向前填充如果第一个时间戳是缺失时仍缺失；均值填充用整个序列的均值来填充缺失值，填充之后不存在缺失值。

1. 定基价格计算。

暂无

1. 季节调整。

首先检查各个时序数据的稳定性，adf单位根检验，若P值<置信度（置信度取0.05），拒绝原假设，即数据是稳定的，否则是非平稳的。若不具有稳定性，通过因素分解，剔除季节因素和不规则因素（因素分解剔除季节项和不规则项，保留趋势和循环项）。

1. 确定是否需要尺度变化。

对数变换、归一化/标准化变换，确定指标单位。

## 指标遴选

确定基准指标。考察其它基准指标的循环与基准指标的关系确定哪些指标属于先行、一致、滞后指标。（需要确定先行、一致、滞后选取标准）

计算出各个指标与基准指标的时差相关系数和K-L信息量(K-L信息量中对进行绝对值处理)。计算不超过最大延迟数（最大延迟数取12）的所有时差相关系数和K-L信息量，然后选择最佳的系数和对应的延迟数，时差相关系数取绝对值的最大值，K-L信息量取最小值。

时差相关系数计算公式：

K-L信息量计算公式：

## 指标精选

## 合成指数计算

将选出的先行指标、一致指标、滞后指标标识分组后统一纳入。（指标分组为：先行组、一致组、滞后组）

1. 求指标的对称变化率并将其标准化

计算各组的指标对称变化率若指标序列是比率序列或指标中存在非正数值时，对该指标取一阶差分作为。将指标对称变化率进行标准化处理，首先计算标准化因子，用标准化因子来标准化得到标准化变化率。

或

1. 求各指标组的标准化平均变化率

计算各指标组的标准化平均变化率，确定权重首先计算各组的平均变化率，然后根据算标准化因子，用标准化因子来标准化各组的平均变化率得到。(确定各组指标的权重)

则：

1. 计算初始合成指数。
2. 进行趋势调整

对一致指标组的每个序列使用复利公式分别计算各自的平均增长率。然后对一致指标组中所有序列的平均增长率求平均得到一致指标组的平均增长率。（循环的确定方法：采用Bry-Boschan算法）

然后对先行、一致、滞后的初始合成指数分别用复利公式求出他们各自的平均增长率。

最后对三个指标组的标准化平均变化率作趋势调整。

1. 计算三个指标组合成指数

将各年份的指标变化率相当于基准年份进行计算合成指数。同第三步。

最后得到，

1. 移动平均

对上一步得到的先行指数、一致指数、滞后指数进行一次三项移动平均。得到最终的合成指数。

## 投资景气预警

确定预警指标

确定预警界线值，并计算各预警指标的预警得分。

将所有预警指标的预警得分求和得到综合指数。

确定综合指数的界限值（红黄蓝信号灯）