

黄金期货基本面量化择时策略研究

——使用主成分分析方法分析基本面数据

2020 年 3 月 3 日

摘要：

- 本文首先提取了影响黄金价格的 10 个基本面变量，分别为美国 10 年期国债收益率、美国 10 年期国债实际收益率、美元指数、VIX 恐慌指数、美国失业率、美国新增非农就业人数、美国 CPI、美国 PPI、美国 ISM 制造业 PMI、各国央行和其他机构黄金需求。并采用 Z-Score 的方法对基本面数据进行处理。
- 得到统一量纲的数据后，我们对数据进行 PCA（主成分分析法）处理，将 10 维数据降维成 2 维数据，提取基本面数据中主要的影响因素。并分析降维后变量和原始变量的关系，得出对基本面影响较大的两个变量为美国 10 年期国债收益率和 CPI。
- 最后使用降维后的变量做成综合指标对黄金期货的历史数据进行量化回测。

量化策略研究团队

分析师：吴广奇

从业资格号：F3048078

投资咨询号：Z0014024

电话：021-50586261

邮箱：wugq@sywgqh.com.cn

分析师：林新杰

从业资格号：F3032999

投资咨询号：Z0014722

电话：021-50586279

邮箱：linxj@sywgqh.com.cn

申银万国期货有限公司

地址：上海东方路 800 号

宝安大厦 7、8、10 楼

邮编：200122

电话：021 5058 8811

传真：021 5058 8822

网址：www.sywgqh.com.cn



申银万国期货
宏观金融研究



申银万国期货研究

1、黄金价格影响因素

黄金作为一种特殊的大宗商品，具有商品、货币和投资避险的多重属性。我们从黄金的供需、经济数据、金融资产三个维度中选一些基本面因子对黄金价格的影响因素进行分析。

1) 供需

作为一个大宗商品，黄金有着作为金属的供给模式，上游供给量受到金矿开采以及废金回收的影响。而主要的下游需求可分为制造业需求（珠宝首饰、工业材料等）和金融投资需求。此外像央行这样的大型市场参与者的售/买金行为也会影响市场供需格局。

2) 经济数据

黄金作为金融投资品，受到所处时期的经济周期影响，比如滞涨时期黄金的表现通常较好，在这些时期投资者也更愿意将资金配置在黄金上。经济数据除了描绘当前的经济图景外，也会影响各国央行的货币政策，而以美联储为代表的大国央行的货币政策也对同时期黄金的价格走势产生比较大的影响。

3) 金融资产

一方面，一些金融资产的变化可以直接影响黄金的价格和投资者偏好，比如在利率较低的时期，黄金的持有成本相对较低，配置价值也相对较高。外汇的波动会直接影响以该货币计价的黄金价格等。另一方面，一些金融资产和黄金受到相同影响因素的影响，价格波动有着较高的相关度，并且能够更好更快地反映当时市场的经济、金融氛围。

由上，我们最终选取十个相关度相对较高的基本面的指标进行回溯：美国 10 年期国债收益率、美国 10 年期国债实际收益率、美元指数、VIX 恐慌指数、美国失业率、美国新增非农就业人数、美国 CPI、美国 PPI、美国 ISM 制造业 PMI、各国央行和其他机构黄金需求。

2、数据处理

商品期货的基本面数据具有种类多、时间离散、公布时间不规律等特点，需要对其进行处理，处理的主要目的是将其转化为可以对比分析的数据。

1) 数据分类

在上述基本指标中，从获得的数据公布频率来看：

日数据：美国 10 年期国债收益率、美国 10 年期国债实际收益率、美元指数、VIX 恐慌指数是日数据。

月数据：美国失业率、美国新增非农就业人数、美国 CPI、美国 PPI、美国 ISM 制造业 PMI。

季度数据：各国央行和其他机构黄金需求。

在数据处理时要注意美国公布时间和国内时间的时区差异，
月度和季度数据公布的滞后性。

2) Z-Score

由于基本面数据量纲不统一，我们需要将其处理成无量纲并且可对比的数据。在基本面分析中，我们一般采用 Z-Score 的方法对基本面数据进行处理。

具体处理方法为，将原始数据减去一段时间的均值再除以这段时间数据的标准差。Z-Score 数据能够真实的反应一个分数距离平均数的相对标准距离，能确实反应原始数据的波动率信息。

3、 量化回测

在模型构建方面，我们首先用 PCA(主成分分析)方法对基本面数据进行降维处理，提取较为精简的基本数据信息。再分析降维后的数据变量和原始数据的关系，寻找数据逻辑，并将降维后的数据变量做成交易指标。

1) PCA

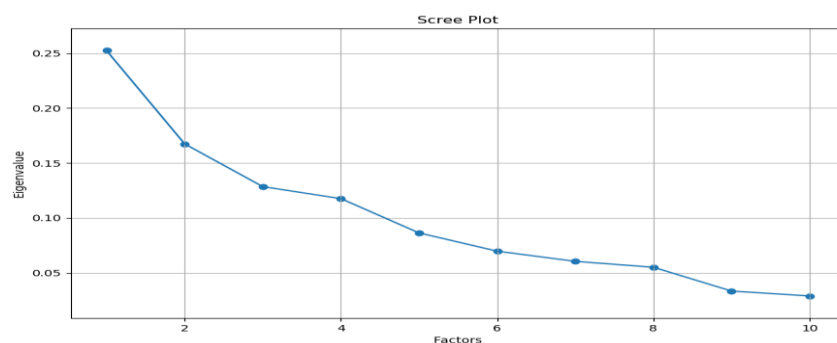
主成分分析法 (Principal Component Analysis, PCA) 是一种运用线性代数的知识来进行数据降维的方法。它将多个变量转换为少数几个不相关的变量来，但转换后的变量能比较全面地反映整个数据集。这是因为数据集中的原始变量之间存在一定的关系，可用较少的综合变量来综合各原始变量之间的信息。

由于黄金的基本面数据本身存在较多的内在关系，比如通胀会影响债券的收益，影响就业率等等，通过数据的降维，我们能提取较为精简的数据信息。

2) 变量选取

我们首先要确定降维后的变量个数，我们通过对其波动变量方差占比进行分析，选出对 10 个指标波动贡献较大的变量个数，从波动的贡献来看：

图 1：变量波动方差贡献分布



资料来源：Wind，申万期货研究所

从图中可以看出，前波动率贡献最大的四个变量对数据整体波动率贡献分别为 25.24%、16.74%、12.85%和 11.76%，均大于 10%，考虑到后面两个影响较少和尽量精简模型输入，我们选取最为重要的两个变量。

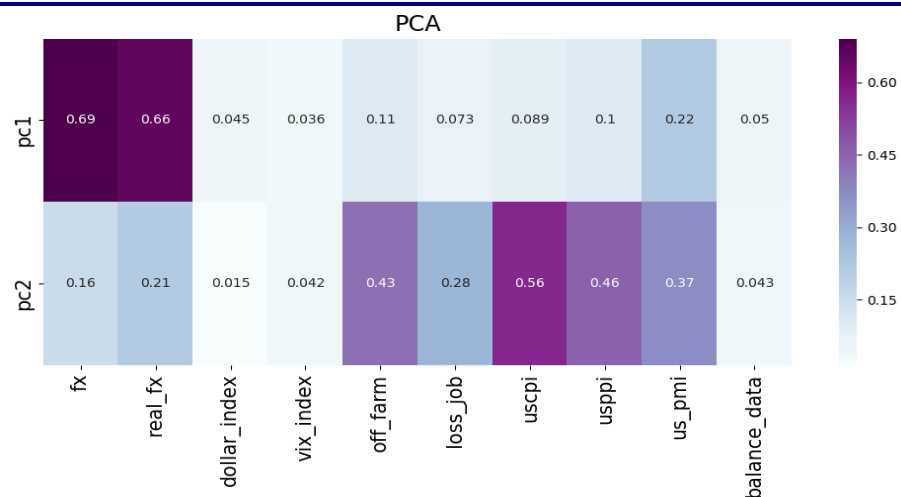
确定好最终的变量个数后，我们对数据进行降维处理，降维后得到两个新的数据序列。

3) 变量分析

为方便识别，降维后的变量序列我们称为 principal_component1 和 principal_component2，简称 pc1 和 pc2。对原始基本面数据变量进行编码，期对应列表如下：

'fx': '美国 10 年期国债收益率',
 'real_fx': '美国 10 年期国债实际收益率',
 'dollar_index': '美元指数',
 'vix_index': 'VIX 恐慌指数',
 'loss_job': '美国失业率',
 'off_farm': '美国新增非农就业人数',
 'uscpi': '美国 CPI 环比',
 'usppi': '美国 PPI 环比',
 'us_pmi': '美国 ISM 制造业 PMI',
 'balance_data': '各国央行和其他机构黄金需求量'。

图 2：原始变量和降维后变量关系（绝对值）



资料来源：Wind，申万期货研究所

从参数的热力图上可以看出，和降维后数据关系较为密切的变量为美国 10 年期国债收益率、美国 10 年期国债实际收益率、美国 CPI 环比和 PPI 环比等数据，这说明在这个 10 个基本面的变量中对单个变量对其他的变量影响较大的为以上四个变量。

在实际分析中，美国 10 年期国债收益率、美国 10 年期国债实际收益具有较高的相关性，CPI 和 PPI 长期走势趋于统一。

因此对基本面影响较大的两个变量可以概括为美国 10 年期国债收益率和 CPI。

部分数据像各国央行和其他机构黄金需求量公布频率较低，经过数据填充处理后，整体波动不大，弱化了它对其它变量的影响。美元指数和 VIX 指数尽管公布频率较高，但整体对基本面数据影响较小。

4) 量化回测

我们将降维后的两个变量进行处理，构建与黄金价格的关系，其较为直接的想法是对两个变量进行赋权，构建成一个包含两变量的线形变量。

我们将综合变量设计成一个与黄金期货价格正相关的变量。从参数的热力图上可以看出，principal_component1 与美国 10 年期国债收益率相关度较高，由此可以判断它与黄金价格为负相关关系；principal_component2 与美国 CPI 和 PPI 相关度较高，由此可以判断它与黄金价格为正相关关系。在构建变量时不能简单的做权重分布，还要考虑方向问题。

图 3：降维后变量走势



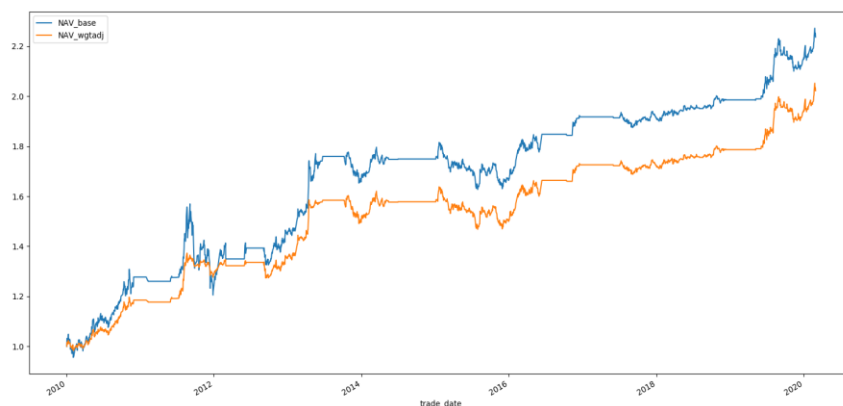
资料来源：Wind，申万期货研究所

由于前期做了数据的 Z-Score 处理，数据本身具有均值回归的特征，类似布林带指标，我们设定一个阈值，当综合指标低于阈值的反数时，做多黄金期货，在综合指标高于阈值时，做空黄金期货。为充分反映趋势，当做多时，综合指标高于阈值时平仓，当做空时，综合指标低于阈值的反数时平仓，平仓和开仓不同时进行，不做止损和止盈。

数据回测采用申万期货黄金期货指数，资金不加杠杆，策略的开平仓费率设为 0.02%，数据结果供投资者参考。

我们用年化波动率对收益净值进行仓位控制，控制目标为年化波动率为 10%，其收益走势如下：

图 4 收益净值和仓位控制后净值



资料来源：Wind，申万期货研究所

策略近 10 年时间共交易 58 次，做多 38 次，做空 20 次，持有期波动较大，最近一次交易为 2019 年 6 月 10 日收盘做多，多单持有至今。策略的收益绩效为：

表 1：策略净值绩效表

Index	NAV_base	NAV_wgtadj
累计收益	123.64%	102.17%
年化收益	8.22%	7.15%
胜率	52.63%	52.63%
盈亏比	1.2081	1.2582
最大回撤	23.20%	10.36%
年化波动率	10.62%	6.89%
年化下侧波动率	10.13%	5.33%
夏普比率	0.4913	0.6019
索提诺比率	0.5147	0.7788
Calmar比率	0.3541	0.6901

资料来源：Wind，申万期货研究所

4、 总结

本文对黄金期货的基本面数据进行分析，提取了对黄金影响较大的 10 个变量，使用 PCA(主成分分析)方法对数据进行降维处理，并使用降维后的数据构建黄金期货的量化择时策略，通过回测，我们发现基本面数据对黄金的择时具有一定的效果，在不加杠杆的情况下，总体年化收益在 7%至 8%。在数据降维后我们发现其基本面数据较为核心的影响因素为美国 10 年期国债收益率和 CPI，而美元指数和 VIX 指数对基本面数据整体影响较弱。

免责声明

本报告的信息均来源于公开资料，本公司对这些信息的准确性和完整性不作任何保证，也不保证所包含的信息和建议不会发生任何变更。我们力求报告内容的客观、公正，但文中的观点、结论和建议仅供参考，报告中的信息或意见并不构成所述品种的买卖出价，投资者据此做出的任何投资决策与本公司无关。

本报告所涵括的信息仅供交流研讨，投资者应合理合法使用本报告所提供的信息、建议，不得用于未经允许的其他任何用途。如因投资者将本报告所提供的信息、建议用于非法目的，所产生的一切经济、法律责任均与本公司无关。

本报告的版权归本公司所有。本公司对本报告保留一切权利，未经本公司书面许可，任何机构和个人不得以任何形式翻版、复制和发布。如引用、刊发，需注明出处为申银万国期货，且不得对本报告进行有悖原意的引用、删节和修改。