arXiv:2602.03325v1 [q-fin.PM] 3 Feb 2026

投资组合构建的新方法

T. Di Matteo[1,2,3],L. Riso* [4]和 M.G. Zoia[4]

[1]伦敦国王学院数学系，斯特兰德，伦敦，WC2R 2LS，英国。[2]维也纳复杂性科学中心，Josefst¨adter Straße 39，1080 维也纳，奥地利。[3]恩里科·费米研究中心，潘尼斯佩尔纳街 89A，罗马，00184，意大利[4]圣心天主教大学经济政策系，米兰，意大利

2026年2月4日

摘要

本文提出一种基于机器学习的资产选择与投资组合构建框架，称为最优路径算法稀疏图模型（BPASGM）。该方法扩展了最佳路径算法（BPA），将依赖关系发现转化为结构化且具有经济学依据的资产选择程序。BPASGM识别大量金融资产之间的线性与非线性依赖关系，并通过满足结构马尔可夫性质的稀疏图形模型加以表示。基于这一表示，BPASGM系统地筛除彼此正相关或存在冗余关联的资产，找出彼此条件独立或负相关的资产子集。此筛选步骤旨在增强分散化，并降低高维投资组合问题中的估计误差。随后，使用标准均值—方差技术，在选定的资产子集上构建投资组合。BPASGM并非旨在超越总体参数已知时的理论均值—方差最优解，而是力求改善有限样本下的已实现投资组合绩效；众所周知，在这种情况下，经典的基于样本的 Markowitz 实现对估计误差高度敏感。蒙特卡洛模拟表明，与标准的基于样本的均值—方差投资组合相比，基于 BPASGM 构建的投资组合具有更稳定的风险—收益特征、更低的实现波动率以及更优的风险调整后绩效。对1990—2025年美国股票、全球股票指数和外汇汇率的实证应用证实了这些发现，并突显了该方法大幅降低投资组合基数的效果。总体而言，BPASGM提供了一个具有统计学依据且计算高效的框架，连接了稀疏图形建模与投资组合理论，并作为原始 BPA 在依赖关系感知型投资组合构建方面的重要方法论扩展。

关键词：资产选择；投资组合优化；有效前沿；稀疏图模型；机器学习；BPASGM。

*JEL*分类：*C63* – 计算技术；仿真建模；*G11* – 投资组合选择；投资决策；*G17* – 金融预测与仿真；*C58* – 金融计量经济学；*C61* – 优化技术；规划模型；动态分析。

1 引言

投资组合选择是现代金融学的核心，也是投资理论中经久不衰且受到深入研究的问题之一。自 Markowitz 的开创性研究以来，Markowitz [1952b], 均值—方差框架一直是构建最优投资组合的经典基础，它通过优化资产权重，将预期收益与风险之间的权衡形式化。尽管均值—方差优化在概念上优雅且影响深远，但在实践中仍面临挑战。投资组合结果对预期收益的估计高度敏感，

*通讯作者。电子邮箱：luigi.riso@unicatt.it

尤其对协方差矩阵的估计和求逆高度敏感。即使估计误差相对较小，也可能导致最优投资组合权重大幅波动 [Braga 等人， 2025], 资产配置不稳定，样本外表现不佳 [Michaud 和 Michaud，2007，Olmo，2023]。在高维情形下，这些困难会进一步加剧，因为资产数量相对于可用的时间序列观测值而言较多。

过去几十年来，大量文献致力于通过提高投资组合输入的统计稳健性，或直接改进优化过程，来缓解这些局限。

主要方法包括协方差矩阵的收缩估计量 [Ledoit 和 Wolf，2003a，b],稳健型和基于重抽样的投资组合方法 [Michaud 和 Michaud，2007，Martin 等，2010年，Gre- gory 等， 2011], 贝叶斯框架，例如 Black–Litterman 模型 [Black 和 Litterman， 1992],以及稀疏型或惩罚型投资组合形式 [Broadie，1993，Lefebvre 等， 2020]。尽管这些方法能够显著提高数值稳定性，但从根本上说，它们仍依赖于二阶矩的估计，而且在许多情况下还依赖于大型协方差矩阵的求逆或谱分解。因此，面对复杂依赖结构、有限样本量和含噪数据时，这些方法仍然面临挑战。

与此同时，人们提出了有效前沿的若干扩展，以减少过拟合并提升投资组合稳健性 [Markowitz 等人，2022年，Hakem， 2025]。这些方法通常依靠限制投资范围、对投资组合权重施加额外约束，或采用正则化技术来稳定估计。
模拟研究和回测的实证证据表明，这类经稳健性增强的前沿在有限样本中可能优于经典马科维茨前沿：在风险水平相同的情况下实现更高收益，或在目标收益相同的情况下实现更低风险 [Bodnar 等人，2022年，Zhou， 2024]。重要的是，这些改进应被视为有限样本稳健性的提升，而非对总体均值—方差前沿所体现的基本风险—收益权衡的违背。

投资组合构建中一个关键却常被忽视的脆弱性来源，在于资产选择过程本身。估计收益、协方差和风险度量的质量，极大程度上取决于资产集合的构成。纳入冗余资产、受相似潜在风险来源驱动的资产，或对增量分散化贡献甚微的资产，可能会放大估计误差并降低投资组合绩效。因此，改进资产选择是鲁棒投资组合构建不可或缺的一环，因为它直接影响估计问题的条件数以及由此产生的有效前沿的稳定性。

资产选择方法因投资目标、数据可用性和投资者偏好而有很大差异。现有方法包括收益导向型选择策略 [Tu 和 Li，2024，Lassance 和 Vrins， 2023], 风险最小化准则 [Campbell 等人，2001，Fabozzi 等人，2008], 以及基于环境、社会和治理（ESG）考量的筛选程序 [Garcia-Bernabeu 等人，2024，M¨uller 和 Joubrel， 2025]。
更广泛而言，资产选择技术涵盖理论驱动的优化、基于因子的筛选，以及启发式规则和自由裁量或主动管理策略 [Brown 和 Smith，2011，Carroll 等人，2017年，Jacobs 等人， 2014]。在实践中，人们通常采用结合多项标准的混合方法。量化资产选择在这一领域中占据核心地位。
量化方法依赖计量经济学模型、统计技术和计算算法，根据可观测特征对资产进行排序或评分。

常用因子包括估值指标，例如市盈率 [Block，1995年，Ra-madan， 2009], 市净率 [Nezlobin 等人，2016], 企业价值倍数 [Mauboussin，2018年，Olbert， 2025], 动量指标、质量指标以及波动率、贝塔和回撤等风险度量指标。
与自主型策略不同，量化方法明确以数据为驱动，旨在减少资产选择中的主观判断。近年来，高维金融数据集的快速增长以及计算方法的进步，推动了人们对运用机器学习技术进行资产选择和投资组合构建的兴趣日益浓厚 [Mirete-Ferrer 等人， 2022]。

许多机器学习方法侧重于收益预测、特征提取或因子发现[Barfuss 等人，2016年，Raddant 和 Di Matteo， 2023]。然而，相比之下，较少研究关注对资产间依赖结构进行显式建模，而这种结构在分散化和风险管理中发挥着核心作用。尤其是，了解资产之间的条件关联方式——超越简单的两两相关性——对于识别冗余和提高投资组合稳健性至关重要。

本文 提出了一种直接应对这一缺口的新颖方法论。具体而言，我们引入了 e

最优路径算法稀疏图模型 （BPASGM）这是一种依赖关系感知框架，用于资产选择，融合了信息论、图模型和机器学习的理念。BPASGM扩展了最初由 Riso 等人 [2023]提出的最佳路径算法（BPA）。

图模型为表示随机变量之间复杂的依赖结构提供了一种有原则的方法，其中节点对应变量，边则编码条件依赖关系[Lauritzen 和 Wermuth，1989年，Jordan 等人， 2004]。

在高维环境中，稀疏性至关重要，因为潜在相互关系的数量会随着变量数量呈二次方增长 [Barfuss 等人， 2016]。稀疏图模型，例如图形套索 [Friedman 等人， 2008], 贝叶斯正则化方法 [Gan 等人， 2019], 以及相关方法已广泛应用于遗传学、气候科学和金融等领域[Bleik，2013年，Zerenner 等人，2014年，Aste 和 Di Matteo，2017年，Aste 等人，2010年，Musmeci 等人， 2014]。在金融应用中，这些模型增强了可解释性，并缓解了大型协方差矩阵带来的计算难题。然而，许多现有方法依赖于精度矩阵估计，局限于线性依赖，并且主要在成对层面评估关系。BPASGM 在几个重要方面不同于这些方法。首先，它不会将协方差矩阵的求逆或特征分解作为优化对象。其次，它利用信息论度量，同时处理线性和非线性依赖。第三，它构建有向无环图表示，并通过设计满足条件独立性（马尔可夫型）性质。通过为每项资产识别出一组能够共同解释其行为的最小预测变量，BPASGM 系统地剔除冗余或正向连接的资产，并分离出条件独立或负相关的资产子集。需要强调的是，BPASGM 并非作为替代偏好模型提出，也不是为了取代均值—方差优化。相反，应将其视为一种预处理和筛选工具，在投资组合构建之前重塑资产集合。在已知真实均值向量和协方差矩阵的总体情形下，在相同的二次效用下，限制资产集合无法优于理论上的 Markowitz 最优解，因为不存在估计问题，而且较小的可行集不可能优于较大的可行集。然而，BPASGM 并不在这种预言机环境中运行，也不以固定估计值在缩小后的可行集上求解同一个样本均值—方差问题。相反，BPASGM 引入数据依赖型筛选步骤，从而重塑估计问题本身：根据资产的依赖结构进行选择，在所选子集上重新估计投资组合矩，随后使用这些经过选择条件化的估计值进行均值—方差优化。通过剔除冗余和正依赖资产，并在依赖关系特征更有利的子集上重新估计投资组合矩，BPASGM 降低了有效维数，改善了矩估计量的条件数，并增强了样本外已实现投资组合结果的稳定性和稳健性，同时不改变基本的风险—收益权衡。

本文展示的蒙特卡洛模拟和实证应用表明，与基于完整资产集合构建的投资组合相比，经过 BPASGM 资产选择后构建的投资组合具有更好的分散化特性、更稳定的已实现夏普比率和索提诺比率，以及对估计误差更低的敏感性。因此，诸如已实现风险—收益斜率局部为负等表面特征，应被视为有限样本稳健性效应，而非理论优势或偏好逆转的证据。此外，BPASGM 自然能够显著降低投资组合基数，从而在不对投资组合权重施加临时性稀疏性约束的情况下，构建更简单、更易于解释的投资组合。本文的贡献兼具方法论和实践意义。在方法论方面，本文提出了一种以图模型和信息论为基础的依赖感知型资产选择框架。在实践方面，本文提供了一种计算效率高、模块化的工具，可与标准均值—方差优化相结合，也可用于兼顾下行风险或高阶矩的其他投资组合目标。本文其余部分安排如下。第 2节介绍 BPASGM 方法论及其理论特性，包括该算法所形成的条件独立结构。第3 节通过蒙特卡洛模拟，展示基于 BPASGM 选择的资产构建有效前沿。第 4 节介绍用于评估所得投资组合分散化和波动性特征的指标。第5 节报告基于美国股票、全球股票指数和汇率的实证结果，包括样本外比较和稳健性检验。第 6 节总结全文，并讨论其启示及未来研究方向。

2 稀疏图模型的最佳路径算法

2.1 BPASGM程序概述及其相较于BPA的扩展

BPASGM 被设计为一个两阶段程序。第一阶段，对大规模资产集合应用依赖感知筛选程序，以识别资产之间潜在的依赖结构。此步骤依托图论和信息论工具，刻画资产收益率之间的条件关系和带符号关系，从而得到依赖关系的稀疏有向表示。根据这一结构，剔除冗余性强或呈正向依赖的资产，最终得到依赖关系特征更优的精简资产子集。重要的是，这一筛选步骤会重塑资产集合本身，并不涉及投资组合优化。

第二阶段是在所选子集上，使用标准均值—方差准则构建投资组合权重。针对缩减后的资产集合，重新估计预期收益和协方差，并通过求解经典的均值—方差优化问题来确定投资组合权重，同时遵守全文采用的相同可行性约束。因此，BPASGM并不改变投资者偏好或投资组合目标，而是通过依据数据驱动的资产选择进行条件化，改变估计问题。

这一区别对于解读本文结果至关重要。BPASGM并非旨在改进基于已知总体矩定义的理论马科维茨最优解，也不等同于在固定估计值下、于缩减后的可行集上求解相同的样本均值—方差问题。相反，筛选步骤通过降低维度并依据观测到的依赖结构进行条件化，改变了投资组合构建所用估计量的统计性质。因此，相关比较应针对在不同估计问题下构建的投资组合，而不是针对具有不同可行集的相同优化问题。

第一阶段所观察到的资产集合大幅缩减，与低维依赖结构的存在相一致，而因子模型通常正是利用这一特征来正则化协方差估计。然而，BPASGM 与基于因子的方法存在根本差异。因子模型施加全局线性结构以稳定协方差矩阵，同时保留完整的资产集合；BPASGM 则进行数据驱动的依赖感知型资产选择，明确利用条件关系与带符号关系，识别并剔除共同变异的冗余来源。从这个意义上说，BPASGM 是在优化之前重塑资产集合，而非仅仅正则化二阶矩估计。为从实证上区分这些效应，本文在相同的滚动样本外评估方案下，纳入了与基于因子的协方差基准的比较。

BPASGM 也以实质性方式扩展了原始 BPA。BPA 如 Riso 等人[2023], 所述，是一种逐节点程序，针对给定的目标变量，依据信息论准则识别出能够共同解释其行为的最小预测变量子集。一个自然的扩展是将 BPA 对称地应用于资产集合中的每一项资产，从而得到预测关系的有向图。虽然这种对称应用是必要步骤，但对于投资组合构建而言并不足够。具体而言，所得图可能包含强正向依赖的资产稠密簇，反映出共同变异来源，并从分散化角度导致冗余。

BPASGM 超越了对称 BPA，通过引入额外机制，主动利用所检测到的依赖结构进行资产选择。具体而言，BPASGM 构建了一种有符号图表示，用于区分正依赖关系与负依赖关系，并引入明确的剪枝规则，以打破经济意义上不理想的正依赖连接。这些规则旨在消除冗余资产，同时保留能够贡献不同风险来源的资产。因此，BPASGM 不仅能够检测依赖关系，还利用这些关系以一种与投资组合构建直接相关的方式重塑资产集合。

最后，由于投资组合矩必须在有限样本中进行估计，已实现投资组合结果可能不同于总体均值—方差前沿的理论性质。因此，诸如负的已实现风险—收益斜率等表面特征，应被解读为有限样本稳健性效应，而不是对经典风险—收益权衡的违反。从这个意义上说，BPASGM 旨在通过依赖感知型资产选择减轻估计误差，在不改变均值–方差投资组合理论基本经济原则的情况下，提高已实现投资组合绩效的稳定性与稳健性。

在概述 BPASGM 的目标和结构后，我们现在回顾最佳路径算法（BPA）的核心思想 Riso 等人 [2023], 它构成了所提出的稀疏图框架的基本构建模块，其扩展和调整是

构建 BPASGM 的核心。考虑一个由 n 个观测值和 p 个变量组成的数据集 $\boldsymbol{X}$，[1]记为 X_1、X_2、...、X_p。BPA 构建变量之间关系的初始图形表示 $\boldsymbol{X}$ [参见 Edwards 等人（2010年）、Riso 和 Guerzoni（ 2022]）等人的研究。从概念上讲，该算法属于序贯前向搜索家族 [Patil 和 Sane（ 2014], ），它从一个空变量集开始，逐步添加可能作为目标变量决定因素的变量，例如 X_i。BPA 利用 Edwards 等人提出的算法所揭示的依赖结构 [2010], 该算法依据 Akaike 信息准则和贝叶斯信息准则（AIC 和 BIC），在高维混合图形模型（HDGM）中识别最大似然树。最优预测变量集——定义为与给定目标变量相连的变量子集，该变量通常代表模型中的响应变量或因变量 [Manual（ 2013]）——通过互信息的概念来确定，互信息量化变量之间共享信息的数量。形式上，对于每个变量 X_i，我们将 $\boldsymbol{X}_{ps_i}$ 记为最优预测变量集，称为最优路径步，即在信息论意义上，为 X_i 提供最高联合解释力的变量子集 $\boldsymbol{X}$。BPA 在检测变量关系方面非常有效，因为它有助于解决应用当前一些最常用方法时出现的三个关键问题。首先，尽管使用协方差估计来确定依赖关系的符号，BPASGM 并不依赖协方差求逆或惩罚精度矩阵估计，而这正是高维投资组合问题中不稳定性的主要来源 [参见 Friedman 等人（2008年）、Riccobello 等人（2023年）、Scheinberg 等人（ 2010]）等人的研究；其次，它不需要指定适当的正则化参数，而某些优化问题（如 GLASSO）会用到该参数；最后，它能够检测变量之间并非仅呈线性的关系。有关 BPA 的更多方法细节见附录A。

2.2 通过 BPA 的推广进行网络构建

可以将 BPA 算法迭代应用于数据集中的每个变量，以构建具有特定属性的稀疏图形模型（SGM）。SGM 的构建方式如下。设数据集 $\boldsymbol{X}$ 由 p 个变量 $\{X_1, X_2, ..., X_p\}$组成。将 BPA 依次应用于这 p个变量，得到一个混合图（MG），下文称为 BPASGM，其中仅包含箭头和弧（即双向边），不包含线[2]。这一结果可以这样解释：算法选出的变量之间的关系可能是单向的，也可能是双向的。两个变量之间的关系（例如 X_j 和 X_i （其中 $i \neq j$））在前者解释后者时是单向的，即 $X_j \rightarrow X_i$；反之亦然，即 $X_j \leftarrow X_i$；而在$X_j \leftrightarrow X_i$时则是双向的，也就是说，一个变量影响另一个变量，反之亦然。

应用 BPA 时，最优路径步 $\boldsymbol{X}_{ps_i}$ of X_i 中所有对后者具有解释作用的变量，都表现出指向后者的单向关系。从图形上看，这一信息通过一组箭头表示：箭头从与 X_i 相关的 $\boldsymbol{X}_{ps_i}$ 中的节点出发，指向后者。其中一些关系实际上是双向的，因为 X_i 可能反过来被纳入其某个预测变量的最佳路径步骤中，或者当 X_i 的最优路径步与其某个预测变量的最优路径步共享一些共同变量时，也会出现这种情况。

在前一种情况下，X_j 与 X_i 之间的联系属于直接类型；而在后一种情况下，则属于间接类型，因为 X_i 与其某个预测变量之间的联系，是通过一些（一个或多个）其他共同预测变量建立的。需要注意的是，与依赖于使用协方差矩阵逆矩阵（精度矩阵）的方法的图相比，BPASGM 并不是无向图 [Friedman 等人，2008年，Meinshausen 和 B¨uhlmann， 2006]。这一点凸显了基于精度矩阵的方法与 BPASGM 之间的重要区别。基于精度矩阵的方法通过对单个变量进行成对比较，在局部层面评估依赖关系——尽管所评估的依赖关系仅限于线性关系。相比之下，BPASGM 在全局层面考察线性和非线性依赖关系，从而能够将单个变量与数据集中的所有其他变量进行比较。

此外，BPASGM 的节点满足马尔可夫性质，如下引理所示：

引理 2.1 (马尔可夫性质)。 设 X_i 为 *BPASGM* 中的一个节点， $\boldsymbol{X}_{ps_i}$ 为其最优预测变量集，$\boldsymbol{X}/\{\boldsymbol{X}_{ps_i}\}$ 为所有未包含在 $\boldsymbol{X}_{ps_i}$ 中的变量集

$$\boldsymbol{X}/\{\boldsymbol{X}_{ps_i}\}: \quad \{X_j \notin \boldsymbol{X}_{ps_i} \quad j = 1, ..., p\} \tag{1}$$

[1]其中 p 可以是 $p > n$ 或 $p < n$[2]众所周知，混合图包含三种类型的边，分别用箭头、弧（双向边）和线（实线）表示（参见 Sadeghi 和 Lauritzen [2014]）。该混合图仅包含前两种类型的边。

那么，可以证明以下性质成立：

$$X_i \perp\!\!\!\perp \boldsymbol{X}/\{\boldsymbol{X}_{ps_i}\}|\boldsymbol{X}_{ps_i} \tag{2}$$

证明：见附录 C。

这一性质意味着，一旦已知最优预测变量集 $\boldsymbol{X}_{ps_i}$，其余所有变量便都与 X_i 条件独立。这种结构在复杂系统中尤为重要，因为相互依赖关系往往源于数量有限的主导关系。因此，BPASGM框架通过揭示这些最小依赖结构，不仅可应用于金融领域，还可用于气候网络、神经系统或宏观经济相互联系等领域；在这些领域中，识别众多相互作用的组成部分之间的因果或信息路径，对于理解系统的整体动态至关重要。

2.3 通过矩阵 Θ检测连接

在确立了 BPASGM 的结构和概率基础之后，我们现在转向其实际应用。具体而言，我们引入一种基于矩阵的表示方法，以便检测算法识别出的连接并对其进行定量表征。

BPASGM 的应用结果可通过一个 $(p \times p)$ 邻接矩阵表示，记作 $\boldsymbol{\Theta}$；如果变量 X_j 属于变量 X_i 的最佳路径步骤，那么该矩阵的 $(i,\ j)$ 项等于 1，即：

$$\theta_{j,i} = \begin{cases} 1 & if \quad X_j \in \boldsymbol{X}_{ps_i} \\ 0 & otherwise \end{cases}, \quad i,j = 1,\ldots,p, \quad i \neq j. \tag{3}$$

矩阵 $\boldsymbol{\Theta}$是一个空矩阵，其非零元素突出了图中每个变量的最优预测变量，并提供了有关变量之间联系的发生情况及其性质（直接而非间接）的有用信息。

$\boldsymbol{\Theta}$的每一行之和（例如第 i 行）表示第 $i-$ 个变量作为其他变量的预测变量所起作用的次数，即 X_i 出现在其他变量的最佳路径步骤中的次数。每一列之和（例如第 $j-$ 列）表示包含在 X_j 的最佳路径步骤中的变量数量。

为了突出 BPASGM 的工作方式，我们将基于一个由简单数据 et 组成的 $p = 7$ 个变量的示例。

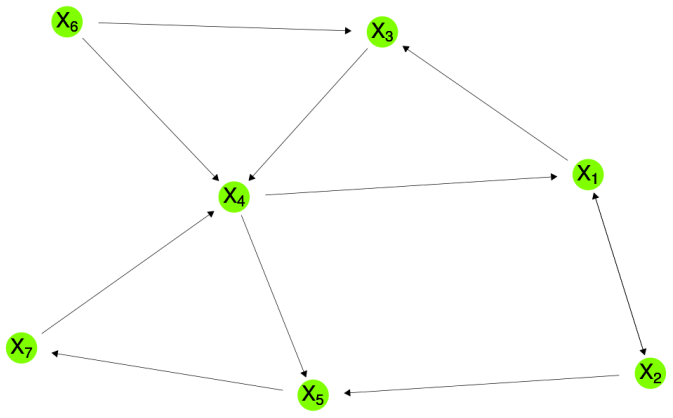


图1：通过 BPASGM 构建网络的示例

该网络的邻接矩阵为

$$\boldsymbol{\Theta} = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \\ X_7 \end{array} \begin{array}{c} \begin{array}{ccccccc} X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

从该矩阵的各行可以看出，数据集中的每个变量最多解释两个变量。具体而言， X_2、 X_4和 X_6 是另外两个变量的预测变量，而 X_3、 X_5 和 X_7仅参与另一个变量的路径步。具体来说，从该矩阵的第一行可以看出， X_i 是 X_2 和 X_3的预测变量。

查看 $\boldsymbol{\Theta}$ 的各列可知， X_3 是预测变量最多的变量，共有三个，而 X_6 没有预测变量。

矩阵 $\boldsymbol{\Theta}$ 可用于确定连接的性质，即连接是直接的，而非间接的或简单的。如前所述，当两个变量都出现在对方的最佳路径步中时，它们之间就存在直接连接。图 1 显示变量 X_1 和 X_2 之间存在直接连接

$$X_1 \leftrightarrow X_2$$

当 X_j 出现在 X_τ 的最优路径步中，且后者又出现在一个或多个作为 X_j 预测变量的变量的最优路径步中时， X_i 与 X_j 之间就会产生间接连接。因此，两个变量之间的间接关联可能经过多个步骤，具体取决于在它们之间起中介作用的变量数量，这些变量构成一个闭合链。事实上，间接关联形成一个由变量组成的闭合序列，其中每个变量都充当链中后续变量的预测变量。

在图1的网络中，间接关联的示例包括

$$X_4 \to X_5 \to X_7 \to X_4, \qquad X_4 \to X_1 \to X_3 \to X_4$$

最后，当 X_j 出现在 X_i预测变量之一的最优路径中，但 X_j 与 X_i之间既不存在直接连接，也不存在间接连接时， X_i 与 X_j 之间会形成简单连接。在图1所示的网络中，简单关联包括

$$X_6 \to X_4, \qquad X_6 \to X_4, \qquad X_3 \to X_4, \qquad X_4 \to X_5, \qquad X_4 \to X_1 \qquad X_2 \to X_3$$

特别值得注意的是，虽然 i列的第 $\boldsymbol{\Theta}$ 列出了对 X_i起解释作用的变量，但其 i行突出显示了由 X_i解释的变量。因此，如果某个变量，例如第X_j 个变量，同时出现在 i行的第 $\boldsymbol{\Theta}$（这意味着该变量由 X_i解释）以及其第i列（这意味着它解释了 X_i）中，则 X_i 与 X_j之间存在直接连接。这一结果为凸显变量之间的直接连接铺平了道路。事实上，可以通过计算 [Horn和Johnson所述的 2012] 哈达玛积，即邻接矩阵与其转置之间的哈达玛积，来检测直接连接

$$\boldsymbol{D} = \boldsymbol{\Theta} \odot \boldsymbol{\Theta}' \tag{4}$$

其中， $\odot$ 表示哈达玛积。 $\boldsymbol{D}$ 是一个对称矩阵，其非零元素突出显示了直接连接的变量对。如果该矩阵的$d_{i,\ j}$ 和$d_{j,\ i}$ 元素均为非零元素，则 X_i 与 X_j 之间存在直接连接。

例如，矩阵 $\boldsymbol{D}$ 对应图1 所示示例的矩阵为：

$$\boldsymbol{D} = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \\ X_7 \end{array} \begin{array}{c} \begin{array}{ccccccc} X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

位置 (1,2) 和 (2,1) 中同时存在非空条目，表明变量 X_1 和 X_2 通过直接连接相连，如下图所示。

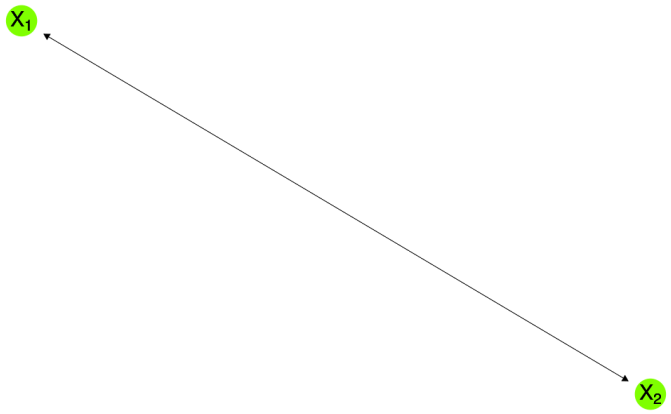


图2：从 $\boldsymbol{D}$ 图1中的网络开始的矩阵1

使用矩阵 $\boldsymbol{D}$ 表示有向连接后，可以通过以下矩阵找出变量之间的间接连接：

$$\boldsymbol{U} = \boldsymbol{\Theta} \odot \sum_{r=2}^{p-1} ((\boldsymbol{\Theta}^b)^r)' \tag{5}$$

$\boldsymbol{U}$ 的非零元素表示参与闭合链的变量对。因此，如果该矩阵的元素$U_{(i,j)}$、 $U_{(j,\gamma)}$ 和 $U_{(\gamma,i)}$ 均非零，则意味着 X_i 是 X_j的预测变量，而 X_j又解释了 X_γ，后者是 X_i的预测变量，从而在 X_i 和 X_j之间形成闭合链。以下矩阵 $\boldsymbol{U}$ 表示图1所示网络中的间接连接。非零元素u_1、3、 u_3、4和u_4、1表明 X_1 解释了 X_3，而 X_3 处于 X_4 的最佳路径步骤中； X_4 又是 X_1的预测变量，因此在这些变量之间形成闭合链。同样，u_7、4、 u_4、5和u_5、7非零，表示 X_7 处于 X_4 的最佳路径步骤中； X_4 解释了 X_5，而 X_5 又是 X_7的预测变量，因此再次形成闭合链。在所研究的情形中，这两条闭合链都连接两个变量：前一条通过第三个变量 X_3 连接 X_1 和 X_4，后一条则通过第三个变量 X_4连接 X_7 和 X_5。

图 3 展示了图 1所提出示例中的间接关联

$$\boldsymbol{U} = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \\ X_7 \end{array} \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

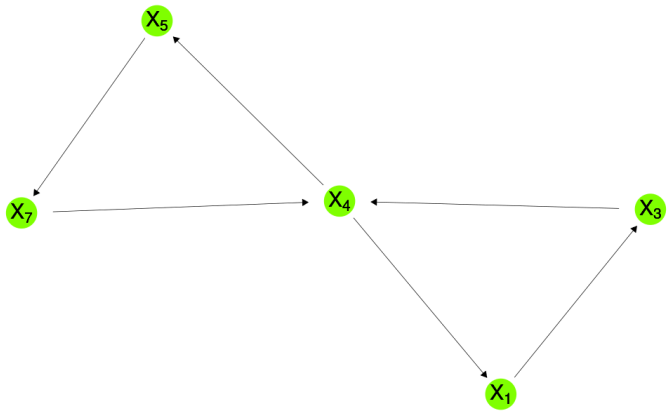


图3：表示 $\boldsymbol{U}$ 从图中网络开始的矩阵1

最后，变量也可以通过简单连接或伪关联连接。这些类型的连接可以通过以下矩阵检测

$$\boldsymbol{S} = \boldsymbol{\Theta} - \boldsymbol{D} - \boldsymbol{U} \tag{6}$$

其非零元素表示成对的变量之间存在连接，因为其中一个解释另一个，但反之不然；或者通过一个或多个变量的中介间接相连，但不形成闭合链。以下矩阵 $\boldsymbol{S}$ 给出了图1所示网络的简单关联或伪关联

$$\boldsymbol{S} = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \\ X_7 \end{array} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

图4 展示了图1所表示网络中变量之间的伪关联。

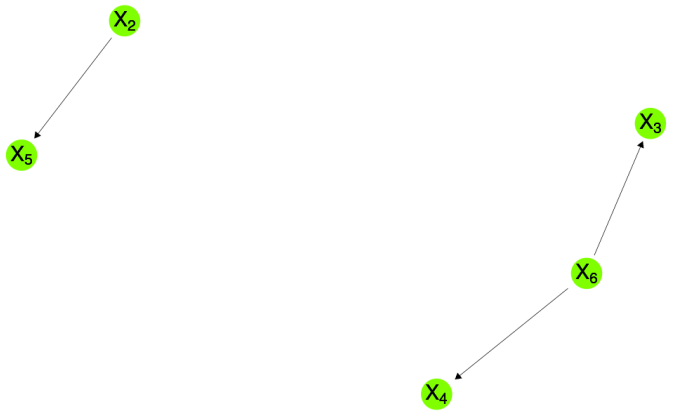


图4： $\boldsymbol{S}$ 矩阵的表示，其由图1中的网络出发

2.4 从BPASGM到最优资产选择

在上一节中，说明了如何通过对矩阵 $\boldsymbol{\Theta}$、 $\boldsymbol{D}$、 $\boldsymbol{U}$ 和 $\boldsymbol{S}$的计算，检测给定数据集中变量之间的直接、间接和简单关联。

在本节中，我们使用BPASGM来构建投资组合。在其开创性研究中，Markowitz[1952a] 证明，这样的投资组合应由负相关或相互独立的资产构成。BPASGM是实现这一目标的有力工具。为说明这一点，我们考虑一个由 p 个对数收益率组成、记为 $\boldsymbol{X}_p$的数据集。首先将BPA应用于 $\boldsymbol{X}_p$中的每个变量（资产）。由此得到BPASGM网络。就此而言，该网络有效地凸显了变量之间的依赖关系，为其相互联系提供了有价值的洞见。尽管它并不直接编码这些依赖关系的方向（符号），但这一方面可通过补充分析进一步探究，这在金融等情境中尤为有意义。由于负相关的股票常被纳入投资组合以降低其方差，因此有必要考虑邻接矩阵 $\boldsymbol{\Theta}$的带符号版本：

$$\boldsymbol{\Theta}_s = \boldsymbol{\Theta} \odot \boldsymbol{\Sigma}^b \tag{7}$$

其中， $\boldsymbol{\Sigma}^b$ 表示资产协方差矩阵的正二次型，定义为： $\boldsymbol{\Sigma} = \mathbb{E}[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})']$ [3]。从现在起，所有由 $\boldsymbol{\Theta}_s$构造的矩阵 $\boldsymbol{D}$、 $\boldsymbol{U}$ 和 $\boldsymbol{S}$均用下标 s表示。

然后， 以下三个步骤将用于构建一个充分考虑资产 et
依赖关系 由 BPASGM 检测出的 es 被提出。

[3]其中， $\boldsymbol{\mu}$ 是数据 数据集。

• **步骤1：** 首先，选择一个目标资产，作为构建适当投资组合的起点。该资产可以根据其在一组 p 资产 [中的最佳表现来选定，衡量指标包括索提诺比率、2001、Vinod与Reagle、 2004], 夏普比率 [Sharpe、 1994], 或信息比率[Bacon、 2023, 2021]。或者，也可以根据投资者的特定偏好来选择表现最佳的资产。选定初始资产后，进行初步筛选，剔除与该资产呈正相关的资产。

用 $j \in \{1, \ldots, p\}$ 表示所选资产的索引后，在此步骤中，所有与节点 j直接相连的节点（无论是源节点还是目标节点）都会被剔除。
这些通过出边或入边与节点 j 直接相连的节点，在二元邻接矩阵 $\boldsymbol{\Theta}_s \in \{0,1\}^{p\times p}$的第 j 行或第 j 列中均以非零元素标示。

$$\boldsymbol{\Theta}_s = (\theta_{s,i,k})_{i,k=1}^{p}$$

因此，剔除集 定义为

$$E_j = \{i \in \{1, \ldots, p\} \mid \theta_{s,j,i} = 1 \ \vee \ \theta_{s,i,j} = 1\},$$

而 缩减索引集

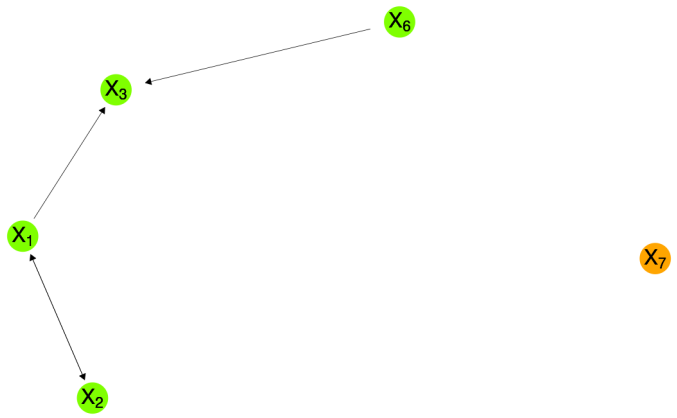
$$I_j = \{1, \ldots, p\} \setminus E_j,$$

结果包含所有满足以下条件的索引 i ： $\theta_{s,j,i} = 0$ 和 $\theta_{s,i,j} = 0$，即所有与节点 j不相邻的节点。

过滤后的带符号邻接矩阵 是以 I_j 为索引从 $\boldsymbol{\Theta}_s$ 中提取的子矩阵：

$$\widetilde{\boldsymbol{\Theta}}_s = (\boldsymbol{\Theta}_s)_{I_j, I_j} \in \{0,1\}^{|I_j|\times|I_j|}. \tag{8}$$

The ~矩阵 $\boldsymbol{\Theta}_s$ 排除了网络中资产 j 与其余节点之间的所有连接 rk
生成 自 BPASGM。
此时，过滤数据集表示为 $\boldsymbol{X}_z$，其中 $z < p$。图5 展示了~过滤后的符号邻接矩阵 $\boldsymbol{\Theta}_s$的结构。该矩阵由图1所示的网络得出，并假设 X_7 为选定的起始资产



~图 5：步骤1所得符号更新矩阵 $\boldsymbol{\Theta}_s$的表示，其起始于图1中的网络，其中 X_7（橙色）为选定的起始资产。

• **步骤 2：** 由 $\boldsymbol{\Theta}_s$~推导出矩阵 $\boldsymbol{D}_s$，~其定义见式 4。通过依据特定标准移除两个直接连接变量中的一个，可以消除直接（正向）连接。例如，对于反馈回路中涉及的每一对变量，可以舍弃在其他变量的预测变量中出现频率较高的那个变量。其他标准包括消除方差最高的变量（σ_{X_j}），或预期收益最低的变量（μ_{X_j}）。此外，还可以依据性能指标来指导筛选过程，例如 索提诺和Satchell [2001]指数 [Vinod与Reagle、 2004], 夏普比率，或信息比率 [1994] [Bacon、 2023, 2021]。移除直接连接后，将更新符号邻接矩阵。

令 $\boldsymbol{\Theta}_{s.u}$ 表示所得矩阵，该矩阵不再包含直接连接。此时，数据集表示为 $\boldsymbol{X}_k$，其中 $k < z < p$。

符号更新矩阵 $\boldsymbol{\Theta}_{s.u}$ 移除 X_1 与 X_2 在 $\boldsymbol{\Theta}_s$ 之间的直接连接后得到的符号更新矩阵如图6所示。该矩阵是通过使用性能指标，在两个节点中选出一个保留（X_2）而得到的。

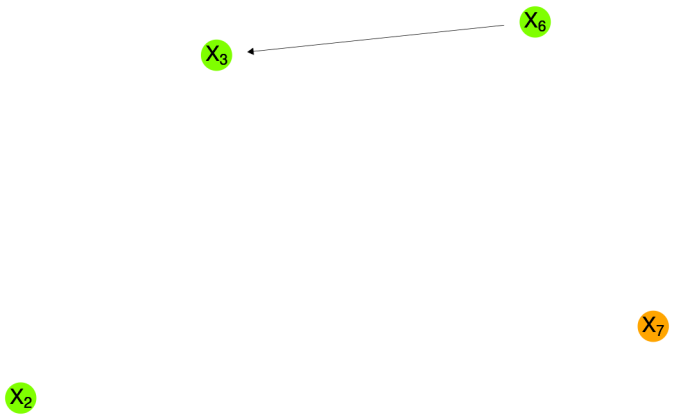


图 6： $\boldsymbol{\Theta}_{s.u}$从图1 中的网络出发，经步骤 2 得到的 X_7（橙色）表示选定的起始资产

- **步骤3：**移除直接连接后，符号邻接矩阵 $\boldsymbol{\Theta}_{s.u}$ 仅保留简单连接或无向连接。通过计算矩阵（$\boldsymbol{U}_{s.u}$）和（$\boldsymbol{S}_{s.u}$），可以检测闭合或开放的因果链。随后，可采用与此前移除直接连接时类似的标准来切断这些连接。此时，数据集缩减为 $\boldsymbol{X}_h$，其中$h < k < z < p$。就目前为止所举的示例而言，在此步骤中，移除从 X_6 到 X_3的连接后，仅剩三个资产，同时保留资产 X_6 （如图7所示）。



图7：步骤**3**，从图1中的网络开始，使用 X_7 （橙色）选定的起始资产

步骤3（$\boldsymbol{X}_h$）中的剩余变量不再表现出直接连接、无向连接或简单连接。然而，仍然可以观察到潜在依赖，并且这些依赖仅涉及步骤3（$\boldsymbol{X}_h$）中 ᔑᔑᔑ下的变量。与前面的步骤一样，可以通过移除 $\boldsymbol{X}_h$ 中共享大量共同预测变量的变量，或删除由特定预测变量共同解释的变量，进一步优化数据集。这些优化遵循前面步骤中概述的相同选择标准。最终数据集由资产的最小子集组成，记为 $\boldsymbol{X}_g$，其中 $g < h < k < z < p$，这些资产要么不共享信息，要么仅共享负依赖。最终，该过程使我们能够推导出一个最优变量子集，这些变量要么凭借其马尔可夫性质而相互独立，

要么呈负相关，从而有助于实施稳健的资产配置策略。由于如图7所示，前面三个步骤的输出只是一组未连接的变量，因此本步骤不适用于目前所考虑的示例。

所有链接移除程序均可借助引理2.1 来执行。具体而言，一旦算法识别出资产之间的关系——这些关系反映了某些资产对其他资产所施加的解释性或预测性影响——剔除特定资产便会破坏这些依赖关系。凭借马尔可夫性质，此类破坏可确保其余资产的条件独立性。

3 BPASGM的实际应用：一项模拟研究

在确立BPASGM的方法论基础之后，我们现在转向基于模拟的验证。本节提出一项蒙特卡洛研究，旨在评估依赖感知的资产选择如何影响在有限样本中、依据标准均值—方差准则构建的投资组合的已实现特性。本分析的目的并非修改Markowitz优化问题本身，而是评估一个重塑资产集合的数据驱动筛选步骤如何影响投资组合矩的估计以及由此产生的投资组合结果。

本研究采用的方法遵循两阶段程序。在第一阶段，将BPASGM应用于大量资产，以便根据其依赖结构和冗余特性选出一个子集。在第二阶段，在所选子集上重新估计预期收益和协方差，并在全文使用的相同可行性约束下，运用经典的马科维茨均值—方差框架确定投资组合权重。因此，本节所考虑的投资组合来自不同的、以选择为条件的估计问题，而非来自可行集缩小且估计值固定的同一优化问题。因此，模拟结果应被解释为有限样本稳健性与稳定性的证据，而非相对于在已知总体矩下定义的理论Markowitz有效前沿的改进。

在其开创性贡献中，Markowitz [1952b] 将投资组合选择形式化为一种权衡：预期收益被视为理想属性，而收益方差则被视为不理想属性。在这一框架下，有效投资组合是在给定风险水平下使预期收益最大化的投资组合，而所有此类投资组合构成了均值—方差空间中的有效前沿。在不存在无风险资产的情况下，该前沿对应于双曲线的上支，通常称为马科维茨子弹曲线。
实践中，当投资组合矩根据有限样本估计时，投资组合相对于这一理论前沿的实际位置可能受到估计误差、维度以及资产间依赖结构的显著影响。这些影响在高维情形下尤为突出，因为冗余和强正依赖可能导致估计不稳定，并造成样本外表现不佳。

有效前沿的几何形态与资产收益率的相关结构密切相关。负相关会使前沿向左移动，从而在给定预期收益水平下降低风险，或在给定风险水平下提高预期收益，进而增强分散化。当资产在统计上相互独立时，分散化收益依然存在，但其来源仅仅是投资组合权重的分布，而非相互抵消的共同变动。随着独立资产数量增加，投资组合方差会随投资组合权重的平方而下降，但这种收益通常小于存在负依赖时所能获得的收益。尽管相关模式对分散化发挥着重要作用，但下文所述影响主要源于依赖感知型资产选择与随后重新估计投资组合矩之间的相互作用，而非仅仅源于相关结构。

遵循Broadie [1993], 的方法，我们通过受控模拟实验评估 BPASGM 的有效性，以考察依赖感知筛选如何影响有限样本估计下得到的实际有效前沿。具体而言，我们分析基于资产依赖结构进行选择，会如何影响投资组合稳定性和分散化特性，尤其是在对重新估计的矩应用标准均值—方差优化时。

在这项模拟研究中，我们生成一组包含十二条人工金融收益序列的数据，以模拟金融市场中常见的复杂依赖结构和噪声特征。模拟涵盖每种资产的 $t = 2520$ 个观测值，约相当于十年的日度数据（假设每年有 252 个交易日）。数据生成过程的详细信息见附录 D。

图 8 展示了数据生成过程所包含的依赖结构，这是应用 BPASGM 的起点。在 Edwards 等人 [2010], 提出的最小BIC树基础上，对每个变量对称地应用 BPA，得到图 9所示的 BPASGM 网络。尽管数据生成过程包含线性和非线性依赖，

BPASGM 网络成功捕捉了变量之间的潜在结构。由于此处 BPASGM 应用于金融收益率，因此采用符号邻接矩阵 $\mathbf{\Theta}_s$，其中由负依赖关系连接的变量被视为相互独立，以避免排除因负向关系而具有分散化潜力的资产。

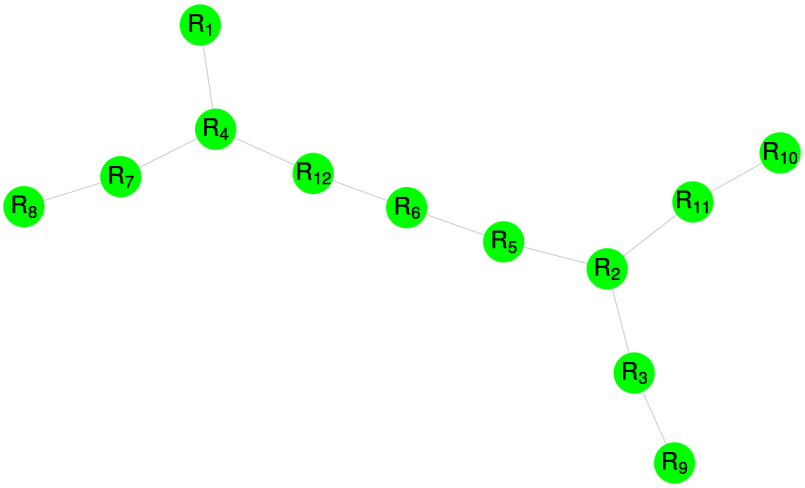


图8：模拟收益的数据集 $\boldsymbol{X}_{simulated}$ 的最小BIC树。

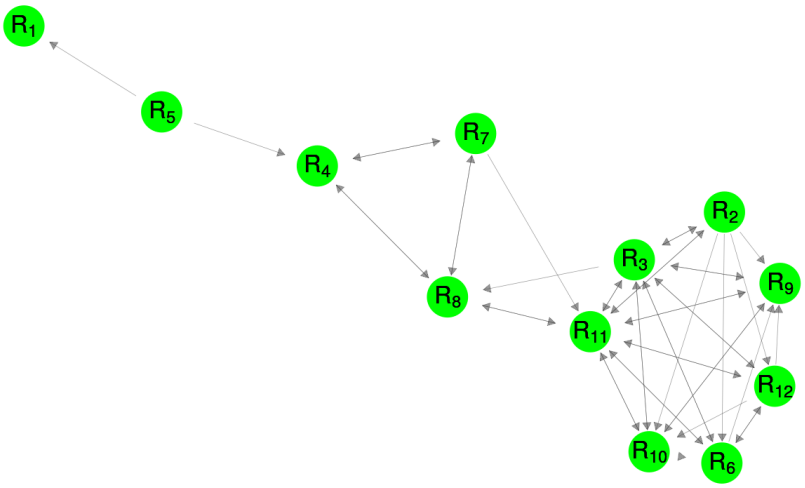


图9： 矩阵的表示 $\mathbf{\Theta}_s$（方程 7），由将 BPASGM 应用于数据集后得到 et 的模拟收益得出的收益率 $\boldsymbol{X}_{simulated}$。

构建 BPASGM 网络后，应用第2.4节所述的步骤来推导最优投资组合。如前所述，可以采用不同的方法来初始化该步骤，具体取决于用于选择起始节点的准则。

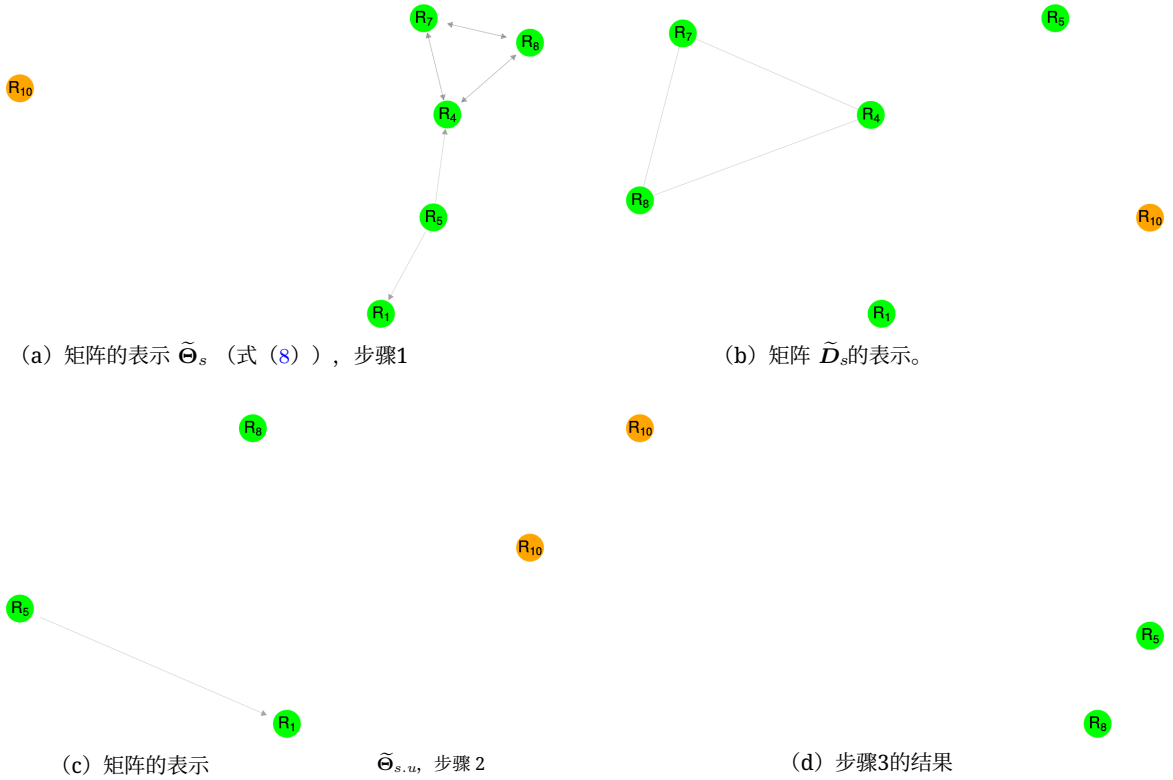
在本应用中，该流程首先将 p 个模拟资产中表现最佳的资产作为起始资产。具体而言，根据索提诺比率和夏普比率，节点 R_{10} 被确定为表现最佳的资产。

接着，执行第 2.4 节所述流程的第一步。这包括计算~ ~符号邻接矩阵 $\mathbf{\Theta}_s$（见公式（8））。图10a 展示了更新后的矩阵 $\mathbf{\Theta}_s$，其中资产 R_{10} 与 BPASGM 网络中其余节点之间的所有连接均已移除。

基于 $\widetilde{\mathbf{\Theta}}_s$，在方程2.4中直接连接通过推导得出。按照预先设定的选择标准，移除两个相连资产中的一个，以消除直接连接。在本应用中，索提诺比率最低的资产 [Sortino andSatchell, 2001]~将被移除。图10b展示了所得矩阵 $\boldsymbol{D}_s$。

按照第2.4节中的步骤 2，随后更新符号邻接矩阵，以排除所有直接~连接。所得矩阵记为 $\boldsymbol{\Theta}_{s.u}$，如图10c所示。

如步骤3所述，间接（或无向）连接也可通过应用类似准则予以去除，该准则可根据投资者特定偏好进行调整。此步骤的结果是一组完全不连通资产，构成投资组合构建的一个有意义子集。这些资产要么不相关，要么呈现负相关性，从而有助于波动率降低。由BPASGM程序得到的最终资产选择如图10d。



（a）矩阵的表示 $\widetilde{\boldsymbol{\Theta}}_s$ （式（8）），步骤1　（b）矩阵 $\widetilde{\boldsymbol{D}}_s$的表示。

（c）矩阵的表示 $\widetilde{\boldsymbol{\Theta}}_{s.u}$，步骤 2　（d）步骤3的结果

图10：为节点 R_{10} 构建充分利用 BPASGM 所提供信息的投资组合的三步程序

图11 展示了基于 BPASGM 的投资组合构建程序在各个阶段有效前沿的演变。各面板呈现了一系列散点图，其中每个灰色点代表一个随机生成的可行投资组合，该组合根据算法在相应阶段选取的收益率构建。红色点标示经验有效前沿上的投资组合，蓝线则表示通过二次拟合对前沿进行的平滑近似。

- **面板 (a)**：这是基准情景，展示了根据模拟收益的完整数据集$\mathbf{X}_{\text{模拟}}$构建的投资组合可行集。前沿较宽，表明分散化程度较高，但也包含许多被支配投资组合。上边界（即有效前沿）覆盖了广泛的风险和回报范围。

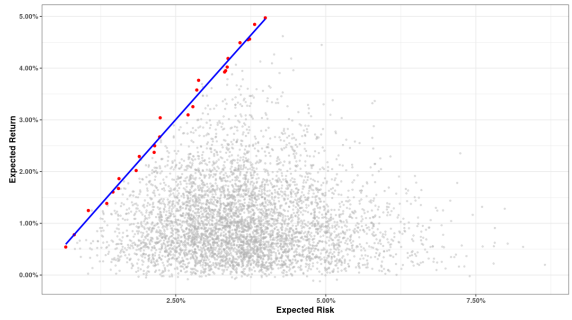
- **面板 (b)**：应用 BPASGM程序并选取 Θ_s 中的资产后，可行集变得更加精细。有效前沿在形状和跨度上仍与原始前沿相近，这表明初始变量选择保留了投资组合的大部分效率。~

- **面板 (c)**：移除直接依赖关系（如 D_s 所示）后，可行集收缩。有效前沿变得更加凹且更短，体现出一种权衡：虽然排除了一些可能获利的组合，但所选集合倾向于不相关资产，因此在给定风险水平下提高了多元化效率。

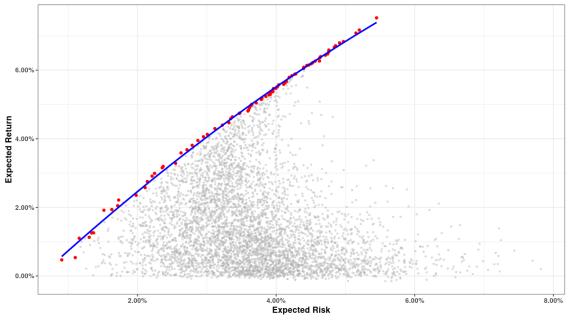
- **面板 (d)**：此最终面板展示了移除无向连接（步骤3）后的结果。可行集受到的约束最为严格，反映出完全断开（或负相关）资产的子集。有趣的是，尽管前沿明显缩短，但仍保持着理想的形状，曲率平滑且清晰。这支持了这样一种观点：BPASGM程序成功筛选出一个精简而高效的资产子集，仍可构建表现优异的投资组合，同时降低复杂性并提升可解释性。

  更重要的是，与该程序早期阶段获得的有效投资组合相比，位于有效前沿上的这组投资组合具有更低的风险水平和更高的预期收益。

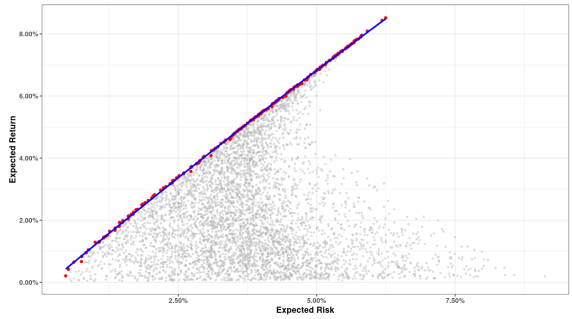
总体而言，资产集合的逐步精炼使可行投资组合空间逐渐缩小。然而，有效前沿在整个过程中保持相对稳定且形态良好，表明该方法论能够有效滤除冗余或有害的资产，同时保留对实现最优投资组合绩效至关重要的资产。



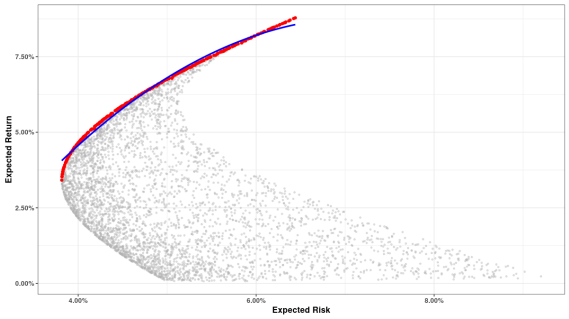
(a) 模拟收益 $\boldsymbol{X}_{simulated}$数据集的随机可行投资组合散点图



(b) 步骤1中 $\boldsymbol{\Theta}_s$所含资产$\tilde{}$的随机可行投资组合散点图



(c) 步骤 2中 $\boldsymbol{\Theta}_{s.u}$所含资产$\tilde{}$的随机可行投资组合散点图



(d) 步骤3中所含资产的随机可行投资组合散点图

图11：由模拟投资组合构建的有效前沿。蓝色曲线表示对经验有效前沿进行的平滑二次拟合，该前沿由风险—收益空间中的非支配投资组合构成。灰色点表示随机生成投资组合的全集，红色点则突出显示位于经验有效前沿上的投资组合。

模拟结果表明，在这一有限样本环境中，与在相同均值—方差准则下评估的非知情资产选择程序所获得的投资组合相比，经过基于 BPASGM 的资产选择后构建的投资组合呈现出更有利的实际风险—收益权衡。这些差异源于 BPASGM 在投资组合优化之前重塑资产集合的方式，而非对马科维茨目标函数本身作出任何修改。为更好地理解导致这些结果的机制——尤其是依赖结构、分散化和估计误差所发挥的作用——下一节将从理论角度探讨有效前沿的统计与结构性质。

在这一模拟环境中，与随机选择的相同数量资产子集相比，BPASGM系统性地选择出能够带来更稳定且表现更佳的实际投资组合的资产子集。

如图 12所示，基于BPASGM选择的资产构建的投资组合（红点）往往位于通过所有可能组合获得的投资组合之上[4]，这些所有可能组合由规模相同的随机选择资产（灰点）构成，并使用经筛选条件化的预期收益和协方差估计值进行评估。这一模式表明，BPASGM能够有效利用资产收益率依赖结构中所蕴含的信息，避免在未经信息引导的选择下可能出现的冗余且分散化程度不足的组合。

重要的是，应将这些证据解读为相较于随机或无信息资产选择的有限样本稳健性结果，而非优于在总体矩已知条件下定义的理论 Markowitz 有效前沿。通过将投资组合矩的估计建立在依赖感知筛选步骤之上，BPASGM 降低了估计误差，并提高了已实现投资组合结果的稳定性，同时并未颠覆均值–方差投资组合理论所依据的基本风险—收益权衡。

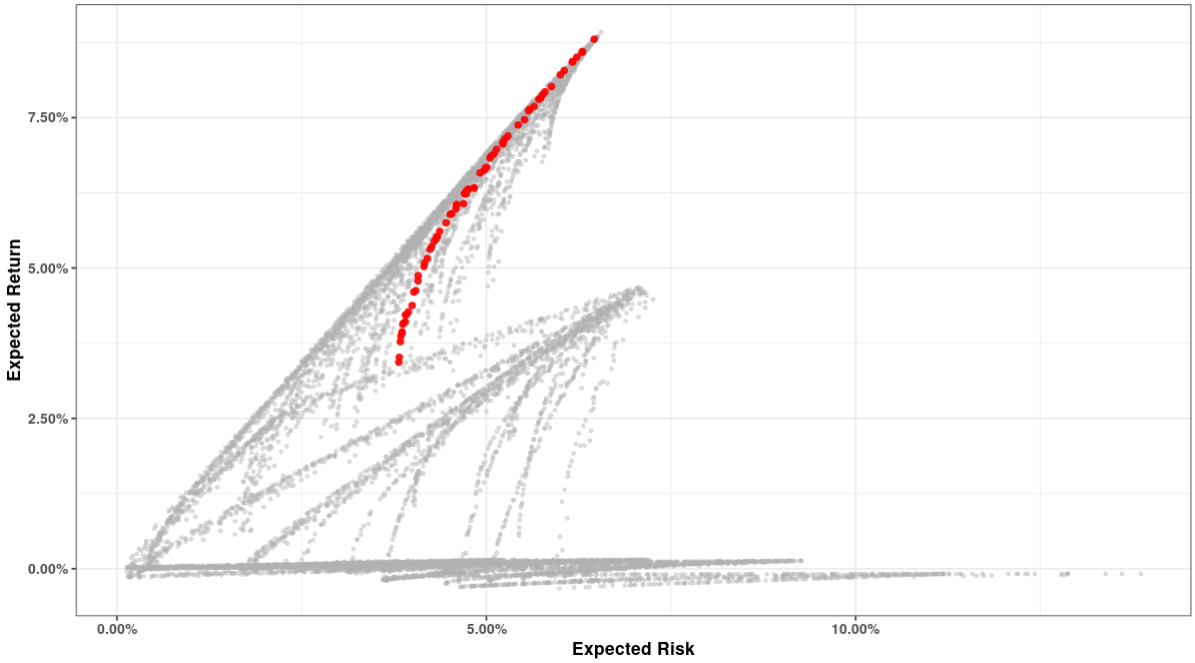


图12：由 $\boldsymbol{X}_{simulated}$ 得到的有效前沿。灰色点表示通过随机选择与过程3所选资产数量相同的资产而得到的所有可能有效前沿。红色点对应于由 BPASGM 选出的资产子集构建的投资组合，展示了由此产生的已实现风险—收益权衡；在有限样本估计下，与从相同数量的资产中随机选取其他资产子集所得结果相比，这些权衡系统性地更为有利。

4 有效前沿的性质

基于上一节呈现的蒙特卡洛证据，我们现在考察依赖感知型资产选择如何影响有限样本条件下依据标准均值—方差准则构建的投资组合的已实现特性。BPASGM 并不改变 Markowitz 优化问题本身，而是通过在优化前重塑资产集合，并使投资组合矩的估计以所选子集为条件，来影响投资组合结果。本节探讨这一选择过程所产生的统计和结构特征，并讨论其对分散化和波动率动态的影响。

众所周知，有效前沿的形状和位置取决于投资组合中所纳入资产的统计性质，以及对投资组合权重施加的约束。具体而言，前沿由资产之间的预期收益、波动率和相关性决定，并进一步受到卖空约束以及是否存在无风险资产等实际因素的影响。当投资组合矩是根据有限样本估计得出时，这些特征会与

[4]我们从十二种资产中计算出了三种资产的 220 种投资组合组合。

估计误差和维度相互作用，从而可能影响优化投资组合的实际表现。

预期收益影响前沿的垂直位置，而资产波动率决定可实现风险水平的范围。在这些决定因素中，依赖结构——由相关性及更一般形式的统计依赖所体现——在塑造前沿曲率方面发挥着核心作用。较低或负向的依赖通过相互抵消的变动降低投资组合方差，从而增强分散化；而强正向依赖则限制分散化机会，并可能导致投资组合过度集中且不稳定。

BPASGM框架旨在通过考虑依赖关系的资产筛选来利用这些机制。通过识别并剔除冗余性强或存在正向依赖的资产，同时保留关联较弱或呈现负向依赖的资产，BPASGM以有利于分散化的方式调整资产集合的构成。重要的是，这种效果并非仅仅源于资产数量的减少，而是源于依据依赖结构筛选资产的具体方式。

此外，BPASGM还采用基于绩效的标准（如夏普比率和索提诺比率），以指导对直接关联和间接关联资产的筛选。这有利于选择预期收益较高的资产，或在回报水平相同的情况下选择下行风险较低的资产。因此，当基于筛选后的子集重新估计投资组合矩并应用标准均值—方差优化时，所得投资组合在有限样本中往往呈现更优的实际风险—收益权衡，这反映的是分散化改善和估计误差减少，而非理论有效前沿发生了根本性改变。

从这个意义上说，本节讨论的性质应被理解为有限样本估计下由依赖感知选择所诱发的结构性特征，而非相较于在矩已知条件下定义的总体均值—方差前沿所取得的改进。

本节其余部分将通过考察依赖感知型资产选择如何影响所选资产的估计收益率、方差和依赖模式，进而改变有效前沿的几何形态，对这些见解加以形式化。

移除正依赖并优先选择相互独立或负相关的资产，会降低分散化比率中的波动率加权平均相关性成分，从而提高比率本身。分散化比率的上升会推动有效前沿上移，并在特定条件下使投资组合预期收益与波动率之间的关系近似呈线性。

为解释这一耐人寻味的结果，请看分散化比率Meucci [2009] 它展示了分散化如何帮助降低风险

$$DR = \frac{\sum_i^N w_i \sigma_i}{\sigma_p} \tag{9}$$

其中， w_i 和 σ_i 分别为投资组合权重及其资产的波动率，而 $\sigma_p = (\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w})^{1/2}$ 为投资组合的波动率。DR 的分子是假设资产完全相关时投资组合的波动率，分母则是实际投资组合的有效波动率。

Choueifaty等人 [2013] 给出了分散化比率的如下分解：

$$DR_{\sigma_P}(\boldsymbol{w}) = [\rho_{MDP}(1 - CR_{MDP}) + CR_{MDP}]^{-1/2} \tag{10}$$

其中， ρ_{MDP} 和 CR_{MDP}分别称为波动率加权平均相关性和波动率加权集中度比率，其定义如下：

$$\rho_{MDP} = \frac{\sum_{i \neq j}^N \rho_{i,j} w_i \sigma_i w_j \sigma_j}{\sum_{i \neq j}^N w_i \sigma_i w_j \sigma_j}, \qquad CR_{MDP} = \frac{\sum_{i=1}^N (w_i \sigma_i)^2}{\left(\sum_{i=1}^N w_i \sigma_i\right)^2} \tag{11}$$

加权平均相关性 ρ_{MDP}将实际投资组合中资产的协方差与由完全相关资产构成的投资组合的协方差进行比较，并表示由资产相关性带来的风险降低百分比。当所有资产互不相关时，该值为零；当所有资产完全相关时，该值为1。

加权集中度比率 CR_{MDP}衡量的是，经过由完全相关资产构成的投资组合的波动率调整后，投资组合权重的集中程度。直观而言， CR_{MDP} 衡量的是投资组合风险集中在少数资产上的程度。对于波动率相近且权重相等的资产组成的投资组合， $CR_{MDP} \approx 1/N$；随着资产数量 N 增加，该值会降低。这意味着

通过降低集中度而获得的分散化增益会随资产数量增加而扩大。相比之下， ρ_{MDP} 反映资产间的平均相关性，但资产数量增加并不一定会使其降低。如果新增资产与现有资产高度相关， ρ_{MDP} 可能仍然较高，从而限制分散化效益。因此，资产数量增加带来的整体分散化比率提升，主要源于集中度降低；相关性成分的贡献则取决于投资组合的相关结构。BPASGM通过选择不相关或负相关资产，有效抵消或降低 ρ_{MDP} 成分的贡献，从而最大化整体分散化比率。因此，采用BPASGM构建的投资组合由于分散化比率较高，在给定回报水平下往往风险较低。这些投资组合使前沿向左下方移动，从而实现更优的风险调整后收益。

从本质上说，BPSGM 通过消除或降低分散化成分 ρ_{MDP} 的作用（该成分反映资产间线性依赖的贡献），生成最大分散化投资组合。这使投资组合能够到达有效前沿的左上区域，实现较高的单位风险回报。

这一特性可以通过以下形式化解释：投资组合的预期收益可以表示为其波动率、分散化程度以及单项资产平均夏普比率的函数，如下所示

$$\mu_p = \bar{s}_v DR(w)\sigma_p. \tag{12}$$

此处， $\bar{s}_v = \frac{\sum_{i=1}^{N} v_i s_i}{\sum_{i=1}^{N} v_i}$ 中， $s_i := \frac{\mu_i}{\sigma_i}$ $i = 1, \ldots, N$, 表示第 i 项资产的夏普比率， $v_i = w_i\sigma_i$，并采用了以下对分散化比率的重新表述：

$$DR = \frac{\sum_i^N v_i\sigma_i}{\sigma_p} \tag{13}$$

其中

$$\sigma_p = (\boldsymbol{v}'\boldsymbol{R}\boldsymbol{v}) = (\sum_{i=1}^{N} v_i)^2(\rho_{MDP}(1 - CR) + CR) \tag{14}$$

其中， $\boldsymbol{R}$ 表示资产相关矩阵，且

$$\rho_{MDP} = \frac{\rho_{i,j}}{v_i v_j}\sum_{i\neq j} v_i v_j, \quad CR = \frac{\sum_{i=1}^{N} v_i^2}{(\sum_{i=1}^{N} v_i)^2} \tag{15}$$

公式（12）揭示了采用 BPASGM 选择证券时所得到的有效前沿形状的另一个原因。事实上，该算法会选择表现出色的节点，也就是说，夏普比率较高的节点。因此，所选资产的夏普比率差异往往较小，其均值 $\bar{s}_v$ 也较为稳定。与此同时，排除相关资产实际上消除了波动率加权平均相关性成分 ρ_{MDP}，因此投资组合方差几乎与成分 $\sum_{i=1}^{N} v_i^2$重合，且 DR 函数随权重变化而变化不大。尤其是当混合项包含负值时，投资组合方差会进一步降低，从而使前沿趋于平坦。 $\bar{s}_v$ 和 DR 的低差异性因此使 μ_p与 σ_p 之间的关系稳定且近乎线性。

以下段落概述了将采用的方法，以验证该算法能否有效筛选不相关资产，从而生成一系列有效前沿投资组合，其夏普比率平均值稳定，且分散化特征一致。

4.1 对 BPASGM 筛选的股票所构成投资组合的特征进行评估。

BPASGM方法生成由低风险、不相关或负相关资产构成的投资组合的能力，可以通过比较所得最优投资组合的波动率进行评估：该最优投资组合由算法选出的收益率构成，并与假设这些资产相互独立时同一投资组合所呈现的波动率进行比较。

1. 由独立资产构成的虚拟投资组合的波动率，是将通过应用于其各项资产的单变量 GARCH 模型确定的各组成部分的波动率相加计算得出的 [Bollerslev， 1986] 。2. 而BPASGM方法所得最优投资组合的波动率考虑了资产之间的时变相关性，因此通过多变量动态条件相关（DCC）-GARCH 模型进行评估 [Engle， 2002]。

计算这两个投资组合波动率所需的投资组合权重，采用Markowitz [1952提出的]全局最小方差策略确定。

这一比较有助于评估使用BPASGM构建的投资组合波动率是否低于由独立资产组成的投资组合。如果后者的波动率高于通过BPASGM确定、并使用DCC-GARCH模型进行评估且考虑潜在资产相关性的最优投资组合波动率，则可以认为BPASGM执行的资产选择是有效的。事实上，这一结果表明，BPASGM会隐含地选择独立资产或具有负相关性的资产。相反，如果DCC-GARCH模型揭示出显著且随时间变化的共同变动，则说明BPASGM未能有效选择投资组合中的资产。

虚拟投资组合的波动率需要对每个收益序列$r_{i进行t}$单变量GARCH(p,q)建模，以刻画条件方差$\sigma^2_{i,\ t}$ 具体如下

$$r_{i,t} = \mu_i + \varepsilon_{i,t}, \quad \varepsilon_{i,t} = \sigma_{i,t}\, z_{i,t}, \quad z_{i,t} \sim \mathcal{N}(0,1), \tag{16}$$

$$\sigma^2_{i,t} = \gamma_i + \sum_{j=1}^{p} \alpha_{i,j}\, \varepsilon^2_{i,t-j} + \sum_{k=1}^{q} \beta_{i,k}\, \sigma^2_{i,t-k}, \tag{17}$$

其中， μ_i 是r_i的条件均值，$_t$ 且 $\gamma_i > 0$、 $\alpha_{i,\ j} \geq 0$ 和$\beta_{i,\ k} \geq 0$ 是可通过最大似然法估计的参数。对于每项资产，使用贝叶斯信息准则（BIC）选择最优 GARCH 阶数（p， q），以平衡拟合效果与简约性。假设各资产之间条件独立，则时点 t 的投资组合方差可按如下方式计算

$$\sigma^{2,\text{uni}}_{p,t} = \sum_{i=1}^{N} w_i^2\, \sigma^2_{i,t}, \tag{18}$$

其中， w_i 表示资产 i 的投资组合权重（通过Markowitz [1952a] 策略获得），且 N 表示资产数量。

采用 BPASGM 得到的最优投资组合波动率，已通过多变量 DCC-GARCH 模型 [Engle， 2002], 进行了估计。该模型将条件协方差矩阵分解为各项资产的波动率和随时间变化的相关矩阵。对于 N收益率 $\boldsymbol{r}_t = (r_{1,\ t}, \ldots\ldots, r_{N,\ t})^\top$，模型定义如下

$$\boldsymbol{r}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(0, \boldsymbol{H}_t), \tag{19}$$

$$\boldsymbol{H}_t = \boldsymbol{D}_t\, \boldsymbol{R}_t\, \boldsymbol{D}_t, \tag{20}$$

其中 $\boldsymbol{\mu}_t$ 是条件 $\boldsymbol{r}_t$的期望值向量， $\boldsymbol{H}_t$ 是 $\boldsymbol{\epsilon}_t$的条件方差矩阵，其设定如式（20）。此处， $\boldsymbol{D}_t = \text{diag}(\sigma_{1,\ t}, \ldots\ldots, \sigma_{N,\ t})$是 $\boldsymbol{\epsilon}_t$ 的条件标准差对角矩阵，由应用于各资产的单变量 GARCH 模型估计得到，且 $\boldsymbol{R}_t$ 是按如下方式设定的条件相关性矩阵（Engle [2002]）

$$\boldsymbol{R}_t = \text{diag}(\boldsymbol{Q}_t)^{-1/2}\, \boldsymbol{Q}_t\, \text{diag}(\boldsymbol{Q}_t)^{-1/2}, \tag{21}$$

with

$$\boldsymbol{Q}_t = (1 - a - b)\, \overline{\boldsymbol{Q}} + a\, \boldsymbol{z}_{t-1} \boldsymbol{z}_{t-1}^\top + b\, \boldsymbol{Q}_{t-1}, \tag{22}$$

其中 $\overline{\boldsymbol{Q}}$是标准化残差 $\boldsymbol{z}_t = \boldsymbol{D}_t^{-1} \boldsymbol{\varepsilon}_t$ 的无条件协方差，且（a， b）是满足 a， $b \geq 0$ 和 $a + b < 1$ 的DCC参数，以保证估计的 $\boldsymbol{H}_t$ 为正。从实际角度来看，资产 i 与 j 之间的条件相关性动态演变为

$$\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t}\, q_{jj,t}}}, \tag{23}$$

其中$q_{ij,\ t}$ 是矩阵 Q_t的第（$i,\ \ j$）个元素。
使用 DCC-GARCH 估计的投资组合方差由下式给出：

$$\sigma_{p,t}^{2,\mathrm{dcc}} = \boldsymbol{w}^{\top}\, \boldsymbol{H}_t\, \boldsymbol{w}, \tag{24}$$

随后对两个投资组合波动率 $\{\sigma_{p,t}^{\mathrm{uni}}\}$ 和 $\{\sigma_{p,t}^{\mathrm{dcc}}\}$ 进行比较。持续较低的$\sigma_{p\,t}^{\mathrm{dcc}}$,将表明存在有意义的、随时间变化的负相关性，从而证实 BPASGM 选择的是独立资产或表现出负依赖的资产，进而增强分散化并降低投资组合波动率。

为评估投资组合绩效的逐步提升，针对算法在每次迭代中识别出的资产所构建的最优投资组合，计算索提诺比率和夏普比率。可通过在该程序的不同选择步骤上计算分散化比率及其组成部分，来评估 BPASGM 生成具有稳定平均夏普比率和一致分散化特征的投资组合有效前沿的能力。

这将使我们能够观察波动率加权平均相关性如何在每一步下降，从而导致分散化比率上升。此外，在每一步还将计算经验前沿上各投资组合的均值和方差，以及线性回归的决定系数

$$\mu_P = \alpha + \beta\sigma_P + \epsilon \tag{25}$$

已计算得出。后者从一个步骤到下一个步骤的增加，可作为衡量 μ_P 与 σ_P之间比例关系的指标。

最后，还将评估有效前沿上各投资组合的夏普比率和分散化比率的稳定性。

5 BPASGM的实证应用

本研究通过一项实证应用评估所提出的资产选择方法的表现。该应用使用三类金融资产的收益率（根据对数收盘价的一阶差分计算得出 [Pan和Liu, 2018]）：（i）美国市场中市值最高的333只股票（通常称为蓝筹股，包括“微软”和“Apple”等公司，通常是在1990年1月1日之前于证券交易所上市的成熟企业）；（ii）16个最重要的全球股票指数（如“标普500”、“STOXX欧洲600指数”和“富时100指数”）；以及（iii）9种交易最频繁的汇率。数据集涵盖1990年1月1日至2025年12月31日期间。数据集取自 *Refinitiv Workspace*。表2 列出了本分析所用资产的识别代码。

5.1 BPASGM所凸显的依赖关系网络

为评估第4.1节所述基于BPASGM的方法的有效性，数据集被划分为训练集和测试集。具体而言，使用1990年1月1日至2024年10月31日的数据进行训练（数据集 $\boldsymbol{X}_{\mathrm{training}}$），使用2024年10月31日至2025年12月31日的数据进行测试（数据集 $\boldsymbol{X}_{\mathrm{test}}$）。在训练阶段，应用BPASGM选择纳入投资组合的收益率。

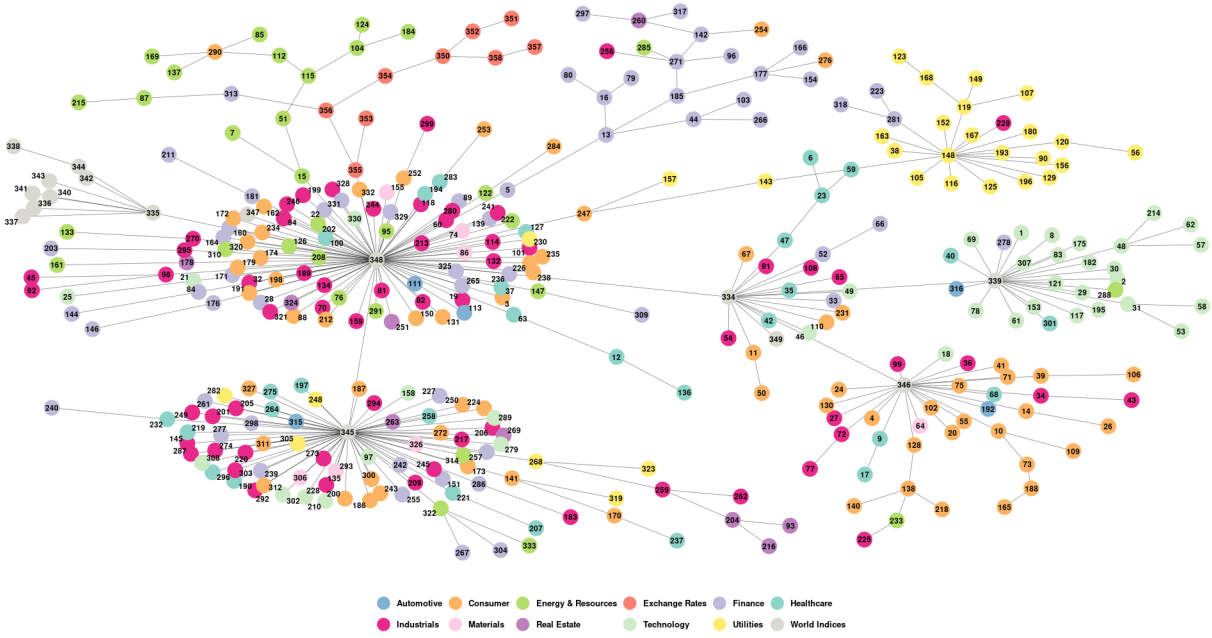


图13：收益率数据集 $\boldsymbol{X}_{training}$ 的最小BIC树，样本期为1990年1月1日至2024年10月30日

图13展示了使用基于BPASGM的方法构建最优投资组合的起点。图中的节点根据其对应收益率所属的行业以不同颜色标示。附录2 B 中的表列出了图中各节点的标签。具体而言，该图展示了算法所识别的数据集中各变量背后的网络结构 $\boldsymbol{X}_{training}$，该算法由Edwards 等人 [2010], 提出，并突显了各项收益率之间的关系。

值得注意的是，这个高维图模型的结构由一个大型连通分量——最小BIC树——构成。在该分量中，可以识别出不同的聚类。有些聚类围绕全球市场指数（以灰色节点表示）组织，另一些则明显与特定经济行业相关。例如，有一个对应公用事业 行业的聚类（黄色节点），另一个对应能源与资源 （浅绿色节点），还有一个由汇率 构成的聚类（深橙色节点）。尤其是，汇率聚类仅通过一条直接连线与金融聚类相连。具体而言，节点355（“美元/加元汇率”）直接连接到节点348（“纽约证券交易所综合指数”），而其余汇率变量之间主要相互连接。

在以全球市场指数为中心的簇中，节点 339（“纳斯达克指数”）是科技行业股票（例如“微软”、“Apple”和“甲骨文”）的关键枢纽节点。节点 346（“道琼斯指数”）主要与消费和工业行业的股票相关联。节点 348（“纽约证券交易所综合指数”）和 345（“罗素2000指数”）是不同类型股票的中心枢纽节点，因为与它们相连的节点涵盖多个行业。最后，节点 335（“STOXX欧洲600价格指数”）是其他国际指数的中心参照点。

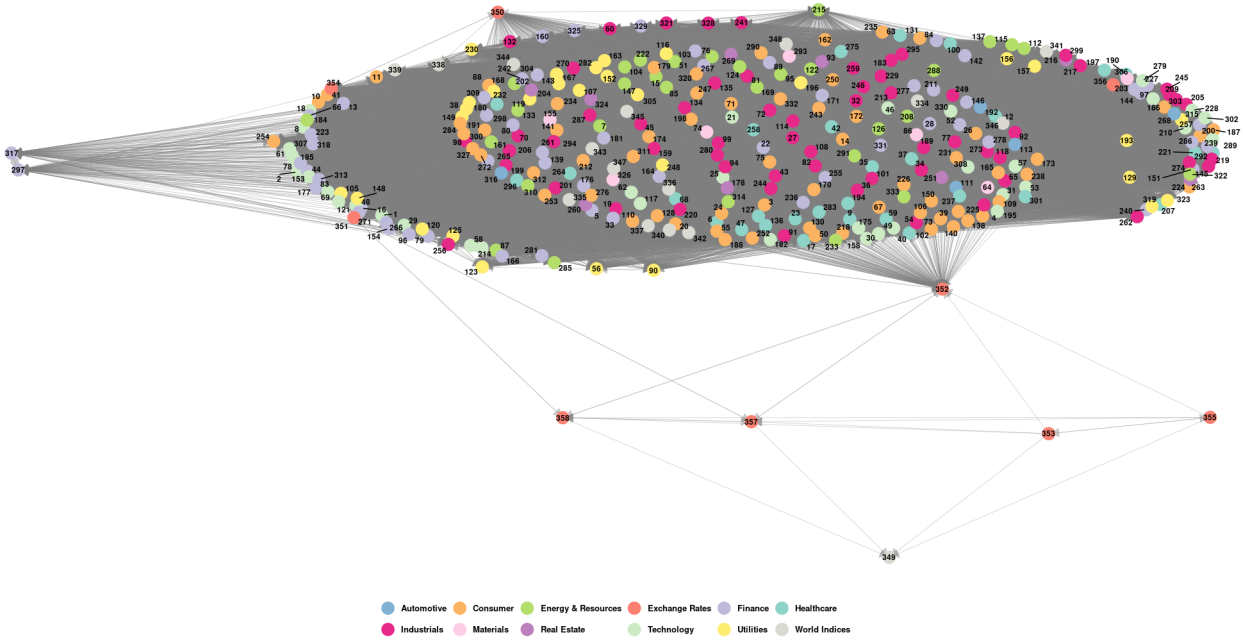


图14：矩阵的表示 $\boldsymbol{\Theta}_s$（方程 7），通过将 BPASGM 应用于收益数据集 $\boldsymbol{X}_{\text{training}}$获得，样本期为1990年1月1日至2024年10月30日。

图14 展示了在最小BIC树的网络中对每个节点迭代应用BPA的结果（见图13）。如第2.2节所述，该过程得出了BPSGM。

对负连接进行调整后（移除这些连接，以便将由负依赖关系连接的变量视为相互独立），BPSGM 识别出的变量间剩余连接可用于构建邻接矩阵 $\boldsymbol{\Theta}_s$（方程7）

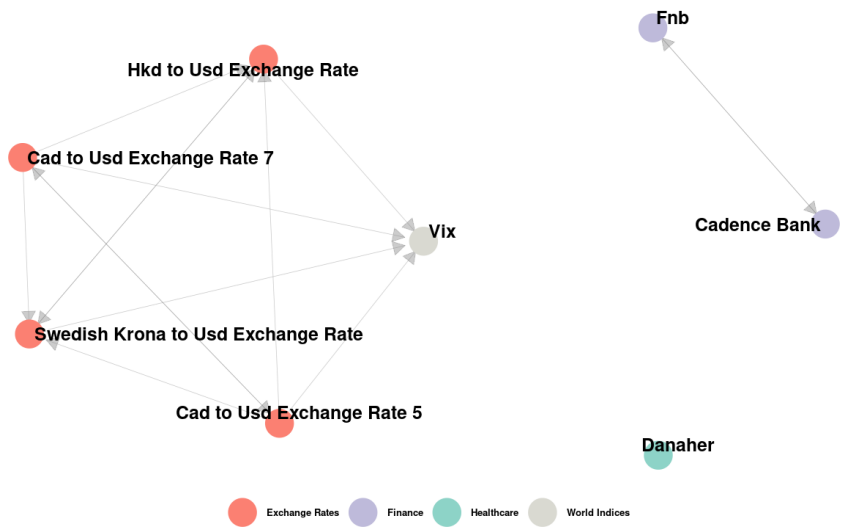
图 13揭示了网络结构中明显的聚类，而图 14 则未显示出任何明显的聚类。不过，它突显了数据集 $\boldsymbol{X}_{\text{training}}$中几乎所有股票之间的重要依赖关系，例外的是四种汇率（节点353、355、356和357，分别对应“CAD兑USD汇率5”“CAD兑USD汇率7”“USD兑AUD汇率”和“HKD兑USD汇率”）以及代表“VIX指数”的节点349。

尤其值得关注后者这一指数的作用，它似乎脱离了由主要网络成分所代表的正依赖核心。它通常被称为“恐慌指数”，反映市场参与者对未来波动率的预期。更具体而言，它反映“标普500”指数期权在30日期限内的隐含波动率，并被广泛用作不确定性水平或感知市场风险的代理指标 [Whaley， 2000]。

5.2 最优资产选择

在构建BPASGM之后，采用第2.4 节所述程序，从风险角度选择用于构建最优投资组合的资产。如上一节所述，可采用不同方法来初始化该程序，即将根据索提诺比率等标准指标表现最佳的资产选为起始节点 [索提诺与Satchell，2001，Vinod与Reagle， 2004] 或夏普比率 [夏普， 1994]，，或者选择反映投资者特定偏好的资产。例如，可以优先选择市值最高的资产，或波动率最低的资产。

在本例中，节点42（“Danaher”）被选为起始资产，因为它根据 [索提诺和Satchell、 2001] 以及夏普 [1994] 比率，均是表现最佳的资产。随~后，邻接矩阵 $\boldsymbol{\Theta}_s$ 按照第2.4 节所述进行了计算（参见公式（8））。



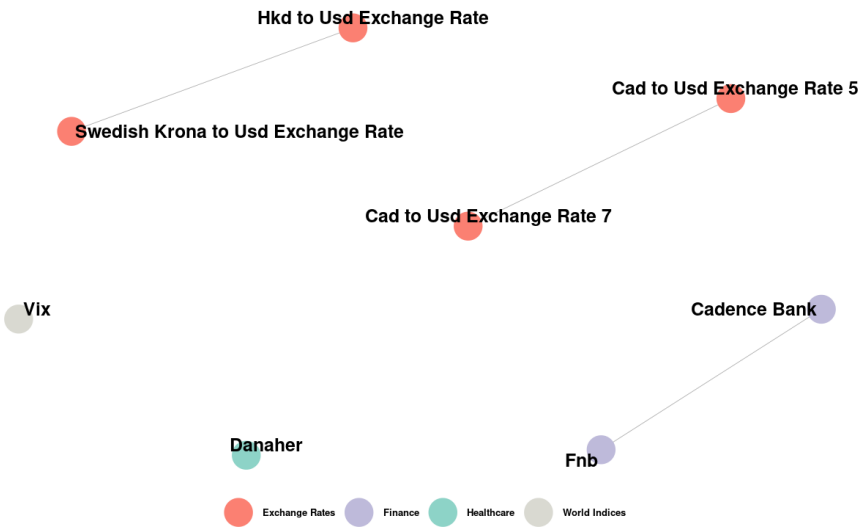
~图15：矩阵的表示 $\boldsymbol{\Theta}_s$ （方程式（8）），从节点42“Danaher”出发得到——步骤1。

~图15展示了更新后的符号邻接矩阵 $\boldsymbol{\Theta}_s$ ，即在移除资产“Danaher”与通过BPASGM生成的网络中其余节点之间的全部连接之后的结果。值得注意的是，剔除所有与所关注节点呈正向连接的变量后，变量数量大幅减少，从358降至7。

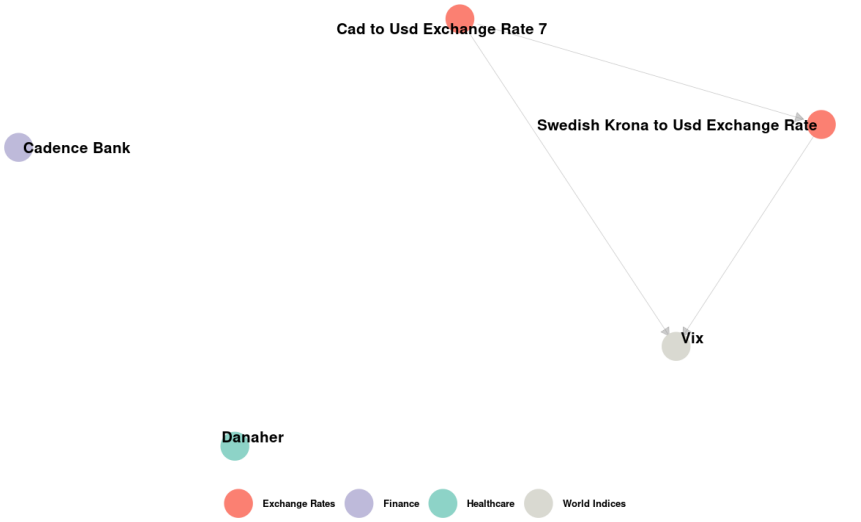
图15 所示的结构保留了汇率与“VIX指数”之间的依赖关系，以及“FNB”和“Cadence Bank”之间的双向依赖关系。

使用 $\boldsymbol{D}_s$，如方程式（9）定义。根据特定标准，删除两个相连资产中的一个即可移除直接连接。在本应用中，移除了表现最差的资产，其表现以索提诺指数衡量 [Sortino 和 Satchell， 2001]，。值得注意的是，使用夏普比率~[Sharpe， 1994]评估资产表现时，最终结果仍保持不变。图16展示了直接连接矩阵 $\boldsymbol{D}_s$。

需要强调的是，“VIX指数”未参与任何正向直接连接。汇率似乎分为两组截然不同的正向直接连接：一组是“CAD兑USD汇率7”和“CAD兑USD汇率5”；另一组是“HKD兑USD汇率”和“瑞典克朗兑美元汇率”。此外，“FNB”与“Cadence Bank”之间存在正向直接连接。根据 [索提诺和 Satchell， 2001]指数表现最差、因而被剔除的资产为“FNB”、“HKD兑USD汇率”和“CAD兑USD汇率5”。



~图16：从节点42“Danahe开始获得的矩阵 $\boldsymbol{D}_s$ 表示 r”.



~图17：从节点42“Danaher”开始获得的矩阵 $\boldsymbol{\Theta}_{s.u}$ 表示——步骤 2。

如第2.4节的步骤 2 所述，移除直接连接后，带符号的邻接~矩阵随之更新。图17展示了更新后的矩阵 $\boldsymbol{\Theta}_{s.u}$，其中突出显示了两个孤立节点（“Cadence Bank”和起始节点“Danaher”），以及一个由简单关联连接而成的节点簇（定义见第2.3节），其中包括“VIX指数”、“瑞典克朗兑美元汇率”和“CAD兑USD汇率7”。

如第 2.4 节的步骤3所述，可根据投资者的具体偏好或业绩指标，采用与此前筛选有向连接类似的标准，剔除简单连接或无向连接。此阶段的结果是一组彼此无连接或负相关的资产，因此适合构建波动率较低的投资组合。

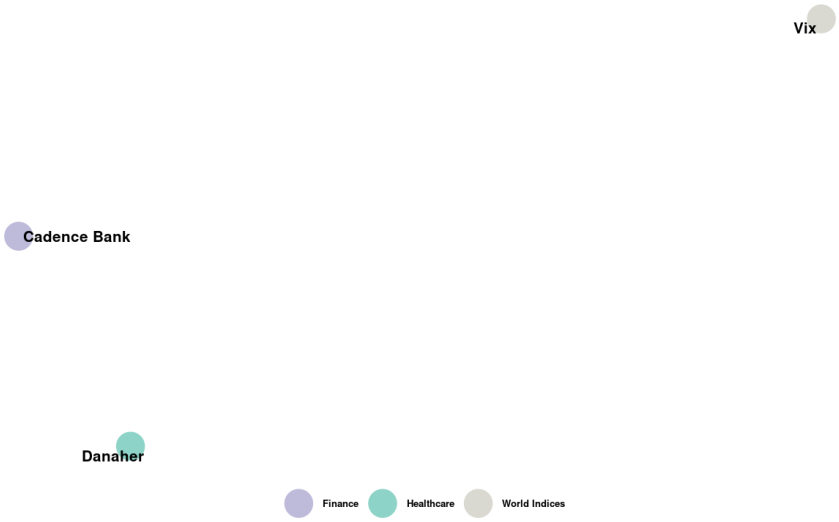


图18：从节点42“Danaher”开始执行步骤3的结果。

图 17 展示了此阶段得到的最终资产集合，其中除了节点“Danaher”外，还包括节点“Cadence Bank”和“VIX指数”。下一节将分析该投资组合的特性。

5.3 最优投资组合的性质

分析采用BPASGM方法得到的最优投资组合在不同配置下的性质。首先，识别并分析由算法选出的收益率所能构成的所有可行投资组合。随后，计算最优投资组合的夏普比率和Sortino指数。这些投资组合中的资产按Markowitz系数加权，并由算法在程序的不同阶段选出。

此外，按照第2.4节所述的方法计算最优投资组合的方差。具体而言，使用DCC-GARCH模型估算最优投资组合的波动率，并将其与假设相同投资组合中的各项资产彼此独立时的波动率进行比较。

为进行这项分析，首先通过随机分配权重给BPSGM选出的资产，构建一组可行投资组合。由此得到图19所示的散点图。
本图报告了在不同投资组合构建程序下生成的投资组合所实现的风险—收益特征。经验风险—收益关系中出现斜率为负的区域，不应被解释为对经典风险—收益权衡的违背，也不应被视为对理论均值—方差前沿构成支配的证据。在矩已知的总体设定中，风险厌恶的投资者会理性地选择前沿上帕累托有效区段的投资组合，而被支配投资组合永远不会被选中。

图19中观察到的模式出现在有限样本环境中，此时投资组合矩必须加以估计，因而会受到估计误差的影响。在这一设定下，基于样本的经典均值—方差优化可能导致事后无效率的投资组合，从而产生较高收益率与较低已实现风险相关联的实现结果。BPASGM通过执行依赖感知型资产选择，并在所选子集上重新估计投资组合矩来缓解这一问题，从而改善估计问题的条件数，并带来更为稳定的实现结果。

因此，图19应解读为稳健性提升的体现，而非投资者偏好或风险—收益权衡发生了根本性变化。表面上的改善反映了 BPASGM 降低估计误差并选择具有有利依赖关系特征的资产的能力，从而使已实现投资组合结果较经典的基于样本的均值—方差投资组合更接近有效区域。

这种模式反映了依赖感知型资产选择对有限样本投资组合表现的影响。根据本研究提出的具体选择标准——系统性地剔除冗余且具有正依赖关系的资产——可以构建出相较于使用标准

基于样本的估计、从完整资产集合中构建的投资组合而言，实际风险—收益特征更加稳定的投资组合。重要的是，这并不代表违反了理论风险—收益权衡。相反，这表明以依赖结构为条件进行投资组合构建，能够显著减少估计误差，使实际结果比朴素的基于样本的马科维茨优化所得结果更接近总体有效前沿。这一结果不仅在理论上颇具启发性，也对资产配置策略和约束投资组合优化具有深远影响，因为它凸显了可行投资组合均值—波动率集合中能够实现卓越表现的区域。

需要强调的是，BPASGM 并不是在缩减后的可行集上求解同一个均值—方差优化问题。相反，BPASGM 引入了一个数据依赖型筛选步骤，使投资组合矩的估计以所选资产子集为条件。具体而言，首先根据资产的依赖结构进行筛选，随后在筛选得到的子集上重新估计预期收益和协方差，并使用这些经筛选条件化的估计值进行均值—方差优化。因此，BPASGM 并不试图超越在矩已知时定义的理论马科维茨最优解，也不等同于在固定估计值下缩减可行集。相反，它定义了一个不同的估计问题：通过依赖感知选择和降维来减轻估计误差、改善矩估计量的条件数，从而提高样本外已实现投资组合结果的稳定性和稳健性，而不颠覆风险—收益权衡这一基本关系。

因此，相关的比较并非是在

$$\max_{w\in\mathcal{W}} U(\hat{\mu},\hat{\Sigma}) \quad \text{and} \quad \max_{w\in\mathcal{W}_g} U(\hat{\mu},\hat{\Sigma}),\ \ \mathcal{W}_g \in \mathcal{W}$$

这会机械地偏向更大的可行集；相关的比较应是在两个截然不同的优化问题之间进行，

$$\max_{w\in\mathcal{W}} U(\hat{\mu},\hat{\Sigma}) \quad \text{and} \quad \max_{w\in\mathcal{W}_g} U(\hat{\mu}_g,\hat{\Sigma}_g),$$

其中，$(\hat{\mu}_g,\ \hat{\Sigma}_g)$ 是经过数据驱动筛选步骤后重新估计的矩。由于筛选程序会同时改变投资组合构建所用估计量的维度和统计性质，因此这些问题并不等价。

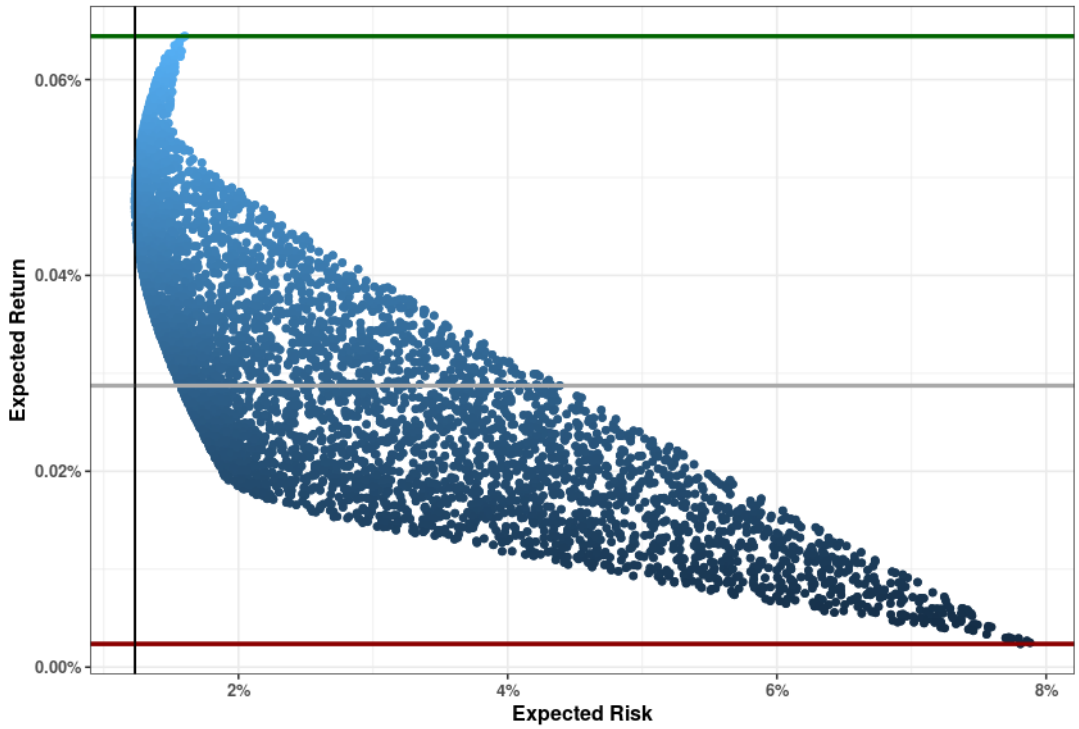
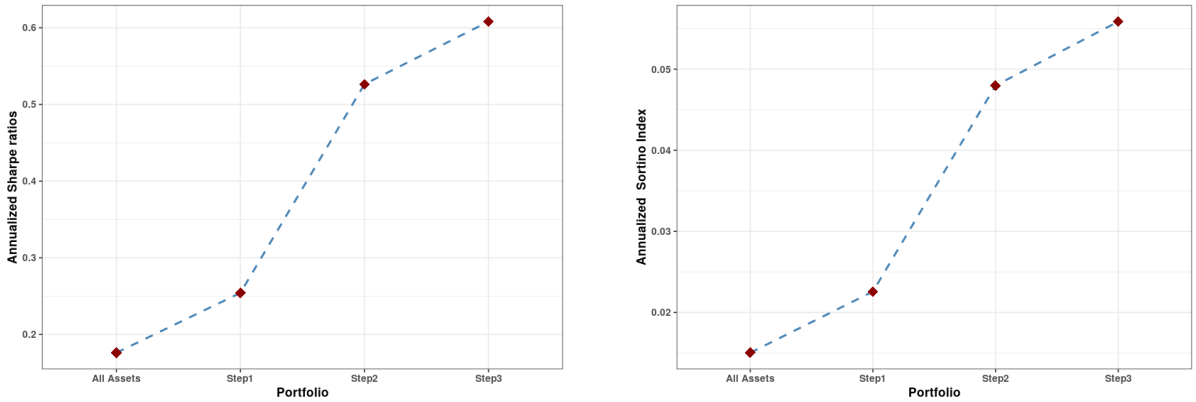


图19：从节点“Danaher”开始，BPSGM 检测到的资产所对应的随机可行投资组合散点图。图中展示了投资组合风险与预期收益之间的权衡。水平线分别标示投资组合收益的最大值（绿色）、中位数（灰色）和最小值（红色），垂直线则表示风险最低的投资组合。

算法每进行一步，都会根据其选出的资产构建最优投资组合，并计算其索提诺比率和夏普比率。资产权重按照最小方差策略确定。图20 展示了年化夏普比率（面板 (a)，20a）和年化索提诺比率（面板 (b)，20b）的变化情况。两种情况下的结果均表明，随着算法逐步推进，纳入投资组合考虑范围的资产集合不断缩小，由此形成的投资组合的财务绩效有所提升。这些证据证实，正如本研究所提出的，适当筛选独立或负相关资产，即使最终投资组合包含的资产数量较少，也能提升投资组合绩效。



(a) 年化夏普比率的演变

(b) 年化索提诺指数的演变

图20：节点“Danaher”各投资组合阶段——“全部资产”、“步骤1”、“步骤 2”和“步骤3”——的年化夏普比率（面板(a)）与年化索提诺指数（面板(b)）。

在每个阶段，多元化指数的相关性依赖成分，即加权平均相关性成分 ρ_{MDP}，由公式10 和11定义，同样进行了计算。图21 展示了 ρ_{MDP} 在BPASGM各个步骤中的演变。可以看出，随着算法逐步推进，逐步排除正相关资产会导致 ρ_{MDP} 的值逐渐下降。

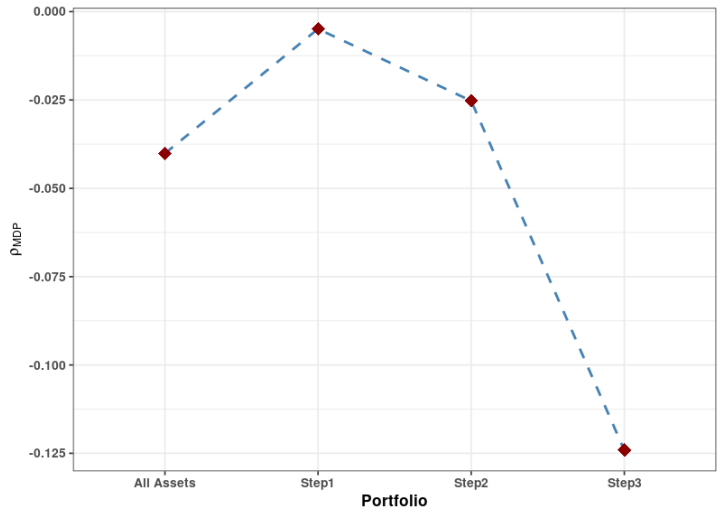


图21： ρ_{MDP} 的演变（定义见公式11 ），对应节点“Danaher”的各投资组合阶段——“全部资产”、“步骤1”、“步骤 2”和“步骤3”。

此外，为考察 BPASGM 选出的股票所构成的最优投资组合的平均收益率与波动率之间是否存在反向关系，我们估计了回归系数

（25）中指定的系数。表1 展示了算法各个阶段基于 BPASGM 选出的股票构建的最优投资组合的回归结果。在初始阶段，投资组合由完整资产池中的资产构成；在后续阶段，则仅由 BPASMG 逐步选出的资产构成。每列报告估计的截距（$\hat{}\alpha$）和斜率$\hat{}$（β），以及它们的标准误、显著性水平和调整后决定系数（$\bar{R}^2$）。纳入全部资产时，$\hat{}\ \beta$ 为正且具有很高的统计显著性，这与投资组合理论预测的风险与预期收益之间经典的正向关系一致。然而，随着算法推进，β 转为负值，同时仍具有统计显著性。这一转变表明，随着冗余或贡献较弱的资产被逐步筛除，传统的风险—收益权衡发生逆转：波动率较低的投资组合开始呈现更高的超额收益。在最终的选择步骤（步骤3）中，β 变为显著负值，凸显出分散化效应相较于传统线性风险—收益机制占据主导地位。与此同时，调整后的 $\bar{R}^2$ 大幅上升——从早期阶段的较低值升至最终选择阶段的 $\bar{R}^2 = 0.596$，表明解释力显著增强，并呈现出更加稳定、界定清晰的经验前沿。

总的来说，这些结果表明，BPASGM程序构建的有效前沿具有逐渐变得更负的风险—收益斜率，但每一步的解释力都更强。随着 $\bar{R}^2$单调增加，同时$\hat{}$ β逐渐变得更负，这表明，随着冗余资产被剔除，BPASGM会筛选出一组投资组合，其中分散化效应系统性地主导了传统风险溢价，从而形成结构上反向的风险—收益关系。解释力的逐步提升与反向风险—收益系数的同步增强，与引理2.1相符；该引理从理论上支撑了算法的筛选机制。具体而言，一旦BPASGM识别出资产之间的依赖结构——捕捉某些资产对其他资产的解释或预测作用——剔除选定资产就会打破这些依赖关系。根据马尔可夫性质，这类破坏会使剩余资产之间形成条件独立，从而改善投资组合的信息结构并增强其解释连贯性；这一点在实证上体现为 $\bar{R}^2$ 不断增加，以及 β的$\hat{}$幅度不断增大。

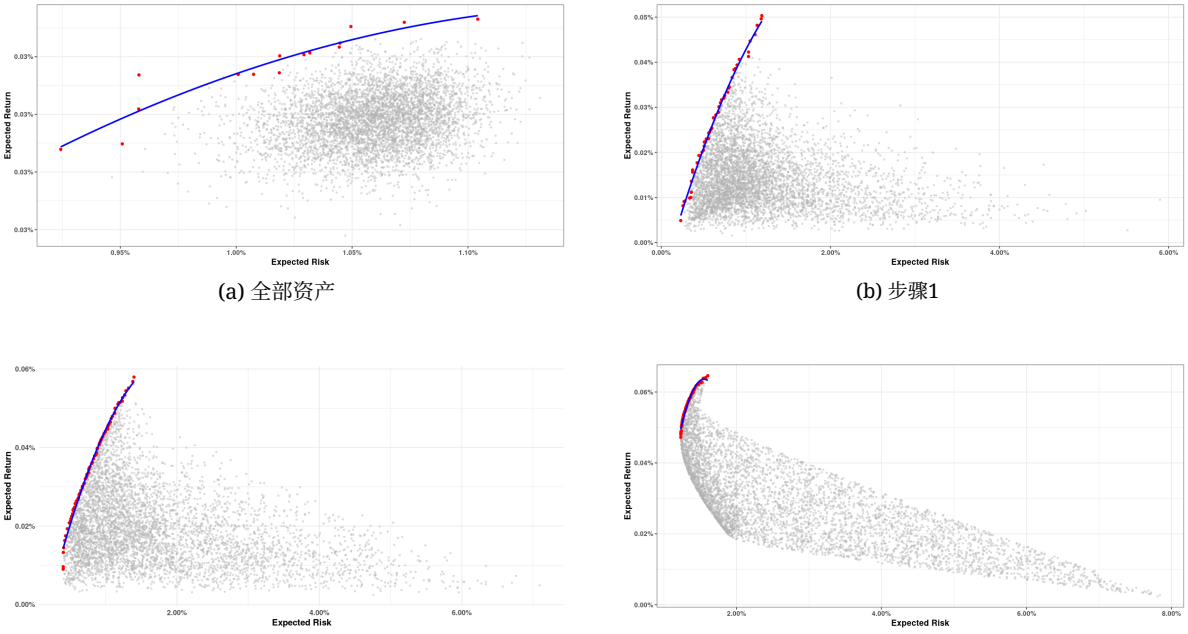
这支持了这样一种解释：该方法增强了分散化并减少了冗余，从而构建出表现特征更稳定的投资组合，如图 20 和 21所示。图22 展示了这些结果；5,000 个投资组合样本以灰色点表示，红色点标出了经验有效前沿上的投资组合，蓝线则表示通过二次拟合对前沿进行的平滑近似。

附录E比较了通过 BPASGM 得到的最优投资组合与 GLASSO 技术的表现。作为补充，附录E报告了基于 BPASGM 的投资组合与使用图形套索（ Glasso ）估计的协方差矩阵构建的投资组合之间的比较。这一比较凸显了两者在方法上的根本区别。Glasso 通过施加 ℓ_1 惩罚，直接在逆协方差矩阵层面强制实现稀疏性，从而在保留完整资产集合的同时，对二阶矩估计进行正则化。相比之下，BPASGM 并不旨在估计稀疏精度矩阵。相反，它会进行依赖感知型资产选择，在投资组合优化之前重塑资产集合，随后在选定的子集上重新估计投资组合矩。

附录中报告的结果表明，尽管稀疏协方差正则化可以缓解估计误差，但在保留完整资产池时，它无法彻底解决强依赖资产带来的冗余问题。BPASGM 通过明确剔除冗余资产，而非仅仅收缩其依赖结构，对这些方法进行了补充。因此，绩效差异应理解为源自两种根本不同的机制——资产选择与协方差正则化，而不是对同一估计问题采用了不同的调参方式。

表 1：沿经验有效前沿的投资组合收益对投资组合波动率的线性回归（方程（25）），针对 BPASGM 程序的不同选择步骤——节点“Danaher”。

	因变量： μ_p			
	全部资产	**步骤1**	**步骤 2**	**步骤3**
$\hat{\beta}$	0.008***	–0.002***	–0.003***	–0.007***
	(5.611e–04)	(1.399e–04)	(1.166e–04)	(7.864e–05)
$\hat{\alpha}$	0.0002***	0.0002***	0.0002***	0.0005***
	(5.936e–06)	(1.892e–06)	(2.418e–06)	(2.515e–06)
观测值	5,000	5,000	5,000	5,000
R^2	0.043	0.027	0.108	0.596
注：			*p<0.1;	**p<0.05; ***p<0.01



(a) 全部资产　(b) 步骤1

(c) 步骤 2　(d) 步骤3

图22：从数据集构建的有效前沿，采用从“Danaher”节点初始化的 BPASGM 程序。蓝色曲线表示经验有效前沿的平滑二次拟合，该前沿由风险收益空间中的非支配投资组合构成。灰色点表示随机生成投资组合的完整集合，红色点则突出显示位于经验有效前沿上的投资组合。

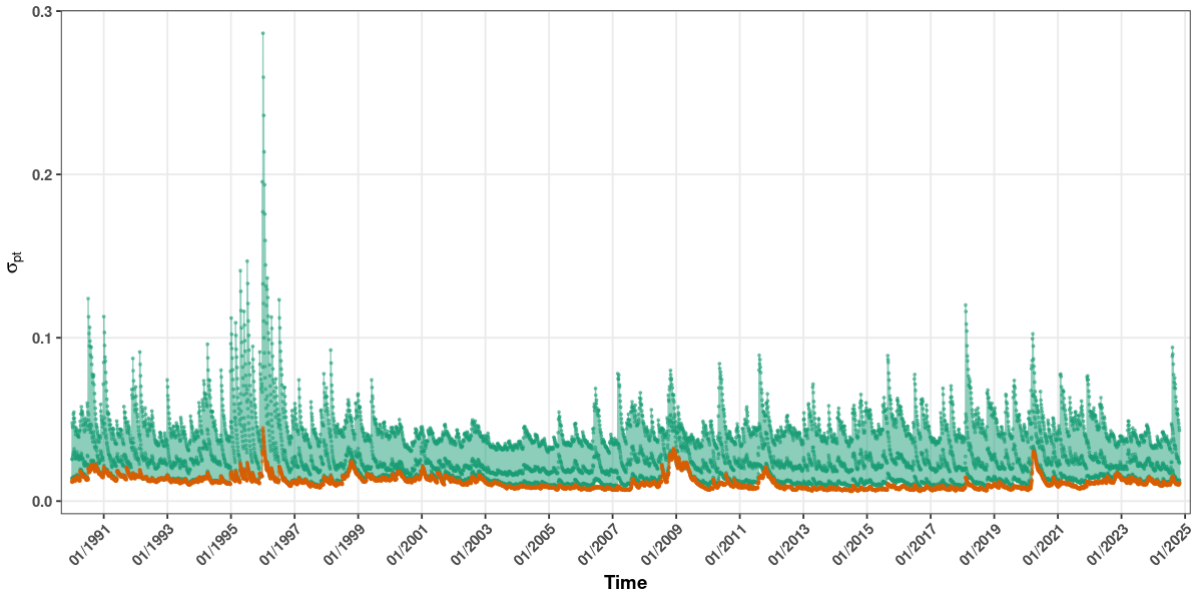
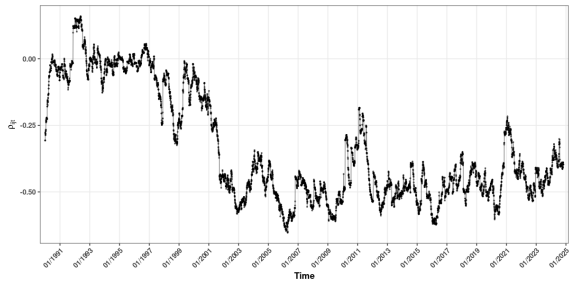


图23：以“Danaher”节点为起点的最优投资组合中，由单变量 GARCH 模型（绿色）和 DCC-GARCH 模型（橙色）估计的样本内波动率轨迹。

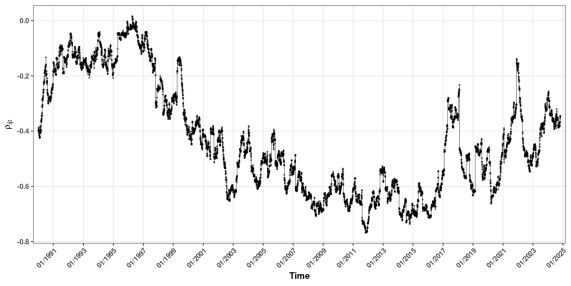
最后，图23展示了样本内波动率轨迹：绿色曲线表示对构成最优投资组合的单项资产应用单变量 GARCH 模型后得到的估计值，并以 Markowitz 系数加权；橙色曲线则表示对源自“Danaher”节点的最优投资组合应用 DCC-GARCH 模型后得到的估计值。图中显示，尽管两种模型都捕捉到了市场波动率升高的共同阶段（例如 20 世纪 90 年代中期和 2008年金融危机），但与单变量 GARCH 模型相比，DCC-GARCH 模型得出的波动率估计值始终更低、更平滑。

这一差异尤其重要，因为它凸显了多变量 DCC-GARCH 模型能够更充分地利用分散化收益。该模型通过纳入动态的（负）条件相关性，考虑了随时间变化的相互依赖关系和跨资产风险溢出，从而降低了整体投资组合波动率。相比之下，分别对各项资产建模的单变量 GARCH 模型无法捕捉这些分散化效应，因此会高估总体波动率。

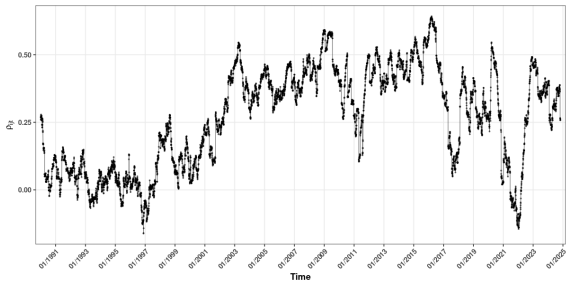
这些发现为第4.1节所概述的理论框架提供了实证支持，表明使用BPASGM构建的投资组合——该方法纳入了时变负相关性——相比资产之间简单视为相互独立的投资组合，整体波动率更低。图24 展示了 ρ_t （方程（23））中所示的时变条件相关性，这些相关性是通过DCC-GARCH模型在样本内期间对源自“Danaher”节点的最优投资组合进行估计得到的。如图所示， ρ_t “VIX”指数与“Cadence Bank”之间的相关性（24a）以及 ρ_t “VIX”指数与“Danaher”之间的相关性（24b）在整个期间始终为负。相比之下，ρ_t“Cadence Bank”与“Danaher”之间的相关性（24c）表现出更大的变异性，在正负两个方向上均出现波动和尖峰。



(a) ρ_t “VIX指数”与“Cadence Bank”之间



(b) ρ_t “VIX指数”与“Danaher”之间



(c) ρ_t “Cadence Bank”与“Danaher”之间

图24：从“Danaher”节点开始的最优投资组合在样本内期间通过 DCC-GARCH模型估计的时变条件相关性 ρ_t（方程式（23））。

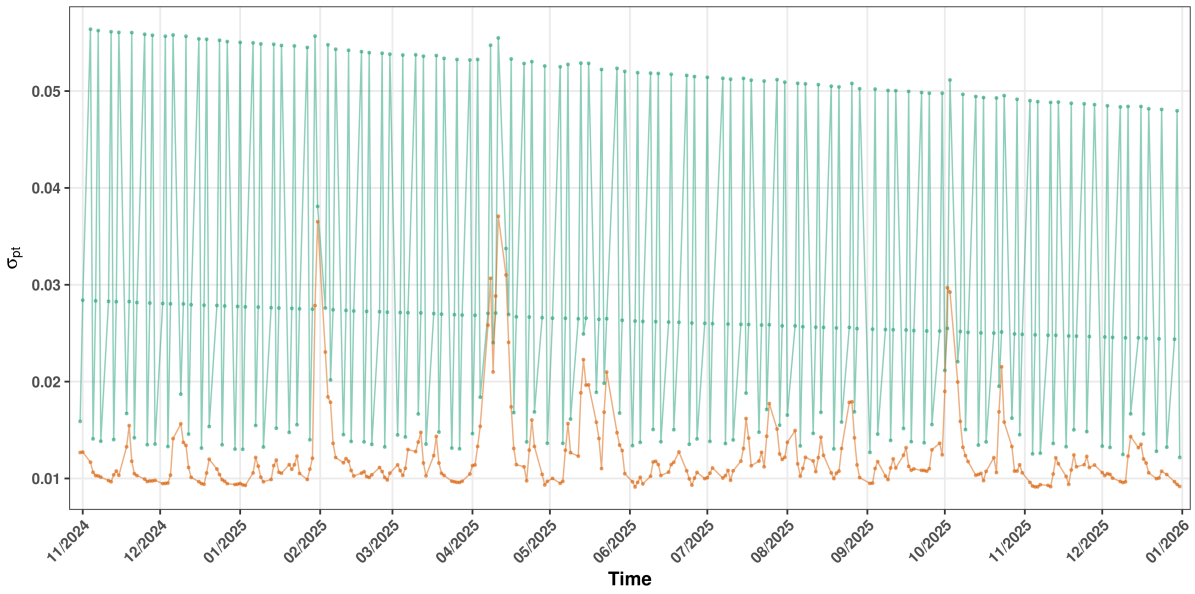
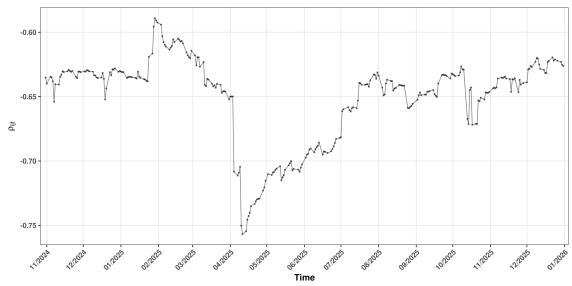


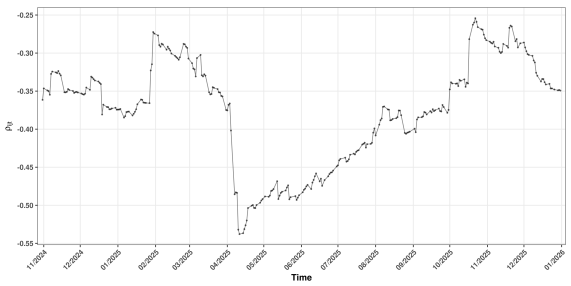
图25：从“Danaher”节点开始的最优投资组合的样本外波动率轨迹，分别由单变量GARCH模型（绿色）和DCC-GARCH模型（橙色）估计。

图25 展示了由单变量GARCH模型（绿色）和DCC-GARCH模型（橙色）估计的、源自“Danaher”节点的最优投资组合的样本外波动率轨迹。该图印证了样本内分析（23）的发现，显示由DCC-GARCH模型估计的最优投资组合波动率始终更低且更加平滑，相比之下，虚拟投资组合的波动率则更高；在虚拟投资组合中，单变量GARCH模型分别应用于被假定相互独立的成分资产。图26 展示了通过DCC-GARCH模型在样本外期间对所选节点估计的时变相关性（ρ_t）。通过捕捉投资组合资产之间不断变化的负向共同变动，基于BPASGM的

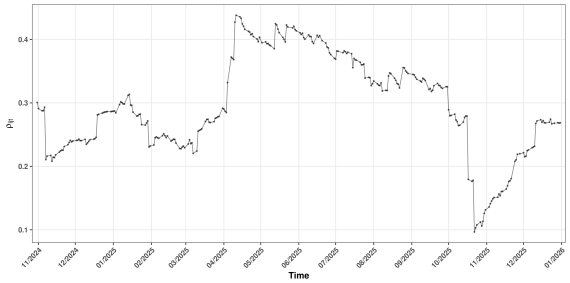
该方法增强了对系统性风险和资产相互依赖关系的建模，从而使投资组合在样本内和样本外均表现出系统性更低的波动率。



(a) ρ_t 在“VIX指数”与“Cadence Bank”之间



(b) ρ_t 在“VIX指数”与“Danaher”之间



(c) ρ_t 在“Cadence Bank”与“Danaher”之间

图 26：时变条件相关性 ρ_t （式（23）），通过DCC-GARCH模型对从节点“Danaher”出发的最优投资组合在样本外期间进行估计。

6 结论

本文提出了最佳路径资产选择图模型（BPASGM），这是一种关注依赖关系的框架，用于资产选择和投资组合构建，并将最佳路径算法从依赖关系发现扩展至可操作的投资组合方法论。BPASGM 将稀疏图模型与结构化剪枝程序相结合，明确考虑资产收益率之间依赖关系的符号与冗余性。

所提出的方法并非旨在改进已知总体参数条件下定义的理论均值—方差最优解，而是针对高维情形下基于样本的均值—方差优化在实践中的局限性；在这类情形中，估计误差和冗余依赖关系起着主导作用。通过依据精心筛选出的资产子集来构建投资组合（该子集的特点是依赖性较弱或呈负依赖），BPASGM 提升了有限样本下的稳健性和实际表现。

蒙特卡洛模拟和实证证据表明，与经典的基于样本的 Markowitz 投资组合相比，基于 BPASGM 的投资组合具有更低的实际波动率、更优的风险调整后绩效，且投资组合基数大幅降低。这些改进既源于估计误差的减少，也源于基于经济考量选择具有有利依赖结构的资产，而非对理论风险—收益权衡的违背。

从方法论角度来看，BPASGM 是对原始 BPA 的重要拓展。BPA 专注于揭示依赖结构，而 BPASGM 则将这些信息转化为可复现的资产选择与投资组合构建流程。该框架计算效率高、易于解释且灵活，并且可以自然地与方差以外的其他风险度量相结合。

总体而言，BPASGM 为稀疏图模型与投资组合理论搭建了桥梁，并为未来关于依赖感知型资产选择、稳健投资组合优化和高维金融决策的研究提供了颇具前景的基础。

参考文献

T. Aste 和 T. Di Matteo。基于短时间序列检索稀疏因果网络。复杂性，2017 (1):4518429，2017。T. Aste、W. Shaw 和 T. Di Matteo。波动性市场中的相关结构与动态。新物理学杂志，12(8):085009，2010。C. R. Bacon。实用风险调整绩效度量。约翰·威利父子公司，2021。C. R. Bacon。实用投资组合绩效度量与归因。约翰·威利父子公司，2023。W. Barfuss、G. P. Massara、T. Di Matteo 和 T. Aste。利用信息过滤网络进行简约建模。《物理评论*E*辑》，94(6):062306，2016。F. Black 和 R. Litterman。全局投资组合优化。金融分析师杂志，48(5):28–43，1992。S. Bleik。概念图：在生物医学文本分类与概念抽取中的应用。2013。F. E. Block。市净率关系研究。金融分析师杂志，51(1):63–73，1995。T. Bodnar、Y. Okhrin 和 N. Parolya。高维环境下基于收缩的最优投资组合选择。商业与经济统计杂志，41(1):140–156，2022。T. Bollerslev。广义自回归条件异方差。计量经济学杂志，31(3): 307–327，1986。M. D. Braga、L. Riso 和 M. G. Zoia。运用投资组合波动率与峰度的创新型基于风险的资产配置策略的理论性质。国际金融分析评论，第104728页，2025。M. Broadie。利用估计参数计算有效前沿。运筹学年鉴，45 (1):21–58，1993。D. B. Brown 和 J. E. Smith。考虑交易成本的动态投资组合优化：启发式方法与对偶界。管理科学，57(10):1752–1770，2011。R. Campbell、R. Huisman 和 K. Koedijk。在险价值框架下的最优投资组合选择。银行与金融杂志，25(9):1789–1804，2001。R. Carroll、T. Conlon、J. Cotter 和 E. Salvador。考虑相关性的资产配置：一种综合权衡。《欧洲运筹学杂志》，262(3):1164–1180，2017。Y. Choueifaty、T. Froidure 和 J. Reynier。最大分散化投资组合的性质。投资策略杂志，2(2):49–70，2013。A. Dembo、T. M. Cover、J. A. Thomas 等。信息论不等式。*IEEE*信息论汇刊，37(6):1501–1518，1991。D. Edwards、G. C. De Abreu 和 R. Labouriau。使用最小AIC或BIC森林选择高维混合图模型。*BMC*生物信息学，11(1):18，2010。R. Engle。动态条件相关：一类简单的多变量广义自回归条件异方差模型。商业与经济统计杂志，20(3):339–350，2002。N. Eshima 和 M. Tabata。用于衡量广义线性模型预测能力的熵相关系数。统计与概率快报，77(6):588–593，2007。N. Eshima 和 M. Tabata。广义线性模型的熵决定系数。计算统计与数据分析，54(5):1381–1389，2010。F. J. Fabozzi、H. M. Markowitz 和 F. Gupta。投资组合选择。金融手册，2:3–13，2008。J. Friedman、T. Hastie 和 R. Tibshirani。使用图形套索进行稀疏逆协方差估计。生物统计学，9(3):432–441，2008。

L. Gan、N. N. Narisetty 与 F. Liang。不等收缩图模型的贝叶斯正则化。美国统计学会期刊，114(527)：1218–1231，2019年。A. Garcia-Bernabeu、A. Hilario-Caballero、F. Tardella 与 D. Pla-Santamaria。投资组合选择中的ESG整合：一种基于偏好的稳健多准则方法。运筹学视角，12：100305，2024年。C. Gregory、K. Darby-Dowman 与 G. Mitra。鲁棒优化与投资组合选择：稳健性的代价。《欧洲运筹学杂志》，212(2)：417–428，2011年。J. Hakem。提升投资组合效率：马科维茨优化模型中的基金分离。博士论文，维尔纽斯大学，2025年。R. A. Horn 与 C. R. Johnson。矩阵分析。剑桥大学出版社，2012年。H. Jacobs、S. M¨uller 与 M. Weber。个人投资者应如何分散化？替代资产配置策略的实证评估。《金融市场杂志》，19：62–85，2014年。M. I. Jordan 等人。图模型。《统计科学》，19(1)：140–155，2004年。A. Kraskov、H. St¨ogbauer 与 P. Grassberger。互信息估计。《物理评论E辑》，69(6)：066138，2004年。N. Lassance 与 F. Vrins。投资组合选择：一种目标分布方法。《欧洲运筹学杂志》，310(1)：302–314，2023年。S. L. Lauritzen 与 N. Wermuth。用于变量间关联的图模型：部分变量为定性变量，部分为定量变量。《统计年鉴》，第31–57页，1989年。O. Ledoit 与 M. Wolf。亲爱的，我缩小了样本协方差矩阵。2003a。O. Ledoit 与 M. Wolf。股票收益协方差矩阵的改进估计及其在投资组合选择中的应用。《实证金融学期刊》，10(5)：603–621，2003b。W. Lefebvre、G. Loeper 与 H. Pham。带跟踪误差惩罚的均值-方差投资组合选择。《数学》，8(11)：1915，2020年。R. N. Mantegna。金融市场中的层级结构。《欧洲物理期刊B辑——凝聚态物质与复杂系统》，11(1)：193–197，1999年。A. B. Manual。统计学习导论及其在R中的应用。2013年。H. Markowitz。投资组合选择。《金融学杂志》，7(1)：77–91，1952a。ISSN 00221082, 15406261。URL http://www.jstor.org/stable/2975974。H. Markowitz。投资组合选择/h. markovitz。《金融学杂志》，7(1)：77–91，1952b。H. M. Markowitz、F. Schirripa 与 N. D. Tecotzky。一条更有效的前沿。《投资组合管理期刊》，第99页，1999年。R. D. Martin、A. Clark 与 C. G. Green。鲁棒投资组合构建。载于《投资组合构建手册》，第337–380页。施普林格，2010年。M. J. Mauboussin。企业价值与息税折旧摊销前利润倍数意味着什么。Bluemont Investment Research，第1–19页，2018年。N. Meinshausen 与 P. B¨uhlmann。高维图与基于套索回归的变量选择。2006年。A. Meucci。管理分散化。《风险》，第74–79页，2009年。R. O. Michaud 与 R. Michaud。估计误差与投资组合优化：一种重采样解法。可于SSRN 2658657获取，2007年。P. M. Mirete-Ferrer、A. Garcia-Garcia、J. S. Baixauli-Soler 与 M. A. Prats。资产管理中机器学习研究综述。《风险》，10(4)：84，2022年。

L. M¨uller 与 M. Joubrel。可持续均值-方差投资组合优化的新方法：考虑ESG相关不确定性。《金融研究快报》，第108056页，2025年。N. Musmeci、T. Aste 与 T. Di Matteo。风险分散：基于过滤相关网络方法的持续性研究。*arXiv*预印本 *arXiv:1410.5621*，2014年。A. Nezlobin、M. V. Rajan 与 S. Reichelstein。市盈率与市净率的结构性质。《会计研究评论》，21(2)：438–472，2016年。L. Olbert。金融分析师对行业特定股票估值模型的使用。《应用会计研究期刊》，26(6)：108–138，2025年。J. Olmo。最优投资组合配置问题中的估计误差。载于《牛津经济学与金融学研究百科全书》。2023年。Z. Pan 与 L. Liu。股票收益波动率预测：短期与长期杠杆效应作用的比较。《物理学*A*：统计力学及其应用》，492：168–180，2018年。M. D. Patil 与 S. S. Sane。降维：综述。《国际计算机应用期刊》*92*，16：23–29，2014年。G. Peralta 与 A. Zareei。投资组合选择的网络方法。《实证金融学期刊》，38：157–180，2016年。F. Pozzi、T. Di Matteo 与 T. Aste。动态金融相关性过滤图中的中心性与边缘性。《复杂系统进展》，11(06)：927–950，2008年。F. Pozzi、T. Di Matteo 与 T. Aste。风险在金融市场中的扩散：投资于边缘更为有利。《科学报告》，3(1)：1665，2013年。M. Raddant 与 T. Di Matteo。从经济物理学与金融经济学视角审视金融依赖关系。《经济互动与协调期刊》，18(4)：701–734，2023年。A. H. Ramadan。资本结构的决定因素与企业财务绩效：在英国资本市场中的应用。萨里大学（英国），2009年。R. Riccobello、G. Bonaccolto、P. J. Kremer、S. Paterlini 与 M. Bogdan。面向最小方差投资组合的稀疏图模型建模。可于*SSRN 4383314*获取，2023年。L. Riso 与 M. Guerzoni。基于图模型的概念漂移估计。《信息科学》，606：786–804，2022年。L. Riso、M. G. Zoia 与 C. R. Nava。高维图模型中基于最佳路径算法的特征选择。《信息科学》，649：119601，2023年。K. Sadeghi 与 S. Lauritzen。混合图的马尔可夫性质。2014年。K. Scheinberg、S. Ma 与 D. Goldfarb。通过交替线性化方法进行稀疏逆协方差选择。《神经信息处理系统进展》，23，2010年。W. F. Sharpe。夏普比率。《投资组合管理期刊》，21(1)：49–58，1994年。F. A. Sortino 与 S. Satchell。管理金融市场中的下行风险。Butterworth-Heinemann，2001年。X. Tu 与 B. Li。结合智能收益预测的稳健投资组合选择。《经济建模》，135：106719，2024年。H. R. D. Vinod 与 D. Reagle。为最坏情况做准备：将下行风险纳入股市投资。约翰·威利父子公司，2004年。R. E. Whaley。投资者恐惧指数。《投资组合管理期刊》，26(3)：12，2000年。J. Yin 与 H. Li。用于遗传基因组学数据分析的稀疏条件高斯图模型。《应用统计学年鉴》，5(4)：2630，2011年。

T. Zerenner、P. Friederichs、K. Lehnertz 和 A. Hense。基于高斯图模型的气候网络研究方法。*Chaos: An Interdisciplinary Journal of Nonlinear Science*，24(2)，2014年。Q. Zhou。具有稳健协方差和条件风险价值约束的投资组合优化。*arXivpreprint arXiv:2406.00610*，2024年。

一种最佳路径算法

BPA 利用 HDGM 的结构进行变量选择 [参见Edwards 等人、2010年、Riso 和 Guerzoni, 2022], 通过利用互信息来定义最终生成的树或森林的拓扑结构。众所周知，互信息（MI） [Kraskov 等人, 2004] 通过衡量一个变量的熵在由另一个变量解释后减少所带来的信息增益，来量化变量之间的关系。因此，两个变量之间的 MI 越大，一个变量对另一个变量的解释作用就越显著。该算法通过测量图中属于不同路径步的变量之间的 MI，来识别 X_i 的可接受预测变量集合 ps_i ，并将这些变量连接到 X_i。路径步是由与目标变量 X_i相连的一组变量构成的。路径步根据变量与 X_i之间的（最大）距离进行排序。最佳路径步骤是指包含与 X_i共享最大 MI 的变量的路径步，也就是说

$$MI(X_i, \boldsymbol{X}_{ps_i}) \geq MI(X_i, \boldsymbol{X}_{ps_j}), i \neq j, i = 1, \ldots, p \tag{A.1}$$

其中 $\boldsymbol{X}_{ps_i}$， $\boldsymbol{X}_{ps_j}$ 表示属于路径步 ps_i 和^{ps}j的变量集合。随后，X_i 与属于不同路径步的变量之间的互信息通过 Kullback-Leibler 信息（KL）指数进行量化。 KL指数衡量两个随机变量的概率密度之间的散度，并与其互信息相关联，因为后者对应于 KL在其密度之间的期望 [Dembo 等人， 1991]。

在 BPA 中， X_i 与某一给定路径步（设为 $\boldsymbol{X}_{ps_i}$）中的变量之间的互信息，是通过衡量 X_i 的分布与其在由 $\boldsymbol{X}_{ps_i}$所生成空间上的投影（记为 $f(X_i|\boldsymbol{X}_{ps_i})$）的分布之间的散度来评估的，即:

$$MI(X_i, \boldsymbol{X}_{ps_i}) = \mathbb{E}_{f(X_i)}[KL^{(s)}(f(X_i), f(X_i|\boldsymbol{X}_{ps_i})]. \tag{A.2}$$

其中 $KL^{(s)}$ 为对称 KL 信息，定义如下

$$KL^{(s)}[f(X_i), f(X_i|\boldsymbol{X}_{ps_i})] = KL[f(X_i), f(X_i|\boldsymbol{X}_{ps_i})] + KL[f(X_i|\boldsymbol{X}_{ps_i}), f(X_i)] \tag{A.3}$$

$KL^{(s)}$ 通过如下关系与熵系数（EC）相关联

$$EC(X_i, \boldsymbol{X}_{ps_i}) = \frac{1}{\{\#X_j \in ps_i\}} \sum_{X_j \in ps_i} KL^{(s)}[f(X_i), f(X_i|\boldsymbol{X}_{ps_i})] \tag{A.4}$$

并进而与熵决定系数相关联，其规定为（Eshima 和 Tabata [2007, 2010]）

$$ECD = \frac{EC}{EC+1} \tag{A.5}$$

ECD 指数是标准决定系数的推广，可视为由 $\boldsymbol{X}_{ps_i}$所解释的 X_i方差/熵的比例。

一旦变量 X_i 的最佳预测变量集被检测出来，便通过应用Kraskov 等人 [2004] 检验来考察其中每一个的解释作用。该检验检查X_i 与最优路径步 X_i中的每个变量之间是否存在关系。若 ps_i 中的某些变量对 X_i 不具有统计显著性，则可通过将其排除来进一步缩减预测变量集。该过程可得到针对 X_i的更为简约预测变量集。

对于最佳最优预测变量集 $\boldsymbol{X}_{ps_i}$ （即经由 BPA 为 X_i 检测得到的预测变量），如下所示的关系 g 可以假定成立

$$X_i \approx \phi(\boldsymbol{X}_{ps_i}) + \epsilon$$

其中 $\phi(\cdot)$ 是属于广义线性模型（广义线性模型）族的一般函数 [Eshima 和 Tabata, 2007], 并且 ϵ 是一个随机项。

B 收益率数据集中的成分资产

表2: 分析中用于标记节点的资产代码（包括股票和汇率） S

标签	收益率名称	标签	收益率名称	标签	收益率名称	标签	收益率名称
1	微软	91	Paychex	181	Loews	271	Comerica
2	Apple	92	Csx	182	Gen Digital	272	Toro
3	伯克希尔·哈撒韦 A 股	93	Public Storage	183	西南航空	273	SPX Technologies
4	沃尔玛	94	联邦快递	184	哈里伯顿	274	Repligen
5	摩根大通	95	ONEOK	185	KeyCorp	275	Light & Wonder
6	礼来	96	Truist Financial	186	卡莱尔公司	276	Zions Bancorporation
7	埃克森美孚	97	Fair Isaac	187	Casey's General Stores	277	Rli
8	甲骨文	98	诺福克南方铁路	188	高乐氏	278	莱迪思半导体
9	强生	99	固安捷	189	实耐宝	279	贝杰计量仪器
10	宝洁	100	碧迪医疗	190	库珀公司	280	Slm
11	家得宝	101	帕卡	191	布朗-福曼 B 股	281	爱达荷电力公司
12	联合健康集团	102	克罗格	192	Genuine Parts 公司	282	伯乐生命医学 A 股
13	美国银行	103	美国国际集团	193	联合能源	283	Bath & Body Works
14	可口可乐	104	斯伦贝谢	194	百特国际	284	HF辛克莱
15	雪佛龙	105	道明尼能源	195	西部数据	285	Synovus金融
16	富国银行	106	罗斯百货	196	Evergy	286	阿拉斯加航空集团
17	雅培实验室	107	Exelon	197	West Pharmsvs	287	Louisiana Pacific
18	International Busmchs	108	Fastenal	198	Omnicom Group	288	Arrow Electronics
19	GE Aerospace	109	Kimberly-Clark	199	Jacobs Solutions	289	Brinker International
20	麦当劳	110	塔吉特	200	百思买	290	Apa
21	Att	111	康明斯	201	柯蒂斯-莱特	291	莱德系统
22	美国运通	112	西方石油	202	Ball	292	纽马克特
23	默克公司	113	福特汽车	203	尤纳姆集团	293	奥什科什
24	华特迪士尼	114	L3Harris Technologies	204	Udr	294	Kirby
25	威瑞森通信	115	Hess	205	Graco	295	林肯国民公司
26	百事公司	116	Xcel能源	206	Idex	296	Dillard's A
27	Rtx	117	艺电	207	泰尼特医疗	297	Cadence Bank
28	Progressive俄亥俄州	118	阿美特克	208	艾利丹尼森	298	穆格A
29	超威半导体	119	Pubserentergp	209	Hunt Jb运输服务公司	299	美泰
30	奥多比	120	联合爱迪生	210	Jack Henry And Associates	300	Teleflex
31	德州仪器	121	康宁	211	CNA Financial	301	InterDigital
32	标普全球	122	瓦莱罗能源	212	Masco	302	Federal Signal
33	嘉信理财	123	Pge	213	德事隆	303	Selective Ingp
34	卡特彼勒	124	贝克休斯 A	214	泰瑞达	304	西南燃气控股公司
35	赛默飞世尔科技	125	恩特吉	215	皇家黄金	305	Balchem
36	波音	126	维肯材料	216	Healthpeak Properties	306	美国蜂窝公司
37	史赛克	127	康德乐	217	Clean Harbors	307	Cirrus Logic
38	NextEra Energy	128	好时	218	J. M. Smucker	308	Gatx
39	Tjx	129	WEC能源集团	219	Encompass Health	309	Commercial Metals
40	安进	130	Sysco	220	Mastec	310	V F
41	霍尼韦尔国际	131	星座品牌A	221	环球健康服务B	311	Champion Homes
42	Danaher	132	艾可飞	222	Crown Holdings	312	联合银行股份公司
43	迪尔	133	Eqt	223	Essential Utilities	313	Coeur Mining
44	花旗集团	134	罗克韦尔自动化	224	Leonardo DRS	314	Gentex
45	联合太平洋铁路	135	HEICO	225	康尼格拉品牌	315	Modine Manufacturing
46	康卡斯特 A 类股	136	Humana	226	富兰克林资源	316	康耐视
47	辉瑞	137	德克萨斯太平洋土地信托	227	相干公司	317	Fnb
48	应用材料	138	通用磨坊	228	诺信	318	德克萨斯新墨西哥电力公司
49	自动数据处理公司	139	雷蒙德·詹姆斯金融公司	229	顶峰西部资本	319	梅里蒂奇家园
50	劳氏公司	140	Kellanova	230	国际服务公司	320	铁姆肯
51	康菲石油	141	莱纳	231	Revvity	321	伊坎企业
52	威达信	142	M&T银行	232	卡朋特科技	322	新泽西资源
53	亚德诺	143	DTE能源	233	金宝汤公司	323	Cousins Properties
54	Fiserv	144	W. R. Berkley	234	摩森康胜饮料（B）	324	White Mountains Insurance Group
55	奥驰亚集团	145	罗林斯	235	杰富瑞金融集团	325	Fmc
56	南方公司	146	道富集团	236	Viatris	326	Lancaster Colony
57	Lam Research	147	Nucor	237	Toll Brothers	327	Robert Half
58	Micron Technology	148	Ameren	238	American Finlgpohio	328	Alliancebernstein Hldg Unt
59	Bristol Myers Squibb	149	Ppl	239	Sei Investments	329	Lumen Technologies
60	废物管理公司	150	嘉年华公司	240	Ati	330	希悦尔公司
61	英特尔	151	First Ctznbcsha	241	环球人寿	331	惠而浦
62	Kla	152	Atmos Energy	242	可口可乐联合公司	332	索诺科产品公司
63	信诺	153	Hp	243	Smith	333	Spire
64	宣伟	154	五三银行	244	史丹利百得	334	标准普尔500综合价格指数
65	Cintas	155	PPG工业	245	Stifel Financial	335	STOXX欧洲600E价格指数
66	亚瑟·J·加拉格尔	156	第一能源	246	宏盟集团	336	富时100价格指数
67	耐克B股	157	森特波因特能源	247	OGE能源	337	DAX绩效价格指数
68	CVS健康	158	泰勒科技	248	瑞格尔雷克斯诺德	338	东证价格指数
69	楷登电子设计系统	159	都福	249	美高梅国际酒店集团	339	纳斯达克综合价格指数
70	派克汉尼汾	160	马克尔集团	250	EastGroup Properties	340	法国CAC 40价格指数
71	3m	161	国际纸业	251	纽约时报A股	341	IBEX 35价格指数
72	通用动力	162	阿彻丹尼尔斯米德兰	252	孩之宝	342	瑞士市场价格指数
73	高露洁棕榄	163	Eversource能源	253	Gap	343	奥斯陆交易所
74	艺康	164	辛辛那提金融	254	商业银行股份	344	日经225指数
75	伊利诺伊工具公司	165	丘奇德怀特公司	255	应用工业技术	345	罗素2000价格指数
76	威廉姆斯	166	亨廷顿银行控股	256	韦伯斯特金融	346	道琼斯指数
77	诺斯罗普·格鲁曼	167	CMS能源	257	Chemed	347	标普/多伦多60价格指数
78	摩托罗拉系统	168	爱迪生国际	258	唐纳森公司	348	纽约证交所综合价格指数
79	PNC金融服务集团	169	德文能源	259	联邦房地产投资信托	349	VIX指数
80	美国合众银行	170	普尔特集团	260	老国民银行	350	美元兑欧元汇率
81	豪美特航空航天	171	普信集团	261	沃茨水务技术	351	日元兑美元汇率
82	艾默生电气	172	味好美公司	262	NNN房地产投资信托	352	瑞士法郎兑美元汇率
83	欧特克	173	威廉姆斯-索诺玛	263	Bio-Techne	353	加元兑美元汇率
84	旅行者公司	174	泰森食品A类股	264	UMB金融	354	美元兑英镑汇率
85	EOG资源公司	175	Ptc	265	H&R Block	355	加元兑美元汇率7
86	空气产品与化学品公司	176	北方信托	266	房利美	356	USD兑AUD汇率
87	纽蒙特公司	177	地区金融新	267	Ugi	357	HKD兑USD汇率
88	怪物饮料公司	178	惠好公司	268	沃纳多房地产信托	358	瑞典克朗兑美元汇率
89	Aflac	179	Intlflavors Frag	269	Middleby		
90	Amerelecpwr	180	NiSource	270	National Fuel Gas		

C 引理 2.1 的证明

证明。令 $X = \{X_1, \ldots, X_p\}$ 表示资产收益变量集，并令 $X_{ps_i} \subset X \setminus \{X_i\}$ 表示由最佳路径算法（BPA）选出的 X_i 的最优预测变量集。根据构造，BPA 按照信息论准则选择能使 X_i 的解释信息最大化的变量子集 X_{ps_i}。

这一选择也可以等价地理解为：选择 X_{ps_i}，使得在给定 X_{ps_i} 的条件下，X_i 与 X_{ps_i} 之外的任意其他变量之间的条件互信息可以忽略不计。形式化地，对于任意 $Z \subseteq X \setminus X_{ps_i}$，BPA 确保

$$I(X_i; Z \mid X_{ps_i}) \approx 0,$$

也就是说，将变量纳入 Z 并不会使 X_i 的解释能力得到显著提升，超出 X_{ps_i} 已捕捉的信息。

利用互信息的链式法则，我们可以将 X_i 与其余全部变量之间的互信息分解为

$$I(X_i; X \setminus \{X_i\}) = I(X_i; X_{ps_i}) + I(X_i; X \setminus X_{ps_i} \mid X_{ps_i}).$$

由于 BPA 选择 X_{ps_i}，旨在捕获 $X \setminus \{X_i\}$ 中关于 X_i 的几乎所有信息，因此分解式中的第二项必定近似为零。因此，

$$I(X_i; X \setminus X_{ps_i} \mid X_{ps_i}) \approx 0.$$

条件互信息消失（或近似消失）等价于条件独立（或近似条件独立）。因此，BPASGM 构造蕴含以下条件独立性：

$$X_i \perp\!\!\!\perp X \setminus X_{ps_i} \mid X_{ps_i},$$

由此确立了所述结果。

□

这种条件独立性是由 BPA 选择机制所诱导的结构性质，不依赖于统计一致性、因果充分性或恢复真实马尔可夫毯的假设。

D 数据生成过程

数据生成过程分为两个主要步骤。

前六个收益序列（R_1 至 R_6）由随机噪声、自回归动态和非线性变换共同生成。具体而言，每项资产 $i \in \{1, \ldots, 6\}$ 的真实均值（μ_i）和标准差（σ_i）分别从区间$[0.0001, 0.0009]$ 和 $[0.0003, 0.08]$上的均匀分布中独立抽样。

为捕捉时间依赖性，噪声生成函数定义如下：

$$\texttt{gen\_noise}(\mu, \sigma) = \text{AR}(1) \text{ process with } \phi = 0.2 \text{ and innovation standard deviation } \sigma + \mu.$$

前六个收益序列随后构造如下：

$$\begin{aligned}
R_1 &= \mathcal{N}(\mu_1, \sigma_1^2) + \texttt{gen_noise}(0, \sigma_1), \\
R_2 &= \mathcal{N}(\mu_2, \sigma_2^2) + \texttt{gen_noise}(0, \sigma_2), \\
R_3 &= -0.003 \cdot \exp(-|R_1|) + 0.5R_2 + \mathcal{N}(\mu_3, \sigma_3^2) + \texttt{gen_noise}(0, \sigma_3), \\
R_4 &= -0.9R_1 + \mathcal{N}(\mu_4, \sigma_4^2) + \texttt{gen_noise}(0, \sigma_4), \\
R_5 &= 0.6R_4 - 0.5R_2 + \mathcal{N}(\mu_5, \sigma_5^2) + \texttt{gen_noise}(0, \sigma_5), \\
R_6 &= \tanh(0.8R_3 - 0.5R_5) + \mathcal{N}(\mu_6, \sigma_6^2) + \texttt{gen_noise}(0, \sigma_6).
\end{aligned}$$

其余六个收益序列（R_7 至 R_{12}）是通过对前六个序列进行非线性和线性变换生成的，从而引入了更为复杂的依赖模式。对于这些资产，均值和标准差分别取自区间 $[0.0001, 0.001]$ 和 $[0.006, 0.016]$。

这些序列的定义如下：

$$
\begin{aligned}
R_7 &= -0.9R_5 - 0.6R_4 + \texttt{gen\_noise}(\mu_7, \sigma_7),\\
R_8 &= -0.8R_7 - 0.4R_3 + \texttt{gen\_noise}(\mu_8, \sigma_8),\\
R_9 &= 0.3R_3 + \texttt{gen\_noise}(\mu_9, \sigma_9),\\
R_{10} &= \log(1 + |R_1 + R_5|) + 0.2R_9 + \texttt{gen\_noise}(\mu_{10}, \sigma_{10}),\\
R_{11} &= 0.2R_8 + \max(R_2, R_1) + \texttt{gen\_noise}(\mu_{11}, \sigma_{11}),\\
R_{12} &= -0.6R_4 - 0.4R_6 + \texttt{gen\_noise}(\mu_{12}, \sigma_{12}).
\end{aligned}
$$

模拟数据集（$\boldsymbol{X}_{simulated}$）捕捉了资产收益率之间符合现实情况的相互依赖关系，包括非互惠效应、非线性交互以及信息内容重叠。这种复杂程度使其特别适合用于评估 BPASGM 算法在近似真实金融数据的条件下，恢复相关依赖关系并选取最优资产子集的能力。这种实验设计为评估该算法在具有复杂结构关系的环境中的特征选择性能提供了严谨的依据[Riso 等人， 2023]。

E 与 GLASSO 的性能比较

本节将比较所提出的 BPASGM 算法与由Friedman 等人 [2008]提出的图形套索（GLASSO）方法的性能。GLASSO 是一种惩罚似然方法，通过应用 ℓ_1 正则化项来诱导稀疏性，从而为连续变量估计稀疏精度矩阵（协方差矩阵的逆矩阵）。

令 $R \in \mathbb{R}^{T\times N}$ 表示经过中心化处理的资产收益率矩阵。GLASSO 优化问题表述为

$$
\hat{\Theta} = \arg\max_{\Theta \succ 0} \left\{ \log\det\Theta - \operatorname{tr}(S\Theta) - \lambda \sum_{i\neq j} |\theta_{ij}| \right\},
$$

其中 $S = \frac{1}{T}R^\top R$ 表示样本协方差矩阵， $\Theta = \Sigma^{-1}$ 表示精度矩阵，而 $\lambda \geq 0$ˆ 控制稀疏程度。 Θ 中的非零元素对应于资产对之间的条件依赖关系，反映在给定所有其他资产条件下的偏相关。 λ 值越高，依赖结构越稀疏；而值越低，则会产生更密集的互联关系。适当正则化水平的选择已在文献中得到广泛讨论， [例如，参见 Yin 和 Li，2011年、Friedman 等人，2008年、Pozzi 等人以及 2013, 2008]。在本研究中，我们不固定单一的最优参数，而是考察可能的正则化水平的完整范围，以基准比较 BPASGM 与 GLASSO 在不同稀疏程度下所隐含的投资组合的表现。

对于每个 λ 值，精度矩阵 $\Theta^{(\lambda)}$ 被计算后再二值化，以识别具有统计显著性的条件依赖关系。每个元素 ˆ$\theta_{ij}^{(\lambda)}$ 反映资产 i 与 j之间的条件关系；仅保留最相关的连接，邻接矩阵 $A^{(\lambda)}$ 定义如下：

$$
A_{ij}^{(\lambda)} = \begin{cases} 1, & \text{if } |\hat{\theta}_{ij}^{(\lambda)}| > \tau, \\ 0, & \text{otherwise,} \end{cases} \qquad A_{ii}^{(\lambda)} = 0,
$$

其中， $\tau > 0$ 决定了被视为具有经济意义的条件依赖关系的最小幅度。矩阵 $A^{(\lambda)}$ 定义了一个无向网络 $G^{(\lambda)} = (V, E^{(\lambda)})$，其节点对应资产，边表示具有统计显著性的条件关系。

在每个网络上，计算标准中心性度量，以量化各单项资产的结构重要性。节点 i 的度中心性为

$$
\deg(i) = \sum_{j=1}^{p} A_{ij},
$$

并反映其直接互联程度。介数中心性，

$$
\operatorname{betw}(i) = \sum_{s\neq i\neq t} \frac{\sigma_{st}(i)}{\sigma_{st}},
$$

衡量节点 i 位于连接其他节点的最短路径上的频率，其中 σ_{st} 表示节点 s 和 t之间的最短路径数量，而 $\sigma_{st}(i)$表示经过 i的路径。最后，接近中心性反映了某项资产与所有其他资产之间的平均测地线邻近度：

$$\text{clos}(i) = \frac{1}{\sum_{j=1}^{p} d(i,j)},$$

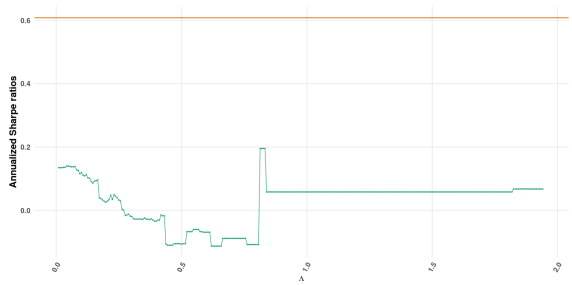
其中， $d(i、 j)$ 是资产 i 和 j之间的最短路径距离。

如果资产彼此直接连接，或在条件依赖网络中处于中心位置，则会将其保留用于投资组合构建。形式上，所选资产子集 S 表示为：

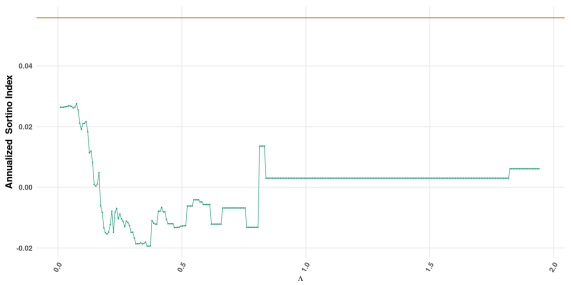
$$S = \{\, i \in V : \deg(i) > 0 \text{ or } \text{betw}(i) > \text{median}(\text{betw}) \text{ or } \text{clos}(i) > \text{median}(\text{clos})\}\,.$$

这种基于网络的筛选方法能够识别出在条件依赖系统中发挥关键结构作用的资产，从而在优化之前，以经济上可解释且数据驱动的方式缩减投资范围。通过结合稀疏图模型 [Friedman 等人， 2008] 与基于网络理论的筛选方法 [Mantegna，1999，Peralta 和 Zareei， 2016]，该方法将统计正则化与拓扑结构相结合，从而提取出最具信息价值的资产子集用于资产配置。

通过在 λ 的整个取值范围内反复执行这一筛选程序，并计算每个由此得到的最小方差投资组合所对应的夏普比率和索提诺比率，我们获得了网络稀疏性与表现之间的详细映射。图 27 显示，在所有稀疏性水平下，BPASGM 选出的投资组合始终优于 GLASSO 方法选出的投资组合。具体而言，基于 BPASGM 的投资组合（橙色线）的年化夏普比率（图 27a）和年化索提诺比率（图 27b）均高于基于 GLASSO 的投资组合（绿色线），证实了所提方法具有更强的稳健性和更高的效率。



(a) 不同 val- 下年化夏普比率的变化 ues λ 的



(b) 不同 val- 下年化索提诺比率的变化 ues λ 的

图27：不同 λ 值下的年化夏普比率（面板 (a)）和年化索提诺比率（面板 (b)）（绿色线）。通过 BPASGM 选出的最优投资组合的相应表现以橙色显示。