arXiv:2611.30998v1 [econ.EM] 23 Nov 2026

大型市场中的虚假信息：

熵、互信息与信息价值的随机矩阵定律

James K Rice*

2026年11月

摘要

熵、Kullback–Leibler散度和互信息已成为衡量一组金融变量能够揭示另一组金融变量多少信息的标准度量。在大型横截面中，它们几乎总是根据样本协方差矩阵计算，而这些矩阵中的噪声正是随机矩阵理论的研究对象。本文发展了高斯信息泛函的随机矩阵理论，并将其应用于市场。我证明，从 n 个观测值中估计由 p 和r 个变量组成的两个区块之间的插件互信息，具有一个精确的期望超额，即虚假信息；该超额是一个与真实依赖关系无关的双伽马望远镜和。在独立性条件下，其分布等同于贝塔变量之积的分布。在比例渐近区间内，它随 n线性增长，增长率为 $\kappa(c_1,\ c_2) = \frac{1}{2}[(1-c_1-c_2)\log(1-c_1-c_2) - (1-c_1)\log(1-c_1) - (1-c_2)\log(1-c_2)]$，且方差为一阶量。同一数量也是Jacobi系综的线性谱统计量，由此可得到Wachter定律对数矩的闭式。加入无信息坐标会使虚假信息严格增加。在CARA–高斯经济中，信号的确定性等价价值等于互信息除以风险厌恶，因此这些结果可一一转化为虚假信息价值。典型相关相变进而定义了最小可检测信息的前沿。对于相关性动态，我证明，矩阵Ornstein–Uhlenbeck市场的Dyson特征值过程足以提供其均值回归率的信息，并具有 N^2阶的信息率。蒙特卡洛实验验证了每一个精确公式和渐近公式。基于1990年至2026年49个美国行业投资组合的日收益率，我记录了市场模态、BBP 相变所蕴含的行业特征向量可靠性、噪声修饰的样本外成本，以及碳密集型行业与金融行业之间的一种信息联结；在进行精确偏差校正并以市场为条件后，该联结依然存在。部分代数和组合步骤已在Lean 4中完成机器验证。

关键词。 随机矩阵理论、Marchenko–Pastur定律、Jacobi系综、Wachter定律、互信息、Kullback–Leibler散度、信息价值、Dyson布朗运动、经济物理学、协方差清洗。

*埃塞克斯大学博士候选人；james.k.rice@essex.ac.uk。包括 Lean 形式化在内的复现材料可在 https://github.com/jameskrice7/RMT获取。

JEL 分类代码. C55, C58, D80, D83, G11, G14.**2020年**
数学主题分类. 60B20, 62H20, 94A17, 91G10.

引言

信息是金融市场所处理的商品，过去二十年来，对信息的定量研究已从文字描述发展到实际测量。熵和互信息出现在理性疏忽模型中（Sims，2003；Matějka 和 McKay，2015；Maćkowiak 等人，2023）、信息获取与投资组合选择理论中（Van Nieuwerburgh 和 Veldkamp，2010；Veldkamp，2011；Kacperczyk 等人，2016）、根据对投资者的价值对信号进行排序的研究中（Cabrales 等，2013；Kadan 和 Manela，2019），以及大量实证文献中；这些文献测量连通性、溢出效应，以及一个市场对另一个市场的信息含量（Billio 等人，2012；Diebold 和 Yılmaz，2014）。Kullback–Leibler散度（Kullback 和 Leibler，1951）和互信息（Shannon，1948；Cover 和 Thomas，2006）已成为报告此类数量的自然度量单位。当变量数量众多时，这些度量单位需要通过协方差矩阵计算，通常使用高斯公式 $I = \frac{1}{2}\log[\det \Sigma_{11} \det \Sigma_{22} / \det \Sigma]$或密切相关的对数行列式表达式。协方差矩阵本身也需要估计，而在现代金融应用中，变量数量与观测值数量相比很少较小。

随机矩阵理论描述了当维度与样本量处于同一数量级时，这类估计会发生什么。自 Laloux 等人（1999）和 Plerou 等人（1999）的研究以来，它一直是经济物理学的核心工具。这些论文表明，股票收益经验相关矩阵的特征值谱 bulk 与纯噪声的 Marchenko–Pastur 定律（Marčenko and Pastur，1967年）接近，少数特征值会脱离谱 bulk，其中最大的特征值刻画了市场模态。后续研究确立了这些离群特征值的统计理论（Johnstone，2001；Baik 等人，2005；Paul，2007年；Benaych-Georges 和 Nadakuditi，2011年），开发了从谱中剔除噪声的估计量（Ledoit 和 Wolf，2004，2012；Bun 等人，2016年、2017年；Ledoit 和 Wolf，2020年），并量化了优化投资组合的风险低估程度（Pafka 和 Kondor，2003；El Karoui，2010；Bodnar等人，2018）。Bai和Silverstein（2010）、Anderson 等（2010）以及 Potters 和 Bouchaud（2020）的专著汇集了相关数学成果。然而，金融领域对样本协方差矩阵信息泛函的随机矩阵理论关注甚少，这些信息泛函不同于矩阵的谱、特征向量或投资组合风险。这一缺口之所以重要，是因为经济理论恰恰会为这些信息泛函定价。

本文建立并应用了这一理论。其核心对象是两个变量块之间的插件高斯互信息，这两个变量块的维度分别为 p 和 r，基于 n 个观测值估计。我证明了其相对于真实值的期望超额（我称之为 虚假信息）仅是（p、r、n）的精确有限样本函数，与真实协方差矩阵无关。它是双伽马函数差的望远镜求和，因此，其正性以及关于每个变量块维度的严格单调性，都是双伽马函数单调性的直接结果。在变量块独立的情况下，插件信息为

Wilks统计量对数的负二分之一。因此，其精确分布就是一组相互独立的对数Beta变量之和，由此可得到精确的零假设方差和精确检验。在比例渐近区间 $p/n \to c_1$、 $r/n \to c_2$ 中，虚假信息以 $\kappa(c_1, c_2)$ 的速率随 n 线性增长，该速率已在摘要中给出，而其标准差仍保持在一阶量级。简单的参数计数会预测经典值 $pr/(2n)$，但这会低估虚假信息；当 $c_1 = c_2 = 0.4$ 时，低估倍数已经接近二。

同一数量还有第二种纯粹谱论的描述。插件信息等于 $-\frac{1}{2}\sum_k \log(1-\hat{\rho}_k^2)$，其中 $\hat{\rho}_k$ 是样本典型相关。在独立性条件下，它们的经验分布收敛于 Wachter（1980）定律，即 Jacobi 系综或 MANOVA系综的极限谱定律（Johnstone，2008；Collins，2005）。将这两种描述相匹配，可得 Wachter 定律对数矩的闭式计算；同一论证还确定了相应线性谱统计量的方差。我还运用典型相关的有限秩理论（Bao 等人，2019）来描述信息的相变。只有当 $t^4 > c_1 c_2/[(1-c_1)(1-c_2)]$时，具有总体典型相关$t$ 的单一依赖方向才会产生分离的样本典型相关。相应的互信息 $I_c(c_1, c_2) = -\frac{1}{2}\log(1 - t_c^2)$ 是一条前沿：低于该前沿时，信息虽存在，却无法通过谱方法定位。

这些结果可以直接转化为经济学结论。在价格接受型CARA–高斯投资组合问题中，我证明，拥有和不拥有联合高斯信号时的事前期望效用之比恰好为$\exp\{-I(v; s)\}$，其中 v 是收益向量， s 是信号。因此，信息的确定性等价价值为 $I(v; s)/\gamma$，即以风险容忍度为单位衡量的互信息。这一恒等式与Grossman和Stiglitz（1980）、Van Nieuwerburgh和Veldkamp（2010）以及Cabrales等（2013）的思想一致，但在这里，它对于任意维度都构成精确等式。这意味着，任何根据数据估计高维信号集信息价值的计量经济学家、投资者或数据供应商，都会将虚假信息继承为信息的虚假价值，并且具有完全相同的期望、零分布和比例极限。由于在加入无信息序列时，虚假价值会严格增加，卖方无需添加任何信息，就可以夸大信号集的测量价值。买方则可以利用双伽马校正精确消除这种膨胀。同样的精确理论还给出了依据插件高斯信念采取行动时，双向情况下预期Kullback–Leibler代价。这些成本就是信息论形式下的Stein损失和熵损失（James 和 Stein，1961；Dey 和 Srinivasan，1985），我还给出了它们的有限样本值和高维极限。

金融相关性并非静态不变，经济物理学强调了其类似状态的动力学特征（Münnix 等人，2012；Schmitt 等人，2013；Allez 和 Bouchaud，2012）。因此，我研究连续观测的矩阵 Ornstein–Uhlenbeck 过程中所包含的信息；该过程的特征值在约束项作用下进行 Dyson 布朗运动（ Dyson，1962 ； Chan，1992 ； Rogers 和 Shi，1993 ）。其 Stieltjes 变换满足一个受迫复 Burgers 方程，该方程的平稳解为半圆分布。根据 Girsanov 定理，两个均值回复状态之间的 Kullback–Leibler 散度率为 $N（N + 1）（b - a）^2/（8a）$。仅凭特征值过程也能得到相同的速率，因此 Dyson 过程对于识别状态已足够，特征向量运动不包含关于该状态的信息。根据 Chernoff–Stein 引理，区分两个状态所需的时间因而按N^{-2}的规律下降。这一充分性结果与潜在随机微分方程的变分推断相关

（Li等人，2020年；Rice，2026年），其中路径空间Kullback–Leibler散度是正则化项。对于旋转等变漂移族，基于特征值过程进行推断不会损失任何信息。

实证应用使用了从1990年1月2日至2026年9月30日，49个Fama–French行业投资组合的9,254个每日价值加权收益（Fama和French，1997年；French，2026年）。这些数据再现了经济物理学中熟悉的图景：Marchenko–Pastur谱 bulk、占主导地位的市场模态，以及若干行业模态。反演BBP离群值映射后，我便能为每个行业特征向量附上其与总体对应特征向量的预测重叠度，从而区分可靠的模态与仅略高于噪声水平的模态。在429个滚动窗口中，市场模态吸收的总方差占比介于0.297和0.820之间。最小方差回测呈现出随机矩阵理论所预测的噪声修饰惩罚，以及清理数据带来的好处；当估计窗口较短时，这种好处尤为显著。信息分析考察八个碳密集型行业的收益能在多大程度上揭示四个金融行业的情况；这一问题的动机来自气候金融文献对转型风险的研究（Battiston等人，2017年；Bolton和Kacperczyk，2021年；Giglio等人，2021年）。对于这一设计，虚假信息的精确值为0.0660纳特，零假设标准差为0.0165。在时间平移安慰剂检验中，插件信息的平均值为0.0653纳特，校正后的安慰剂检验平均值为 -0.0007；在名义显著性水平为5%的情况下，校正检验在6.1%的窗口中拒绝零假设。校正后，真实联系的平均值为0.828纳特，且与市场模态密切相关。对市场模态进行条件化后，联系的平均值降至0.170纳特。它在每个窗口中都显著，与吸收比率不相关，并在2008年10月达到峰值。

本文借鉴并回应了多个领域的文献。在随机矩阵理论方面，本文以高斯系综和Wishart系综的经典理论为基础（Mehta，2004；Forrester，2010；Dumitriu和Edelman，2002；Edelman和Rao，2005），也借鉴了使异常值理论具有一致性的局部定律和各向异性估计（Knowles和Yin，2017），以及Akemann等人（2011）、Livan等人（2018）和Couillet和Liao（2022）的手册与教材论述。后者与本文一样，强调随机矩阵修正在现代数据分析中具有一阶重要性。尖峰模型中的最优特征值收缩与阈值化（Gavish和Donoho，2014；Donoho等人，2018）、通过带状化和阈值化进行正则化（Bickel和Levina，2008）、基于因子的协方差估计（Fan等人，2013），以及采用非线性收缩的动态协方差模型（Engle，2002；Engle等人，2019），构成了估计工具箱；本文的信息专属修正应与这些方法进行比较。Ledoit 和 Wolf（2022）综述了这一工具箱，而其在投资组合中的应用可追溯至Ledoit和Wolf（2003）及Ledoit和Wolf（2017）。在经济物理学领域，Gopikrishnan等人（2001）、Plerou等人（2002）、Utsugi等人（2004）和Burda等人（2005）拓展了对金融相关性的随机矩阵分析，层级描述和网络描述则对其作了补充（Mantegna，1999；Tumminello等人，2010）。相关教材包括Mantegna和Stanley（2000）、Bouchaud和Potters（2003），以及Kwapień和Drożdż（2012）的综述。从Markowitz（1952）到Elton和Gruber（1973）、Michaud（1989）、Jagannathan和Ma（2003）、Pafka 和 Kondor（2002）及Kondor等人（2007）关于估计噪声所导致的投资组合后果的研究，正是最小方差分析的动机。在金融计量经济学中，因子数量的确定（Bai和Ng，2002；Onatski，2009，2010；Ahn和Horenstein，2013）本身就是一个植根于套利

定价理论（Ross，1976；Chamberlain和Rothschild，1983）的随机矩阵问题。大型面板数据和高频数据中的潜在因子估计也是如此（Lettau和Pelger，2020b，a；Pelger，2020；Aït-Sahalia和Xiu，2017）。随机矩阵渐近理论还解释了过参数化回归（Hastie等人，2022）令人意外的行为。在信息论领域，随机矩阵方法长期以来一直用于计算多天线信道的容量；这种容量是本文所研究的那类对数行列式泛函（Telatar，1999；Tulino和Verdú，2004）。已有文献充分记录了高维情形下估计互信息的困难，涉及非参数估计器和神经估计器（Paninski，2003；Kraskov等人，2004；Belghazi等人，2018；McAllester和Stratos，2020）。本文研究的高斯情形，正是能够精确计算这种困难的情形。在经济学中，对信息结构的比较可追溯至Blackwell（1953）。Stiglitz（2000）综述了信息在市场中的作用，而信息如何纳入价格，则是有效性文献和微观结构文献的研究主题（Fama，1970；Kyle，1985）。

本文有四项贡献。第一，基于Bartlett分解、Jacobi系综以及Marchenko–Pastur定律和Wachter定律，推导出关于虚假信息、Kullback–Leibler代价及其波动的一组精确定律和渐近定律，其中包括Wachter定律对数矩的新闭式。第二，通过精确的信息价值恒等式对这些定律进行经济学诠释，并给出由此推出的单调性、修正和前沿结果。第三，对Dyson 动力学进行信息论分析，证明特征值过程的充分性以及状态可检测性的二次标度。第四，开展一项经济物理学应用，将这些工具结合起来分析三十多年的数据。每个公式都通过模拟进行了检验，连接解析论证的组合步骤和代数步骤则在Lean 4定理证明器中得到验证（de Moura and Ullrich，2021年）。

本文的结构如下。下一节回顾全文所使用的随机矩阵工具，即预解式、Stieltjes变换和 S-变换、Marchenko–Pastur定律和Wachter定律、尖峰模型以及Dyson布朗运动。随后，我推导信息定律及其经济学诠释，并给出动态结果。数值部分验证理论，应用部分将其用于数据分析。附录包含证明、Lean形式化的说明以及补充表格。

随机矩阵预备知识

设 $X \in \mathbb{R}^{n\times N}$ 的各行相互独立，且 $x_t^\top \sim$服从 N $(0, \Sigma)$， $t = 1$、……、n，因此 $W = X^\top X$服从Wishart分布 $\mathrm{W}_N(n, \Sigma)$，而 $S = W/n$是均值已知为零时的样本协方差矩阵。当均值需要估计时，将 n替换为 $n-1$后，所有结论仍然成立，因为中心化交叉乘积矩阵为 $\mathrm{W}_N(n-1, \Sigma)$（Muirhead，1982 年，第 3 章）。长宽比为$q = N/n$。下文中， $\lambda_1 \geq \cdots \geq \lambda_N$表示 S的特征值，$\mu_S = N^{-1}\sum_i \delta_{\lambda_i}$表示其经验谱测度。本节汇集了后文将使用的随机矩阵理论工具。其目的是明确约定，并对新结果所依据的少数事实给出自洽的推导。全面的论述可参见 Bai和Silverstein（2010）、Anderson 等（2010）、Tao（2012）、Erdős 和 Yau（2017）；关于金融应用，则可参见 Bouchaud 和 Potters（2011）、Bun 等人（2017）以及 Potters 和 Bouchaud（2020）。

预解式、斯蒂尔切斯变换与 Marchenko–Pastur定律

对于 $z \in \mathbb{C} \setminus \mathbb{R}$，预解式为 $G_S(z) = (z - S)^{-1}$，而 μ_S 的 Stieltjes 变换为

$$g_S(z) = \frac{1}{N} \operatorname{tr} G_S(z) = \int \frac{\mu_S(\mathrm{d}x)}{z - x}.$$

函数 g_S 在谱之外解析，将上半平面映射到下半平面，并且在无穷远处呈现 $1/z$ 的渐近行为，其洛朗展开为 $g_S(z) = \sum_{k \geq 0} m_k z^{-k-1}$，其中涉及矩 $m_k = N^{-1} \operatorname{tr} S^k$。概率密度可通过反演公式 $\rho(x) = \pi^{-1} \lim_{\eta \downarrow 0} \operatorname{Im} g(x - i\eta)$，从其 Stieltjes 变换中恢复。对于每个 $z \in \mathbb{C} \setminus \mathbb{R}$，Stieltjes 变换的收敛等价于相应测度的弱收敛。因此，谱的渐近理论都可以用 g 来表述。

g_S 的自洽方程可由留一法论证推出。我在此记录这一论证，因为同样的代数推导还会出现在离群值和清理公式中。记 $S = n^{-1} \sum_t x_t x_t^\top$ 和 $S_{(t)} = S - n^{-1} x_t x_t^\top$，其预解式为 $G_{(t)} = (z - S_{(t)})^{-1}$。对恒等式 $zG_S - SG_S = I$ 两边取迹，得到 $zN g_S - n^{-1} \sum_t x_t^\top G_S x_t = N$。由谢尔曼–莫里森公式可得 $G_S = G_{(t)} + G_{(t)} x_t x_t^\top G_{(t)} / (n - x_t^\top G_{(t)} x_t)$，于是

$$\frac{1}{n} x_t^\top G_S x_t = \frac{n^{-1} x_t^\top G_{(t)} x_t}{1 - n^{-1} x_t^\top G_{(t)} x_t}.$$

由于 x_t 与 $G_{(t)}$相互独立，二次型 $n^{-1} x_t^\top G_{(t)} x_t$ 集中于$n^{-1}\operatorname{tr}(\Sigma G_{(t)})$附近，而 $G_{(t)}$ 与 G_S 相差一个秩一扰动，其归一化迹为 $O(N^{-1})$。当 $\Sigma = I$ 时，该二次型因此为 $q g_S + o(1)$，将迹恒等式除以 N 可得 $zg - g/(1 - qg) = 1$。写成多项式形式，

$$qz\, g(z)^2 - (z - 1 + q)\, g(z) + 1 = 0. \tag{1}$$

在无穷远处满足 $g(z) \sim 1/z$ 的根为 $g(z) = [z - 1 + q - \sqrt{(z - 1 + q)^2 - 4qz}\,]/(2qz)$。判别式在 $\lambda_\pm = (1 \pm \sqrt{q})^2$, 处为零，而反演公式给出 Marchenko– Pastur 密度

$$\rho_q(x) = \frac{\sqrt{(\lambda_+ - x)(x - \lambda_-)}}{2\pi q x}, \qquad x \in [\lambda_-, \lambda_+], \tag{2}$$

并且在原点处有质量为 $1 - 1/q$的原子，当 $q > 1$ (Marčenko and Pastur，1967年）。边缘也是极端特征值几乎必然的极限（Geman，1980年；BaiandYin，1993年）。图 1 展示了这一族分布。随着 q 增大，谱 bulk 从位于 1 处的点质量扩展为一个在 $q = 1$处触及原点的区间。这种扩展正是本文所讨论的所有维数代价背后的机制。

.....的矩 ρ_q 可由（1）直接推出，无需积分。令 h（x）$= \sum_k m_k x^k = z\, g$（z），其中 $x = 1/z$，则二次式变为 $h = 1 + xh$（$1 - q + qh$），这等价于以下递推关系

$$m_k = (1 - q)\, m_{k-1} + q \sum_{j=0}^{k-1} m_j\, m_{k-1-j}, \qquad m_0 = 1. \tag{3}$$

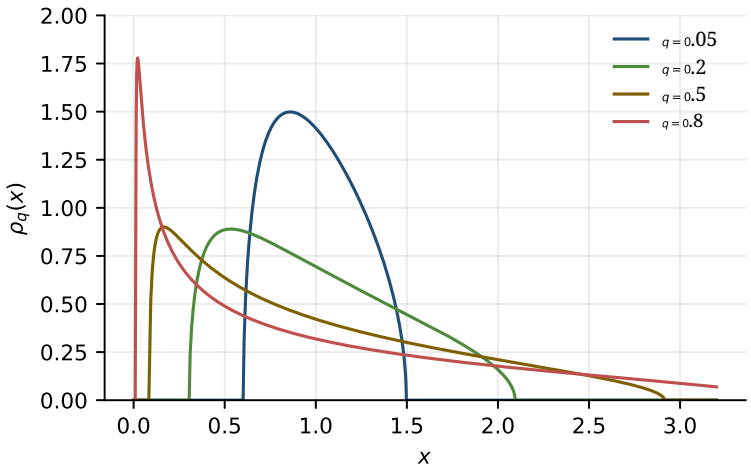

图1： 四种长宽比下的Marchenko–Pastur密度 ρ_q。边缘$(1 \pm \sqrt{q})^2$ 是二次式（1）判别式的零点。左边缘在 $q = 1$处到达原点，此时$\log\det S$ 发散。

它的解是 Narayana多项式 $m_k = \sum_{j=0}^{k-1} \frac{1}{k}\binom{k}{j}\binom{k}{j+1} q^j$，其系数按区块数对集合 $\{1, \dots, k\}$ 的非交叉划分进行计数（Stanley，1999；Nica和Speicher，2006）。前四个分别为 1、$1 + q$、$1 + 3q + q^2$ 和 $1 + 6q + 6q^2 + q^3$。当 $q = 1$ 时，Narayana 数之和等于 Catalan 数；它们也是 Wigner 半圆律的偶阶矩 $\rho_{\mathrm{sc}}(x) = (2\pi)^{-1}\sqrt{4 - x^2}$（1955，1958）。随本文提供的Lean文件通过精确整数算术验证：递推关系（3）生成至十阶的 Narayana 多项式，半圆分布递推生成至十二阶的 Catalan 数，且 Narayana 行之和等于 Catalan 数。在有限 n 下，精确的 Wishart 矩与极限值之间存在显式修正项。利用 $\mathbb{E}W^2 = n\,(n + N + 1)\,I$ 和 $\mathbb{E}\,\mathrm{tr}\,W^3 = nN\,(n^2 + N^2 + 3nN + 3n + 3N + 4)$ 对于 $\Sigma = I$（Muirhead，1982），

$$\frac{1}{N}\mathbb{E}\,\mathrm{tr}\,S^2 = 1 + q + \frac{1}{n}, \qquad \frac{1}{N}\mathbb{E}\,\mathrm{tr}\,S^3 = 1 + 3q + q^2 + \frac{3(1+q)}{n} + \frac{4}{n^2}.$$

这两个公式都在数值部分通过模拟进行了验证。

对于具有极限谱分布 H 的一般 Σ，采用同样的留一法论证，现在跟踪 $n^{-1}\mathrm{tr}(\Sigma G_{(t)})$，即可得到 Marchenko–Pastur–Silverstein方程

$$g(z) = \int \frac{H(\mathrm{d}\tau)}{z - \tau\, w(z)}, \qquad w(z) = 1 - q + qz\, g(z), \tag{4}$$

(Silverstein和Bai，1995）。函数 w 是一个从属函数：在 z 处取值的样本预解式，其平均表现类似于在 $z/w(z)$处取值并经过重新缩放的总体预解式。下文使用的三个事实都可直接由（4）得出。首先，当 $H = \delta_1$ 时，它简化为$g = 1/(z - w)$，即（1）。其次，位于 ℓ 谱 bulk 之外的总体特征值，会在 z 处产生一个样本离群特征值；该值是 $z = \ell\, w(z)$的被积函数变为奇异时对应的值。第三，样本特征向量在总体特征基上的期望投影，是一个在样本特征值处取值的 w 函数。这一事实构成了

预言式非线性收缩公式

$$\xi(\lambda)=\frac{\lambda}{|1-q+q\lambda\, g(\lambda-i0)|^2}=\frac{\lambda}{|w(\lambda-i0)|^2} \tag{5}$$

关于 Ledoit 和 Péché （2011年）以及 Bun 等人 （2016年）提出的旋转不变估计量。在该估计量中，每个样本特征值都被替换为 ξ（λ_i），其中 g 根据样本谱在实轴下方一个很小的距离处进行评估 η。

自由概率

Silverstein 方程是自由乘法卷积的一个实例。对于定义在 $[0,\infty$ 上的概率测度 μ）及其矩生成函数 ψ_μ（z）$=\sum_{k\geq 1} m_k z^k$，令 χ_μ 为 ψ_μ 在原点附近的逆函数，并定义 S-变换 S_μ（w）$=\chi_\mu$（w）（$1+w$）$/w$。Voiculescu（1991）证明，大型旋转不变矩阵的行为类似于自由随机变量，并且自由正元素乘积的谱的 S-变换等于各自 S-变换的乘积（Voiculescu 等人，1992；Mingo 和 Speicher，2017年）。由（3）可知，$\psi=h-1$ 满足 $\psi=x$（$1+\psi$）（$1+q\psi$），因此 χ（w）$=w/[$（$1+w$）（$1+qw$）$]$ 以及

$$S_{\mathrm{MP}}(w)=\frac{1}{1+qw}.$$

.....的谱 $S=\Sigma^{1/2}S_0\Sigma^{1/2}$，与 S_0 白Wishart矩阵的自由乘法卷积 $H\boxtimes\rho_q$，其 S-变换为 $S_H(w)/(1+qw)$。以下两点推论值得说明。由于 S_{MP} 是莫比乌斯映射，清理样本谱相当于进行自由反卷积，这正是（5）逐个特征值实现的操作。ρ_q的自由累积量都等于 q^{k-1}，这正是Narayana公式在自由概率中的表述。加法类比是 R-变换，$R_\mu=g_\mu^{-1}(w)-1/w$。半圆分布满足 $R(w)=\sigma^2 w$，这推动了后文研究的动力学。

典型相关与Jacobi系综

将 $x_t=(x_t^\top,1,x_t^\top,2)^\top$ 划分为维度分别为 $p\leq r$ 的区块，并相应地划分 S。样本典型相关平方 $\hat{}\rho_1^2\geq\cdots\geq\hat{}\rho_p^2$ 是 $S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}$ 的特征值。等价地，若 P_1 和 P_2 是将 $\mathbb{R}^n$ 投影到两个数据块列空间上的正交投影，则它们是 $P_1P_2P_1$ 的非零特征值。当这些区块相互独立时，其列空间均匀分布且相互独立。此时，$\hat{}\rho_k^2$ 构成 Jacobi（MANOVA）系综，其联合密度与 $\prod_k x_k^{(r-p-1)/2}(1-x_k)^{(n-r-p-1)/2}\prod_{j<k}|x_j-x_k|$ 成正比，定义在 $[0,1]^p$ 上（Muirhead，1982；Johnstone，2008）。用自由概率论的术语来说，P_1 和 P_2 是迹分别为 $c_1=p/n$ 和 $c_2=r/n$ 的渐近自由投影，而 $P_1P_2P_1$ 的极限定律是两个伯努利律的自由乘法卷积（Collins，2005）。对于 $c_1\leq c_2$ 和 $c_1+c_2<1$，其绝对连续部分是 Wachter（1980）的定律，

$$w_{c_1,c_2}(x)=\frac{\sqrt{(x-\omega_-)(\omega_+-x)}}{2\pi c_1\, x(1-x)},\qquad \omega_\pm=\left(\sqrt{c_1(1-c_2)}\pm\sqrt{c_2(1-c_1)}\right)^2. \tag{6}$$

该密度在 $[\omega_-, \omega_+]$ 上的总质量为 1，均值为 c_2。图 2 将其与独立高斯数据块的模拟典型相关进行了比较。当且仅当 $\omega_+ < 1$时，上边缘恰好满足 $c_1 + c_2 < 1$，从而使 $\log(1-x)$ 在支撑集上有界。这种有界性推动了下文的信息定律。

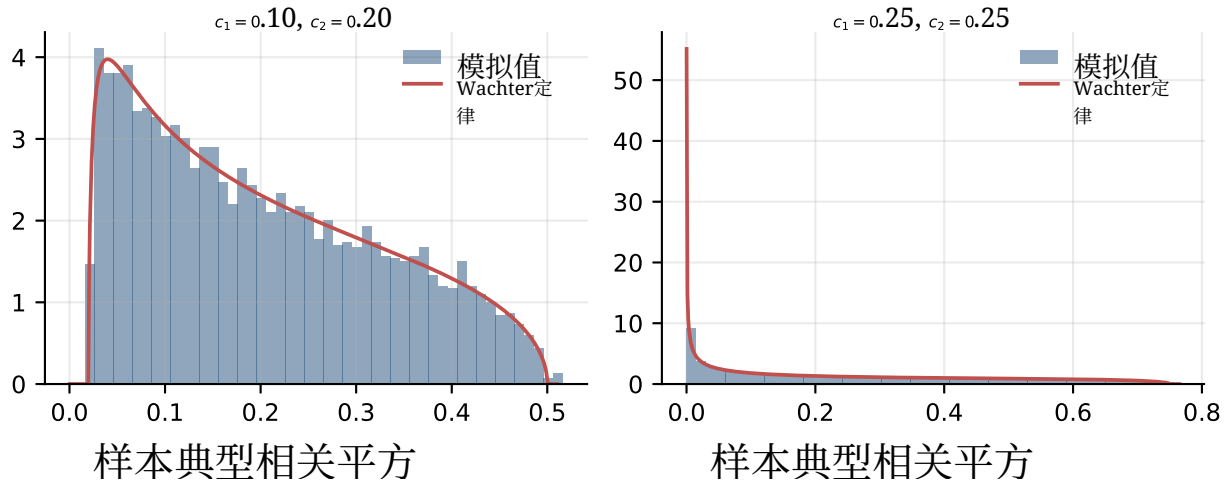

图 2： 独立高斯块的样本典型相关平方（30 次重复实验）与 Wachter定律（6）。左图：$n = 1000$、$p = 100$、$r = 200$。右图：$n = 600$、$p = r = 150$。上边缘 ω_+ 保持在 1 以下，因为 $c_1 + c_2 < 1$。

尖峰模型与相变

实际的协方差矩阵并非白噪声，而金融领域中重要的结构是低秩的：一个市场因子、少数几个行业，以及区块之间的几个依赖方向。尖峰协方差模型 $\Sigma = I + \sum_{k \le K} (\ell_k - 1)\, v_k v_k^\top$ 在固定 K 的情况下捕捉了这一点。将$H = (1 - K/N)\,\delta_1 + N^{-1}\sum_k \delta_{\ell_k}$代入（4），并结合离群值条件 $z = \ell w(z)$，利用 Marchenko–Pastur 从属函数可得到显式关系

$$\lambda(\ell) = \ell\left(1 + \frac{q}{\ell - 1}\right), \qquad \ell > 1 + \sqrt{q}. \tag{7}$$

推导很简短。由 $w = z/\ell$，主体谱关系 $(w - 1 + q)\ (z - w) = qz$ 变为 $(z/\ell - 1 + q)\ (1 - 1/\ell) = q$，整理后得到 (7)。只有当$\lambda(\ell) > \lambda_+$时，解才会离开谱 bulk。由于 $\lambda(1 + \sqrt{q}) = (1 + \sqrt{q})^2$，，阈值为 $\ell_c = 1 + \sqrt{q}$，这就是 Baik–Ben Arous–Péché（BBP）相变（Baik 等人，2005；Baik 和 Silverstein，2006）。超过阈值后，样本特征向量 u_k 与 v_k之间会形成非平凡夹角，

$$\langle u_k, v_k \rangle^2 \to \frac{1 - q/(\ell_k - 1)^2}{1 + q/(\ell_k - 1)}. \tag{8}$$

在阈值以下，重叠会消失（Paul，2007年；Benaych-Georges 和 Nadakuditi，2011年；Bloemendal 等人，2016年）。对（7）进行反演可得二次式 $\ell^2 - (\lambda + 1 - q)\,\ell + \lambda = 0$，其较大的根是观测到离群值 λ所隐含的总体尖峰，而（8）则预测了对应特征向量的可靠程度

。Lean 文件证明了阈值恒等式、重叠分子在阈值处消失，以及反演二次方程。边缘处最大特征值的涨落服从 Tracy–Widom定律（Tracy 和 Widom，1994，1996；Johnstone，2001）。这里不需要用到这些结果。

典型相关的类似转变则不太为人熟悉，但在下文的信息前沿中发挥着核心作用。假设这些区块只有一个非零的总体典型相关 t。Bao 等人(2019年) 的理论表明，当且仅当 $t > t_c$时，最大样本典型相关会从 Wachter谱 bulk 中分离出来；否则会黏附在 ω_+上。在我所采用的参数化方式下（数值部分中的模拟证实了这一点），阈值和分离极限分别为

$$t_c^4 = \frac{c_1 c_2}{(1-c_1)(1-c_2)}, \qquad \hat{\rho}_1^2 \to \gamma(t) = t^2\Big(1 - c_1 + \frac{c_1}{t^2}\Big)\Big(1 - c_2 + \frac{c_2}{t^2}\Big) \quad (t > t_c). \tag{9}$$

在 $t = t_c$ 时，$\gamma(t_c) = c_1(1-c_2) + c_2(1-c_1) + 2t_c^2(1-c_1)(1-c_2) = \omega_+$，因为 $t_c^2(1-c_1)(1-c_2) = \sqrt{c_1 c_2}$ $(1-c_1)(1-c_2)$。在 $t = 1$ 时，$\gamma = 1$。Lean文件在 t_c 的定义约束下证明了阈值恒等式。与（7）一样，映射 $t \mapsto \gamma(t)$ 在阈值以上是可逆的。关于 t^2 的一个二次式可以从观测到的离群值恢复出去偏典型相关。

对数谱统计量

本文的核心泛函是对数型的。Bartlett分解将 $W = LL^\top$ 写成 $\Sigma = I$，其中 L 为下三角矩阵，具有独立的元素，$L_{ii}^2 \sim \chi^2_{n-i+1}$ 和$L_{ij} \sim$服从 N(0, 1)，对应于 $i > j$（Muirhead，1982，第 3 章）。因此，log det $W = \sum_{i=1}^N \log L_{ii}^2$ 是独立的对数卡方变量之和。由于 $\mathbb{E}$ log $\chi_k^2 = \psi(k/2)+\log 2$，且 Var log $\chi_k^2 = \psi'(k/2)$，其中 ψ 为双伽马函数；又由于对于白 Wishart矩阵 W_0，det $S = \det\Sigma \det(W_0/n)$，因此对每个 Σ，都有以下精确恒等式

$$\begin{aligned} \ell_n(N) \equiv \mathbb{E}\log\det S - \log\det\Sigma &= \sum_{i=1}^{N}\Big[\psi\Big(\frac{n-i+1}{2}\Big) - \log\frac{n}{2}\Big], \\ \operatorname{Var}\log\det S &= \sum_{i=1}^{N}\psi'\Big(\frac{n-i+1}{2}\Big). \end{aligned} \tag{10}$$

利用 $\psi(x) = \log x - 1/(2x) + O(x^{-2})$ 和 $\psi'(x) = 1/x + O(x^{-2})$，通过黎曼和论证可得

$$\frac{1}{N}\ell_n(N) \to \varphi(q) \equiv \int \log x\, \rho_q(\mathrm{d}x) = -1 - \frac{1-q}{q}\log(1-q), \qquad \operatorname{Var}\log\det S \to -2\log(1-q), \tag{11}$$

因为 $N^{-1}\sum_{i\le N}\log(1-(i-1)/n) \to q^{-1}\int_0^q \log(1-u)\,\mathrm{d}u$。二阶项同样是显式的。Euler–Maclaurin展开表明，$\ell_n(N) = N\varphi(N/n) + \frac{1}{2}\log(1-N/n) + o(1)$，这正是 Bai 和 Silverstein（2004）线性谱统计量中心极限定理中的均值校正值，其在实数情形下的方差为 $-2\log(1-q)$。因此，对数行列式向下偏移了一个与 N 成正比的量，而其波动仍保持在一阶量级。该偏差就是对数函数相对于谱 bulk 的扩散进行积分所产生的 Jensen间隙

Marchenko–Pastur主体，而它是下一节每条信息定律的种子。公式（10）及其极限构成了以下高维熵估计量的基础：Ahmed 和 Gokhale （1989），Misra 等（2005）以及Cai 等（2015），还有以下文献中的校正值似然比检验：Bai 等（2009）和Jiang 和 Yang（2013年）。

信息的精确谱定律

对于协方差为 Σ 的中心化高斯向量，其微分熵为 $h = \frac{1}{2}\log\det(2\pi e\Sigma)$。划分 $x = (x_1,\ x_2)$ 所对应的两个区块的维度为 $(p,\ r)$，它们之间的互信息为

$$I(\Sigma) = \frac{1}{2}\log\frac{\det\Sigma_{11}\det\Sigma_{22}}{\det\Sigma} = -\frac{1}{2}\sum_{k=1}^{p}\log(1-\rho_k^2),$$

其中 ρ_k 表示总体典型相关（Cover 和 Thomas，2006）。两个中心化高斯分布之间的 Kullback–Leibler 散度为 $\mathrm{KL}(\mathcal{N}_A\|\mathcal{N}_B) = \frac{1}{2}[\mathrm{tr}(B^{-1}A) - N - \log\det(B^{-1}A)]$。这三者都是协方差矩阵的光滑函数，通常通过将 S 代入 Σ来进行估计。现在我推导这些插件估计量的精确定律和渐近定律。正文中未给出的证明见附录 A。

伪互信息

定义双伽马部分和 $\Psi_n(k) = \sum_{i=1}^{k}\psi((n-i+1)/2)$ 以及虚假信息

$$B_n(p,r) = \tfrac{1}{2}\Big[\Psi_n(p) + \Psi_n(r) - \Psi_n(p+r)\Big] = \tfrac{1}{2}\sum_{i=1}^{r}\Big[\psi\Big(\frac{n-i+1}{2}\Big) - \psi\Big(\frac{n-p-i+1}{2}\Big)\Big]. \quad (12)$$

第二个表达式由第一个表达式通过裂项相消得到，并且它关于（$p,\ r$）是对称的，尽管表面上并不像。

命题 （精确的虚假信息）。 设 $n \geq p+r$，并令 $\hat{}I = I$（S） 表示根据 $S = W/n$ 并使用 $W \sim W_{p+r}$（n, Σ）计算得到的插件互信息。对于任意正定的 Σ,

$$\mathbb{E}\hat{I} = I(\Sigma) + B_n(p,r).$$

当 $p, r \geq 1$ 时，超额量 B_n（p, r） 严格为正，且与 Σ 无关。偏差校正估计量 $\tilde{I} = \hat{}I - B_n$（$p, r$） 对于 I（Σ） 是完全无偏的。

证明只需两行。对角区块满足 $nS_{11} \sim W_p(n, \Sigma_{11})$ 和 $nS_{22} \sim W_r(n, \Sigma_{22})$，且自由度相同（Muirhead，1982，第 3 章）。将 (10) 应用于 $\hat{}I$ 中的三个行列式，得到 $\mathbb{E}\hat{}I - I = \frac{1}{2}[\ell_n(p) + \ell_n(r) - \ell_n(p+r)]$,，而 $\log(n/2)$ 项相互抵消，因为 $p + r - (p+r) = 0$。根据 (12) 的第二种形式，以及 ψ 在 $(0,\infty)$ 上的严格单调性，可知结果为正。Lean文件证明了任意严格反单调序列的裂项恒等式和正性。分析性内容仅限于 (10) 和双伽马函数的单调性。 B_n 与 Σ 的独立性使这一修正成为可能。分析者无需了解.....之间的依赖关系

便能消除维数给其估计带来的偏差。此论证不要求 p 和 r 很小。用 $n-1$ 替代 n 进行去均值处理后，同一命题仍然成立；将两个区块都对 k 个公共回归变量进行回归，并用 $n-k$ 替代 n 后，命题也仍然成立。第二种情形是应用中用于部分互信息的版本。

当 $\Sigma_{12}=0$ 完整零分布可用时。在这种情况下， $\Lambda=\det S/(\det S_{11}\det S_{22})$ $=\prod_k 1-\rho_k^2$ $I=-\frac{1}{2}$ Λ _ （$\hat{}$）是 Wilks统计量，而 $\hat{}\log$。

命题 (零分布)。 如果 $\Sigma_{12}=0$ 且 $n\geq p+r$，那么 $\hat{}I \overset{d}{=} -\frac{1}{2}\sum_{i=1}^{p}\log\beta_i$，其中含有独立的 $\beta_i\sim$贝塔分布$\Big(\frac{n-r-i+1}{2},\frac{r}{2}\Big)$。因此

$$\mathbb{E}_0\hat{I}=B_n(p,r),\qquad \operatorname{Var}_0\hat{I}=V_n(p,r)\equiv\frac{1}{4}\sum_{i=1}^{p}\Big[\psi'\Big(\frac{n-r-i+1}{2}\Big)-\psi'\Big(\frac{n-i+1}{2}\Big)\Big].$$

Wilks统计量的乘积表示是经典结果（Anderson，2003，第8章），其矩可由 $\mathbb{E}\log\beta=\psi(a)-\psi(a+b)$ 以及对 $\beta\sim\mathrm{Beta}(a,b)$ 有 $\operatorname{Var}\log\beta=\psi'(a)-\psi'(a+b)$ 得出。$\mathbb{E}_0\hat{}I$ 的两个表达式因式（12）中的望远镜消去而一致，这提供了内部一致性检验。零分布给出了信息独立性的精确单侧检验；当 $(\hat{}I-B_n)/V_n^{1/2}$ 超过正态分位数或精确分位数时，即拒绝零假设。它还为零假设下的校正后估计量提供了精确标准误。

比例渐近与 Wachter定律

当两个区块与样本一起增长时，有限样本公式具有简单的极限。令

$$\kappa(c_1,c_2)=\tfrac{1}{2}\Big[(1-c_1-c_2)\log(1-c_1-c_2)-(1-c_1)\log(1-c_1)-(1-c_2)\log(1-c_2)\Big],\tag{13}$$

$$\nu(c_1,c_2)=\tfrac{1}{2}\log\frac{(1-c_1)(1-c_2)}{1-c_1-c_2}.\tag{14}$$

命题（比例极限）。令 $p/n\to c_1$ 和 $r/n\to c_2$ 满足 c_1、$c_2>0$ 以及$c_1+c_2<1$。则 $B_n(p,r)/n\to\kappa(c_1,c_2)$，且 $V_n(p,r)\to\nu(c_1,c_2)$。在独立性成立的条件下，$(\hat{}I-B_n(p,r))/V_n(p,r)^{1/2}$ 依分布收敛于标准正态分布。对于每个 Σ，$\hat{}I/n-I(\Sigma)/n-\kappa(c_1,c_2)\to 0$ 依概率。

T展开式 $(1-c)\log(1-c)=-c+\sum_{k\geq 2}c^k/[k(k-1)]$ 给出级数

$$\kappa(c_1,c_2)=\frac{1}{2}\sum_{k\geq 2}\frac{(c_1+c_2)^k-c_1^k-c_2^k}{k(k-1)}=\frac{c_1c_2}{2}\Big(1+\frac{c_1+c_2}{2}+\cdots\Big),$$
$$\nu(c_1,c_2)=\frac{1}{2}\sum_{k\geq 2}\frac{(c_1+c_2)^k-c_1^k-c_2^k}{k}=\frac{c_1c_2}{2}\Big(1+c_1+c_2+\cdots\Big).$$

所有系数均为正，因此 κ和 ν 均为正，并且分别随各自参数增大而增大。首项还原了经典的固定维数近似 $2n\hat{}I\approx\chi^2_{pr}$，其均值为

$pr/(2n)$，方差为 $pr/(2n^2)$。高阶项即为随机矩阵修正。它们并不小。当 $c_1 = c_2 = 0.4$ 时，比值 $n\kappa/(pr/2n)$ 等于 1。82，因此经典公式遗漏了近一半的虚假信息。图 3 展示了 κ 在可行区域上的分布。等高线在原点附近呈双曲线状，此时 $\kappa \approx c_1 c_2/2$；当 $c_1 + c_2 \to 1$ 时，曲线会急剧变陡，此时 $(1 - c_1 - c_2)\log(1 - c_1 - c_2)$ 具有无穷导数。

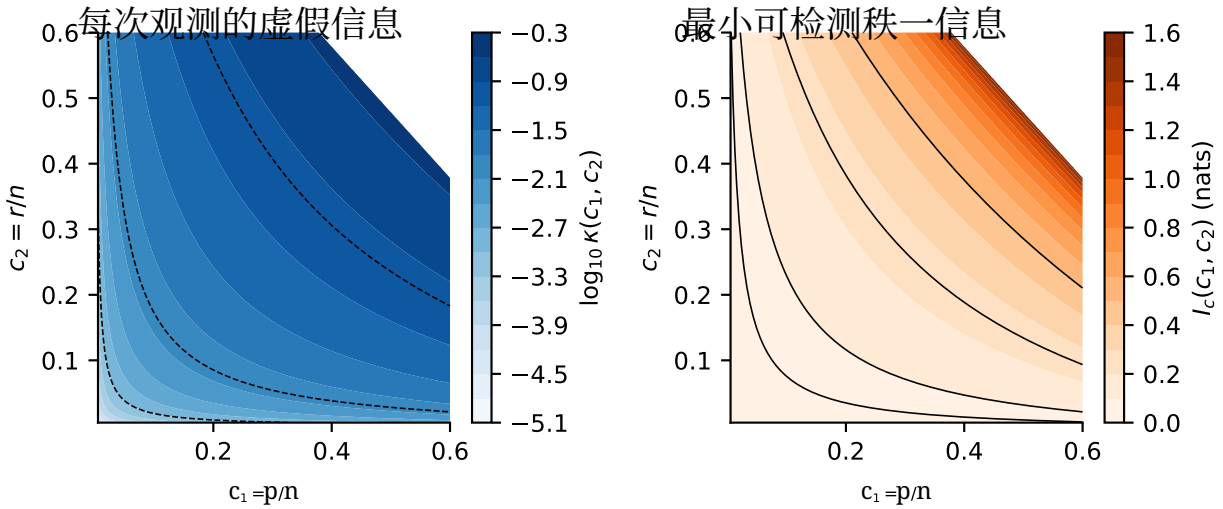

图3：左：由式 (13) 得到的每次观测虚假信息，$\kappa(c_1, c_2)$，采用对数颜色刻度，等高线位于 10^{-3}、10^{-2} 和 10^{-1}。右：由式 (16) 得到的最小可检测秩一信息 $I_c(c_1, c_2)$，等高线位于 0.05、0.1、0.25 和 0.5 纳特。两者均定义在 $c_1 + c_2 < 1$ 上。

$\hat{}$插件信息也是关于典型相关的线性谱统计量，$I = -\frac{p}{2}\int \log(1-x)\,\hat{}F_p(\mathrm{d}x)$，其中 $\hat{}F_p$ 是$\hat{}\rho_1^2$、……、$\hat{}\rho_p^2$的经验分布。在独立性成立时，$\hat{}F_p$ 弱收敛于 Wachter 定律（6），其支撑集始终位于 $[0, \omega_+]$ 内，且 $\omega_+ < 1$。结合前一个命题，可得出 Wachter 定律的一个对数矩，而此前似乎尚未有人记录过它。

定理（Wachter定律的对数矩）。对于 $0 < c_1 \le c_2$ 满足 $c_1 + c_2 < 1$，

$$\int_{\omega_-}^{\omega_+} \log(1-x)\, w_{c_1,c_2}(x)\,\mathrm{d}x = -\frac{2}{c_1}\,\kappa(c_1, c_2)$$
$$= \frac{1}{c_1}\Big[(1-c_1)\log(1-c_1) + (1-c_2)\log(1-c_2) - (1-c_1-c_2)\log(1-c_1-c_2)\Big].$$

附录 A 中的证明通过两种方式计算 $\hat{}I/n$ 的相同几乎必然极限：一种借助 Bartlett 分解，另一种借助 Wachter定律。它只需要极端典型相关收敛到支撑集的边缘。数值求积证实，在 10^{-13} 的网格上，(c_1, c_2)内该恒等式成立，而图 2 中的比较则展示了经验谱很快就会趋近 Wachter定律。对方差 V_n应用相同的推理，可得 $\nu(c_1, c_2)$ 是线性统计量 $\sum_k \log(1 - \hat{}\rho_k^2)/2$的极限方差，这与 Jiang 和 Yang Jiang 和 Yang (2013年) 以及Yang 和 Pan (2015) 所采用的 Jacobi 系综中心极限定理相符。

无信息坐标

一个重要的实际后果是，虚假信息可以被人为制造出来。假设某位分析者向第二个数据块添加 m 个与其他所有内容都独立的坐标，那么真实的互信息保持不变。

命题 （无信息坐标下的单调性）。 对于 $m \geq 1$， $n \geq p + r + m$ 以及任意 Σ，

$$\begin{aligned}\mathbb{E}\hat{I}_{p,r+m} - \mathbb{E}\hat{I}_{p,r} &= B_n(p, r+m) - B_n(p, r) \\ &= \frac{1}{2}\sum_{i=1}^{m}\Big[\psi\Big(\frac{n-r-i+1}{2}\Big) - \psi\Big(\frac{n-p-r-i+1}{2}\Big)\Big] > 0.\end{aligned}$$

在 $m/n \to c_3$ 的比例渐近区间内，每次观测的增量趋于 $\kappa(c_1, c_2 + c_3) - \kappa(c_1, c_2)$。

该恒等式再次是一个望远镜求和：对于任意数列，其证明见 Lean文件；对于严格反单调数列，还可证明其严格为正。下一节将阐述其经济含义。这里需要注意的是，双伽马校正估计量按构造在m 上是平坦的，而正则化代入估计量并非如此。正如 Ledoit 和 Wolf（2004）所做的那样，将协方差矩阵向单位矩阵的倍数收缩，或应用旋转不变估计量（5），都会去除虚假信息。同时也会去除真实信息；数值部分表明，随着 m 的增长，他们对 I 的估计会趋近于零。协方差清洗针对 Σ上的二次损失进行了优化，而信息并非二次泛函。

代入信念的信息成本

采用代入分布 $\mathcal{N}(0, S)$ 而非 $\mathcal{N}(0, \Sigma)$ 的高斯决策者会产生两种自然的信息成本。正向散度 $\mathrm{KL}(\mathcal{N}_\Sigma \| \mathcal{N}_S)$ 是指每个未来观测使用估计密度进行预测时的期望额外对数损失，或以纳特计的码长。反向散度 $\mathrm{KL}(\mathcal{N}_S \| \mathcal{N}_\Sigma)$ 是 S (James 和 Stein，1961) Stein损失的一半。记 $\ell_n(N)$ 为精确的对数行列式偏差（10）。

命题 （精确信息成本）。 对于 $n \geq N + 2$ 以及任意 Σ，

$$\mathbb{E}\,\mathrm{KL}(\mathcal{N}_\Sigma \| \mathcal{N}_S) = \frac{1}{2}\Big[\frac{N(N+1)}{n-N-1} + \ell_n(N)\Big], \qquad \mathbb{E}\,\mathrm{KL}(\mathcal{N}_S \| \mathcal{N}_\Sigma) = -\frac{1}{2}\ell_n(N).$$

当 $N/n \to q \in (0,\ 1)$时，每个维度的成本分别收敛于 $\frac{1}{2}[q/(1-q) + \varphi(q)]$和 $-\frac{1}{2}\varphi(q)$，其中 φ来自 (11)。

正向公式使用 $\mathbb{E}W^{-1} = \Sigma^{-1}/(n-N-1)$ (Muirhead，1982，第3章）以及代数恒等式 $Nn/(n-N-1) - N = N(N+1)/(n-N-1)$，该恒等式已在 Lean 中得到证明。两种成本的每维展开形式均为 $q/4 + O(q^2)$，因此当 q较小时，总成本接近参数数量 $N(N+1)/2$ 除以 $2n$，即熟悉的赤池惩罚。对于更大的 q

正向成本按 $(1-q)^{-1}$ 的形式爆炸增长，因为预测密度继承了最小的样本特征值，而这些特征值在左侧 Marchenko–Pastur边界处趋近于零。反向成本在 $q \to 1$ 时保持有界，其极限为每维$_{1\ 2}$ 。对于后文使用的实证设计，其中包含 $N = 49$ 个行业和 $n = 249$ 个去均值日观测值，精确的正向成本为 3.515 纳特，反向成本为 2.641 纳特。在六十日窗口期（$n = 59$）下，它们分别升至 120.05 和 16.07 纳特。按每个未来日期表示时，与金融领域试图测量的大多数信号的每日信息量相比，正向成本要高得多。正向成本是最小方差投资组合样本外惩罚的信息论对应物。对于 $w = S^{-1}\mathbf{1}/\mathbf{1}^\top S^{-1}\mathbf{1}$ ，样本内方差和样本外方差相对于真实最小方差的精确比值，其期望为

$$\mathbb{E}\frac{R_{\text{in}}}{R} = \frac{n-N+1}{n}, \qquad \mathbb{E}\frac{R_{\text{out}}}{R} = \frac{n-1}{n-N}, \tag{15}$$

其乘积趋于 $(1-q)^{-2}$ (Kan and Zhou，2007年；El Karoui，2010；Bodnar等人，2018）。第一个恒等式由 $\mathbf{1}^\top\Sigma^{-1}\mathbf{1}/\mathbf{1}^\top W^{-1}\mathbf{1} \sim \chi^2_{n-N+1}$ (Muirhead，1982，第3章）推出，且以下数值结果证实了这两个恒等式。

可检测信息的前沿

比例极限描述了谱 bulk。尖峰理论（9）描述了我们能够从单个依赖方向中学到什么。假设跨区块依赖的秩为一，典型相关为 t，因此 $I(\Sigma) = -\frac{1}{2}\log(1-t^2)$。只有当 t 满足 $t > t_c$ 时，最大样本典型相关才包含关于它的信息。定义信息前沿

$$I_c(c_1, c_2) = -\frac{1}{2}\log\bigl(1-t_c^2\bigr) = -\frac{1}{2}\log\Bigl(1-\sqrt{\frac{c_1 c_2}{(1-c_1)(1-c_2)}}\Bigr). \tag{16}$$

低于 I_c 的秩一信息是真实存在的，因为它会改变联合分布并具有正的经济价值，但它在谱中不可见。最大典型相关收敛到 ω_+ ，无论 t 取何值。聚合统计量 I 仍会有所响应，因为其均值偏移 $I(\Sigma)$，而其零假设标准差为 $\nu^{1/2}$，但它本身无法定位这些信息。定位信息是首个典型变量对的任务，而该变量对在阈值以下会失效。图3右面板显示，当两个区块相对于样本都较小时， I_c 可以忽略；当任一区块较大时，它就会变得显著。例如，在 (c_1, c_2) 取 $= (0.1, 0.2)$ 时，信息为 $I_c = 0.091$ 纳特；当区块大小相当且 $c_1 + c_2$ 接近 0.8 时，I_c 会超过半纳特。信息前沿为这样一种直觉赋予了精确含义：在大型横截面中，有些信息确实存在，却无法据此采取行动。

信息的价值及其虚假成分

信息定律通过一个简单而精确的恒等式获得经济含义。考虑一位具有常数绝对风险厌恶程度 $\gamma > 0$ 、初始财富为 w_0 的价格接受者投资者，他可以按价格 $p \in \mathbb{R}^N$ 交易 N 种风险资产，以及一种零净收益的无风险资产。终值

收益为 $v \sim \mathcal{N}(\mu,\ \Sigma_v)$。交易之前，投资者可以观测到信号 $s \in \mathbb{R}^K$，该信号与 v 联合高斯。令 $\Sigma_{v|s}$ 表示给定 s 时 v 的条件协方差，它是非随机的；并回忆 $I(v; s) = \frac{1}{2}\log \det(\Sigma_v \Sigma_{v|s}^{-1})$。头寸为 x 时的终值财富为$w_0 + x^\top(v - p)$。

定理 (信息的精确价值)。 设 U_s 和 $U_\varnothing$ 分别为在$u(w) = -e^{-\gamma w}$下，采用有信号与无信号时最优策略所产生的事前期望效用。那么

$$\frac{U_s}{U_\varnothing} = \exp\{-I(v; s)\},$$

因此，信号的确定性等价价值——也就是投资者在观察信号之前愿意为其支付的最高金额——恰好为 $I(v; s)/\gamma$。该价值与 μ、 p 和 w_0无关，并且在对 s进行可逆线性变换后保持不变。

给定 s，最优需求为 $x^* = (\gamma\Sigma_{v|s})^{-1}(m - p)$，其中 $m = \mathbb{E}[v \mid s]$；条件期望效用为 $-\exp\{-\gamma w_0 - \frac{1}{2}(m - p)^\top \Sigma_{v|s}^{-1}(m - p)\}$。事前，$m - p \sim \mathcal{N}(\mu - p, \Sigma_v - \Sigma_{v|s})$。对于 $y \sim \mathcal{N}(c, B)$ 和半正定的 A，高斯积分 $\mathbb{E}\exp(-\frac{1}{2}y^\top Ay) = \det(I + BA)^{-1/2}\exp\{-\frac{1}{2}c^\top(A^{-1} + B)^{-1}c\}$ 成立。令 $A = \Sigma_{v|s}^{-1}$和 $B = \Sigma_v - \Sigma_{v|s}$，则有 $I + BA = \Sigma_v\Sigma_{v|s}^{-1}$ 和 $A^{-1} + B = \Sigma_v$。因此，指数因子恰好等于无信息时的期望效用，而行列式因子就是 $e^{-I(v;s)}$。两个矩阵恒等式的标量形式已在 Lean 中得到证明，附录 A 给出了详细内容。

该恒等式表明，在信息下投资组合选择的典型模型中，互信息不仅仅是价值的一种类比。它就是价值，只不过以风险容忍度为单位计量。它明确了基于熵的度量如何为投资者排序信号（Cabrales 等，2013），并且构成了信息选择模型中对数行列式目标的基础（Van Nieuwerburgh and Veldkamp，2010；Veldkamp，2011；Kacperczyk 等人，2016）。它还解释了为什么每纳特的线性成本——如理性疏忽模型中的设定（Sims，2003；Maćkowiak 等人，2023）——会在 CARA–高斯情形下产生开关式获取决策：当且仅当 $1/\gamma > \lambda$ 时，以每纳特 λ 的价格购买信号才是值得的。另一个类比来自 Kelly（1956）：他表明，拥有辅助信息的对数效用赌徒能够以恰好等于互信息的幅度更快地实现财富增长。该定理是 CARA 对应形式，只是以确定性等价取代了增长率。该恒等式属于局部均衡范畴。在含噪理性预期均衡中，价格本身携带信息（Grossman and Stiglitz，1980；Hellwig，1980；Verrecchia，1982），相关量则变为条件互信息 $I(v; s \mid p)$。该定理同样逐字适用于这一量，因为以联合高斯价格为条件只会将 Σ_v 替换为 $\Sigma_{v|p}$。

虚假信息价值

在实践中，为信号估值需要估计 $I(v; s)$。拥有 $(v,\ s)$ 的 n 联合观测值的投资者、计量经济学家或数据供应商，其维度分别为 $v = N$ 和 $s = K$，若使用代入法插件估计量，则会报告 $\hat{}V = \hat{}I/\gamma$。上一节的所有结果均适用。期望

报告值相对于真实值的超额量为

$$\mathbb{E}\hat{V} - V = \frac{B_n(N,K)}{\gamma} > 0,$$

确切地说，校正值 $\tilde{V} = (\hat{} I - B_n(N,K))/\gamma$是无偏的。在 ^信号毫无价值的零假设下， γV服从零命题的对数贝塔分布，因此无论样本量多大，都可以检验正价值。在比例渐近区间内，虚假价值按$n\kappa(N/n, K/n)/\gamma$的速度增长。因此，它并非会随着数据增多而消失的小样本困扰。如果资产和信号的数量随样本量一同增长，那么它就是一个永久性的差距，其每次观测的规模由 κ 决定，并显示于图 3中。在风险容忍度为一百万美元的情况下，应用中的实证设计包含八项资产、四个信号和一年每日数据，其虚假价值为 $0.0660 百万美元，即每次决策约 66,000 美元。

单调性命题对信息被出售的市场具有更明确的含义。假设一位供应商提供 K 个信号，买方使用共同样本中的插件价值对其进行评估。加入 m 个与收益及原始信号相互独立的序列，不会改变真实值，却会使测得值的期望增加

$$\begin{aligned}\Delta_m &= \frac{B_n(N,K+m) - B_n(N,K)}{\gamma} \\ &= \frac{1}{2\gamma}\sum_{i=1}^{m}\left[\psi\left(\frac{n-K-i+1}{2}\right) - \psi\left(\frac{n-N-K-i+1}{2}\right)\right] > 0.\end{aligned}$$

因此，未对维度进行校正的买方面临着Akerlof （1970年）所说的柠檬市场问题：填充数据在私人层面有利可图，且平均而言无法被察觉。校正能在期望意义上恢复正确的激励机制，却无法恢复精确度。零假设方差 V_n（N， $K+m$）随 m严格递增，因为 ψ' 递减；因此，即使在无偏评估的情况下，经过填充的信号集合仍然更难评估。这一机制为数据市场和金融数据技术相关研究中反复出现的一种担忧提供了形式化解释（Farboodi 和 Veldkamp，2020年）：随着可用序列数量增加，数据的测得信息含量会机械性地上升，但其经济相关内容未必随之增加。这也精确说明了为何无信息或受污染的信息源会削弱价格发现。分析者进行校正后，这些信息源平均而言不会造成误导，但会消耗自由度，并抬高未经校正的评估结果。

无法定位的信息

前沿（16）又增添了一个层面。假设收益与信号之间的依赖关系集中在一个方向上，因此真实值为 $-\frac{1}{2\gamma}\log(1-t^2)$。知道这一方向的投资者可以据此交易并获取该值。必须从 n 个关于 N 收益和 K 个信号的观测值中学习这一方向的统计学家，则会面临典型相关相变。当 $t < t_c$时，无论 ω_+ 取何值，最大样本典型相关都会收敛至谱带边缘 t，因此无法揭示有关 t的任何信息。类比 BBP 相变中关于特征向量的部分，估计出的典型方向也被认为渐近地不含有关真实方向的信息。下文将通过模拟验证特征值这一结论

，并将特征向量这一结论作为由类比支持的猜想留待讨论。因此，值 $V_c = I_c(N/n,\ K/n)/\gamma$ 是给定维度下，仍无法通过数据学习获得的最大秩一值。低于这一水平的优势可以持续存在，而不被统计套利消除。这为 Martin 和 Nagel(2022年)关于高维环境中进行学习的投资者面临估计摩擦，导致收益在事后看来似乎可预测的观察，提供了信息论层面的对应解释。这也与 Kelly 等人(2024) 以及 Kozak 等人(2020年)的证据相呼应：资产定价中相关的信号数量众多，且各自都很弱。在下文采用的实证设计中，该前沿很小，$I_c = 0.0118$ 纳特，且 $t_c = 0.153$，因为八和四相较于一整年的日度数据都很小。而在涵盖数百个特征的个股完整截面尺度上，它就成为一阶效应。

谱动力学与信息率

金融收益的相关矩阵会发生波动，经济物理学的大部分研究都关注其谱的动力学。例如，市场模态的运动、行业特征向量的重新排列，以及相关性状态的交替变化（Allez 和 Bouchaud，2012；Münnix 等人，2012；Schmitt 等人，2013年；Reigneron 等人，2011年）。具有正确对称性的最简单模型是对称矩阵 Ornstein–Uhlenbeck 过程，

$$\mathrm{d}H_t = -a\,H_t\,\mathrm{d}t + \frac{1}{\sqrt{N}}\,\mathrm{d}B_t, \tag{17}$$

其中，B 是对称矩阵布朗运动，其非对角元素 B_{ij}、$i<j$ 是相互独立的标准布朗运动，而对角元素是 $\sqrt{2}$ 倍的相互独立标准布朗运动。我将 H_t 解释为一个 $N\times N$ 相关矩阵或协方差矩阵相对于其长期水平的偏差，其中 $a>0$ 是相关性冲击消散的速度。Bru（1991）和 Gouriéroux 等（2009）研究了这类矩阵值扩散过程及其 Wishart 对应过程。其特征值动力学可追溯至 Dyson（1962），而其谱收敛至半圆律则可追溯至 Chan（1992）以及 Rogers 和 Shi（1993）。式（17）的平稳分布是经过缩放的高斯正交系综，使得非对角元素的方差为 $1/(2aN)$，对角元素的方差为 $1/(aN)$。

Dyson 动力学与 Burgers 方程

特征值 $\lambda_1(t) > \cdots > \lambda_N(t)$ 的 H_t 满足 Dyson 系统

$$\mathrm{d}\lambda_i = \sqrt{\frac{2}{N}}\,\mathrm{d}w_i + \Big[\frac{1}{N}\sum_{j\neq i}\frac{1}{\lambda_i-\lambda_j} - a\lambda_i\Big]\mathrm{d}t, \tag{18}$$

与独立的布朗运动 w_i。对于正交对称类，特征值在零时刻之后不会发生碰撞（Anderson 等，2010，第 4.3 节），因此（18）具有唯一的强解。排斥项是其他特征值对每个特征值产生的二阶扰动。约束项是矩阵漂移的映像，因为 $-aH$ 与每一个

旋转对易，因此仅作用于谱上。令 $g_t(z) = N^{-1}\,\mathrm{tr}(z - H_t)^{-1}$。将 Itô 公式应用于 $(z - \lambda_i)^{-1}$，通过 $\sum_{i\neq j} f'(\lambda_i)/(\lambda_i - \lambda_j)$ $\frac{1}{2}\sum_{i\neq j}[f'(\lambda_i) - f'(\lambda_j)]/(\lambda_i - \lambda_j)$ $O(N^{-1})$ 对排斥求和进行对称化，并注意到伊藤修正项为 $-g\,\partial_z g$，便可在大-N 极限下得到受迫复 Burgers 方程

$$\partial_t g = -g\,\partial_z g + a\,\partial_z (z g). \tag{19}$$

第一项是自由布朗运动的无黏 Burgers 流，其特征线是直线。用 R-变换表示，即谱的自由累积量随时间线性增加，$R_t(w) = R_0(w) + tw$，这相当于自由概率论中独立增量下方差相加的对应结果。第二项是确定性收缩。平稳解必须满足 $g\,\partial_z g = a\,\partial_z(zg)$。积分一次并利用无穷远处的 $g \sim 1/z$ 可得 $g^2 = 2a(zg - 1)$，因此

$$g_\infty(z) = az - \sqrt{a^2 z^2 - 2a}, \qquad \rho_\infty(x) = \frac{1}{\pi}\sqrt{2a - a^2 x^2}, \quad |x| \le \sqrt{2/a},$$

半径为 $\sqrt{2/a}$ 的半圆律。对于 $a = \frac{1}{2}$，这是 $[-2, 2]$ 上的标准半圆律。从 $H_0 = 0$ 出发，方程（19）的解在每个时刻都是方差为 $\sigma_t^2 = (1 - e^{-2at})/(2a)$ 的半圆分布，它在小时间尺度下的自由布朗运动与平稳分布之间进行插值。图 4 展示了 $N = 40$ 时 Dyson 系统的模拟路径，以及 $N = 400$ 时的平稳谱。左面板中可见的能级排斥，是样本特征值避免交叉的机制，也是 Marchenko–Pastur 展宽的动态对应机制。

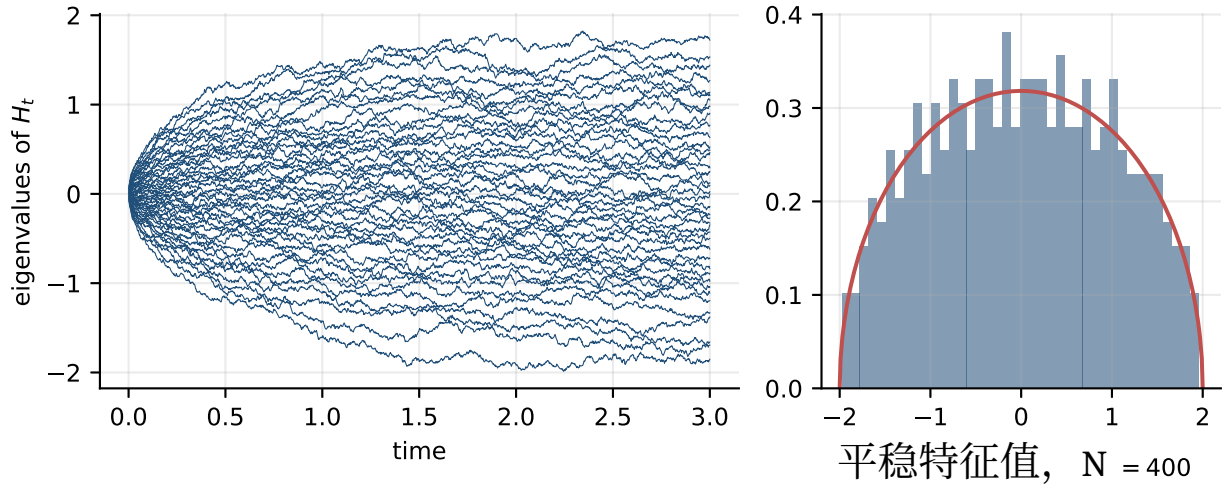

图 4： 左：矩阵 Ornstein–Uhlenbeck 过程（17）的特征值轨迹，$a = \frac{1}{2}$ 和 $N = 40$，初始于 $H_0 = 0$，并采用长度为 0.002的欧拉步进进行模拟。轨迹彼此排斥且永不相交。右：绘制 400×400 的平稳谱，并与半径为 $\sqrt{2/a} = 2$的半圆分布进行对比，该半圆分布是 Burgers 方程（19）的平稳解。

信息率与谱的充分性

假设一位观察者在整个 $[0,T]$ 期间连续监测 H，并且必须判断相关性冲击是以 a 的速率消散，还是以 b 的速率消散。令 P_a^T 和 P_b^T 分别表示路径 $(H_t)_{t\le T}$ 的相应分布，二者均从 a 制度的平稳分布出发。令 Λ_a^T 和 Λ_b^T 分别表示单独针对特征值路径 $(\lambda(t))_{t\le T}$ 的分布。

定理 (Dyson 过程的信息率与充分性)**.** 对于每个 $N \ge 1$ 和$T > 0$,

$$\mathrm{KL}(P_a^T \| P_b^T) = \mathrm{KL}(\Lambda_a^T \| \Lambda_b^T) = T\,\frac{N(N+1)(b-a)^2}{8a}.$$

因此，特征向量过程不携带关于耗散率的信息。根据 *Chernoff–Stein* 引理，从连续观测中区分这两种状态的误差指数为 $N(N+1)(b-a)^2/(8a)$ 每单位时间。

矩阵部分是将 Girsanov 定理应用于 $N(N+1)/2$ 个独立坐标的 17 式（Girsanov，1960）。每个非对角坐标的扩散方差为 $1/N$ 每单位时间，漂移差为 $(b-a)H_{ij}$。每个对角坐标的扩散方差为 $2/N$，漂移差为 $(b-a)H_{ii}$。Kullback–Leibler散度等于漂移差平方的期望积分的一半，并以逆扩散度量衡量：

$$\mathrm{KL}(P_a^T \| P_b^T) = \frac{N}{4}(b-a)^2 \int_0^T \mathbb{E}_a \,\mathrm{tr}\, H_t^2 \,\mathrm{d}t, \qquad \mathbb{E}_a \,\mathrm{tr}\, H_t^2 = \frac{N(N-1)}{2aN} + \frac{N}{aN} = \frac{N+1}{2a}.$$

特征值部分将 Girsanov 定理应用于（18）式。两种状态具有相同的排斥项，且仅在约束项上不同，漂移差为 $(b-a)\lambda_i$，扩散方差为$2/N$。由此得到 $\frac{N}{4}(b-a)^2 \int \mathbb{E}_a \sum_i \lambda_i^2 \mathrm{d}t$，这与前述积分相同，因为 $\sum_i \lambda_i^2 = \mathrm{tr}\ H^2$。因此两个散度相等。由于特征值路径是矩阵路径的可测函数，数据处理不等式表明，散度相等恰恰说明，舍弃特征向量不会损失关于 a 的信息。Lean文件证明了这两个闭式及其相等关系，且均为有理数域中的恒等式。Chernoff–Stein 结论是经典结果的连续时间版本：在第一类错误固定时，最优第二类错误指数等于 Kullback–Leibler 率（Chernoff，1952；Cover 和 Thomas，2006）。

对实证研究而言，有两个重要推论。第一，在给定误差水平下，区分两种耗散状态所需的时间按 $8a\log(1/\varepsilon)/[N(N+1)(b-a)^2]$ 缩放。随着资产数量增加，该时间呈二次方下降，因为每个 $N(N+1)/2$ 个独立的相关性坐标都会贡献各自的证据，而旋转不变性使所有这些证据都集中在 N 特征值中。对于 $N = 10$、$a = \frac{1}{2}$ 和 $b = 1$，速率为每单位时间 6.88 纳特；对 Euler对数似然比进行模拟所得结果也在表 1 中重现了这一速率。第二，这一结果是关于对称性的一项充分性结论。如果漂移破坏了旋转不变性，例如 $-a(H-M)$ 中有一个固定的非标量目标 M，那么特征向量过程就包含关于 M 的信息。在共同的 a 下区分两个目标 M 和 M' 的速率为$\frac{a^2 N}{4}\|M-M'\|_F^2$，其中包括 M 的旋转；这些旋转不会改变其谱，因此无法通过 $\lambda(t)$ 观察到。

定理中的路径空间散度，正是用于正则化潜在随机微分方程变分推断的同一对象。在该方法中，证据下界会减去$\frac{1}{2}\mathbb{E}\int\|u_t\|^2\mathrm{d}t$，其中 u识别漂移以扩散为标准的坐标表示（Li等人，2020年；Rice，2026年）。因此，这项充分性结果对于矩阵值潜在状态有直接的计算推论。若生成漂移和识别漂移都是关于 H 的旋转等变谱函数，那么 Kullback–Leibler 项以及任何旋转不变观测的似然都只取决于特征值路径。这样一来，变分推断便可在 N 维的 Dyson 系统上进行，而无需在 $N(N+1)/2$ 个矩阵坐标上计算，且不会造成信息损失。在有限时间步长 Δ 下，两种状态的 Euler 转移律具有相同的高斯协方差，仅均值不同。因此，根据相对熵链式法则，在当前状态下，矩阵链的每步散度为$\frac{N}{4}(b-a)^2\Delta\mathrm{tr}\ H^2$；对于特征值链，则以 $\sum_i \lambda_i^2$ 代替 $\mathrm{tr}\,H^2$，得到相同结果。两种离散速率的差异仅为两条 Euler 链平稳二阶矩之间的 $O(\Delta)$ 差异。当两种状态分别从各自的平稳分布出发时，平稳初始分布还会贡献一个与时间跨度无关的额外散度 $\frac{N}{(N+1)\,4}\ [b/a-1-\log(b/a)]$。由于该项也按 N^2 增长，因此，无论在快照数据还是在动力学中，截面上的相关性状态信息都十分丰富。

数值验证

前文各节中的每个精确公式和每个极限，都通过使用固定种子的蒙特卡洛模拟进行了检验。表 1 报告了比较结果。除尖峰行（在 $N=400, n=1600$ 处进行 400 次抽样）和动态行（200 条 Euler 路径）外，每一行均使用 4,000 次独立的 Wishart 抽样。允许一般 Σ 的各行中，协方差矩阵是随机的，条件数为 10 或 20，因此精确公式中缺少 Σ这一点本身也正在接受检验。最后一列报告了在存在比例极限时的相应值，从而展示了在所用维度下，有限样本值与其极限相距多远。精确的第二和第三 Wishart 矩一致到小数点后四位，而 Narayana 极限 $1+3q+q^2$ 在 $n=400$处相差约百分之一。这个差距正是 $3(1+q)/n+4/n^2$ 修正项。若以极限而非精确值为基准，模拟得到的第三阶矩看起来会偏离十多个标准误，这一点恰好提醒我们注意本文的主题。对数行列式这一行表明，需要加入二阶项 $\frac{1}{2}\log(1-q)$，才能连接极限 $N\varphi(q)$与精确值。其他每个精确结果都处于其模拟对应值的两个蒙特卡洛标准误以内。这包括强依赖下的虚假信息、其零均值和 Wilks 方差、两个 Kullback–Leibler 代价、两个最小方差比、两种尖峰强度下的 BBP 离群值及重叠，以及矩阵 Ornstein–Uhlenbeck 信息率。

图 5 左面板绘制了等大小独立块的插件信息零均值 $p=r$ 在 $n=250$时的结果，并显示了 ± 2 标准差带。精确的双伽马公式和比例极限 $n\kappa$ 在此尺度下无法区分，且与模拟均值吻合。经典近似 $pr/(2n)$ 在 $p\le 20$ 时准确，随后逐渐落后，在 55% 处仅达到真实虚假信息的 $p=r=100$。

表1： 精确公式和极限公式的蒙特卡洛验证。“精确值或极限值”是用于比较的解析值。对于尖峰行和动态行，该值分别是大-N 极限值或连续时间值。最后一列给出了不同情况下的比例极限。维度：在 $N = 100$、 $n = 400$ 处的矩和对数行列式。信息行在 $(p$、 r、 $n) = (10$、 20、 $100)$。库尔贝克–莱布勒行在 $N = 20$、 $n = 100$ 处。投资组合行在$N = 50$、 $n = 200$ 处。BBP 行在 $N = 400$、 $n = 1600$ 处。矩阵奥恩斯坦–乌伦贝克过程在 $N = 10$、$a = \frac{1}{2}$、$b = 1$ 处，欧拉步长为 0。01，时间跨度为 20 个时间单位。

数量	精确值或极限值	模拟值	蒙特卡洛标准误	极限
谱矩和对数行列式				
$N^{-1}\mathbb{E}\,\mathrm{tr}\ S^2$（精确值）	1.2525	1.2521	0.0003	1.2500
$N^{-1}\mathbb{E}\,\mathrm{tr}\ S^3$（精确值）	1.8219	1.8212	0.0007	1.8125
$\mathbb{E}\log\det S - \log\det\Sigma$（精确值）	-13.8391	-13.8515	0.0119	-13.6954
虚假信息				
$\mathbb{E}\hat{I}\ (p,\ r,\ n) = (10,\ 20,\ 100)$	1.1904	1.1885	0.0037	1.1833
$\mathbb{E}_0\hat{I}$ （独立性）	1.1904	1.1871	0.0019	–
$\mathrm{Var}_0\,\hat{I}$（Wilks）	0.0143	0.0140	0.0003	0.0141
信息成本与投资组合风险				
$\mathbb{E}$KL 散度$(\mathcal{N}_\Sigma\|\mathcal{N}_S)$	1.5283	1.5320	0.0035	1.4257
$\mathbb{E}$KL 散度$(\mathcal{N}_S\|\mathcal{N}_\Sigma)$	1.1299	1.1312	0.0018	1.0743
$\mathbb{E}R_{\mathrm{in}}/R$	0.7550	0.7535	0.0014	0.7500
$\mathbb{E}R_{\mathrm{out}}/R$	1.3267	1.3279	0.0012	1.3333
尖峰协方差（*BBP*）				
BBP 离群值, $\ell = 2$	2.5000	2.5043	0.0031	–
BBP重叠², $\ell = 2$	0.6000	0.5985	0.0022	–
BBP 离群值, $\ell = 4$	4.3333	4.3263	0.0071	–
BBP重叠², $\ell = 4$	0.8974	0.8974	0.0004	–
矩阵*Ornstein–Uhlenbeck*动力学				
矩阵OU KL速率	6.8750	6.9439	0.0873	–

这些带状区域也展示了比例极限的第二个含义。标准差从 0.006 （当 $p = 2$ 时）增至仅 0.54 （当 $p = 100$时），因此随着维度增大，虚假信息变得极其显著。未经校正的检验会以趋近于一的概率拒绝独立性。附录 C 中的表 6 列出了底层数据，包括精确标准差和模拟标准差。

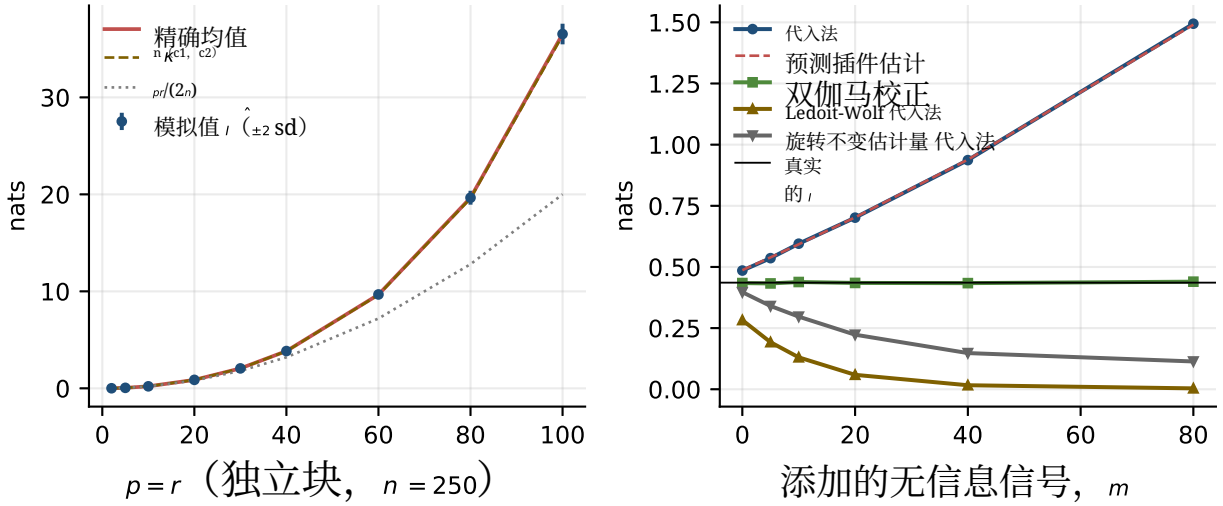

图5： 虚假信息。左：独立区块的插件互信息，其中$p = r$ 和 $n = 250$，精确均值为 $B_n(p, p)$，比例极限为 $n\kappa(p/n,\ p/n)$，经典值为 $pr/(2n)$。右：包含五种收益、五个典型相关为 0.4 的信息信号，以及 m 个附加无信息信号的设计，在 $n = 250$下。插件估计量如单调性命题所预测的那样上升，经 digamma 校正的估计量保持在真实值，而 Ledoit–Wolf 和旋转不变估计量的插件估计量则向零偏移。

图5右面板和表7展示了单调性命题及其经济解读。五种收益分别与五个信息信号相关联，典型相关系数均为 0.4，因此真实信息为 0.436 纳特。加入 m 个独立信号不会改变真实值。然而，代入法估计量却从 $m = 0$ 时的 0.485 上升至 $m = 80$时的 1.494，在每个 m处都与精确预测 $I + B_n(5, 5 + m)$ 相符，误差仅在模拟误差范围内。经 digamma 校正的估计量则保持在 0.433 至 0.440 之间。两种清洗估计量的表现截然不同。Ledoit–Wolf 收缩消除了虚假信息，也消除了大量真实信息，因此其估计值在 $m = 80$时降至 0.004 纳特。旋转不变估计量的收缩力度较弱，但其估计值也逐渐下降至 0.114 纳特。向单位阵收缩会将所有典型相关系数拉向零，而增加噪声坐标会提高最优收缩强度。对于信息估计，精确偏差校正优于协方差清洗。

图 6 检验了信息前沿背后的典型相关转变。当$c_1 = 0.1$ 且 $c_2 = 0.2$时，阈值为 $t_c = 0.408$，Wachter 边缘为 $\omega_+ = 0.500$。低于阈值时，最大的样本典型相关平方会停留在边缘，约为0.49 这是由于负的有限N 边缘偏移，无论总体值是多少。高于阈值后，模拟均值会紧密贴合 $\gamma(t)$ 。当 $t = 0.6$ 时，公式给出的值为 0.5748，模拟值为 0.5732；当 $t = 0.8$ 时，两者分别为 0.7521 和 0.7524。阈值的信息含量

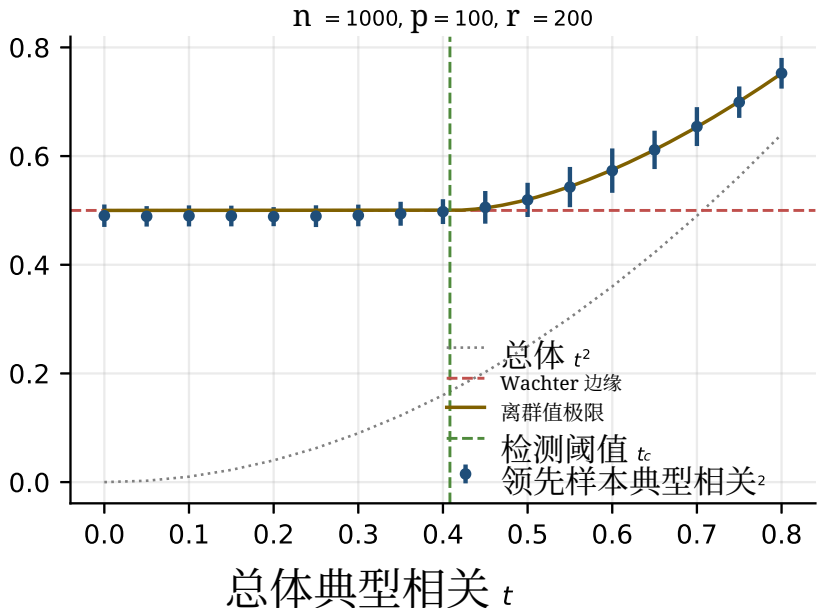

图6： 在 $n = 1000$， $p = 100$， $r = 200$处的典型相关相变。领先样本典型相关平方（160次抽样的均值 ± 2 标准差）在低于阈值 $\omega_+ = 0.500$ 时贴近Wachter 边缘 $t_c = 0.408$，超过阈值后则遵循离群值极限 $\gamma(t)$，即式9)所示的结果。低于阈值时，有限次N 抽样略低于 ω_+，这正如Tracy–Widom标度所预测的那样。

方向为 $I_c = 0.0912$ 纳特。强度低于该值的秩一联系，在这些维度下无法从最大的典型相关中显现出来；即便是知道其方向的投资者，其价值也约相当于九分之一单位的风险容忍度。

应用：美国行业收益的谱与信息

该应用使用1990年1月2日至2026年9月30日期间49个Fama–French行业投资组合的每日价值加权收益，共计9,254个交易日；数据取自Kenneth French数据库的CRSP 202609版本（Fama和French，1997年；French，2026年）。在此期间，全部49个序列均完整无缺。行业投资组合是经济物理学的天然实验室。它们的数量足够多，因此在常见窗口期长度下，随机矩阵效应会产生影响；其特征向量的经济含义易于解读；同时，它们具有足够的分散化程度，因此主导个股的极端特质事件不会驱动该谱。收益呈厚尾且具有异方差特征（Cont，2001），因此，前文各节中的高斯恒等式在此处只是基准，而非对实际情况的字面描述。本应用的部分目的，就是考察这些恒等式在多大程度上仍然成立。所有统计量均在滚动窗口中计算，窗口长度为 $T = 250$ 个交易日，以21天为步长推进，由此得到429个窗口，长宽比为 $q = 49/250 = 0.196$，用于相关分析。

行业相关性的谱

图7 展示了在2025日历年期间估计的相关矩阵的特征值。按照Laloux 等人（1999年），Marchenko–Pastur基准将主体方差重新缩放为$\hat{\sigma}^2 = 1 - \lambda_1/N = 0.502$，即未被市场模态吸收的方差，从而将上边缘置于1.046。最大特征值为24.38，因此市场模态承载了总方差的一半。总共有5个特征值超过缩放边界。与白噪声Marchenko–Pastur定律的预测相比，主体谱在零附近更加集中。这正是与非平坦总体谱进行自由乘法卷积的特征，$H \boxtimes \rho_q$ 按照预备知识部分的记号。行业投资组合具有异质的残差方差，并且存在一些近共线对，因此总体残差谱 H 本身就是分散的，而样本主体谱也继承了这种分散性。因此，对于检测上边缘处的离群值而言，白噪声基准是保守的；而对于下边缘而言，则缺乏信息。

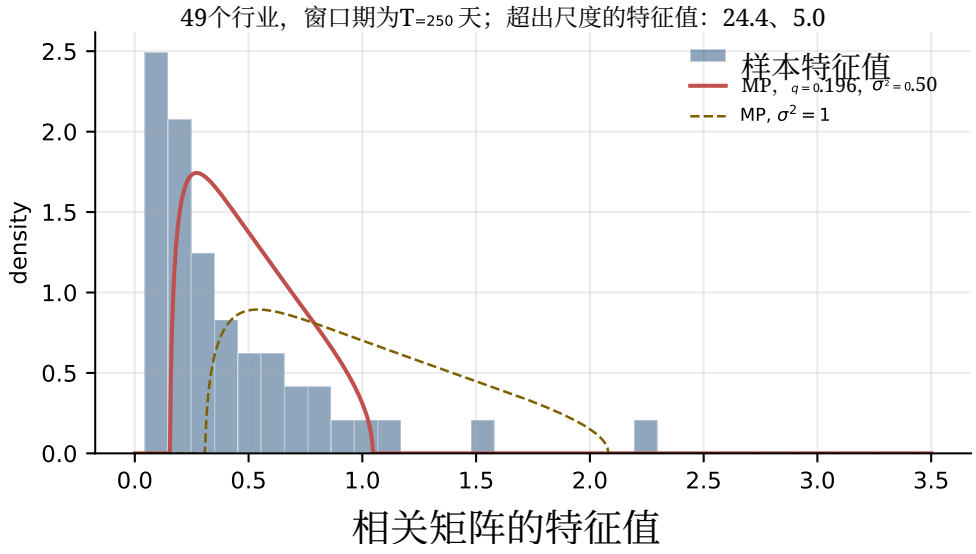

图 7： 2025日历年内 49 个行业相关矩阵的特征值（$T = 250$、$q = 0.196$），以及按残差方差 $\hat{\sigma}^2 = 1-\lambda_1/N$ 缩放后的 Marchenko–Pastur密度和未缩放的密度，以供参考。市场模态（$\lambda_1 = 24.38$）和第二个特征值超出了图表的尺度范围。

BBP反演将每个离群值转化为关于可靠性的判断。表2 报告了 2025 年窗口期最大八个特征值的相关结果：经$\hat{\sigma}^2$缩放后的特征值、总体$\hat{}$尖峰 ℓ_k（由公式（7）推导得出），以及样本特征向量与其总体对应特征向量之间的预测平方重叠度（公式（8））。市场模态几乎可以被完美估计。第二模态在必需消费品（软饮料、食品、家居用品和啤酒）上的载荷最大，其预测重叠度为0.98。第三和第四模态分别突出烟草、娱乐和国防，以及贵金属、煤炭和采矿，仍然可以得到可靠识别，重叠度分别为0.93和0.84。第五个模态则处于临界状态。其推导出的尖峰值为1.85，仅略高于阈值 $1 + \sqrt{q} = 1.44$；其预测重叠度为 0.59，意味着估计方向的方差中约有百分之四十是噪声。其载荷混合了农业、保险业和饮料，因此不应

从经济角度加以解读。第六个及更高的特征值位于缩放边界以内，理论预测它们的特征向量不包含有用信息。这种分类筛选正是异常值理论的实际价值所在。它不再只问特征值是否高于边界，而是以渐进的尺度衡量特征向量的可信程度。

表 2：对 2025年行业相关矩阵（$q = 0.196$，$\hat{\sigma}^2 = 0.502$）的八个最大特征值进行尖峰模型筛查。缩放特征值为 $\lambda_k/\hat{\sigma}^2$，其中 $k \geq 2$。隐含尖峰满足 $\ell^2 - (\lambda + 1 - q)\ell + \lambda = 0$，预测重叠度为（8）。破折号表示缩放特征值位于 Marchenko–Pastur主体内。载荷列出特征向量权重绝对值最大的四个行业。

k	λ_k	$\lambda_k/\hat{\sigma}^2$	隐含 ℓ_k	预测 $\lvert\langle u_k,\ v_k\rangle\rvert^2$	最大载荷
1	24.38	24.38	24.18	0.99	BldMt ，Chems ，Fin，Banks
2	5.02	9.98	9.76	0.98	Soda，Food，家庭用品，Beer
3	2.30	4.57	4.32	0.93	烟草 ，Fun，Guns，纺织品
4	1.55	3.08	2.77	0.84	Gold，Coal，矿山 ，零售
5	1.14	2.28	1.85	0.59	农业、保险、啤酒、汽水
6	1.04	2.07	–	0.00	枪支、煤炭、石油、药品
7	0.90	1.80	–	0.00	Gold，Oil，Fun，Aero
8	0.85	1.69	–	0.00	枪支、药品、石油、医疗器械

在整个样本期内，整体结构保持稳定，但强度有所变化。图8展示了吸收比率 λ_1/N，即市场模态所承载的方差份额，也是Kritzman 等人(2011年提出的系统性风险指标；同时还展示了市场特征向量与等权方向之间的平方重叠，以及高于边缘的特征值数量。吸收比率的平均值为0.545，范围从截至2026年9月的窗口期中的0.297，到截至2012年6月的窗口期中的0.820。在2001年、2008年和2020年的市场失序前后，该比率急剧上升，并在欧元区危机期间维持高位。市场特征向量始终没有明显偏离等权方向，其平方重叠至少为0.798。正如BBP重叠公式对较弱尖峰的预测，其重叠往往会在吸收比率下降时降低。离群特征值的数量在4至9之间变化，中位数为6，并且与吸收比率呈正相关（相关性为0.41）。危机会强化市场模态，也会使额外的行业模态从谱 bulk中分离出来。这与Münnix 等人(2012)的状态划分图景一致，也与上文的动态分析一致；在该分析中，相关性冲击属于谱事件，会使多个特征值一起移动。

噪声修饰的样本外成本

最小方差投资组合是最能直接观察估计噪声成本的情境，因为它仅取决于协方差矩阵。我每隔 21 个交易日，利用长度为 $T \in \{60,\ 120,\ 250,\ 500\}$ 天的估计窗口，构建由 49 个行业组成的全局最小方差投资组合，并持有接下来的 21 天。每种估计量都在相同的样本外期间进行评估，该期间从样本中的第 500 个交易日起算。使用的估计量包括样本协方差矩阵、对特征值进行截断的

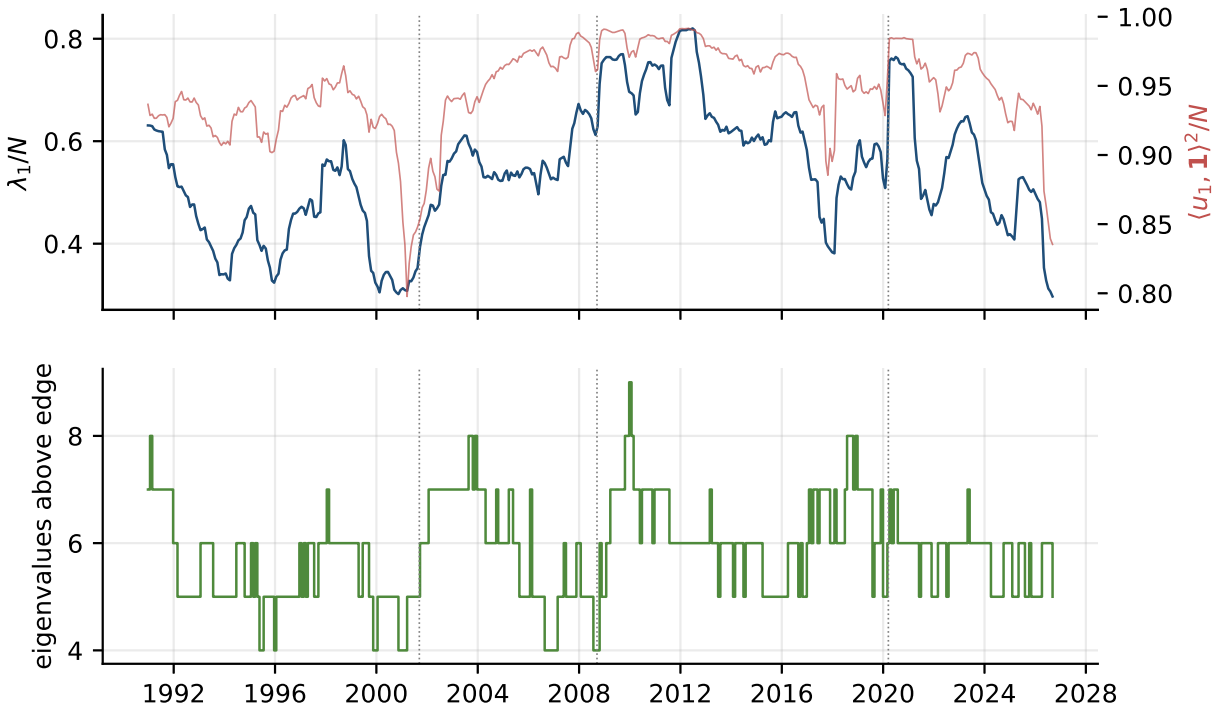

图8： 滚动谱统计量，采用250天窗口，每次向前推进21天。上图：吸收比率 λ_1/N（左轴）以及市场特征向量与归一化等权向量的重叠平方（右轴）。下图：高于重标度的马尔琴科–帕斯捷尔边缘的特征值数量。虚线标记2001年9月11日、2008年9月15日和2020年3月16日。

重标度的 Marchenko–Pastur边界处的相关矩阵（Laloux 等人，2016年）、向缩放单位矩阵进行线性收缩（Ledoit 和 Wolf，2004），以及 Bun 等人（2016年）提出的旋转不变估计量（5）$\eta = N^{-1/2}$。表 3 报告了实现的年化波动率。对于样本估计量，表中还报告了实现的投资组合方差与预测的投资组合方差之比，并在旁边列出高斯基准 $(1-q)^{-2}$，该基准由（15）推导得出。

表 3： 49 个行业的全局最小方差投资组合，在共同的样本外期间每 21 天再平衡一次。第 3–6 列：实现的年化波动率（百分比）。第 7 列：对于样本估计量，实现方差与样本内预测方差之比。第 8 列：其高斯基准 $(1-q)^{-2}$，即（15）中样本内因子与样本外因子的乘积。

T	q	样本	截断法	Ledoit–Wolf	RIE	$\hat{R}_{\text{out}}/\hat{R}_{\text{in}}$	$(1-q)^{-2}$
60	0.817	23.51	11.87	11.99	15.37	44.07	29.75
120	0.408	13.02	11.84	11.59	12.06	3.77	2.86
250	0.196	11.76	11.72	11.36	11.39	2.06	1.55
500	0.098	11.59	11.93	11.46	11.47	1.63	1.23
等权重（$1/N$）： 17.93							

结果与随机矩阵逻辑高度一致。在六十天窗口期（$q = 0.82$）下，样本估计量产生了23.51%的实现波动率，差于等权重投资组合的17.93%；这正是 Pafka 和 Kondor（2003）以及 DeMiguel 等人（2009）所记录的噪声装饰失败。截断法和线性收缩将该波动率削减了一半，分别降至11.87%和11.99%。在这一极端比率下，旋转不变估计量的表现稍差（15.37%），因为它所依赖的样本斯蒂尔切斯变换在仅有49个特征值的情况下解析度较低。随着窗口期延长，各估计量逐渐趋同。在 $T = 250$，最佳值与最差值之间的差距不到半个百分点；而在 $T = 500$，四者均介于11.59%和11.93%之间。在该窗口长度下，截断法略差于样本估计量，因为它舍弃了边缘以下真实存在的行业结构。样本估计量的实现方差与预测方差之比在每个窗口期都超过高斯基准。例如，在 $T = 250$，两者分别为2.06和1.55；在 $T = 500$，分别为1.63和1.23。该超额是非平稳性和厚尾共同作用的结果，它们会夸大每个估计量的样本外风险。收缩和旋转不变估计量不受 $(1-q)$ 因子的影响，在 $T = 500$的实现值与预测值之比也约为1.5。相对于这一共同因子进行衡量，样本估计量在 $T = 250$的超额约为1.23，在 $T = 500$的超额约为1.09。这低于高斯基准，因此在这些数据中，两种样本外误差来源是重叠的，而不是以乘法方式叠加。

碳密集型产业与金融产业之间的信息联结

信息分析考察碳密集型行业的日收益率能在多大程度上揭示金融行业的收益率。这是气候金融领域所研究的转型风险

渠道的一种简化形式（Battiston等人，2017年；Engle等人，2020年；Bolton和 Kacperczyk，2021年；Giglio等人，2021年；Pástor等人，2022年）。第一个区块包含 $p = 8$行业：石油和天然气、煤炭、公用事业、非金属和工业金属采矿、钢铁制造、化工、运输和汽车。第二个区块包含$r = 4$ 行业：银行业、保险业、房地产和交易业。在每个250天的窗口期内，收益率均经过去均值处理，因此相关自由度为 $n = 249$，并基于 12×12 样本协方差矩阵计算代入法信息 Iˆ。对于这一设计，精确的虚假信息为 B_{249}(8, 4) $= 0.0660$ 纳特，精确的零假设标准差为 $V_n^{1/2} = 0.0165$。根据价值信息定理，将相同数值除以 γ，即可衡量金融区块的投资者在交易前若能观察到碳区块，会赋予其的确定性等价价值。我报告了三个变体。˜ ˆ第一个是校正后的信息 $I = I - B_n$。第二个是安慰剂：在每个窗口期内，将碳区块按至少21天的随机偏移量进行循环移位，这会破坏同期依赖关系，同时保留各区块自身的动力学和边际分布。第三个是在给定等权市场收益率条件下的部分信息，其精确的虚假信息为 $B_{n-1}(p, r) = 0.0663$ 纳特，因为该回归额外使用了一个自由度。

图9 展示了结果。安慰剂是使用真实数据检验该理论最干净的方法。各窗口期的插件信息平均为 0.0653 纳特，而精确预测值为 0.0660。其校正值平均为 -0.0007，各窗口期的标准差为 0.0171，而精确值为 $V_n^{1/2} = 0.0165$。在名义显著性水平为 5% 时，单侧校正检验在 6.1% 的安慰剂窗口期中拒绝零假设。尽管存在厚尾、波动率聚集和序列依赖，高斯虚假信息有限样本理论在这些数据上的精度仍达到小数点后三位。原因之一是，零分布仅取决于两个区块的列空间。在任一列空间服从均匀分布时，该分布就是精确的；而在这些数据中，厚尾和波动率聚集所导致的偏离似乎只是二阶效应。

真实联系很强，且变化很大。校正信息平均为 0.828 纳特，而插件估计量为 0.894 纳特。每个窗口期的结果都在 5% 的水平上显著（100.0%），数值从 1996年1月的 0.284 纳特到 2020年4月的 1.638 纳特不等，且与吸收比率的相关性为 0.82。因此，碳密集型收益所包含的关于金融收益的大部分信息都是市场层面的信息。对市场进行条件化处理后，约五分之四的信息被剔除。部分信息占总信息的比例中位数为 18.6%。其余部分在经济意义和统计意义上均有所不同。校正后的部分信息平均为 0.170 纳特，在每个窗口期都达到 5% 的显著性水平（ 100.0% ）；即使在 1996 年 2 月的最低值 0.034 纳特时也是如此。该值在 2008年10月达到峰值 0.490 纳特，与吸收比率的相关性仅为 0.05。部分联系具有自身的动力学：它在金融危机前夕及期间上升，在大宗商品下行期（2015—2016年）之后的 2016—2017年再次上升，并在 2020年之后以及 2025—2026年上升。这些时期均未出现异常高的总体相关性。因此，碳密集型行业与金融之间存在行业特定信息渠道，贯穿整个样本期，并非市场模态的副产品。

最后，附加噪声实验在真实数据上重现了单调性命题。向金融区块添加八个独立高斯序列，会使

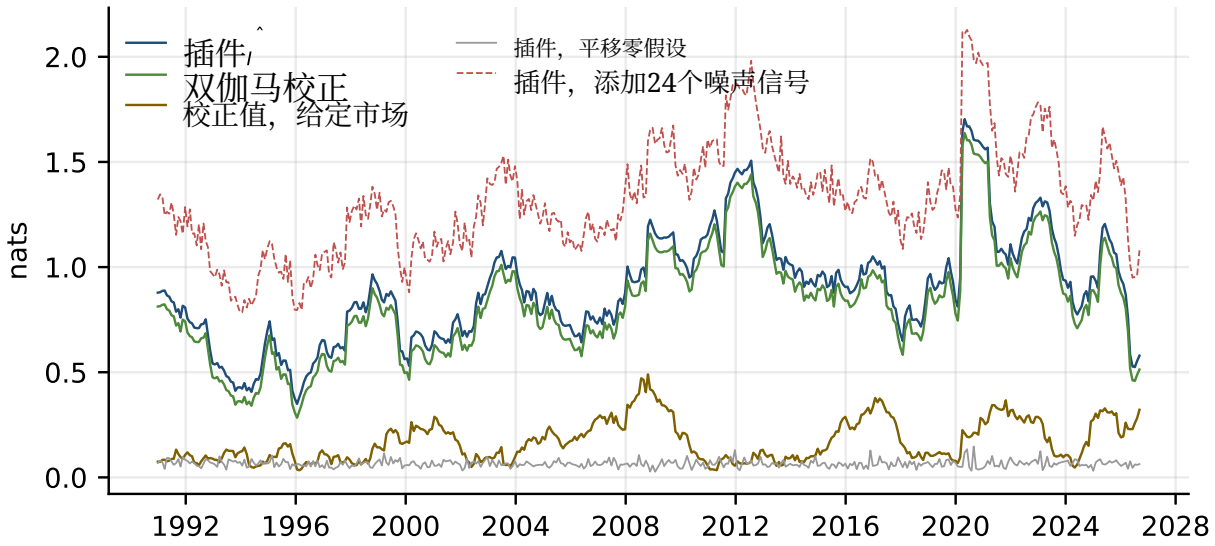

图9：在250日滚动窗口中，八个碳密集型行业与四个金融行业之间的互信息。插件估计值和经 digamma 校正的估计量、给定等权市场收益率的校正部分信息、循环平移安慰剂中的插件估计量，以及在金融区块中添加24个独立高斯序列后的插件估计量。

平均插件信息从 0.894 纳特升至 1.028 纳特；添加二十四个则升至 1.319 纳特。相应的精确虚假信息分别为 0.201 和 0.487 纳特，预测增量与观测增量精确到小数点后两位一致。校正估计值仍为 0.827 和 0.833，与原始设计的 0.828 几乎无法区分。就价值而言，未进行校正的评估者会把额外的确定性等价价值归因于扩充后的信号集 $0.42/\gamma$ 。这些价值并不存在。

讨论

这些结果共享同一个机制。高斯信息泛函是对数行列式之差，样本协方差矩阵的对数行列式是对数线性谱统计量，而对数函数是凹函数。每个样本谱都依据 Marchenko–Pastur 定律或 Wachter定律围绕其总体对应结果展开，因此每个对数行列式都会向下偏移，而且偏移量会随维度增加而增大。互信息是三个对数行列式的带符号组合，其维度之和为零。领先项相互抵消，但凹性并未消失，剩下的正是虚假信息。Kullback–Leibler代价将一个对数行列式与逆矩阵的迹结合起来，并继承后者在谱的左边缘处的发散性。信息价值只是以其他单位表示的互信息。因此，随机矩阵理论不仅给出了这些效应的大小，也揭示了它们的结构：有限 n下精确的双伽马望远镜求和、比例极限中的 Wachter积分，以及低秩结构与谱 bulk相遇时出现的相变。

需要作出若干限定。精确结果假设观测值彼此独立且服从高斯分布，而金融收益既不独立，也不服从高斯分布。应用中的安慰剂证据表明，对于中等规模的区组设计，零分布具有非凡的稳健性。对于更厚的尾部或更高的维度，不应直接假设这种稳健性；在这些情况下，谱的顶部可能会被少数极端观测值支配（Biroli 等人，2007年）。基于稳健散布估计量，或基于带随机尺度的一般 Silverstein 方程构建的椭圆分布与厚尾版本理论，是自然的扩展。信息价值恒等式属于局部均衡，并依赖于 CARA 偏好和联合正态性。在其他偏好下，对于某些类别的决策问题，互信息仍然是信息价值的上界（Cabrales 等人，2013年），但等式关系不再成立。信息前沿建立在典型相关转变的特征值部分之上；该部分已经得到确立，而关于特征向量对应结果的猜想，本文仅通过与 BBP 情形的类比提供支持。动态分析采用最简单的旋转不变模型。实证相关性动态具有一个破坏旋转不变性的市场模态，因此，在投影除去市场方向后，应将充分性结果应用于谱 bulk。

尽管如此，这些方法论启示是直接明确的。凡是估计高维高斯信息测度之处，无论是熵、互信息、转移熵，还是拟合高斯模型之间的 Kullback–Leibler散度，都可以计算出精确的有限样本偏差，并应予以消除。这项校正是精确的，无需了解真实协方差，而且在偏差相对于信号最大的情形下仍然有效。它也不同于协方差清洗：协方差清洗适用于二次损失，却会使信息估计向零偏移。在关于金融中的连通性、溢出效应和信息传递的实证研究中，校正后的统计量与精确的零分布提供了一种有原则的方法，可将真实信息与维度带来的机械性贡献区分开来。对于信息经济学而言，互信息与确定性等价价值之间的恒等关系意味着，在高维情形下看似有价值的数据，其表面价值很大一部分可能源于维数。虚假成分可以精确计算；每当加入无信息的序列时，它都会增大。

结论

本文构建了高斯信息泛函的随机矩阵理论，并将其应用于金融数据。两个变量块之间的插件互信息包含一个精确的虚假成分，该成分仅取决于维度和样本量。在独立性条件下，它服从精确的对数贝塔分布；在比例渐近区间内，它以速率$\kappa(c_1,\ c_2)$随样本量线性增长，并且每增加一个无信息坐标，其值都会严格上升。其谱表示为 Wachter 定律的对数矩提供了一个闭式，而典型相关相变则给出了一个前沿，低于该前沿便无法定位信息。精确的 CARA–高斯恒等式将上述所有结论转化为关于信息的确定性等价价值的结论。对于相关性动态，矩阵 Ornstein–Uhlenbeck 市场的 Dyson 特征值过程包含了关于其耗散率的充分信息，而耗散率可以通过学习资产数量的二次式速率来估计。在

三十六年的美国行业收益率数据中，这些工具以不同程度的可靠性恢复出市场和行业模态，量化噪声修饰带来的样本外成本，并识别出碳密集型行业与金融行业之间的信息联结；经过精确校正并在给定市场条件后，该联结依然存在。手稿、数据、复现代码、模拟输出和 Lean 证明均可在随附的代码仓库中获取。

参考文献

Nabil Ali Ahmed和D. V. Gokhale。多元分布的熵表达式及其估计量。《*IEEE*信息论汇刊》，35(3):688–692，1989年。doi: 10.1109/18.30996。Seung C. Ahn和Alex R. Horenstein。用于确定因子数量的特征值比率检验。《计量经济学》，81(3):1203–1227，2013年。doi: 10.3982/ECTA8968。Yacine Aït-Sahalia和Dacheng Xiu。利用主成分分析，结合高频数据估计高维因子模型。《计量经济学杂志》，201(2):384–399，2017年。doi: 10.1016/j.jeconom.2017.08.015。Gernot Akemann、Jinho Baik和Philippe Di Francesco（编）。《牛津随机矩阵理论手册》。牛津大学出版社，牛津，2011年。doi: 10.1093/oxfordhb/ 9780198744191.001.0001。George A. Akerlof。“柠檬”市场：质量不确定性与市场机制。《经济学季刊》，84(3):488–500，1970年。doi: 10.2307/1879431。Romain Allez和Jean-Philippe Bouchaud。特征向量动力学：一般理论及若干应用。《物理评论*E*》，86(4):046202，2012年。doi: 10.1103/PhysRevE.86.046202。Greg W. Anderson、Alice Guionnet和Ofer Zeitouni。《随机矩阵导论》，《剑桥高等数学研究》第118卷。剑桥大学出版社，剑桥，2010年。doi: 10.1017/CBO9780511801334。T. W. Anderson。《多元统计分析导论》。约翰·威立父子出版公司，霍博肯，新泽西州，第3版，2003年。Jushan Bai和Serena Ng。确定近似因子模型中的因子数量。《计量经济学》，70(1):191–221，2002年。doi: 10.1111/1468-0262.00273。Zhi-Dong Bai和Jack W. Silverstein。大维样本协方差矩阵线性谱统计量的中心极限定理。《概率年刊》，32(1A):553–605，2004年。doi: 10. 1214/aop/1078415845。Zhi-Dong Bai和Yong-Qua Yin。大维样本协方差矩阵最小特征值的极限。《概率年刊》，21(3):1275–1294，1993年。doi: 10.1214/aop/1176989118。

Zhidong Bai和Jack W. Silverstein。大维随机矩阵的谱分析。《Springer统计学系列》。Springer，纽约，第2版，2010年。doi: 10.1007/ 978-1-4419-0661-8。Zhidong Bai、Dandan Jiang、Jian-Feng Yao和Shurong Zheng。基于随机矩阵理论对大维协方差矩阵似然比检验的修正。《统计年刊》，37(6B):3822–3840，2009年。doi: 10.1214/09-AOS694。Jinho Baik和Jack W. Silverstein。尖峰总体模型的大样本协方差矩阵特征值。《多元分析杂志》，97(6):1382–1408，2006年。doi: 10.1016/j.jmva.2005.08.003。Jinho Baik、Gérard Ben Arous和Sandrine Péché。非零复样本协方差矩阵最大特征值的相变。《概率年刊》，33(5):1643–1697，2005年。doi: 10.1214/009117905000000233。Zhigang Bao、Jiang Hu、Guangming Pan和Wang Zhou。高维高斯向量的典型相关系数：有限秩情形。《统计年刊》，47(1):612–640，2019年。doi: 10.1214/18-AOS1704。Stefano Battiston、Antoine Mandel、Irene Monasterolo、Franziska Schütze和Gabriele Visentin。金融系统的气候压力测试。《自然·气候变化》，7(4):283–288，2017年。doi: 10.1038/nclimate3255。Mohamed Ishmael Belghazi、Aristide Baratin、Sai Rajeswar、Sherjil Ozair、Yoshua Bengio、Aaron Courville和Devon Hjelm。互信息神经估计。载于第*35*届国际机器学习会议论文集，《机器学习研究论文集》第80卷，第531–540页，2018年。Florent Benaych-Georges和Raj Rao Nadakuditi。大型随机矩阵的有限低秩扰动的特征值与特征向量。《数学进展》，227(1):494–521，2011年。doi: 10.1016/j.aim.2011.02.007。Peter J. Bickel和Elizaveta Levina。大维协方差矩阵的正则化估计。《统计年刊》，36(1):199–227，2008年。doi: 10.1214/009053607000000758。Monica Billio、Mila Getmansky、Andrew W. Lo和Loriana Pelizzon。金融与保险部门连通性和系统性风险的计量经济学度量。《金融经济学杂志》，104(3):535–559，2012年。doi: 10.1016/j.jfineco.2011.12.010。Giulio Biroli、Jean-Philippe Bouchaud和Marc Potters。关于重尾随机矩阵的最大特征值。*EPL*（《欧洲物理学快报》），78(1):10001，2007年。doi: 10.1209/0295-5075/78/10001。David Blackwell。实验的等价比较。《数学统计年刊》，24(2):265–272，1953年。doi: 10.1214/aoms/1177729032。

Alex Bloemendal、Antti Knowles、Horng-Tzer Yau 和 Jun Yin。论样本协方差矩阵的主成分。《概率论及相关领域》，164(1):459–552，2016年。doi: 10.1007/s00440-015-0616-x。Taras Bodnar、Nestor Parolya 和 Wolfgang Schmid。高维下全局最小方差投资组合的估计。《欧洲运筹学杂志》，266(1): 371–390，2018年。doi: 10.1016/j.ejor.2017.09.028。Patrick Bolton 和 Marcin Kacperczyk。投资者是否关注碳风险？《金融经济学杂志》，142(2):517–549，2021年。doi: 10.1016/j.jfineco.2021.05.008。Jean-Philippe Bouchaud 和 Marc Potters。《金融风险与衍生品定价理论：从统计物理到风险管理》。剑桥大学出版社，剑桥，第2版，2003年。doi: 10.1017/CBO9780511753893。Jean-Philippe Bouchaud 和 Marc Potters。随机矩阵理论的金融应用：简要综述。载于 Gernot Akemann、Jinho Baik 和 Philippe Di Francesco（编），《牛津随机矩阵理论手册》。牛津大学出版社，牛津，2011年。Marie-France Bru。Wishart过程。《理论概率杂志》，4(4):725–751，1991年。doi: 10.1007/BF01259552。Joël Bun、Romain Allez、Jean-Philippe Bouchaud 和 Marc Potters。一般含噪矩阵的旋转不变估计量。《*IEEE*信息论汇刊》，62(12): 7475–7490，2016年。doi: 10.1109/TIT.2016.2616132。Joël Bun、Jean-Philippe Bouchaud 和 Marc Potters。清理大型相关矩阵：随机矩阵理论工具。《物理学报告》，666:1–109，2017年。doi: 10.1016/j.physrep. 2016.10.005。Zdzisław Burda、Jerzy Jurkiewicz 和 Bartłomiej Wacław。相关Wishart矩阵的谱矩。《物理评论*E*》，71(2):026111，2005年。doi: 10.1103/PhysRevE.71. 026111。Antonio Cabrales、Olivier Gossner 和 Roberto Serrano。投资者的信息熵与信息价值。《美国经济评论》，103(1):360–377，2013年。doi: 10.1257/aer.103. 1.360。T. Tony Cai、Tengyuan Liang 和 Harrison H. Zhou。样本协方差矩阵对数行列式的分布律，以及高维高斯分布微分熵的最优估计。《多元分析杂志》，137:161–172，2015年。doi: 10.1016/ j.jmva.2015.02.003。Gary Chamberlain 和 Michael Rothschild。在大型资产市场中的套利、因子结构与均值—方差分析。《计量经济学》，51(5):1281–1304，1983年。doi: 10.2307/1912275。Terence Chan。Wigner半圆定律与矩阵值扩散过程的特征值。《概率论及相关领域》，93(2):249–272，1992年。doi: 10.1007/BF01195231。

Herman Chernoff。基于观测值之和的假设检验渐近效率度量。《数学统计年刊》，23(4):493–507，1952年。doi: 10.1214/aoms/1177729330。Benoît Collins。随机投影的乘积、Jacobi系综以及自由概率引发的普适性问题。《概率论及相关领域》，133(3):315–344，2005年。doi: 10.1007/s00440-005-0428-5。Rama Cont。资产收益的实证特性：典型事实与统计问题。《定量金融》，1(2):223–236，2001年。doi: 10.1080/713665670。Romain Couillet 和 Zhenyu Liao。《用于机器学习的随机矩阵方法》。剑桥大学出版社，剑桥，2022年。doi: 10.1017/9781009128490。Thomas M. Cover 和 Joy A. Thomas。《信息论要义》。Wiley-Interscience，霍博肯，新泽西州，第2版，2006年。doi: 10.1002/047174882X。Leonardo de Moura 和 Sebastian Ullrich。Lean 4定理证明器与编程语言。载于自动演绎——*CADE 28*，《计算机科学讲义》第12699卷，第625–635页，Cham，2021年。Springer。doi: 10.1007/978-3-030-79876-5_37。Victor DeMiguel、Lorenzo Garlappi 和 Raman Uppal。最优分散化与朴素分散化：1/N投资组合策略的效率有多低？《金融研究评论》，22(5): 1915–1953，2009年。doi: 10.1093/rfs/hhm075。Dipak K. Dey 和 C. Srinivasan。Stein损失下协方差矩阵的估计。《统计年刊》，13(4):1581–1591，1985年。doi: 10.1214/aos/1176349756。Francis X. Diebold 和 Kamil Yılmaz。论方差分解的网络拓扑：衡量金融企业的连通性。《计量经济学杂志》，182(1):119–134，2014年。doi: 10.1016/j.jeconom.2014.04.012。David L. Donoho、Matan Gavish 和 Iain M. Johnstone。尖峰协方差模型中特征值的最优收缩。《统计年刊》，46(4):1742–1778，2018年。doi: 10.1214/17-AOS1601。Ioana Dumitriu 和 Alan Edelman。β系综的矩阵模型。《数学物理学杂志，43(11):5830–5847，2002年。doi: 10.1063/1.1507823。Freeman J. Dyson。随机矩阵特征值的布朗运动模型。《数学物理学杂志》，3(6):1191–1198，1962年。doi: 10.1063/1.1703862。Alan Edelman 和 N. Raj Rao。随机矩阵理论。《数值学报》，14:233–297，2005年。doi: 10.1017/S0962492904000236。Noureddine El Karoui。Markowitz问题及其他带线性约束的二次规划中的高维效应：风险低估。《统计年刊》，38(6):3487–3566，2010年。doi: 10.1214/10-AOS795。

Edwin J. Elton 和 Martin J. Gruber。股票价格依赖结构的估计——对投资组合选择的启示。《金融学杂志》，28(5)：1203–1232，1973年。doi：10.1111/j.1540-6261.1973.tb01451.x。Robert Engle。动态条件相关性：一类简单的多元广义自回归条件异方差模型。《商业与经济统计学杂志》，20(3)：339–350，2002年。doi：10.1198/073500102288618487。Robert F. Engle、Olivier Ledoit 和 Michael Wolf。大型动态协方差矩阵。《商业与经济统计学杂志》，37(2)：363–375，2019年。doi：10.1080/07350015. 2017.1345683。Robert F. Engle、Stefano Giglio、Bryan Kelly、Heebum Lee 和 Johannes Stroebel。对冲气候变化新闻。《金融研究评论》，33(3)：1184–1216，2020年。doi：10.1093/rfs/hhz072。László Erdős 和 Horng-Tzer Yau。《随机矩阵理论的动力学方法》，《库朗数学讲义》第28卷，《库朗数学讲义》。美国数学学会，罗得岛州普罗维登斯，2017年。doi：10.1090/cln/028。Eugene F. Fama。有效资本市场：理论与实证研究综述。《金融学杂志》，25(2)：383–417，1970年。doi：10.2307/2325486。Eugene F. Fama 和 Kenneth R. French。行业股权成本。《金融经济学杂志》，43(2)：153–193，1997年。doi：10.1016/S0304-405X(96)00896-3。Jianqing Fan、Yuan Liao 和 Martina Mincheva。通过主正交补阈值化进行大型协方差估计。《皇家统计学会会刊*B*辑：统计方法论》，75(4)：603–680，2013年。doi：10.1111/rssb.12016。Maryam Farboodi 和 Laura Veldkamp。金融数据技术的长期增长。《美国经济评论》，110(8)：2485–2523，2020年。doi：10.1257/aer.20171349。Peter J. Forrester。《对数气体与随机矩阵》，《伦敦数学学会专著》第34卷。《伦敦数学学会专著》。普林斯顿大学出版社，新泽西州普林斯顿，2010年。Kenneth R. French。数据资料库：49个行业投资组合，日收益率。https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html，2026年。CRSP 202609版本，访问日期为2026年11月23日。Matan Gavish 和 David L. Donoho。奇异值的最优硬阈值为 $4/\sqrt{3}$。《*IEEE*信息论汇刊》，60(8)：5040–5053，2014年。doi：10.1109/TIT.2014. 2323359。Stuart Geman。随机矩阵范数的一个极限定理。《概率年刊》，8(2)：252–261，1980年。doi：10.1214/aop/1176994775。Stefano Giglio、Bryan Kelly 和 Johannes Stroebel。气候金融。《金融经济学年度评论》，13：15–36，2021年。doi：10.1146/annurev-financial-102620-103311。

–

伊戈尔·弗拉基米罗维奇·吉尔萨诺夫。通过测度的绝对连续替换变换某类随机过程。《概率论及其应用》，5 (3)：285–301，1960年。doi：10.1137/1105027。
Parameswaran Gopikrishnan、Bernd Rosenow、Vasiliki Plerou 和 H. Eugene Stanley。量化并解释金融市场中的集体行为。《物理评论E》，64(3)：035106，2001年。doi：10.1103/PhysRevE.64.035106。克里斯蒂安·古里耶鲁、乔安·亚夏克和拉兹万·苏法纳。多元随机波动率的Wishart自回归过程。《计量经济学杂志》，150(2)：167–181，2009年。doi：10.1016/j.jeconom.2008.12.016。桑福德·J·格罗斯曼和 Joseph E. Stiglitz。论信息有效市场的不可能性。《美国经济评论》，70(3)：393–408，1980年。特雷弗·哈斯蒂、安德烈亚·蒙塔纳里、萨哈龙·罗塞特和瑞安·J·蒂布希拉尼。高维无岭最小二乘插值中的意外现象。《统计年刊》，50(2)：949–986，2022年。doi：10.1214/21-AOS2133。马丁·F·赫尔维希。论竞争市场中的信息聚合。《经济理论杂志》，22(3)：477–498，1980年。doi：10.1016/0022-0531(80)90056-3。拉维·贾甘纳坦和马同书。大型投资组合中的风险降低：为什么施加错误的约束反而有帮助。《金融学杂志》，58(4)：1651–1683，2003年。doi：10.1111/ 1540-6261.00580。W. 詹姆斯和查尔斯·斯坦。二次损失下的估计。载于《第四届伯克利数学统计与概率论研讨会论文集》，第1卷，第361–379页。加利福尼亚大学出版社，加利福尼亚州伯克利，1961年。蒋铁峰和杨帆。高维正态分布下经典似然比检验的中心极限定理。《统计年刊》，41(4)：2029–2074，2013年。doi：10.1214/13-AOS1134。Iain M. Johnstone。主成分分析中最大特征值的分布。《统计年刊》，29(2)：295–327，2001年。doi：10.1214/aos/1009210544。
Iain M. Johnstone。多元分析与Jacobi系综：最大特征值、Tracy–Widom极限及收敛速率。《统计年刊》，36(6)：2638–2716，2008年。doi：10.1214/08-AOS605。
Marcin Kacperczyk、Stijn Van Nieuwerburgh 和 Laura Veldkamp。共同基金注意力分配的理性理论。《计量经济学》，84(2)：571–626，2016年。doi：10.3982/ECTA11412。奥哈德·卡丹和阿萨夫·马内拉。信息价值的估计。《金融研究评论》，32(3)：951–991，2019年。doi：10.1093/rfs/hhy087。

Raymond Kan 和 Guofu Zhou。参数不确定性下的最优投资组合选择。《金融与定量分析杂志》，42(3):621–656，2007年。doi：10.1017/ S0022109000004129。Bryan Kelly、Semyon Malamud 和 Kangying Zhou。收益预测中复杂性的价值。《金融学杂志》，79(1):459–503，2024年。doi：10.1111/jofi.13298。Jr. Kelly、John L. 信息率的新解释。《贝尔系统技术杂志》，35(4):917–926，1956年。doi：10.1002/j.1538-7305.1956.tb03809.x。Antti Knowles 和 Jun Yin。随机矩阵的各向异性局部定律。《概率论及相关领域》，169(1):257–352，2017年。doi：10.1007/s00440-016-0730-4。Imre Kondor、Szilárd Pafka 和 Gábor Nagy。不同风险度量下投资组合选择的噪声敏感性。《银行与金融杂志》，31(5):1545–1573，2007年。doi：10.1016/ j.jbankfin.2006.12.003。Serhiy Kozak、Stefan Nagel 和 Shrihari Santosh。收缩截面。《金融经济学杂志》，135(2):271–292，2020年。doi：10.1016/j.jfineco.2019.06.008。Alexander Kraskov、Harald Stögbauer 和 Peter Grassberger。互信息估计。《物理评论*E*》，69(6):066138，2004年。doi：10.1103/PhysRevE.69.066138。Mark Kritzman、Yuanzhen Li、Sebastien Page 和 Roberto Rigobon。作为系统性风险度量的主成分。《投资组合管理杂志》，37(4):112–126，2011年。doi：10.3905/jpm.2011.37.4.112。Solomon Kullback 和 Richard A. Leibler。论信息与充分性。《数学统计年刊》，22(1):79–86，1951年。doi：10.1214/aoms/1177729694。Jarosław Kwapień 和 Stanisław Drożdż。复杂系统的物理学方法。《物理学报告》，515(3–4):115–226，2012年。doi：10.1016/j.physrep.2012.01.007。Albert S. Kyle。连续拍卖与内幕交易。《计量经济学》，53(6):1315–1335，1985年。doi：10.2307/1913210。Laurent Laloux、Pierre Cizeau、Jean-Philippe Bouchaud 和 Marc Potters。金融相关矩阵的噪声修饰。《物理评论快报》，83(7):1467–1470，1999年。doi：10.1103/PhysRevLett.83.1467。Laurent Laloux、Pierre Cizeau、Marc Potters 和 Jean-Philippe Bouchaud。随机矩阵理论与金融相关性。《理论与应用金融国际杂志》，3(3):391–397，2000年。doi：10.1142/S0219024900000255。Olivier Ledoit 和 Sandrine Péché。某些大样本协方差矩阵系综的特征向量。《概率论及相关领域》，151(1):233–264，2011年。doi：10.1007/ s00440-010-0298-3。

Olivier Ledoit 和 Michael 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。《实证金融学杂志》，10(5): 603–621，2003年。doi：10.1016/S0927-5398(03)00007-0。Olivier Ledoit 和 Michael Wolf。大维协方差矩阵的良态估计量。《多元分析杂志》，88(2):365–411，2004年。doi：10.1016/S0047-259X(03)00096-4。Olivier Ledoit 和 Michael Wolf。大维协方差矩阵的非线性收缩估计。《统计年刊》，40(2):1024–1060，2012年。doi：10.1214/12-AOS989。Olivier Ledoit 和 Michael Wolf。用于投资组合选择的协方差矩阵非线性收缩：Markowitz 遇见 Goldilocks。《金融研究评论》，30(12):4349–4388，2017年。doi：10.1093/rfs/hhx052。Olivier Ledoit 和 Michael Wolf。大维协方差矩阵的解析非线性收缩。《统计年刊》，48(5):3043–3065，2020年。doi：10.1214/19-AOS1921。Olivier Ledoit 和 Michael Wolf。非线性收缩的力量：协方差矩阵估计的综述与指南。《金融计量经济学杂志》，20(1):187–218，2022年。doi：10.1093/jjfinec/nbaa007。Martin Lettau 和 Markus Pelger。潜在资产定价因子的估计。《计量经济学杂志》，218(1):1–31，2020a。doi：10.1016/j.jeconom.2019.08.012。Martin Lettau 和 Markus Pelger。拟合股票收益时间序列与截面的因子。《金融研究评论》，33(5):2274–2325，2020b。doi：10.1093/rfs/ hhaa020。Xuechen Li、Ting-Kam Leonard Wong、Ricky T. Q. Chen 和 David Duvenaud。随机微分方程的可扩展梯度。载于第二十三届国际人工智能与统计会议论文集，第108卷，《机器学习研究论文集》，第3870–3882页，2020年。Giacomo Livan、Marcel Novaes 和 Pierpaolo Vivo。《随机矩阵导论：理论与实践》，第26卷，*Springer*数学物理简明丛书。Springer，Cham，2018年。doi：10.1007/978-3-319-70885-0。Bartosz Maćkowiak、Filip Matějka 和 Mirko Wiederholt。理性疏忽：综述。经济文献杂志，61(1):226–273，2023年。doi：10.1257/jel.20211524。Rosario N. Mantegna。金融市场的层级结构。《欧洲物理学杂志 *B*》，11(1):193–197，1999年。doi：10.1007/s100510050929。Rosario N. Mantegna 和 H. Eugene Stanley。《经济物理学导论：金融中的相关性与复杂性》。剑桥大学出版社，剑桥，2000年。doi：10.1017/ CBO9780511755767。

Vladimir A. Marčenko 和 Leonid A. Pastur. 若干随机矩阵集合的特征值分布。《苏联数学汇刊》，1(4):457–483，1967年。doi: 10.1070/ SM1967v001n04ABEH001994. Harry Markowitz. 投资组合选择。《金融学杂志》，7(1):77–91，1952。doi: 10.1111/j.1540-6261.1952.tb01525.x. Ian W. R. Martin 和 Stefan Nagel. 大数据时代的市场效率。《金融经济学杂志》，145(1):154–177，2022。doi: 10.1016/j.jfineco.2021.10.006. Filip Matějka 和 Alisdair McKay. 对离散选择的理性疏忽：多项式逻辑模型的新基础。《美国经济评论》，105(1):272–298，2015。doi: 10.1257/aer.20130047. David McAllester 和 Karl Stratos. 互信息测量的形式化局限。载于第二十三届人工智能与统计国际会议论文集，第108卷《机器学习研究论文集》，第875–884页，2020年。Madan Lal Mehta. 随机矩阵，第142卷纯粹数学与应用数学。Elsevier/Academic Press，阿姆斯特丹，第3版，2004年。Richard O. Michaud. Markowitz优化之谜：“优化”是否就是最优？《金融分析师杂志》，45(1):31–42，1989年。doi: 10.2469/faj.v45.n1.31. James A. Mingo 和 Roland Speicher. 《自由概率与随机矩阵》，第35卷菲尔兹研究所专著。Springer，纽约，2017年。doi: 10.1007/978-1-4939-6942-5. Neeraj Misra、Harshinder Singh 和 Eugene Demchuk. 多元正态分布熵的估计。《多元分析杂志》，92(2):324–342，2005年。doi: 10.1016/j.jmva.2003.10.003. Robb J. Muirhead. 《多元统计理论概论》。约翰·威立父子出版公司，纽约，1982年。doi: 10.1002/9780470316559. Michael C. Münnix、Takashi Shimada、Rudi Schäfer、Francois Leyvraz、Thomas H. Seligman、Thomas Guhr 和 H. Eugene Stanley. 金融市场状态识别。《科学报告》，2:644，2012年。doi: 10.1038/srep00644. Alexandru Nica 和 Roland Speicher. 《自由概率组合学讲义》，第335卷伦敦数学学会讲义系列。剑桥大学出版社，剑桥，2006年。doi: 10.1017/CBO9780511735127. Alexei Onatski. 大型因子模型中因子数量的假设检验。《计量经济学》，77(5):1447–1479，2009年。doi: 10.3982/ECTA6964. Alexei Onatski. 根据特征值的经验分布确定因子数量。《经济学与统计学评论》，92(4):1004–1016，2010年。doi: 10.1162/REST_a_00043.

Szilárd Pafka 和 Imre Kondor. 含噪协方差矩阵与投资组合优化。《欧洲物理学杂志 *B*》，27(2):277–280，2002年。doi: 10.1140/epjb/e20020153. Szilárd Pafka 和 Imre Kondor. 含噪协方差矩阵与投资组合优化 II。《物理学 *A*：统计力学及其应用》，319:487–494，2003年。doi: 10.1016/ S0378-4371(02)01499-1. Liam Paninski. 熵与互信息的估计。《神经计算》，15(6): 1191–1253，2003年。doi: 10.1162/089976603321780272. Ľuboš Pástor、Robert F. Stambaugh 和 Lucian A. Taylor. 剖析绿色收益。《金融经济学杂志》，146(2):403–424，2022年。doi: 10.1016/j.jfineco.2022.07.007. Debashis Paul. 大维度尖峰协方差模型中样本特征结构的渐近性质。《中华统计学志》，17(4):1617–1642，2007年。Markus Pelger. 理解系统性风险：一种高频方法。《金融学杂志》，75(4):2179–2220，2020年。doi: 10.1111/jofi.12898. Vasiliki Plerou、Parameswaran Gopikrishnan、Bernd Rosenow、Luís A. Nunes Amaral 和 H. Eugene Stanley. 金融时间序列交叉相关的普适与非普适特性。《物理评论快报》，83(7):1471–1474，1999年。doi: 10.1103/PhysRevLett.83. 1471. Vasiliki Plerou、Parameswaran Gopikrishnan、Bernd Rosenow、Luís A. Nunes Amaral、Thomas Guhr 和 H. Eugene Stanley. 金融数据交叉相关的随机矩阵方法。《物理评论*E*》，65(6):066126，2002年。doi: 10.1103/PhysRevE.65.066126. Marc Potters 和 Jean-Philippe Bouchaud. 《随机矩阵理论入门：面向物理学家、工程师和数据科学家》。剑桥大学出版社，剑桥，2020年。doi: 10.1017/9781108768900. Pierre-Alain Reigneron、Romain Allez 和 Jean-Philippe Bouchaud. 主回归分析与指数杠杆效应。《物理学 *A*：统计力学及其应用》，390(17):3026–3035，2011年。doi: 10.1016/j.physa.2011.04.007. James K. Rice. 随机深度学习：用于结构化时间数据不确定性建模的概率框架。arXiv预印本 arXiv:2601.05227，2026年。L. C. G. Rogers 和 Z. Shi. 相互作用的布朗粒子与维格纳定律。《概率论及相关领域》，95(4):555–570，1993年。doi: 10.1007/BF01196734. Stephen A. Ross. 资本资产定价的套利理论。《经济理论杂志》，13(3):341–360，1976年。doi: 10.1016/0022-0531(76)90046-6. Simo Särkkä 和 Arno Solin. 《应用随机微分方程》。数理统计学会教材丛书。剑桥大学出版社，剑桥，2019年。doi: 10.1017/9781108186735.

Thilo A. Schmitt, Desislava Chetalova, Rudi Schäfer, Tomas Kaumann。金融时间序列的非平稳性：一般特征与尾部行为。*EPL*（《欧洲物理学快报》），103 (5):58003, 2013年。doi: 10.1209/0295-5075/103/58003. Claude E. Shannon。《通信的数学理论》。《贝尔系统技术杂志》, 27(3):379–423, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x. Jack W. Silverstein 和 Zhi-Dong Bai。论一类大维随机矩阵的特征值经验分布。《多元分析杂志》, 54(2):175–192, 1995. doi: 10.1006/jmva.1995.1051. Christopher A. Sims。理性疏忽的影响。《货币经济学杂志》, 50(3):665–690, 2003. doi: 10.1016/S0304-3932(03)00029-1. Richard P. Stanley。《组合学计数》，第2卷，剑桥高级数学研究丛书第*62*卷。剑桥大学出版社，剑桥，1999。doi: 10.1017/ CBO9780511609589. Joseph E. Stiglitz。信息经济学对20世纪经济学的贡献。《经济学季刊》, 115(4):1441–1478, 2000. doi: 10.1162/ 003355300555015. Terence Tao。《随机矩阵理论专题》，第132卷，数学研究生教材。美国数学学会，罗得岛州普罗维登斯，2012。doi: 10.1090/gsm/132. Emre Telatar。多天线高斯信道的信道容量。《欧洲电信学报》, 10(6):585–595, 1999. doi: 10.1002/ett.4460100604. Craig A. Tracy 和 Harold Widom。能级间距分布与艾里核。《数学物理通讯》, 159(1):151–174, 1994. doi: 10.1007/BF02100489. Craig A. Tracy 和 Harold Widom。论正交与辛矩阵系综。《数学物理通讯》, 177(3):727–754, 1996. doi: 10.1007/BF02099545. Antonia M. Tulino 和 Sergio Verdú。随机矩阵理论与无线通信。《通信与信息论基础与趋势》, 1(1):1–182, 2004. doi: 10.1561/0100000001. Michele Tumminello、Fabrizio Lillo 和 Rosario N. Mantegna。金融市场中的相关性、层级结构与网络。《经济行为与组织杂志》, 75(1):40–58, 2010. doi: 10.1016/j.jebo.2010.01.004. Akihiko Utsugi、Kazusumi Ino 和 Masaki Oshikawa。金融市场交叉相关性的随机矩阵理论分析。《物理评论*E*》, 70(2):026110, 2004. doi: 10.1103/PhysRevE.70.026110. Stijn Van Nieuwerburgh 和 Laura Veldkamp。信息获取与投资组合多元化不足。《经济研究评论》, 77(2):779–805, 2010. doi: 10.1111/ j.1467-937X.2009.00583.x.

Laura L. Veldkamp。《宏观经济学与金融中的信息选择》。普林斯顿大学出版社，新泽西州普林斯顿，2011年。Robert E. Verrecchia。噪声理性预期经济中的信息获取。《计量经济学》, 50(6):1415–1430, 1982. doi: 10.2307/1913389. Dan Voiculescu。随机矩阵与自由乘积的极限定律。《数学发明》, 104(1):201–220, 1991. doi: 10.1007/BF01245072. Dan V. Voiculescu、Ken J. Dykema 和 Alexandru Nica。自由随机变量，第1卷，CRM专著丛书第1卷。美国数学学会，罗得岛州普罗维登斯，1992。doi: 10.1090/crmm/001. Kenneth W. Wachter。多重判别比的极限经验测度。《统计年刊》, 8(5):937–957, 1980年。doi: 10.1214/aos/1176345134. Eugene P. Wigner。无限维增广矩阵的特征向量。《数学年刊》, 62(3):548–564, 1955. doi: 10.2307/1970079. Eugene P. Wigner。论某些对称矩阵根的分布。《数学年刊》, 67(2):325–327, 1958. doi: 10.2307/1970008. Yanrong Yang 和 Guangming Pan。基于正则化典型相关系数的高维数据独立性检验。《统计年刊》, 43(2):467–500, 2015. doi: 10.1214/14-AOS1284.

附录 A。证明

全文中，$\psi = \Gamma'/\Gamma$ 表示双伽马函数，ψ' 表示三伽马函数。所使用的事实包括：ψ 严格递增，而 ψ' 在 $(0,\infty)$ 上严格递减且为正；当 $x \to \infty$时，$\psi(x) = \log x - 1/(2x) + O(x^{-2})$，且 $\psi'(x) = 1/x + 1/(2x^2) + O(x^{-3})$；对于 $\beta \sim \mathrm{Beta}(a,b)$，有 $\mathbb{E}\log\beta = \psi(a) - \psi(a+b)$，且 $\mathrm{Var}\log\beta = \psi'(a) - \psi'(a+b)$。当$Y \sim \chi^2_k$时，$\mathbb{E}\log Y = \psi(k/2) + \log 2$，且 $\mathrm{Var}\log Y = \psi'(k/2)$。

对数行列式

设 $W \sim \mathrm{W}_N(n,\Sigma)$，其中 $n \ge N$。则 $W = \Sigma^{1/2} W_0 \Sigma^{1/2}$，且 $W_0 \sim \mathrm{W}_N(n, I)$，因此 $\log\det(W/n) - \log\det\Sigma = \log\det W_0 - N\log n$。根据 Bartlett分解，$W_0 = LL^\top$，其中 L 为下三角矩阵，且 $L_{ii}^2 \sim \chi^2_{n-i+1}$ 相互独立（Muirhead，1982，第 3 章），因此 $\log\det W_0 = \sum_i \log L_{ii}^2$。取期望和方差可得（10）。对于极限，写成

$$\ell_n(N) = \sum_{i=1}^{N} \log\Big(1 - \frac{i-1}{n}\Big) - \sum_{i=1}^{N} \frac{1}{n-i+1} + O\Big(\sum_{i\le N} (n-i+1)^{-2}\Big).$$

当 $N/n \to q < 1$ 时，最后一项为 $O(1/n)$。第二个求和等于 $-\log(1-q) + o(1)$。对第一个求和应用梯形法则，其被加数 $f(u) = \log(1-u)$ 在 $[0,q]$ 上是光滑的，可得 $n\int_0^{N/n} \log(1-u)\,\mathrm{d}u + \tfrac{1}{2}[f(0) - f(N/n)] + O(1/n)$。由于 $\int_0^q \log(1-u)\,\mathrm{d}u = -(1-q)\log(1-q) - q = q\varphi(q)$，结果为 $\ell_n(N) = N\varphi(N/n) + \tfrac{1}{2}\log(1-N/n) + o(1)$。特别地，$\ell_n(N)/N \to \varphi(q)$。方差极限可由同样的方法从 $\psi'((n-i+1)/2) = 2/(n-i+1) + O(n^{-2})$ 得出，从而得到 $\sum_{i\le N} 2/(n-i+1) \to -2\log(1-q)$。

精确的虚假信息

如果 $nS \sim \mathcal{W}_{p+r}(n,\ \Sigma)$，则对角区块满足 $nS_{11} \sim \mathcal{W}_p(n,\ \Sigma_{11})$ 以及 $nS_{22} \sim \mathcal{W}_r(n,\ \Sigma_{22})$，因为 $X^\top X$ 的主子矩阵就是相应列的交叉乘积矩阵 X。因此，根据（10），

$$2(\mathbb{E}\hat{I} - I) = \ell_n(p) + \ell_n(r) - \ell_n(p+r) = \Psi_n(p) + \Psi_n(r) - \Psi_n(p+r) - (p + r - p - r)\log\frac{n}{2},$$

即 $2B_n(p,\ r)$。对于望远镜形式，将 $\Psi_n(p+r) = \Psi_n(p) + \sum_{i=1}^{r}\psi((n-p-i+1)/2)$ 拆开，使得 $2B_n = \sum_{i=1}^{r}[\psi((n-i+1)/2) - \psi((n-p-i+1)/2)]$。每个加项都严格为正，因为 ψ 严格递增且 $p \ge 1$。从第一种形式可以明显看出 $(p,\ r)$ 的对称性。Lean 定理 `twiceBias_eq`、`twiceBias_pos` 和 `twiceBias_symm` 证明了望远镜性质、正性和对称性；其中可用任意数列 c_i 替代 $\psi((n-i+1)/2)$，且证明正性时只需假设该数列严格反单调。对 $n-1$ （去均值）或 $n-k$（对 k 公共回归变量进行回归），同样的结论也成立，因为此时残差交叉乘积矩阵为 $\mathcal{W}_{p+r}(n-k,\ \Sigma)$，且 I 不受 S 归一化的影响。

零分布

在 $\Sigma_{12}=0$ 条件下，给定第二个数据块 X_2，令 P_2 为投影到其列空间上的投影算子。记 $E=X_1^\top(I-P_2)X_1$ 和 $H=X_1^\top P_2 X_1$。给定 X_2，X_1 的各行相互独立，且服从 $\mathcal{N}(0,\ \Sigma_{11})$，因此 $E\sim\mathcal{W}_p(n-r,\ \Sigma_{11})$ 与 $H\sim\mathcal{W}_p(r,\ \Sigma_{11})$ 相互独立，且其分布不依赖于 X_2。由于 $\det S=\det S_{22}\det(S_{11}-S_{12}S_{22}^{-1}S_{21})$ 且 $n(S_{11}-S_{12}S_{22}^{-1}S_{21})=E$、$nS_{11}=E+H$，我们得到

$$\Lambda=\frac{\det S}{\det S_{11}\det S_{22}}=\frac{\det E}{\det(E+H)},$$

即为 Wilks 的 $\Lambda(p,n-r,r)$。应用 Bartlett分解于 $E+H$并进行变换可得，$\Lambda\overset{d}{=}\prod_{i=1}^{p}\beta_i$具有独立的 $\beta_i\sim$贝塔分布$((n-r-i+1)/2,r/2)$（Anderson，2003，第 8 章）。$\hat{}\ I=-\frac{1}{2}\log\Lambda$的矩可由对数贝塔矩得到。再次化为望远镜形式可得，$\frac{1}{2}\sum_{i\le p}[\psi((n-i+1)/2)-\psi((n-r-i+1)/2)]=B_n(p,r)$。

–

比例极限

令 $p/n\to c_1$, $r/n\to c_2$, $c_1+c_2<1$。由望远镜形式以及 $\psi(x)-\psi(y)=\log(x/y)+O(1/\min(x,y))$，

$$\frac{2B_n(p,r)}{n}=\frac{1}{n}\sum_{i=1}^{r}\log\frac{n-i+1}{n-p-i+1}+O\Big(\frac{1}{n}\Big)\to\int_0^{c_2}\Big[\log(1-u)-\log(1-c_1-u)\Big]\mathrm{d}u.$$

设 $\int_0^{c_2}\log(1-u)\,\mathrm{d}u=-(1-c_2)\log(1-c_2)-c_2$ 且 $\int_0^{c_2}\log(1-c_1-u)\,\mathrm{d}u=(1-c_1)\log(1-c_1)-(1-c_1-c_2)\log(1-c_1-c_2)-c_2$，则极限为 $2\kappa(c_1,c_2)$。同理，$\psi'(x)=1/x+O(x^{-2})$ 给出

$$V_n(p,r)=\frac{1}{2}\sum_{i=1}^{p}\Big[\frac{1}{n-r-i+1}-\frac{1}{n-i+1}\Big]+O\Big(\frac{1}{n}\Big)\to\frac{1}{2}\int_0^{c_1}\Big[\frac{1}{1-c_2-u}-\frac{1}{1-u}\Big]\mathrm{d}u=\nu(c_1,c_2).$$

在独立性条件下，$\hat{}\ I-B_n$是 p个独立中心化变量 $-\frac{1}{2}$（$\log\beta_i-\mathbb{E}\log\beta_i$）之和。由于$\beta_i$ 的两个形状参数均为 n 阶，且 $\log\beta_i$ 近似服从方差为$\psi'(a)-\psi'(a+b)\asymp 1/n$的正态分布，因此每个变量的方差为 $O(1/n)$，绝对三阶矩为 $O(n^{-3/2})$，且关于 i一致。李雅普诺夫条件以比值 $O(pn^{-3/2})/O(1)^{3/2}=O(n^{-1/2})$ 成立，从而得到中心极限定理。对于一般的 Σ，根据 (10) 和 (11)，$\hat{}\ I$中的三个对数行列式各自的方差均为 $O(1)$，因此由柯西–施瓦茨不等式可得 $\mathrm{Var}\,\hat{}I=O(1)$。于是，$(\hat{}\ I-\mathbb{E}\,\hat{}I)/n\to 0$在 L^2中，且 $\mathbb{E}\hat{}I/n=I/n+B_n/n\to I/n+\kappa$。

Wachter定律的对数矩

在独立性条件下，上一步得到 $\hat{}\ I/n\to\kappa(c_1,\ c_2)$ 依概率成立。另一方面，$\hat{}\ I/n=-\frac{p}{2n}\int\log(1-x)\,\hat{}\ F_p(\mathrm{d}x)$。根据 Wachter (1980年)，$\hat{}\ F_p$ 几乎处处弱收敛

到具有密度（6）的分布；当 $c_1 \le c_2$ 且 $c_1 + c_2 < 1$时，该分布没有原子。最大样本典型相关系数平方几乎处处收敛于 ω_+ (Johnstone，2008；Bao 等人，2019年），并且 $\omega_+ < 1$ 因为

$$1-\omega_+ = (1-c_1)(1-c_2)+c_1c_2-2\sqrt{c_1c_2(1-c_1)(1-c_2)} = \left(\sqrt{(1-c_1)(1-c_2)}-\sqrt{c_1c_2}\right)^2 > 0,$$

其中，最后一个不等式是严格的，因为 $(1-c_1)(1-c_2) - c_1c_2 = 1-c_1-c_2 > 0$。因此，以趋近于 1 的概率，所有 $\hat{\rho}_k^2$ 均属于 $[0,(1+\omega_+)/2]$，在该集合上，$\log(1-x)$ 有界且连续。根据弱收敛，有 $\int \log(1-x)\,\hat{F}_p(\mathrm{d}x) \to \int \log(1-x)\, w_{c_1,\,c_2}(x)\,\mathrm{d}x$。令 $\hat{I}/n$的两个极限相等，即得该定理。

信息成本与投资组合比率

对于反向散度，$\mathbb{E}\,\mathrm{tr}(\Sigma^{-1}S) = N$以及 $\mathbb{E}\log\det(\Sigma^{-1}S) = \ell_n(N)$。对于正向散度，根据 $n > N+1$ (Muirhead, 1982, Ch. 3)，有 $\mathbb{E}S^{-1} = n\Sigma^{-1}/(n-N-1)$，因此 $\mathbb{E}\mathrm{tr}(S^{-1}\Sigma) = Nn/(n-N-1)$，且 $\mathbb{E}\log\det(S^{-1}\Sigma) = -\ell_n(N)$。根据 $\ell_n(N)/N \to \varphi(q)$ 和 $n/(n-N-1) \to (1-q)^{-1}$，可得这些极限。

对于最小方差投资组合，白化将问题简化为 $\Sigma = I$ with$a = \Sigma^{-1/2}\mathbf{1}$，且两个比率都不随 a 的尺度和旋转而改变，因此我们可以取$a = e_1$and $W \sim \mathsf{W}_N(n, I)$，$S = W/n$。于是 $R_{\mathrm{in}}/R = 1/(n\,e_1^\top W^{-1}e_1)$ and $1/e_1^\top W^{-1}e_1 = w_{11} - w_{12}^\top W_{22}^{-1}w_{21} \sim \chi^2_{n-N+1}$ (Muirhead, 1982, Ch. 3)，从而得到 $\mathbb{E}R_{\mathrm{in}}/R = (n-N+1)/n$。对于样本外比率，$R_{\mathrm{out}}/R = e_1^\top W^{-2}e_1/(e_1^\top W^{-1}e_1)^2$。由 $V = W^{-1}$，分块逆矩阵公式给出 $V_{21}/V_{11} = -W_{22}^{-1}w_{21}$，因此比率等于 $1 + \|W_{22}^{-1}w_{21}\|^2$。给定 W_{22},$w_{21} \sim \mathsf{N}(0, W_{22})$，因此 $\mathbb{E}[\|W_{22}^{-1}w_{21}\|^2 \mid W_{22}] = \mathrm{tr}\ W_{22}^{-1}$。由于 $W_{22} \sim \mathsf{W}_{N-1}(n, I)$，$\mathbb{E}\ \mathrm{tr}\ W_{22}^{-1} = (N-1)/(n-N)$，从而得到 $\mathbb{E}R_{\mathrm{out}}/R = 1 + (N-1)/(n-N) = (n-1)/(n-N)$。

信息价值

设 $m = \mathbb{E}[v \mid s]$ 和 $\Sigma_{v|s} = \mathsf{Var}(v \mid s)$。给定 s，头寸为 x 时的终值财富服从均值为 $w_0 + x^\top(m-p)$、方差为 $x^\top\Sigma_{v|s}x$的正态分布，因此 $\mathbb{E}[-e^{-\gamma w} \mid s] = -\exp\{-\gamma w_0 - \gamma x^\top(m-p) + \frac{\gamma^2}{2}x^\top\Sigma_{v|s}x\}$。该式在$x^* = (\gamma\Sigma_{v|s})^{-1}(m-p)$处取得最大值，其值为 $-\exp\{-\gamma w_0 - \frac{1}{2}(m-p)^\top\Sigma_{v|s}^{-1}(m-p)\}$。事前，$y = m - p \sim \mathcal{N}(c, B)$，根据全方差定律，$c = \mu - p$且 $B = \mathsf{Var}(m) = \Sigma_v - \Sigma_{v|s}$。对于 A正定且 B半正定的情形，配方可得

$$\mathbb{E}e^{-\frac{1}{2}y^\top Ay} = \det(I+BA)^{-1/2}\exp\Big\{-\tfrac{1}{2}c^\top(A^{-1}+B)^{-1}c\Big\}.$$

为说明这一点，先假设 B 可逆。被积函数的指数为 $-\frac{1}{2}[y^\top Ay + (y-c)^\top B^{-1}(y-c)]$。对其进行高斯积分，得到因子 $\det(B)^{-1/2}\det(A+B^{-1})^{-1/2} = \det(I+BA)^{-1/2}$。根据伍德伯里恒等式，剩余的二次型为 $c^\top[B^{-1} - B^{-1}(A+B^{-1})^{-1}B^{-1}]c = c^\top(A^{-1}+B)^{-1}c$。一般情形可由连续性得出。由于 $A = \Sigma_{v|s}^{-1}$、$I + BA = \Sigma_v\Sigma_{v|s}^{-1}$ 和 $A^{-1} + B = \Sigma_v$，因此

$$U_s = -e^{-\gamma w_0}\det(\Sigma_v\Sigma_{v|s}^{-1})^{-1/2}\exp\Big\{-\tfrac{1}{2}(\mu-p)^\top\Sigma_v^{-1}(\mu-p)\Big\} = U_\varnothing\, e^{-I(v;s)},$$

因为 $U_{\varnothing}$ 是在 $\Sigma_{v|s}=\Sigma_v$ 下的同一表达式。为信号支付 c 会使 U_s 乘以 $e^{\gamma c}$，因此在 $c=I(v;\,s)/\gamma$ 处达到无差异。若将 s 替换为 Ts，且 T 可逆，则 $\Sigma_{v|s}$ 不变，值也不变。

矩阵 Ornstein–Uhlenbeck 信息率

坐标 $\{H_{ij}\}_{i\le j}$ （17）是具有漂移 $-aH_{ij}$和扩散系数 $N^{-1/2}$（非对角）以及 $(2/N)^{1/2}$（对角）的独立标量 Ornstein–Uhlenbeck 过程。对于线性漂移，Girsanov 密度在有限时间区间上是真鞅（Särkkä and Solin，2019年），在漂移 $-aH$ 和 $-bH$的路径律之间、相对于漂移 $-aH$的路径律，其相对熵为

$$\mathrm{KL}=\frac{1}{2}\mathbb{E}_a\int_0^T\Big[\sum_{i<j}\frac{(b-a)^2H_{ij}^2}{1/N}+\sum_i\frac{(b-a)^2H_{ii}^2}{2/N}\Big]\mathrm{d}t=\frac{N}{4}(b-a)^2\int_0^T\mathbb{E}_a\,\mathrm{tr}\,H_t^2\,\mathrm{d}t,$$

因为 tr $H^2=2\sum_{i<j}H_{ij}^2+\sum_i H_{ii}^2$ $\mathbb{E}H_{ij}^2=1/\,2aN$ $i\neq j$ $\mathbb{E}H_{ii}^2=$。在 $1/\,aN$ $\mathbb{E}\quad H^2=\ N+1\ /\ 2a$ `mou_trace_moment` 平稳状态下，( ) 对于 和 ( )，因此 tr ( ) ( )，这由 Lean 定理 加以证明。对于特征值过程，Dyson 系统（18）在正交类中具有不相交解，并且两个状态共享扩散 $\sqrt{2/N}$ 和排斥项。漂移差为 $(b-a)\,\lambda_i$，因此 Girsanov 给出 $\frac{1}{2}\mathbb{E}_a\int\sum_i(b-a)^2\lambda_i^2/(2/N)\,\mathrm{d}t$，由于 $\sum_i\lambda_i^2=\mathrm{tr}\,H^2$，这正是同一个量。Lean 定理 `mou_rate_closed`、`mou_eigen_closed`和 `mou_eigen_sufficient` 证明了相关代数。两个状态的平稳分布在 $N(N+1)/2$ 个独立坐标上为中心高斯分布，方差比为~~b/a~~，从而初始散度为 $\frac{N(N+1)}{4}\,[b/a-1-\log(b/a)]$。

受迫 Burgers 方程

令 $f_z(\lambda)=(z-\lambda)^{-1}$，因此 $g_t=N^{-1}\sum_i f_z(\lambda_i)$其中 $f_{\prime z}=(z-\lambda)^{-2}$且 $f_{\prime\prime z}=2(z-\lambda)^{-3}$。Itô 公式和（18）给出

$$\mathrm{d}g_t=\frac{1}{N}\sum_i f_z'(\lambda_i)\Big[\frac{1}{N}\sum_{j\neq i}\frac{1}{\lambda_i-\lambda_j}-a\lambda_i\Big]\mathrm{d}t+\frac{1}{N^2}\sum_i f_z''(\lambda_i)\,\mathrm{d}t+\mathrm{d}M_t,$$

其中有一个鞅 M，其二次变差为 $O(N^{-2})$。将双重求和对称化并使用

$$\frac{f_z'(\lambda_i)-f_z'(\lambda_j)}{\lambda_i-\lambda_j}=\frac{1}{(z-\lambda_i)(z-\lambda_j)^2}+\frac{1}{(z-\lambda_i)^2(z-\lambda_j)}$$

得到 $N^{-2}\sum_{i\neq j}(z-\lambda_i)^{-2}(z-\lambda_j)^{-1}=-g\,\partial_z g+O(N^{-1})$，因为 $\partial_z g=-N^{-1}\sum_i(z-\lambda_i)^{-2}$。约束项为 $-aN^{-1}\sum_i\lambda_i(z-\lambda_i)^{-2}=a\,\partial_z(zg)$，因为 $\partial_z(zg)=N^{-1}\sum_i[(z-\lambda_i)^{-1}-z(z-\lambda_i)^{-2}]=-N^{-1}\sum_i\lambda_i(z-\lambda_i)^{-2}$。Itô 项为 $O(N^{-1})$。令 $N\to\infty$，即可得到（19）。

附录 B. Lean 形式化

文件 `lean/RMTChecks.lean`位于复现代码库中，包含 28 个定理，仅使用标准发行版即可在 Lean 4.29.0 下通过编译（de Moura 和 Ullrich，2021年）。它不依赖 mathlib，也不包含 `sorry`，也不包含 `native_decide`。该文件使用 `lean RMTChecks.lean`进行检查。形式化的目的有限且明确。Lean 验证了附录 A 中解析论证所涉及的组合与代数步骤，但不验证测度论陈述本身，即 Wishart 矩、Bartlett分解、digamma 渐近性质、向 Wachter定律的弱收敛以及 Girsanov 定理。表 4按组列出了这些定理。

关于虚假信息的核心论断是针对任意数列$c:\ \mathbb{N}\to\mathbb{Q}$证明的，该数列起到 $i\mapsto\psi((n-i)/2)$的作用。唯一的假设是该数列为反单调，因此一旦认可 ψ 的单调性，这些形式化结果就适用于 digamma 项。源文档中的两项关键论断如下。

```
def twiceBias (c : Nat → Rat) (p r : Nat) : Rat :=
  sumTo c p + sumTo c r - sumTo c (p + r)

theorem twiceBias_eq (c : Nat → Rat) (p r : Nat) :
    twiceBias c p r = sumTo (fun i => c i - c (p + i)) r

theorem twiceBias_noise_pos (c : Nat → Rat)
    (anti : ∀ i j, i < j → c j < c i) (p r m : Nat)
    (hp : 0 < p) (hm : 0 < m) :
    twiceBias c p r < twiceBias c p (r + m)
```

信息率结果明确写出了平稳矩阵元方差和扩散系数，从而可以将闭式与 Girsanov 计算中采用的相同记账方式进行核对：

```
def mouRateEntries (N a b : Rat) : Rat :=
  (1 / 2) * (b - a) ^ 2 *
    (N * (N - 1) / 2 * ((1 / (2 * a * N)) / (1 / N))
      + N * ((1 / (a * N)) / (2 / N)))

theorem mou_eigen_sufficient (N a b : Rat) (hN : N ≠ 0) (ha : a ≠ 0) :
    mouRateEigen N a b = mouRateEntries N a b
```

表 4： `lean/RMTChecks.lean`中的定理。“内核计算”表示由`decide +kernel`证明，它通过精确整数算术在 Lean 内核中计算等式两边。“符号计算”表示通过归纳法或交换环与域的 `grind` 决策过程证明该结论对所有变量值均成立。

定理	陈述	方法
矩组合学		
`mp_moments_are_narayana`	递推（3）得到 Narayana 多项式，$k \le 10$	核
`mp_low_moments`	m_1，……，m_4 如文中所引	核
`catalan_recursion_matches`	Catalan 递推等于 $\binom{2n}{n}/(n+1)$，$n \le 12$	核
`narayana_rows_sum_to_catalan`	Narayana 行的和为 Catalan 数，$k \le 10$	核
`semicircle_moments_catalan`	Wigner矩递推给出 Catalan 数	核
虚假信息		
`sumTo_add`, `sumTo_sub`	有限和的拆分与线性性	符号计算
`sumTo_nonneg`, `sumTo_pos`	非负项或正项之和的符号	符号计算
`twiceBias_eq`	（12）的望远镜形式	符号计算
`twiceBias_nonneg`, `twiceBias_pos`	反单调和严格反单调的正性	符号计算
`twiceBias_noise`, `twiceBias_noise_pos`	序列在无关坐标中的精确增量和严格增长格式化坐标	符号计算
`twiceBias_symm`	$(p,\ r)$ 中的对称性	符号计算
相变		
`mp_edges`	边缘的和与积 $(1 \pm \sqrt{q})^2$	符号计算
`bbp_outlier_at_threshold`	$\lambda(1+\sqrt{q}) = (1+\sqrt{q})^2$	符号计算
`bbp_overlap_at_threshold`	重叠分子在阈值处消失	符号计算
`bbp_inverse`	离群值满足反演二次方程	符号计算
`cca_outlier_at_threshold`	$\gamma(t_c) = \omega_+$ 中（9）	符号计算
`cca_outlier_at_one`	$\gamma(1) = 1$	符号计算
动力学、成本和值		
`mou_trace_moment`	$\mathbb{E}\mathrm{tr}\, H^2 = (N+1)/(2a)$ 基于条目方差的符号计算	
`mou_rate_closed`, `mou_eigen_closed`	逐条目速率和特征值速率相等 $N(N+1)(b-a)^2/(8a)$	符号计算
`mou_eigen_sufficient`	两个比率相等	符号计算
`stein_trace`	$Nn/(n-N-1) - N = N(N+1)/(n-N-1)$ 符号计算	
`gmv_ratio`	投资组合比率的乘积形式	符号计算
`cara_scalar`	标量行列式与逆矩阵恒等式	符号计算

附录 C。补充表格

表 5 是常见设计下精确虚假信息的查找表。每个条目以纳特为单位表示为$B_n(p,r)$，括号内为精确的零假设标准差 $V_n(p,r)^{1/2}$。当数据经过去均值处理时， n 应理解为样本量减一；当两个区块均对 k 共同控制变量进行回归时，则应理解为样本量减 $k+1$。校正估计值通过从插件互信息中减去该条目获得，括号内的值给出了零信息单侧检验的尺度。这里有两个显著特征。当两个区块都较小时，虚假信息大致与 pr/n 成正比，正如经典近似所预测的那样。一旦（$p+r$）$/n$ 变得较大，它的增长速度会快得多。当 $n=120$ 且（p,r）$=$ (25, 50)时，它几乎达到 8 纳特。

表 5： 精确虚假信息 $B_n(p,\ r)$（单位：纳特），括号内为零假设标准差 $V_n(p,\ r)^{1/2}$，对应区块维度（$p,\ r$）和自由度 n。短横线表示$n < p+r+2$。

n	(2, 2)	(5, 5)	(5, 20)	(10, 10)	(10, 40)	(20, 20)	(25, 50)
60	0.035 (0.025)	0.230 (0.065)	1.081 (0.154)	1.017 (0.144)	6.711 (0.516)	5.304 (0.385)	–
120	0.017 (0.012)	0.109 (0.031)	0.469 (0.066)	0.457 (0.065)	2.151 (0.153)	2.024 (0.144)	7.945 (0.325)
250	0.008 (0.006)	0.051 (0.014)	0.211 (0.030)	0.209 (0.030)	0.893 (0.063)	0.873 (0.062)	2.965 (0.119)
500	0.004 (0.003)	0.025 (0.007)	0.103 (0.015)	0.102 (0.014)	0.422 (0.030)	0.417 (0.030)	1.354 (0.054)
1000	0.002 (0.001)	0.013 (0.004)	0.051 (0.007)	0.051 (0.007)	0.205 (0.015)	0.204 (0.014)	0.650 (0.026)
2500	0.001 (0.001)	0.005 (0.001)	0.020 (0.003)	0.020 (0.003)	0.081 (0.006)	0.081 (0.006)	0.254 (0.010)

表6 报告了图5左面板背后的数值，表7 则报告了右面板背后的数值。

表6： 独立区块在 $p=r$ 处的插件互信息的零均值和标准差：精确双伽马值、比例极限、经典近似和模拟值（每行抽样1,000次）。

$p=r$	精确 $\mathbb{E}_0\,\hat{I}$	$n\kappa$	$pr/(2n)$	模拟值	精确标准差	模拟标准差
2	0.008	0.008	0.008	0.008	0.006	0.006
5	0.051	0.051	0.050	0.051	0.014	0.014
10	0.209	0.208	0.200	0.208	0.030	0.029
20	0.873	0.871	0.800	0.874	0.062	0.061
30	2.057	2.052	1.800	2.057	0.097	0.093
40	3.842	3.833	3.200	3.845	0.136	0.135
60	9.664	9.638	7.200	9.680	0.230	0.234
80	19.651	19.588	12.800	19.655	0.355	0.358
100	36.534	36.388	20.000	36.534	0.545	0.531

表7： 附加 m 无信息信号后的互信息估计（$n = 250$，五个收益、五个典型相关为0.4的信息信号），对1,000次抽样取平均值。“预测值”为 $I + B_n$（5，$5 + m$）。

m	真实值 I	代入法	预测值	校正值	Ledoit–Wolf	RIE
0	0.436	0.485	0.487	0.434	0.283	0.397
5	0.436	0.536	0.539	0.433	0.192	0.340
10	0.436	0.595	0.593	0.438	0.130	0.297
20	0.436	0.702	0.703	0.435	0.059	0.223
40	0.436	0.937	0.939	0.434	0.016	0.148
80	0.436	1.494	1.490	0.440	0.004	0.114