

量化投资体系中的多因子融合与投资组合构建：理论、算法与实证分析

1. 宏观背景：从单因子挖掘到系统工程的范式转移

金融市场的演进史，在很大程度上是一部关于“定价效率”与“信息套利”博弈的历史。早期的量化投资主要依赖于少数几个经典的风格因子，如威廉·夏普(William Sharpe)在资本资产定价模型(CAPM)中提出的市场Beta，以及随后的Fama-French三因子模型中引入的规模(Size)和价值(Value)因子¹。在这个阶段，因子的发现(Alpha Generation)几乎等同于策略的全部：投资者只需识别出具有长期正期望收益的特征，构建简单的线性加权组合，即可获得显著的超额收益。

然而，随着全球金融数据维度的爆炸式增长和算法交易技术的普及，市场环境发生了根本性的变化。现代量化交易已经进入了“因子动物园”(Factor Zoo)时代，学术界和业界挖掘出了数以百计的潜在Alpha源³。与此同时，因子的有效性周期显著缩短，单一因子的预测能力(Information Coefficient, IC)变得微弱且不稳定。研究表明，自2007年以来，基于经典风格溢价的多因子指数的夏普比率(Sharpe Ratio)已大幅下降，甚至在统计上与零无异⁴。这标志着量化投资的核心竞争力已经从单纯的“因子挖掘”转移到了“精细化因子工程”与“稳健投资组合构建”的系统工程上来。

在现代量化框架中，投资组合构建不再是简单的资产堆叠，而是一个复杂的凸优化问题(Convex Optimization Problem)。它要求在最大化预期收益的同时，严格控制风险暴露、交易成本和市场冲击³。这一过程面临着多重共线性(Multicollinearity)、过拟合(Overfitting)、协方差矩阵估计误差(Estimation Error)以及非线性市场微观结构等多重挑战³。本报告将深入剖析应对这些挑战的理论框架与数学算法，详细阐述从因子处理、动态融合、风险建模到组合优化的全流程解决方案。

2. 因子工程体系：数据清洗、中性化与正交化处理

因子工程是量化交易系统的基石。原始数据(Raw Data)往往包含大量的噪声、异常值以及非目标风险暴露。直接将原始因子输入预测模型或优化器，会导致模型对某些特征的过度反应，甚至将被动承担的风险误判为主动获取的Alpha⁷。

2.1 数据的标准化与去极值处理

金融时间序列数据通常呈现显著的“尖峰厚尾”分布(Leptokurtic Distribution)，极端值的存在会严重扭曲均值和方差的估计，进而影响线性模型的回归系数。因此，在标准化之前，必须进行去极值处理(Winsorization)。

常用的去极值方法包括中位数去极值法(Median Absolute Deviation, MAD)。与基于均值和标准差的 3σ 法相比，MAD法对异常值更具鲁棒性。其计算公式如下：

$$MAD = \text{median}(|x_i - \text{median}(X)|)$$

我们将超过 n 倍 MAD 的数值截断：

$$x_i^{new} = \begin{cases} \text{median}(X) + n \cdot MAD & \text{if } x_i > \text{median}(X) + n \cdot MAD \\ \text{median}(X) - n \cdot MAD & \text{if } x_i < \text{median}(X) - n \cdot MAD \\ x_i & \text{otherwise} \end{cases}$$

处理完极值后，需对因子进行标准化 (Z-Score Standardization)，使其服从均值为0、标准差为1的标准正态分布。这一步对于消除不同因子 (如价格因子与换手率因子) 之间的量纲差异至关重要⁹。

$$z_i = \frac{x_i - \mu}{\sigma}$$

若忽略标准化，波动率较高的因子将在组合构建中占据主导地位，导致隐含权重的失衡⁹。

2.2 风格与行业中性化 (Neutralization)

在多因子体系中，许多因子的表现与特定的市场风格或行业板块高度相关。例如，动量因子 (Momentum) 在科技股牛市中可能表现为单纯的行业Beta，而非真正的选股Alpha。为了剥离这些非特异性风险，必须对因子进行中性化处理⁸。

中性化在数学上等价于提取线性回归的残差。对于目标因子 \mathbf{F} ，我们构建如下回归模型：

$$\mathbf{F} = \beta_0 + \sum_{k=1}^K \beta_{Ind,k} \cdot \mathbf{Ind}_k + \beta_{Size} \cdot \mathbf{Size} + \epsilon$$

其中：

- \mathbf{Ind}_k 是第 k 个行业的哑变量 (Dummy Variable)。
- \mathbf{Size} 是对数市值因子。
- ϵ 是回归残差。

经过中性化后的因子 $\mathbf{F}_{neutral} = \epsilon$ 剔除了行业 and 市值的影响，代表了纯粹的选股能力。Barra 风险模型中的特异性收益 (Specific Return) 正是基于这一逻辑分离出来的¹¹。这一步骤对于防止组合在单一行业或风格上过度暴露至关重要，确保了超额收益来源的多元化。

2.3 因子正交化 (Orthogonalization)

多重共线性 (Multicollinearity) 是多因子模型的大敌。在高度相关的因子集合中，参数估计的方差会显著膨胀，导致预测结果极不稳定⁶。例如，估值因子 (BP) 与反转因子 (Reversal) 在某些市场环

境下可能高度相关。若不处理，优化器可能会对这些相似信息重复下注。

正交化的目标是寻找一组变换，使得新因子之间互不相关(协方差矩阵为对角阵)。

2.3.1 施密特正交化(Gram-Schmidt Process)

施密特正交化是一种序列处理方法，它按照预定的顺序依次剔除前序因子的投影。假设原始因子向量组为 $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k$ ，正交化后的因子 \mathbf{u}_k 计算如下：

$$\mathbf{u}_k = \mathbf{f}_k - \sum_{j=1}^{k-1} \text{proj}_{\mathbf{u}_j}(\mathbf{f}_k)$$

该方法的特点是结果依赖于因子的输入顺序¹²。在量化实践中，通常将风险解释度最高或最稳定的因子(如市值、行业)放在前面，以保留其原始的经济含义，而将后续因子的重叠部分剔除。

2.3.2 对称正交化(Symmetric Orthogonalization)

为了克服施密特正交化的顺序依赖性，并尽可能保留原始因子的信息结构，对称正交化(Lowdin Orthogonalization)被广泛采用。其目标是寻找一组正交基 \mathbf{F}_{orth} ，使其与原始因子矩阵 \mathbf{F} 的距离平方和最小。

数学上，这是通过对因子相关系数矩阵 Σ 进行奇异值分解(SVD)实现的：

$$\mathbf{F}_{orth} = \mathbf{F}\Sigma^{-1/2} = \mathbf{F}\mathbf{U}\mathbf{D}^{-1/2}\mathbf{U}^T$$

其中 $\Sigma = \mathbf{U}\mathbf{D}\mathbf{U}^T$ 。对称正交化在Alpha因子挖掘中尤为重要，它能在去除相关性的同时，最大限度地保持每个因子的信号特征，常用于构建“纯Alpha”(Orthogonal Alpha)组合¹⁴。

3. 多因子融合机制：从线性加权到机器学习

因子融合(Factor Combination)是将处理后的多个因子信号合成为单一预测值(Composite Signal)的过程。这一步直接决定了策略的预测精度。

3.1 线性融合策略

传统的融合方法基于线性假设，即认为各因子对收益的贡献是可加的。

加权方法	数学表达与逻辑	优缺点分析
------	---------	-------

等权重法 (Equal Weighting)	$w_i =$	<p>优点：鲁棒性极强，不易过拟合。在参数估计噪声极大时，往往跑赢复杂模型¹⁶。</p> <p>缺点：忽略了因子预测能力的差异。</p>
IC加权法 (IC Weighting)	$w_i = \frac{IC_i}{\sum IC_j}$	IC_j
IC_IR加权法 (IC_IR Weighting)	$w_i = \frac{IR_i}{\sum IR_i}, \quad IR_i =$	<p>优点：引入了风险调整概念，降低了高波动因子的权重，提升了组合的长期稳定性 (Sharpe Ratio)¹⁸。</p> <p>缺点：对历史窗口的选择敏感。</p>
最大化复合IC (Max Composite IC)	$\mathbf{w}^* = \Sigma_{IC}^{-1} \cdot$	<p>优点：理论上的最优解，不仅考虑预测能力，还利用逆相关矩阵 Σ_{IC}^{-1} 自动剔除冗余信息¹⁹。</p> <p>缺点：Σ_{IC} 往往病态，求逆不稳定，需引入正则化 (Ridge Regression)。</p>

3.2 机器学习驱动的非线性融合

随着市场复杂度的提升，因子之间的关系不再局限于线性。例如，估值因子 (Value) 在小市值 (Small Cap) 股票中的有效性可能显著高于大市值股票，这种“交互效应” (Interaction Effect) 是线性模型难以捕捉的。机器学习 (ML) 算法因此成为现代量化融合的主流³。

3.2.1 树模型 (XGBoost/LightGBM) 的应用

梯度提升决策树 (GBDT) 及其变体 XGBoost 在处理截面数据 (Cross-sectional Data) 时表现卓越。它通过构建大量弱分类器 (决策树) 并进行加权投票，能够自动进行特征选择并捕捉高阶交互作用³。

- 模型构建：
 - 输入 (Features)：数百个标准化、正交化后的因子。
 - 目标 (Label)：未来一期的超额收益率排序 (Rank) 或分类 (Top/Bottom Quantile)。
 - 损失函数：通常采用 Pairwise Ranking Loss 或多分类 Log-loss，直接优化排序能力而非绝对误差²³。

- 优势实证：研究表明，在沪深300指数期货和A股全市场的实证中，XGBoost模型通过引入“最优权重标签融合”(Optimal Weights)，在收益率预测和风险控制上均显著优于传统线性回归模型²²。其通过特征重要性 (Feature Importance) 分析，还能为研究员提供因子的可解释性依据³。

3.2.2 深度学习与序列建模

对于包含显著时序特征的数据 (如量价高频因子)，循环神经网络 (LSTM) 和 Transformer 架构展现出了独特优势。AlphaForge 等框架利用生成式-预测式网络，不仅挖掘因子，还通过深度学习的非线性映射能力生成复合 Alpha，通过动态权重调整机制适应市场风格切换²⁵。

3.2.3 机器学习中的过拟合与对抗

金融数据的信噪比 (Signal-to-Noise Ratio) 极低，ML 模型极易过拟合。在融合过程中必须采取严格的防御措施：

- **Purged K-Fold Cross Validation**：在划分训练集和验证集时，必须剔除两个集合接壤处的数据 (Purging)，以防止由于序列相关性导致的信息泄露²⁶。
- 特征降维：利用 PCA 或自编码器 (Autoencoder) 压缩特征空间，减少噪声输入²⁷。
- 正则化：在目标函数中加入 L1/L2 惩罚项，限制模型复杂度。

4. 风险模型与协方差矩阵的稳健估计

投资组合构建不仅取决于收益预测 (Alpha)，同样取决于风险控制。在均值-方差优化 (MVO) 框架下，协方差矩阵 (Covariance Matrix) 的估计质量直接决定了权重的分布。

4.1 样本协方差的维数灾难

假设投资域有 $N = 3000$ 只股票，协方差矩阵包含 $N(N + 1)/2 \approx 4.5 \times 10^6$ 个参数。

若使用历史样本直接计算 (Sample Covariance)，由于样本量 T 往往远小于 N ，估计矩阵将是奇异的 (Singular) 或充满噪声。优化器会利用这些噪声，将权重分配给那些看起来相关性极低 (实为估计误差) 的资产，导致“误差最大化” (Error Maximization)¹¹。

4.2 结构化因子风险模型 (Barra Model)

Barra 模型通过引入因子结构，将高维资产相关性降维为低维因子相关性，大幅提升了估计的稳定性。其核心假设是：资产收益由少量共同因子 (Common Factors) 和特异性收益 (Specific Return) 组成。

$$\mathbf{r} = \mathbf{Xf} + \mathbf{u}$$

协方差矩阵分解为：

$$\Sigma_{asset} = \mathbf{X}\Sigma_{factor}\mathbf{X}^T + \Delta$$

其中：

- \mathbf{X} 是 $N \times K$ 的因子暴露矩阵(已知)。
- Σ_{factor} 是 $K \times K$ 的因子协方差矩阵(需估计, 但 $K \ll N$)。
- Δ 是 $N \times N$ 的对角矩阵, 代表特异性风险。

这种结构化处理不仅减少了参数数量, 还通过经济学逻辑(行业、风格)解释了风险来源, 消除了虚假相关性¹¹。

4.3 协方差矩阵收缩 (Ledoit-Wolf Shrinkage)

即使使用了因子模型, 因子协方差矩阵 Σ_{factor} 的估计仍可能存在误差。Ledoit和Wolf(2004)提出了一种非参数化的收缩估计方法, 是目前业界的金标准²⁹。

该方法构建一个线性组合, 将样本协方差矩阵 \mathbf{S} 向一个偏差较大但方差极小的目标矩阵 \mathbf{F} (通常为常相关矩阵或单因子模型矩阵) 进行收缩:

$$\Sigma_{shrunk} = \delta^*\mathbf{F} + (1 - \delta^*)\mathbf{S}$$

这里的核心贡献在于推导出了最优收缩强度 δ^* 的解析解 (Analytical Solution), 使得估计误差 (MSE) 最小化。

- 当样本长度 T 较短或资产数 N 较大时, δ^* 自动增大, 更偏向结构化目标 \mathbf{F} 。
- 当数据充足时, δ^* 减小, 更信任样本数据 \mathbf{S} 。

这一方法在数学上保证了协方差矩阵的正定性 (Positive Definiteness) 和良态性 (Well-conditioned), 显著改善了优化结果的样本外表现³⁰。

5. 投资组合优化理论与工程化落地

拥有了Alpha信号和风险模型后, 投资组合构建演变为一个在约束条件下的凸优化问题。

5.1 均值-方差优化 (MVO) 与 Black-Litterman 模型

经典的Markowitz MVO模型试图求解以下目标函数:

$$\max_{\mathbf{w}} \left(\mathbf{w}^T \boldsymbol{\mu} - \frac{\lambda}{2} \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} \right)$$

然而，MVO对输入参数 $\boldsymbol{\mu}$ 极其敏感。为了缓解这一问题，Black-Litterman(BL)模型引入了贝叶斯框架³²。

- 先验分布(Prior): 假设市场是有效的，当前市值加权组合即为最优组合。通过反向优化(Reverse Optimization)推导出隐含均衡收益率 $\boldsymbol{\Pi} = \lambda \boldsymbol{\Sigma} \mathbf{w}_{mkt}$ 。
- 观点(Views): 投资者对部分资产或因子的主观预测。例如，“动量因子收益率将比市场高2%”。这被建模为线性约束 $\mathbf{PE}[\mathbf{r}] = \mathbf{Q} + \boldsymbol{\epsilon}$ 。
- 后验分布(Posterior): 结合先验与观点，计算出新的预期收益向量：

$$E =^{-1}$$

BL模型巧妙地将量化模型(Exotic Beta/Factors)作为“观点”融入市场均衡，产生的投资组合既偏离基准以获取Alpha，又在整体结构上保持了与市场的一致性，避免了极端持仓³²。

5.2 风险平价(Risk Parity)

与MVO不同，风险平价策略不依赖于难以预测的预期收益 $\boldsymbol{\mu}$ ，而是专注于风险分配。其目标是使得每个资产(或因子)对组合总风险的边际贡献(Marginal Risk Contribution, MRC)相等⁵。

数学上，这是求解权重 \mathbf{W} 使得：

$$w_i \cdot \frac{\partial \sigma_p}{\partial w_i} = w_j \cdot \frac{\partial \sigma_p}{\partial w_j} \quad \forall i, j$$

在股票组合中，由于各股票波动率差异巨大，风险平价通常会显著低配高波动的股票，高配低波动的股票。实证研究显示，基于风险平价构建的组合在市场危机时期表现出更强的抗跌性，长期夏普比率往往优于市值加权组合³²。

5.3 约束管理与凸优化求解

实际交易中存在大量物理和合规约束，这些约束将优化问题复杂化³³。

约束类型	数学形式	经济意义与求解难点
权重箱型约束 (Box Constraint)	$L_i \leq w_i \leq$	防止单只股票持仓过重，满足公募基金“双十规定”。属

		于线性约束, 求解容易。
基数约束 (Cardinality)	$\ \mathbf{w}\ _0 <$	限制持仓股票数量。这是非凸 (Non-convex) 约束, 属于 NP-hard 问题。通常需通过 L1 范数正则化 (Lasso) 或混合整数规划 (MIP) 进行近似求解 ³³ 。
换手率约束 (Turnover)	$\ \mathbf{w}_t - \mathbf{w}_{t-1}\ _1 <$	限制调仓量以控制成本。属于凸约束, 但增加了计算复杂度。
风格暴露约束	$\mathbf{X}_{style}^T (\mathbf{w} -$	确保组合相对于基准在行业或风格上是中性的, 保证 Alpha 的纯度 ¹⁰ 。

6. 交易成本、市场冲击与执行算法

在纸面回测中表现优异的策略, 往往在实盘中因忽略交易成本而失效。现代组合构建理论强调将交易成本作为内生变量纳入优化目标³⁷。

6.1 市场冲击模型: 平方根法则

交易成本由显性成本(佣金、税费)和隐性成本(滑点、市场冲击)组成。其中, **市场冲击 (Market Impact)** 是最关键的隐性成本, 它描述了交易行为本身对价格的推移作用。

经典的 Almgren-Chriss 模型及后续研究表明, 市场冲击遵循平方根法则 (**Square Root Law**)³⁹:

$$\text{Impact Cost} \approx \sigma \cdot \sqrt{\frac{Q}{V}}$$

其中:

- σ 是资产的日波动率。
- Q 是交易数量。
- V 是资产的日均成交量 (ADV)。

这意味着, 随着交易规模 Q 的增加, 单位交易成本并非线性增长, 而是呈现边际递减的凹函数形

态(但在总成本上表现为 $Q^{3/2}$ 关系)。

6.2 带有惩罚项的组合优化

为了内化这些成本, 优化目标函数应修正为:

$$\max_{\mathbf{w}} \left(\mathbf{w}^T \boldsymbol{\mu} - \lambda \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} - \sum_{i=1}^N \text{TC}_i(|\Delta w_i|) \right)$$

其中 TC_i 是交易成本函数。

- 对于线性成本(佣金), $\text{TC}_i \propto |\Delta w_i|$ 。这相当于L1正则化, 会导致稀疏的交易向量(只调整部分股票)。
- 对于非线性冲击成本, $\text{TC}_i \propto |\Delta w_i|^{3/2}$ 或 Δw_i^2 。

引入非线性成本惩罚项后, 优化器会变得“懒惰”: 只有当Alpha收益足以覆盖边际交易成本时, 才会进行换仓。这种机制能自动平滑持仓路径, 显著降低换手率, 虽然理论上的Alpha捕获率下降, 但实盘的净费后收益(Net Return)通常会提升⁴¹。

7. 绩效归因与策略迭代

投资组合构建的闭环在于基于实盘结果的归因分析。这不仅是对过去的总结, 更是模型迭代的依据。

7.1 Brinson归因与因子归因的比较

- **Brinson归因**: 传统的归因方法, 将超额收益分解为“资产配置效应”(Allocation Effect, 如超配了表现好的行业)和“个股选择效应”(Selection Effect)。这种方法简单直观, 但无法解释通过风格暴露(如长期持有小盘股)带来的收益⁴⁴。
- **因子归因(Risk-based Attribution)**: 基于多因子模型的归因体系。它将组合收益分解为:

$$R_p = \sum_k \beta_k \cdot R_{factor,k} + \alpha_{specific}$$

这种方法能清晰地指出: 策略赚钱是因为真正的选股能力(Specific Alpha), 还是仅仅因为暴露了高Beta或动量因子。对于量化基金, 因子归因是识别“伪Alpha”和监控风格漂移(Style Drift)的核心工具⁴⁴。

7.2 因子衰减与Regime Switching

所有因子都会经历从发现、拥挤(Crowding)到失效(Decay)的过程⁴。

- **拥挤度监测**: 通过分析市场上同类策略的资金流向、估值价差(Valuation Spread)和相关性

变化, 预警因子拥挤风险。

- 状态转换(Regime Switching): 市场在不同宏观状态(如高波/低波, 通胀/通缩)下, 因子的表现截然不同。现代模型采用隐马尔可夫模型(HMM)或在线学习算法(Online Learning), 根据最新的市场状态动态调整因子权重, 及时剔除衰减因子⁴⁷。

8. 结论

量化交易中的投资组合构建与多因子融合, 已不再是简单的统计套利, 而是一项集成了金融经济学、统计学习、凸优化与微观结构理论的复杂系统工程。

从因子层面看, 非线性机器学习方法(XGBoost, Deep Learning)结合严格的正交化处理, 正在取代简单的线性模型, 成为挖掘深层Alpha的主力; 从组合层面看, 基于收缩估计的协方差矩阵、Black-Litterman观点融合以及内嵌非线性交易成本的优化算法, 是保证策略从“回测完美”走向“实盘盈利”的关键。

未来的演进方向将是端到端的深度强化学习——不再分割预测与优化, 而是直接以最终的风险调整后净收益为奖励函数, 让AI在模拟环境中自我博弈, 自动学会在Alpha捕捉、风险控制和成本管理之间找到动态最优解。这标志着量化投资正从“工匠式的因子挖掘”迈向“工业化的人工智能决策”新时代。

Works cited

1. Quant's Guide to Factor Investing: Theory, Practice, and Code | by Jakub Polec | Medium, accessed January 27, 2026, https://medium.com/@jpolec_72972/quants-guide-to-factor-investing-theory-practice-and-code-09ce1c06c3e8
2. AlphaForge: A Framework to Mine and Dynamically Combine Formulaic Alpha Factors - arXiv, accessed January 27, 2026, <https://arxiv.org/html/2406.18394v2>
3. Machine Learning Enhanced Multi-Factor Quantitative Trading: A Cross-Sectional Portfolio Optimization Approach with Bias Correction - arXiv, accessed January 27, 2026, <https://arxiv.org/html/2507.07107>
4. The Factor Mirage: How Quant Models Go Wrong - CFA Institute Enterprising Investor, accessed January 27, 2026, <https://blogs.cfainstitute.org/investor/2025/10/30/the-factor-mirage-how-quant-models-go-wrong/>
5. Portfolio Optimization - MATLAB & Simulink - MathWorks, accessed January 27, 2026, <https://www.mathworks.com/discovery/portfolio-optimization.html>
6. Dealing with Multicollinearity in Factor Analysis: The Problem, Detections, and Solutions, accessed January 27, 2026, <https://www.scirp.org/journal/paperinformation?paperid=125846>
7. quant trading strategies - Alpha generation and factor models, accessed January 27, 2026, <https://quant.stackexchange.com/questions/49634/alpha-generation-and-factor->

models

8. Market Factor: Not Just For Show - MSCI, accessed January 27, 2026, <https://www.msci.com/documents/10199/ed6e42a3-c1fa-4430-89ba-efd5a5b52558>
9. Barra model: why standardize the fundamental risk factors?, accessed January 27, 2026, <https://quant.stackexchange.com/questions/46824/barra-model-why-standardize-the-fundamental-risk-factors>
10. Target Exposure: Investment applications and solutions | LSEG, accessed January 27, 2026, https://www.lseg.com/content/dam/ftse-russell/en_us/documents/research/target-exposure-investment-applications-solutions.pdf
11. Barra Risk Model Handbook, accessed January 27, 2026, https://roycheng.cn/files/riskModels/barra_risk_model_handbook.pdf?ref=quany_macro.com
12. Regression 1: eliminating multicollinearity from the Hald data, accessed January 27, 2026, <https://rip94550.wordpress.com/2011/11/14/regression-1-eliminating-multicollinearity-from-the-hald-data/>
13. Gram-Schmidt Orthogonalization and Regression • matlib - Michael Friendly, accessed January 27, 2026, <https://friendly.github.io/matlib/articles/gramreg.html>
14. Synergistic Formulaic Alpha Generation for Quantitative Trading based on Reinforcement Learning - arXiv, accessed January 27, 2026, <https://arxiv.org/html/2401.02710v2>
15. Orthogonal sources of risk and return - Quantitative Finance Stack Exchange, accessed January 27, 2026, <https://quant.stackexchange.com/questions/30945/orthogonal-sources-of-risk-and-return>
16. Portfolio Constraints: An Empirical Analysis - MDPI, accessed January 27, 2026, <https://www.mdpi.com/2227-7072/10/1/9>
17. 【Factor analysis】 Vol. 4. Factor score - Michael's blog, accessed January 27, 2026, <https://mikelhsia.github.io/2021/02/27/2021-02-27-factor-analysis-4/>
18. Analyzing Factor Performance with Alphas: The Value Factors Edition - TEJ, accessed January 27, 2026, <https://www.tejwin.com/en/insight/analyzing-factor-performance-with-alphas-the-value-factors-edition/>
19. Quantitative Trading: How to Build Your Own Algorithmic Trading Business (Wiley Trading), accessed January 27, 2026, <https://www.myquant.cn/uploads/default/original/1X/4c7037365a4bf1623734c1c899baed7855061ace.pdf>
20. Portfolio Weights to Maximize Information Ratio (Finding Alphas), accessed January 27, 2026, <https://quant.stackexchange.com/questions/35076/portfolio-weights-to-maximize-information-ratio-finding-alphas>
21. Machine Learning in Multi-Factor Investing: A Deep Dive | by Leo Mercanti |

- Medium, accessed January 27, 2026,
<https://leomercanti.medium.com/machine-learning-in-multi-factor-investing-a-deep-dive-9ec8a64690bb>
22. Predicting Chinese stock market using XGBoost multi-objective optimization with optimal weighting - PMC, accessed January 27, 2026,
<https://pmc.ncbi.nlm.nih.gov/articles/PMC10936758/>
 23. A Sustainable Quantitative Stock Selection Strategy Based on Dynamic Factor Adjustment, accessed January 27, 2026,
<https://www.mdpi.com/2071-1050/12/10/3978>
 24. Research on Quantitative Investment Strategy of Stock Index Futures Based on XGBoost Model - SciTePress, accessed January 27, 2026,
<https://www.scitepress.org/Papers/2022/118257/118257.pdf>
 25. AlphaForge: A Framework to Mine and Dynamically Combine Formulaic Alpha Factors - arXiv, accessed January 27, 2026, <https://arxiv.org/html/2406.18394v1>
 26. Top 6 Challenges in Quantitative Trading and How to Overcome Them for Long-Term Success | by Quant Hacker | Medium, accessed January 27, 2026,
<https://medium.com/@english111026/top-6-challenges-in-quantitative-trading-and-how-to-overcome-them-for-long-term-success-fc974ff58986>
 27. Developing High-Frequency Equities Trading Models - DSpace@MIT, accessed January 27, 2026,
<https://dspace.mit.edu/bitstream/handle/1721.1/59122/658860705-MIT.pdf>
 28. Covariance matrix shrinkage: Ledoit and Wolf (2004) - YouTube, accessed January 27, 2026, <https://www.youtube.com/watch?v=heXLyuCTP1Y>
 29. Problems in mean-variance optimization. - Olivier Ledoit and Michael Wolf - | Department of Economics | UZH, accessed January 27, 2026,
https://econ.uzh.ch/dam/jcr:8a18d37f-3238-4c14-a276-66392e82961b/jpm_2004..pdf
 30. The Power of (Non-)Linear Shrinking: A Review and Guide to Covariance Matrix Estimation - Olivier Ledoit, accessed January 27, 2026,
http://www.ledoit.net/Review_Paper_2020_JFEc.pdf
 31. Shrinkage of the Sample Correlation Matrix of Returns Towards a Constant Correlation Target: A Pedagogic Illustration Based on D - Spreadsheets in Education, accessed January 27, 2026,
<https://sie.scholasticahq.com/article/4665-shrinkage-of-the-sample-correlation-matrix-of-returns-towards-a-constant-correlation-target-a-pedagogic-illustration-based-on-dow-jones-stock-returns/attachment/13957.pdf>
 32. Black-Litterman, Exotic Beta, and Varying Efficient ... - CME Group, accessed January 27, 2026,
<https://www.cmegroup.com/education/files/black-litterman-exotic-betas-risk-parity-manuscript.pdf>
 33. 6.2 Portfolio Constraints | Portfolio Optimization - Bookdown, accessed January 27, 2026,
<https://bookdown.org/palomar/portfoliooptimizationbook/6.2-portfolio-constraints.html>
 34. Optimization Tutorial - Defining Constraints - Frontline Solvers, accessed January

- 27, 2026, <https://www.solver.com/defining-constraints>
35. Constrained optimization - Wikipedia, accessed January 27, 2026, https://en.wikipedia.org/wiki/Constrained_optimization
 36. Penalty method for the sparse portfolio optimization problem, accessed January 27, 2026, <https://www.aimsociences.org/article/doi/10.3934/jimo.2024031>
 37. 6 Transaction costs — MOSEK Portfolio Optimization Cookbook 1.6.0, accessed January 27, 2026, <https://docs.mosek.com/portfolio-cookbook/transaction.html>
 38. Building and Evaluating a Transaction Cost Model: A Primer, accessed January 27, 2026, <https://www.pm-research.com/content/ijitrade/5/2/57>
 39. Three models of market impact - Baruch MFE Program, accessed January 27, 2026, <https://mfe.baruch.cuny.edu/wp-content/uploads/2012/09/Chicago2016OptimalExecution.pdf>
 40. 120-Crossover-from-linear-to-square-root-market-impact.pdf - Capital Fund Management, accessed January 27, 2026, <https://www.cfm.com/wp-content/uploads/2022/12/120-Crossover-from-linear-to-square-root-market-impact.pdf>
 41. Transaction costs and portfolio strategies - Macrosynergy, accessed January 27, 2026, <https://macrosynergy.com/research/transaction-costs-and-portfolio-strategies/>
 42. Multiperiod Portfolio Optimization with General Transaction Costs, accessed January 27, 2026, <https://optimization-online.org/wp-content/uploads/2013/07/3962.pdf>
 43. Portfolio Optimization with Transaction Costs and Preconceived Portfolio Weights Jeremy D. Myers - DSpace@MIT, accessed January 27, 2026, <https://dspace.mit.edu/bitstream/handle/1721.1/61296/702644615-MIT.pdf?sequence=2>
 44. Risk-based or Brinson attribution - SimCorp, accessed January 27, 2026, <https://www.simcorp.com/resources/insights/industry-articles/2024/Risk-based-or-Brinson-attribution>
 45. Multi-Period Performance Attribution: Framework for an Allocation Effect Taking Active Weight Drift into Account - Ortec Finance, accessed January 27, 2026, https://www.ortecfinance.com/-/media/project/ortec/shared/files/whitepapers/ip-multi-period-performance-attribution-ortec-finance_bas-leerink.pdf
 46. How a Multi-Factor Attribution Framework Can Provide a Deeper Insight Into the Sources of Relative Performance, accessed January 27, 2026, <https://insight.factset.com/how-a-multi-factor-attribution-framework-can-provide-a-deeper-insight-into-the-sources-of-relative-performance>
 47. Learning about Regime Change - Federal Reserve Bank of San Francisco, accessed January 27, 2026, <https://www.frbsf.org/wp-content/uploads/wp2020-15.pdf>
 48. Regime Changes and Financial Markets - Rady School of Management, accessed January 27, 2026, https://rady.ucsd.edu/_files/faculty-research/timmermann/regime_changes_June_22.pdf