

Integrated Proposal Idea

Title

Integrated High-Dimensional Methods for Portfolio Optimization: Combining POET, Spatial Autoregression, and LASSO Regression

1. Introduction

The challenge of high-dimensional data has become central in modern financial econometrics. With tens to hundreds of assets, macroeconomic indicators, and cross-market interactions, traditional mean–variance portfolio theory (Markowitz, 1952) often fails due to unstable covariance estimation, omitted network dependencies, and the curse of dimensionality in regression models.

Three strands of methodology have emerged to address these challenges:

1. **High-dimensional covariance estimation (POET; Fan, Liao & Mincheva, 2013):**

Factor decomposition plus sparse residual covariance to stabilize risk estimation.

2. **Spatial autoregression (SAR; Cliff & Ord, 1981; Baltagi et al., 2014):**

Capturing cross-sectional and network dependence structures among assets or markets.

3. **Regularized regression (LASSO; Tibshirani, 1996):**

Sparse selection of predictors from a high-dimensional set of macroeconomic and financial variables.

While these methods have been studied separately, we propose a unified framework that integrates **LASSO for factor selection, POET for covariance estimation, and Spatial AR for network dependence**. This integration allows for robust portfolio construction that accounts for latent and observable factors, cross-asset contagion, and parsimonious predictive structures.

2. Methodological Framework

2.1. Step 1: Macro-Factor Selection via LASSO

We consider a large set of potential macroeconomic and market predictors (e.g., interest rates, exchange rates, volatility indices, credit spreads, commodity benchmarks). To avoid overfitting and spurious correlations, we apply LASSO regression:

$$\hat{\beta} = \arg \min_{\beta} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1$$

where y is the asset return or factor proxy, and X is the predictor matrix. LASSO selects a sparse subset of variables, forming the observed macro-finance factor block f_t^{Macro} .

2.2. Step 2: Covariance Estimation via POET

The p -dimensional excess returns R_t are modeled as:

$$R_t = Bf_t + u_t$$

where $f_t = (f_t^{Macro}, f_t^{Latent})$ includes both selected macro factors and latent statistical factors. The covariance decomposes as:

$$\Sigma = B\Sigma_f B^\top + \Sigma_u$$

We estimate this via the **POET estimator**:

$$\hat{\Sigma}_{POET} = \hat{\Sigma}_{Factor} + \mathcal{T}_\omega(\hat{\Sigma}_u)$$

where $\hat{\Sigma}_{Factor}$ is obtained by PCA on residual returns (after macro factors) and \mathcal{T}_ω applies adaptive thresholding to enforce sparsity in $\hat{\Sigma}_u$.

2.3. Step 3: Residual Network Dependence via Spatial AR

While POET assumes sparsity in the idiosyncratic covariance Σ_u , empirical evidence suggests that **idiosyncratic shocks may propagate via network spillovers** (e.g., financial contagion, sectoral interdependence). We model this via a **Spatial AR** process:

$$u_t = \rho W u_t + \varepsilon_t$$

where W is a pre-specified or estimated spatial weight matrix (e.g., correlation-based, geographic, or sectoral linkage). This step allows us to go beyond unstructured sparsity and capture structured dependencies among residuals.

2.4. Step 4: Portfolio Optimization

With the integrated covariance estimator $\hat{\Sigma}_{Integrated}$ combining POET and Spatial AR adjustments, and with expected returns $\hat{\mu}$ informed by LASSO-selected macro predictors, we construct portfolios:

- **Global Minimum Variance (GMV):**

$$w_{GMV} = \frac{\hat{\Sigma}^{-1} \mathbf{1}}{\mathbf{1}^\top \hat{\Sigma}^{-1} \mathbf{1}}$$

- **Mean-Variance Optimal Portfolio:**

$$w_{MV} = \frac{\hat{\Sigma}^{-1} \hat{\mu}}{\mathbf{1}^\top \hat{\Sigma}^{-1} \hat{\mu}}$$

- **Risk Budgeting Portfolio:**

Equalizing marginal contributions to risk subject to ℓ_1 gross-exposure constraints.

3. Data and Implementation

We employ a cross-asset dataset (2015.01.01–2024.12.31) spanning approximately 70–80 series:

- **Equities:** Global indices (S&P 500, MSCI World, Nikkei, KOSPI, DAX).
- **Fixed Income:** U.S. Treasury yields, corporate bond spreads.
- **Commodities:** Gold, oil, industrial metals, agricultural futures.
- **Currencies:** USD, JPY, EUR, CNY, KRW.
- **Cryptocurrencies:** Bitcoin, Ethereum, XRP.
- **Volatility Measures:** VIX.

Steps include frequency harmonization, return computation, stationarity testing, and standardization. LASSO-selected macro factors are combined with PCA-based latent factors in POET. Residuals are modeled with Spatial AR using correlation-based network matrices.

4. Expected Contributions

1. Methodological Integration

- First to integrate **POET + Spatial AR + LASSO** into a unified framework.
- Provides joint treatment of **factor structure, network dependence, and predictor selection** in high-dimensional finance.

2. Financial Implications

- **Portfolio stability:** More stable allocations under gross-exposure constraints.
- **Tail-risk management:** Improved VaR and ES performance during crises.
- **Cross-asset diversification:** Capturing contagion and cross-market linkages for multi-asset allocation.

3. Pedagogical Value

- Directly connects to high-dimensional regression, factor models, sparsity, and network econometrics as taught in Financial Big Data Analysis courses.
-

5. References

- Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1), 191–221.
- Bickel, P. J., & Levina, E. (2008). Covariance regularization by thresholding. *Annals of Statistics*, 36(6), 2577–2604.
- Brodie, J., Daubechies, I., De Mol, C., Giannone, D., & Loris, I. (2009). Sparse and stable Markowitz portfolios. *PNAS*, 106(30), 12267–12272.
- Cliff, A., & Ord, J. (1981). *Spatial Processes: Models & Applications*. Pion.
- DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *RFS*, 22(5), 1915–1953.
- Fan, J., Liao, Y., & Mincheva, M. (2013). Large covariance estimation by thresholding principal orthogonal complements. *JRSSB*, 75(4), 603–680.
- Ledoit, O., & Wolf, M. (2004). Honey, I shrunk the sample covariance matrix. *J. Multivariate Analysis*, 91(1), 1–18.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. Royal Statistical Society: Series B*, 58(1), 267–288.