# High-dimensional Expected Utility Portfolios under the Spiked Covariance Model

## 3. Simulation

In this section, we will conduct a simulation study to assess the finite sample properties of the suggested I-OEUP and OP-OEUP, and compare its behavior with existing approaches.

### 3.1. Comparison of limits and empirical values

This section employs the Monte Carlo simulation method to numerically evaluate the performance of two parameterized portfolios: the Out-of-Sample Expected Utility Portfolio (P-OEUP) and the Global Minimum Variance Portfolio (P-GMVP). The simulation procedure consists of six key phases: (1) Initialization of the covariance structure $\mathbf{\Sigma}$ and random parameter $\mathbf{\Theta}$; (2) Estimation of portfolio weights using training samples; (3) Evaluation of out-of-sample returns and risk on a large test set; (4) Averaging over $N = 500$ independent trials to obtain empirical (EMP) values; (5) Calculation of theoretical limiting (LIM) values; and (6) Analysis of asymptotic behavior by proportionally increasing the portfolio dimension $p$ and sample size $n$. This structured workflow facilitates a systematic comparison between empirical performance and theoretical predictions, with the simulation results summarized in Table 1. The detailed steps are outlined below.

- **Step 1:** Set $\sigma^2 = 1$, $k_1 = 3$, and $k_2 = 1$. Construct mutually orthogonal vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_p$ with corresponding weights $\lambda_1 = 20$, $\lambda_2 = 10$, $\lambda_3 = 5$, and $\lambda_p = -0.99$, where $p/n = c$. Obtain the covariance matrix $\mathbf{\Sigma}$ from Eq. (6), and define the overall mean vector as $\boldsymbol{\mu} = \mathbf{1}_p/\sqrt{p}$.
- **Step 2:** Generate $n$ training samples and compute the classical sample covariance and the corresponding eigenvalues and eigenvectors.
- **Step 3:** For the parameter $\Theta = \left(\theta_1, \theta_2, \theta_3, \theta_p\right) \in \mathbf{\Theta}$ in Eq. (11), draw its values randomly from a uniform distribution over $[-0.999, 100]$. Derive the parameterized inverse of the sample covariance matrix $\widehat{\mathbf{C}}^{-1}(\Theta)$ as in Eq. (7). Compute the portfolio weight vectors $\widehat{\mathbf{w}}_{\text{P-OEUP}}(\Theta)$ and $\widehat{\mathbf{w}}_{\text{P-GMVP}}(\Theta)$ according to Eq. (9).
- **Step 4:** Generate $m = 100$ test samples and evaluate the out-of-sample return mean and risk for $\widehat{\mathbf{w}}_{\text{P-OEUP}}(\Theta)$ and $\widehat{\mathbf{w}}_{\text{P-GMVP}}(\Theta)$. Compute the out-of-sample $\text{EU}_{\text{P-OEUP}}(\Theta)$ and $\text{Risk}_{\text{P-GMVP}}(\Theta)$.
- **Step 5:** Repeat Steps 2–4 $N = 500$ times. Compute the average values of $\text{EU}_{\text{P-OEUP}}(\Theta)$ and $\text{Risk}_{\text{P-GMVP}}(\Theta)$, which are reported as "EU" and "Risk", respectively, in Table 1.
- **Step 6:** Calculate the theoretical values of $\overline{\text{EU}}_{\text{P-OEUP}}(\Theta)$ and $\overline{\text{Risk}}_{\text{P-GMVP}}(\Theta)$, which are reported as "$\overline{\text{EU}}$" and "$\overline{\text{Risk}}$", respectively, in Table 1.
- **Step 7:** Increase both $p$ and $n$ proportionally, incrementing $p$ by 50 each time, and repeat Steps 2–6.

### 3.2. Synthetic Data Simulation

This paper uses Monte Carlo simulation to evaluate the effectiveness of these estimation methods in portfolio selection. The specific steps are as follows:

- **Step 1** and **2:** Same as Step 1 and 2 in Section 3.1. Compute the expected utility of $\mathbf{w}_{\text{OEUP}}$ in (4) and the risk of $\mathbf{w}_{\text{GMVP}}$ in (5).
- **Step 3:** Optimize using the gradient descent method with initial values $\Theta = \Theta_0$, where $\Theta_0$ is given in (10). Additionally, perform the following operations respectively:
  - Maximize $\overline{\text{EU}}_{\text{P}-\text{OEUP}}(\Theta)$ in (21), to obtain a corresponding set of optimal parameters $\Theta^*_{\text{OEUP}}$;

- Minimize $p\overline{\text{Risk}}_{\text{P-GMVP}}(\Theta)$ in (22), to obtain a corresponding set of optimal parameters $\Theta^*_{\text{GMVP}}$.

- **Step 4:** Compute $\hat{\mathbf{w}}_{\text{I-OEUP}}(\hat{\mathbf{w}}_{\text{I-GMVP}})$ and $\hat{\mathbf{w}}_{\text{OI-OEUP}}(\hat{\mathbf{w}}_{\text{OI-GMVP}})$ by (10) and (13).

- **Step 5:** Compute the classical portfolio $\hat{\mathbf{w}}_{\text{SDE-OEUP}}$ and $\hat{\mathbf{w}}_{\text{SDE-GMVP}}$ by plugging the sample mean and the Sample Covariance Matrix Estimator in (4) and (5). Meanwhile, for the other estimation methods:
  - For LSHE, estimate the population covariance matrix and derive the corresponding portfolio weights—$\hat{\mathbf{w}}_{\text{LSHE-OEUP}}$ and $\hat{\mathbf{w}}_{\text{LSHE-GMVP}}$;
  - For NLSH, estimate the population covariance matrix and derive the corresponding portfolio weights—$\hat{\mathbf{w}}_{\text{NLSH-OEUP}}$ and $\hat{\mathbf{w}}_{\text{NLSH-GMVP}}$.

- **Step 6:** Generate $m = 10000$ test samples and evaluate the out-of-sample return mean and risk for all portfolio estimations $\widehat{\mathbf{w}}$s.

- **Step 7:** Repeat steps 2-6 500 times to calculate the average out-of-sample expected utility and risk of the portfolio for each of the five methods.