$\mathrm{tnreason}$ FOR PORTFOLIO OPTIMIZATION

RESEARCH NOTES IN THE ENEXA PROJECT

February 14, 2025

We here investigate the portfolio optimization problem described in the tutorial https://qiskit-community.github.io/qiskit-finance/tutorials/01_portfolio_optimization.html.

1 The portfolio optimization problem

Given n products and the covariance matrices $\Sigma[L_0, L_1] \in \mathbb{R}^{d\times d}$, the expectations $\mu[L_0]$, the risk appetite parameter $q > 0$ and the budget $B \in \mathbb{N}$, we state the portfolio optimization problem

$$\mathrm{argmin}_{x_{[d]}} q \left(\sum_{l_0,l_1\in[n],\, l_0\neq l_1} \Sigma[L_0 = l_0, L_1 = l_1] \cdot e_1\left[X_{l_0}\right] \cdot e_1\left[X_{l_1}\right] \right) - \sum_{l\in[n]} \mu[L = l] \cdot e_1\left[X_l\right] \,, \tag{1}$$

$$\text{subject to} \quad \sum_{k\in[d]} x_k = B \,. \tag{2}$$

Remark 1 (Connection with $\mathrm{tnreason}$)**.** *A few comments on this problem to be elaborated on in the following:*

- *The portfolio optimization problem is a mode search in an exponential family of atomic variables and terms of pairs.*
- *The objective of the optimization is the (negative of its) energy tensor E.*
- *It has a monomial decomposition with order constraint $r = 2$.*

2 Connection with $\mathrm{tnreason}$

Let us now show the connection with $\mathrm{tnreason}$ in more precision.

2.1 Exponential family of features

The portfolio optimization problem is a mode search in an exponential family, which statistic is a set of formulas consisting in

- atomic variables: To capture the expectations, i.e. the coordinates of μ
- terms of variable pairs: To capture the correlations, i.e. the coordinates of Σ

2.2 Energy Tensor and corresponding distribution

The energy tensor to be maximized is

$$E = -q \left(\sum_{l_0,l_1\in[n],\, l_0\neq l_1} \Sigma[L_0 = l_0, L_1 = l_1] \cdot e_1\left[X_{l_0}\right] \cdot e_1\left[X_{l_1}\right] \right) + \sum_{l\in[n]} \mu[L = l] \cdot e_1\left[X_l\right] \,.$$

We can reconstruct a member of the corresponding Markov Logic Network to the by the temperature $\beta > 0$ scaled energy as

$$\mathbb{P}^{\beta}\left[X_{[d]}\right] = \langle \exp\left[\beta \cdot E\right]\rangle \left[X_{[d]}|\varnothing\right] \, .$$

Sampling from these members in the low temperature limit $\beta \to 0$ is equivalent to the portfolio optimization problem.

2.3 Representation by a monomial selection architecture

The energy tensor E has a monomial decomposition with order constraint $r = 2$, see Section 17.3 in the main report. For the representation we take the slice selection tensor to $d = n$ and order $r = 2$ (see Definition 77 in 17.3) and define a canonical parameter $\theta\left[L_{0,0}, L_{1,0}, L_{0,1}, L_{1,1}\right]$ by

$$\begin{aligned}\theta\left[L_{0,0}, L_{0,1}, L_{1,0}, L_{1,1}\right] = & \sum_{l_0,l_1 \in [n],\, l_0 \neq l_1} -q \cdot \Sigma[L_0 = l_0, L_1 = l_1] \cdot e_{1,1}\left[L_{0,0}, L_{1,0}\right] \otimes e_{l_0,l_1}\left[L_{0,1}, L_{1,1}\right] \\ & + \sum_{l \in [n]} \mu[L = l] \cdot e_{1,2}\left[L_{0,0}, L_{1,0}\right] \otimes e_{l,0}\left[L_{0,1}, L_{1,1}\right]\end{aligned}$$

Then we have

$$E\left[X_{[d]}\right] = \left\langle \theta\left[L_{0,0}, L_{1,0}, L_{0,1}, L_{1,1}\right], \mathcal{H}_{\wedge,d,r}\left[X_{[d]}, L_{0,0}, L_{1,0}, L_{0,1}, L_{1,1}\right]\right\rangle \left[X_{[d]}\right] \, .$$

3 Outlook

We can use the theory developed in the main documentation for

- **Approximation by sparse energies:** Use either the cross-entropy or the squares risk approach in combination with sparsity enhancing algorithms (here the greedy algorithms implemented in tnreason)
- **Extension to higher order porfolio models:** Instead of restricting to $r = 2$ we could allow for larger orders of correlation models. To enhance (numerical and statistical) efficiency and maintain explainability, we could describe influences based on generic propositional formulas instead of the term restriction done here (which is more a brute force way).
- **Usage of classical optimization algorithms:** We could use the algorithms implemented in the tnreason library to solve the optimization problem or more general sample from the tempered distribution. Here, variational inference algorithms such as mean field approaches and classical optimization such as integer linear programs are available.