Reduction from Mean-Variance to ReHLine

aorazalin

July 2024

1 Introduction

This document shows how ReHLine algorithm can be applied to the Mean-Variance problem with an addition of piece-wise linear convex cost functions.

2 Problem statement

Suppose we are given a task of a classic Markowitz portfolio optimization problem with additional piece-wise linear transaction cost function (which is also convex). One could write the original problem as

$$\min_{\mathbf{w}\in\mathbb{R}^n} \frac{\alpha}{2}\mathbf{w}^T\mathbf{G}\mathbf{w} - \mu^T\mathbf{w} + \sum_{i=1}^{n}\phi_i(w_i), \quad \text{s.t. } \mathbf{A}\mathbf{w} + \mathbf{b} \geq 0$$

where $\phi_i(w_i) = p_{il}w_i + q_{il}$ for $w_i \in [d_{il}, d_{il+1}]$ for $l = 1, 2, ..., L_i$. Here, d_{i0} and d_{iL_i+1} are defined as $-\infty$ and ∞, respectively.

3 Reduction

Let's transform the problem into a ReHLine-like framework. Let $\mathbf{G} = \mathbf{L}\mathbf{L}^{\mathbf{T}}$ be Cholesky decomposition of a positive-definite matrix $\mathbf{G}$. Mapping the problem above by transforming $\mathbf{w}' = \sqrt{\alpha}\mathbf{L}^T\mathbf{w}$, we get

$$\min_{\mathbf{w}\in\mathbb{R}^n} \frac{1}{2}\mathbf{w}^T\mathbf{w} - \tilde{\mu}^T\mathbf{w} + \sum_{i=1}^{n}\phi_i(\mathbf{x_i}^T\mathbf{w}), \quad \text{s.t. } \tilde{\mathbf{A}}\mathbf{w} + \mathbf{b} \geq 0$$

where $\tilde{\mu} = \frac{1}{\sqrt{\alpha}}\mathbf{L}^{-1}\mu$, $\mathbf{x_i} = \frac{1}{\sqrt{\alpha}}\mathbf{L}^T_{i,:}$, and $\tilde{\mathbf{A}} = \frac{1}{\sqrt{\alpha}}\mathbf{A}(\mathbf{L}^T)^{-1}$.

In a similar spirit to the ReHLine paper, let's rewrite piece-wise linear functions ϕ_i in terms of the sum of ReLU functions:

$$\phi_i(w) = q_{i0} + p_{i0}w + \sum_{l=1}^{L_i}\text{ReLU}((p_{il} - p_{il-1})(w - d_{il}))$$

Omitting constant terms, the problem becomes

$$\min_{\mathbf{w}\in\mathbb{R}^n} \frac{1}{2}\mathbf{w}^T\mathbf{w} - (\tilde{\mu}-\mathbf{p_0})^T\mathbf{w} + \sum_{i=1}^{n}\sum_{l=1}^{L_i} \mathrm{ReLU}((p_{il}-p_{il-1})(\mathbf{x_i}^T\mathbf{w}-d_{il})), \quad \text{s.t. } \tilde{\mathbf{A}}\mathbf{w}+\mathbf{b} \geq 0$$

Finally, one could easily get a ReHLine minimization problem (equation #4 in the original paper) by completing the square for $\mathbf{w}^T\mathbf{w}$, which can be optimized as usual.