Literature Review

1 Sparse Portfolio for Index Tracking

Index funds are popular investment tools used in modern portfolio management. They are classified into active and passive portfolios,The Index fund strategy is based on passive investment technique. Several papers report the superior performance of passive index funds[EGB96] compared to the other actively managed portfolios. Index tracking is one of the most popular portfolio management techniques available. Even though driven by the financial industry, it is basically a signal processing problem. The most straightforward approach to index tracking is full replication, that is, to buy appropriate amounts of all the assets to analyze the index. Even though it provides perfect tracking, it has a lot of drawbacks because of the illiquid stocks and High Transaction costs that come from allocating capital to all the assets.

Another popular index tracking approach is purchasing an exchange traded fund (ETF), Which uses full replication of the index for tracking, or sparse tracking using aggressive sampling[San15]. An elegant method to deal with the trade-off that arises with transaction costs and tracking efficiency present in the previous method is to use a small number of assets to replicate an index, called the sparse index tracking portfolio, which is the main focus of this paper. Sparse index tracking is similar to sparsity used in signal processing areas like "LASSO" in [Tib96].

empirical tracking error(ETE) for sparse index tracking is widely used in literature [JD02; BMC03; MO07; Sco+12]. Solving the ETE problem involves two steps stock selection and capital allocation. There are many approaches for stock selection; one way is to select the largest possible K <N(Number of assets) according to their market capitalization as given in[MO07] using a genetic algorithm. Another approach involves selecting K most co-related assets to index discussed in[DC05; BG11].

There are different methods for capacity allocation when selected stocks are known. One very naive method is to distribute the capital with respect to the weight of the original index, which is quite challenging to obtain, so a different model is devised in[MO07], which provides minimization of the tracking error to enhance the sparsity;by selecting few stocks randomly every time for a few iterations to increase sparsity. However, this method also concerns with both (i) asset selection and (ii) capacity allocation.However,the complexity can be reduced by Co-integration based approaches discussed for index tracking in [PSG16] and [LJ13], where they compare the results of goal programming and the co-integration method.

Instead,In [MO07]sparsity is acheived with the help of an l0-norm implementation,but since the l0-norm is not convex it needs approximation. The norm is converted to an l-p norm in[XXX15], while in[Rol92] it is converted to a concave function that acts convex till a certain limit.Furthermore, [Che+22] that uses a time-weighted condition for variable selection consistency, which is an extension of the lasso condition. They are both irrepresentable.[Li20] uses a composite quantile regression method instead of LS regression. In [SS22], they consider the index tracking problem as nothing but a sampling problem and devise a Simple Monte Carlo algorithm that does not need approximation like sparse Index Tracking. Other forms of penalizing cardinality involve introducing a non-convex cardinality constraint.Apart from this [Lee+20] has provided substitutes to l0-norm by using BOOST-MIP for asset selection instead.

Other index tracking methods implementing genetic algorithms like [BMC03],[Sco+12],[MO07], [And+13], [Cha+00],[San+17],similarly in [Día+19] a hybrid genetic algorithm along with MINLP is used while GA is used for asset selectionand the MINLP for distribution of weights among the selected assets. differential evolution heuristics are also devised like in [MO07],[And+13],[RS09],[Ten+17] and [Zhe+14], where they utilize a piecewise DC function to penalize the cardinality constraints.The use of DC approaches and coordinate descent algorithms for solving the non-convex norm shown in [Giu17].. In[Ale05], the authors consider a general formulation and propose an algorithm that determines whether or not to rebalance a given portfolio. In[Rol92], a mean-variance analysis or the index tracking portfolios is provided based on the classical Markowitz mean-variance framework [Mar52]. In contrast, in [Bro+09], Brodie et al.consider the problem of portfolio selection within the Markowitz framework, including a 1-penalty. In [Tak+13], Takeda et al. formulate an MIP problem. Since it is hard to solve in practice due to the prohibiting running time, they propose a greedy algorithm. In [SK10], they consider the index tracking problem as a two-stage stochastic optimization problem that is decomposed and solved using sub-problems.

Recent literatures concentrate both on diversity and sparsity like [Kre+22] use l-1 norm which is sorted depending on the weight of the penalty factor that reduce risks without concentrating on sparsity.Another method of reducing risk is by diversification in [ZHY18] a clustering problem is introduced to provide the algorithm with the knoweldge of choosing diversified assets other than choosing randomly.In [KB19], they combine the l0 and l1 norm in constraints, which favors sparsity and diversity.Introducing general cardinality constraint can make the problem NP-hard which is why [WZ22] has discussed about relaxing the constraint with the help of semi-definite programming.Similarly,in [Lee+20] they use a semi-definite relaxation to provide sparse mean-variance portfolio and in [Kim+16] semi-definite relation has been to a cardinality constrained optimal tangent portfolio selection model.In [WY14] instead of usingboth l1 and l2 norm for regularizationthey implemented a non elastic Irrepresentable condition for better variable consistency.

Taking inspiration from [Kre+22] we try to mitigate the hardness of the sorted l-1 norm aiming to produce diverse and sparse index by substituting with an l-2 norm and derive a more tailored sorting technique, we aim to base our research on these lines and improve them further.

References

[Ale05] Nico van der Wijst Alexei A. Gaivoronski Sergiy Krylov. “Optimal portfolio selection with benchmark tracking”. In: *European Journal of Operational Research* 163 (May 2005), pp. 115–131. DOI: 10.1016/j.ejor.2003.12.001.

[And+13] Kostas Andriosopoulos et al. “Portfolio optimization and index tracking for the shipping stock and freight markets using evolutionary algorithms”. In: *Transportation Research Part E: Logistics and Transportation Review* 52 (2013). Special IssueI: Maritime Financial Management, pp. 16–34. ISSN: 1366-5545. DOI: https://doi.org/10.1016/j.tre.2012.11.006. URL: https://www.sciencedirect.com/science/article/pii/S1366554512000968.

[BG11] Daniele Bianchi and Antonio Gargano. “High-Dimensional Index Tracking with Cointegrated Assets Using an Hybrid Genetic Algorithm”. In: *Capital Markets: Asset Pricing & Valuation eJournal* (Mar. 2011). DOI: 10.2139/ssrn.1785908. URL: https://doi.org/10.2139/ssrn.1785908.

[BMC03] J.E. Beasley, N. Meade, and T.-J. Chang. “An evolutionary heuristic for the index tracking problem”. In: *European Journal of Operational Research* 148.3 (2003), pp. 621–643. ISSN: 0377-2217. DOI: https://doi.org/10.1016/S0377-2217(02)00425-3. URL: https://www.sciencedirect.com/science/article/pii/S0377221702004253.

[Bro+09] Joshua Brodie et al. “Sparse and stable Markowitz portfolios”. In: *Proceedings of the National Academy of Sciences* 106.30 (2009), pp. 12267–12272. DOI: 10.1073/pnas.0904287106. URL: https://www.pnas.org/doi/abs/10.1073/pnas.0904287106.

[Cha+00] T.-J. Chang et al. “Heuristics for cardinality constrained portfolio optimisation”. In: *Computers Operations Research* 27.13 (2000), pp. 1271–1302. ISSN: 0305-0548. DOI: https://doi.org/10.1016/S0305-0548(99)00074-X. URL: https://www.sciencedirect.com/science/article/pii/S030505489900074X.

[Che+22] Qi-an Chen et al. “A kind of new time-weighted nonnegative lasso index-tracking model and its application”. In: *The North American Journal of Economics and Finance* 59 (2022), p. 101603. ISSN: 1062-9408. DOI: https://doi.org/10.1016/j.najef.2021.101603. URL: https://www.sciencedirect.com/science/article/pii/S106294082100200X.

[DC05] Christian Dose and Silvano Cincotti. “Clustering of financial time series with application to index and enhanced index tracking portfolio”. In: *Physica A: Statistical Mechanics and its Applications* 355.1 (2005). Market Dynamics and Quantitative Economics, pp. 145–151. ISSN: 0378-4371. DOI: https://doi.org/10.1016/j.physa.2005.02.078. URL: https://www.sciencedirect.com/science/article/pii/S0378437105002864.

[Día+19] Juan Díaz et al. “Index fund optimization using a hybrid model: genetic algorithm and mixed-integer nonlinear programming”. In: *The Engineering Economist* 64.3 (2019), pp. 298–309. DOI: 10.1080/0013791X.2019.1633450. eprint: https://doi.org/10.1080/0013791X.2019.1633450. URL: https://doi.org/10.1080/0013791X.2019.1633450.

[EGB96] Edwin J. Elton, Martin J. Gruber, and Christopher R. Blake. “Survivorship Bias and Mutual Fund Performance”. In: *The Review of Financial Studies* 9.4 (1996), pp. 1097–1120. ISSN: 08939454, 14657368. URL: http://www.jstor.org/stable/2962224 (visited on 10/08/2022).

[Giu17] Margherita Giuzio. “Genetic algorithm versus classical methods in sparse index tracking”. In: *Decisions in Economics and Finance* 40.1 (Nov. 2017), pp. 243–256. ISSN: 1129-6569. DOI: 10.1007/s10203-017-0191-y. URL: https://doi.org/10.1007/s10203-017-0191-y.

[JD02] Roel Jansen and Ronald van Dijk. “Optimal Benchmark Tracking with Small Portfolios”. In: *The Journal of Portfolio Management* 28.2 (2002), pp. 33–39. ISSN: 0095-4918. DOI: 10.3905/jpm.2002.319830. eprint: https://jpm.pm-research.com/content/28/2/33.full.pdf. URL: https://jpm.pm-research.com/content/28/2/33.

[KB19] Nick Koning and Paul Bekker. *Sparse Unit-Sum Regression*. 2019. DOI: 10.48550/ARXIV.1907.04620. URL: https://arxiv.org/abs/1907.04620.

[Kim+16] Min Jeong Kim et al. “Sparse tangent portfolio selection via semi-definite relaxation”. In: *Operations Research Letters* 44.4 (2016), pp. 540–543. ISSN: 0167-6377. DOI: https://doi.org/10.1016/j.orl.2016.05.012. URL: https://www.sciencedirect.com/science/article/pii/S0167637716300396.

[Kre+22] Philipp J. Kremer et al. “Sparse index clones via the sorted 1-Norm”. In: *Quantitative Finance* 22.2 (2022), pp. 349–366. DOI: 10.1080/14697688.2021.1962539. eprint: https://doi.org/10.1080/14697688.2021.1962539. URL: https://doi.org/10.1080/14697688.2021.1962539.

[Lee+20] Yongjae Lee et al. “Sparse and robust portfolio selection via semi-definite relaxation”. In: *Journal of the Operational Research Society* 71.5 (2020), pp. 687–699. DOI: 10.1080/01605682.2019.1581408. eprint: https://doi.org/10.1080/01605682.2019.1581408. URL: https://doi.org/10.1080/01605682.2019.1581408.

[Li20] Ning Li. “Efficient sparse portfolios based on composite quantile regression for high-dimensional index tracking”. In: *Journal of Statistical Computation and Simulation* 90.8 (2020), pp. 1466–1478. DOI: 10.1080/00949655.2020.1731750. eprint: https://doi.org/10.1080/00949655.2020.1731750. URL: https://doi.org/10.1080/00949655.2020.1731750.

[LJ13] Weng Siew Lam and Saiful Hafizah Hj. Jamaan. “Comparison between goal programming and cointegration approaches in enhanced index tracking”. In: *Research in Mathematical Sciences: A Catalyst for Creativity and Innovation*. Ed. by Anuar Ishak et al. Vol. 1522. American Institute of Physics Conference Series. Apr. 2013, pp. 1111–1115. DOI: 10.1063/1.4801255.

[Mar52] Harry Markowitz. “Portfolio Selection”. In: *The Journal of Finance* 7.1 (1952), pp. 77–91. ISSN: 00221082, 15406261. URL: http://www.jstor.org/stable/2975974 (visited on 09/09/2022).

[MO07] Dietmar Maringer and Olufemi Oyewumi. “Index tracking with constrained portfolios”. In: *Intelligent Systems in Accounting, Finance and Management* 15.1-2 (2007), pp. 57–71. DOI: https://doi.org/10.1002/isaf.285. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/isaf.285. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/isaf.285.

[PSG16] Damián Pastor, Tomas Sabol, and Jozef Glova. “Some ideas for improving quality of the index tracking based on cointegration”. In: *Business: Theory and Practice* 17 (Nov. 2016), pp. 325–333. DOI: 10.3846/btp.17.11128. URL: https://doi.org/10.3846/btp.17.11128.

[Rol92] Richard Roll. “A Mean/Variance Analysis of Tracking Error”. In: *The Journal of Portfolio Management* 18.4 (1992), pp. 13–22. ISSN: 0095-4918. DOI: 10.3905/jpm.1992.701922. eprint: https://jpm.pm-research.com/content/18/4/13.full.pdf. URL: https://jpm.pm-research.com/content/18/4/13.

[RS09] Rubén Ruiz-Torrubiano and Alberto Suárez. “A hybrid optimization approach to index tracking”. In: *Annals of Operations Research* 166.1 (Feb. 2009), pp. 57–71. ISSN: 1572-9338. DOI: 10.1007/s10479-008-0404-4. URL: https://doi.org/10.1007/s10479-008-0404-4.

[San+17] Leonardo Riegel Sant’Anna et al. “Index tracking with controlled number of assets using a hybrid heuristic combining genetic algorithm and non-linear programming”. In: *Annals of Operations Research* 258.2 (Nov. 2017), pp. 849–867. ISSN: 1572-9338. DOI: 10.1007/s10479-016-2111-x. URL: https://doi.org/10.1007/s10479-016-2111-x.

[San15] André A.P. Santos. “Beating the market with small portfolios: Evidence from Brazil”. In: *EconomiA* 16.1 (2015), pp. 22–31. ISSN: 1517-7580. DOI: https://doi.org/10.1016/j.econ.2014.11.003. URL: https://www.sciencedirect.com/science/article/pii/S1517758014000368.

[Sco+12] Andrea Scozzari et al. “Exact and heuristic approaches for the index tracking problem with UCITS constraints”. In: *Annals of Operations Research* 205 (May 2012). DOI: 10.2139/ssrn.2066672. URL: https://doi.org/10.2139/ssrn.2066672.

[SK10] Stephen J. Stoyan and Roy H. Kwon. “A two-stage stochastic mixed-integer programming approach to the index tracking problem”. In: *Optimization and Engineering* 11.2 (June 2010), pp. 247–275. DOI: 10.1007/s11081-009-9095-1. URL: https://doi.org/10.1007/s11081-009-9095-1.

[SS22] Tanmay Satpathy and Rushabh Shah. “Sparse index tracking using sequential Monte Carlo”. In: *Quantitative Finance* 22.9 (2022), pp. 1579–1592. DOI: 10.1080/14697688.2022.2057353. eprint: https://doi.org/10.1080/14697688.2022.2057353. URL: https://doi.org/10.1080/14697688.2022.2057353.

[Tak+13] Akiko Takeda et al. “Simultaneous pursuit of out-of-sample performance and sparsity in index tracking portfolios”. English. In: *Computational Management Science* 10.1 (Feb. 2013), pp. 21–49. ISSN: 1619-697X. DOI: 10.1007/s10287-012-0158-y.

[Ten+17] Yue Teng et al. “A penalty PALM method for sparse portfolio selection problems”. In: *Optimization Methods and Software* 32.1 (2017), pp. 126–147. DOI: 10.1080/10556788.2016.1204299. eprint: https://doi.org/10.1080/10556788.2016.1204299. URL: https://doi.org/10.1080/10556788.2016.1204299.

[Tib96] Robert Tibshirani. “Regression Shrinkage and Selection via the Lasso”. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 58.1 (1996), pp. 267–288. ISSN: 00359246. URL: http://www.jstor.org/stable/2346178 (visited on 09/20/2022).

[WY14] Lan Wu and Yuehan Yang. “Nonnegative Elastic Net and application in index tracking”. In: *Applied Mathematics and Computation* 227 (2014), pp. 541–552. ISSN: 0096-3003. DOI: https://doi.org/10.1016/j.amc.2013.11.049. URL: https://www.sciencedirect.com/science/article/pii/S0096300313012174.

[WZ22] Angelika Wiegele and Shudian Zhao. “Tight SDP Relaxations for Cardinality-Constrained Problems”. In: *Operations Research Proceedings 2021*. Ed. by Norbert Trautmann and Mario Gnägi. Cham: Springer International Publishing, 2022, pp. 167–172. ISBN: 978-3-031-08623-6.

[XXX15] Fengmin Xu, Zongben Xu, and Honggang Xue. *Sparse Index Tracking Based On $L_{1/2}$ Model And Algorithm*. 2015. DOI: 10.48550/ARXIV.1506.05867. URL: https://arxiv.org/abs/1506.05867.

[Zhe+14] Xiaojin Zheng et al. “Successive convex approximations to cardinality-constrained convex programs: a piecewise-linear DC approach”. In: *Computational Optimization and Applications* 59.1 (Oct. 2014), pp. 379–397. ISSN: 1573-2894. DOI: 10.1007/s10589-013-9582-3. URL: https://doi.org/10.1007/s10589-013-9582-3.

[ZHY18] Yu Zheng, Timothy M. Hospedales, and Yongxin Yang. *Diversity and Sparsity: A New Perspective on Index Tracking*. 2018. DOI: 10.48550/ARXIV.1809.01989. URL: https://arxiv.org/abs/1809.01989.