

Structure-Aware Distributionally Robust Portfolio Optimization with Energy-Based Anisotropic Ambiguity Sets

진영봉

FRE LAB

Apr 13, 2026

Background

- Portfolio Optimization under Distributional Uncertainty

- 현실의 포트폴리오 최적화에서는 미래 자산 수익률 ξ 를 정확하게 알 수 없음
- MVO(Mean-Variance Optimization)는 기대수익률 $\hat{\mu}$, 공분산 $\hat{\Sigma}$ 를 점추정하여 그대로 사용
- 따라서, 추정오차에 민감하며, 이로 인해 out-of-sample(OOS) 성능이 낮음 [1]

Formulation: $\min_w \frac{\lambda}{2} w^T \hat{\Sigma} w - w^T \hat{\mu}$

최적화 방법	아이디어	장점	한계
Stochastic Programming (SP)	특정 분포 $\hat{\mathbb{P}}$ 을 신뢰하고 기대손실 최소화	확률적 해석이 자연스러움	분포 가정이 틀리면 취약
Robust Optimization (RO)	불확실성 집합 내부의 최악 상황을 방어	보수적 안정성 확보	지나치게 보수적
Distributionally Robust Optimization (DRO)	$\hat{\mathbb{P}}$ 주변의 분포들 중 최악의 분포를 방어	데이터 기반 + worst-case 방어	Ambiguity set설계가 핵심

- Distributionally Robust Optimization (DRO)

- DRO는 경험적 분포를 믿지만 확신하지는 않음 (SP와 RO의 절충안)
- 포트폴리오 최적화에서 DRO를 활용하는 다수의 연구 존재 [2-6]

Formulation: $\min_w \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}[L(w, \xi)]$

Literature Review

- Mohajerin Esfahani & Kuhn (2018) [2] – WDRO의 표준 방법 제시

- $\inf_{x \in X} \sup_{Q \in \mathbb{B}_\epsilon(\hat{P}_N)} \mathbb{E}_Q[L(x, \xi)], \mathbb{B}_\epsilon(\hat{P}_N) := \{Q: W(Q, \hat{P}_N) \leq \epsilon\}$
- 경험분포 \hat{P}_N 주변의 Wasserstein ball 안에서 worst-case 기대 손실을 최소화하는 WDRO의 표준 틀
- Wasserstein ambiguity set 은 finite-sample 을 보장과 tractable reformulation 을 제공

- Fonseca & Junca (2021) [3] – 제약조건이 있는 WDRO

- $\min_{x \in X} \sup_{Q \in \mathbb{B}_\epsilon(\hat{P}_N)} \mathbb{E}_Q[f(x, \xi)] \quad \text{s.t.} \quad \inf_{Q \in \mathbb{B}_\epsilon(\hat{P}_N)} \mathbb{E}_Q[g(x, \xi)] \geq \eta$
- WDRO를 목적함수 뿐 아니라 기대값 제약까지 포함하는 현실적 최적화 문제로 확장

- Blanchet, Chen, & Zhou (2022) [4] – WDRO 포트폴리오

- $\min_w \sup_{Q \in \mathbb{B}_\epsilon(\hat{P}_N)} w^T \text{Var}_Q(\xi) w \quad \text{s.t.} \quad \inf_{Q \in \mathbb{B}_\epsilon(\hat{P}_N)} \mathbb{E}_Q[w^T \xi] \geq \rho \Leftrightarrow \min_w \sqrt{w^T \text{Var}_{\hat{P}_N}(\xi) w + \sqrt{\epsilon} \|w\|_p} \quad \text{s.t.} \quad \mathbb{E}_{\hat{P}_N}[w^T \xi] \geq \rho + \sqrt{\epsilon} \|w\|_p$
- WDRO를 포트폴리오 이론과 직접 연결, MVO 문제를 DRO 형태로 다시 쓰고 regularized MVO 문제와 동일함을 증명
- 위험이 작은 포트폴리오를 찾되, 한쪽에 쏠리거나 추정오차에 약한 포트폴리오에 제약 조건



무한차원 min-max 문제는 경험적 목적함수 + 제약항 형태로 변환 가능
분포 Q 를 직접 최적화하지 않고, 거리 제약을 라그랑주화

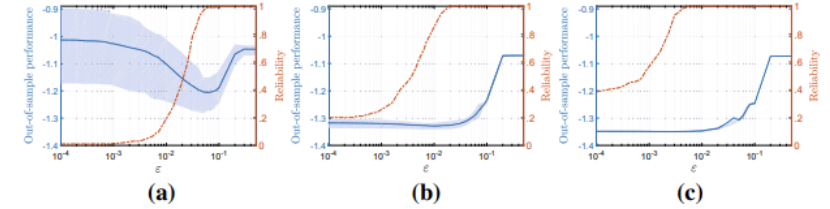


Fig. 5 Out-of-sample performance $J(\hat{x}_N(\epsilon))$ (left axis, solid line and shaded area) and reliability $\mathbb{P}^N[J(\hat{x}_N(\epsilon)) \leq \hat{J}_N(\epsilon)]$ (right axis, dashed line) as a function of the Wasserstein radius ϵ and estimated on the basis of 200 simulations. (a) $N = 30$ training samples. (b) $N = 300$ training samples. (c) $N = 3000$ training samples

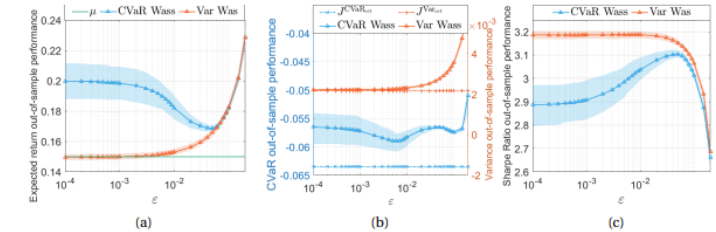
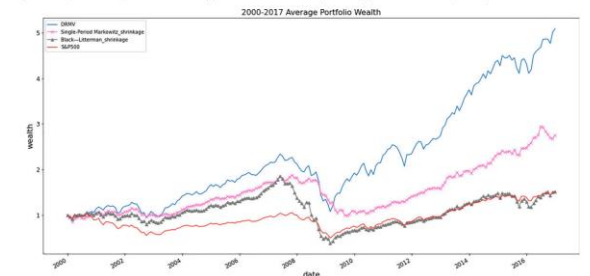


Fig. 6. This figure presents the out-of-sample performance of expected returns $R(\hat{x}_N^{\text{VarEst}}(\epsilon))$ (orange) and $R(\hat{x}_N^{\text{CVaR}}(\epsilon))$ (blue), the out-of-sample performance of variance $V(\hat{x}_N^{\text{VarEst}}(\epsilon))$ (orange), and the Conditional Value at Risk $\text{CVaR}_{\alpha, \xi \sim P}(-\langle \hat{x}_N^{\text{CVaR}}(\epsilon), \xi \rangle)$ (blue), as well as Sharpe Ratios as functions of the Wasserstein radius ϵ . These are estimated based on 200 simulations for $N = 300$. In this instance, $\mu = 0.15$ and $\alpha = 0.05$.

Figure 23. (Color online) Portfolios' Average Wealth Processes Using Shrinkage Estimators from January 2000 to December 2016



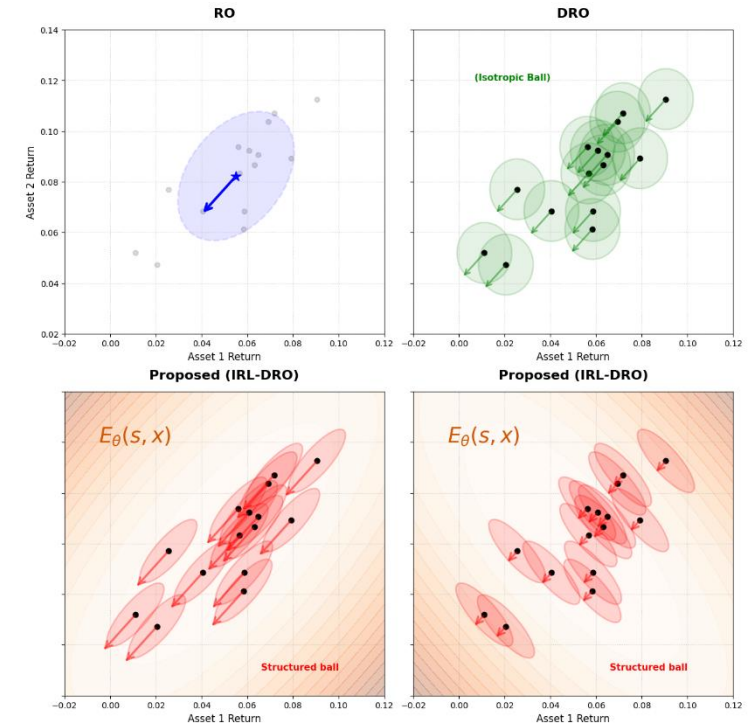
Notes. All the portfolios except S&P 500 consist of 100 stocks, and the averages are calculated over 100 numerical experiments. The x-axis indicates the time in months (from 1 to 204), and the y-axis indicates the portfolio wealth. Initial wealth is set to be one.

Comparative Analysis

- Research gap

- 기존 Wasserstein 거리 기반 DRO는 불확실성 집합이 모든 방향으로 균일한 Isotropic 형태를 갖도록 설계됨
- 금융 자산간 상관관계는 비대칭/비등방성을 띄며, 위기 시 특정 방향으로 위험이 증폭 되지만, Isotropic-sphere 형태의 방어막은 시장의 이러한 구조적 위험을 반영하지 않고 있음

Feature	MVO (SP)	RO	DRO	SA-DRO (proposed)
Uncertainty / Ambiguity set	Point Estimate ($\hat{\mu}, \hat{\Sigma}$)	Fixed set (Ellipsoid/Box)	Wasserstein Ball	Structure-Aware Ellipsoid
Geometry / Metric	Sample Statistics	Worst-case Bounds	Euclidean Distance (l_p family)	(Learned) Mahalanobis Metric
Ambiguity Shape	N/A (point)	Static (Pre-defined)	Isotropic	Anisotropic
Data Usage	Parameter Estimation	Bound Estimation	Empirical Distribution	Empirical Distribution With Geometry



Methodology

- Research goal: 시장의 내재적 정보를 학습하여 불확실성 집합 구조를 반영하는 SA-DRO (Structure-Aware DRO)를 제안

Standard WDRO

$$\min_w \sup_{Q \in \mathcal{U}_\epsilon(P_n)} \mathbb{E}_Q[-w^\top \xi]$$

$$\begin{aligned} \mathcal{U}_\epsilon(P_n) &= \{Q: d_w(Q, P_n) \leq \epsilon\} \\ d_w(Q, P_n) &= \inf_{\pi \in \Pi(Q, P_n)} \mathbb{E}_\pi[\|\xi - \xi'\|] \end{aligned}$$



SA DRO

$$\min_w \sup_{Q \in \mathcal{U}_\epsilon^{SA}(P_n, s)} \mathbb{E}_Q[-w^\top \xi]$$

$$\begin{aligned} \mathcal{U}_\epsilon^{SA}(P_n, s) &= \{Q: d_{M(s)}(Q, P_n) \leq \epsilon\} \\ d_{M(s)}(Q, P_n) &= \inf_{\pi \in \Pi(Q, P_n)} \mathbb{E}_\pi[\sqrt{(\xi - \xi')^\top M(s) (\xi - \xi')}] \\ M(s) &= \nabla_\xi^2 E_\theta(s, \xi) \Big|_{\xi = \bar{\xi}(s)} + \rho I, \rho > 0 \\ p_\theta(\xi | s) &\propto \exp(-E_\theta(s, \xi)) \end{aligned}$$

- Ambiguity set 크기는 ϵ 이 결정
- d_w 는 사전에 고정되어있어 isotropic geometry 위에서 작동
- 결과적으로 robust penalty는 norm-based regularization으로 해석가능

- Ambiguity set 크기는 ϵ 이 결정
- Ambiguity set 형상과 방향성은 $M(s)$ 가 결정
- $M(s)$ 는 시장 데이터에서 학습된 state-dependent local geometry
- 결과적으로 robust penalty는 anisotropic geometry 위에서 작동

Methodology

- Step1. Market Geometry 학습

- 각 훈련 구간에서 수익률 데이터 x 를 사용해 Energy-based model을 학습

$$p_{\theta}(x) \propto \exp(-E_{\theta}(x))$$

- 시장을 예측하는 것이 아닌 시장 분포의 local curvature 학습이 목표
- 에너지 함수의 Hessian이 local curvature이며 ambiguity set의 메트릭으로 사용가능 [7]
- 학습을 위해서 Denoising score matching (DSM) loss 사용

$$L_{DSM} = \mathbb{E} \left[\left\| \nabla_{\tilde{x}} E_{\theta}(\tilde{x}) - \frac{\tilde{x} - x}{\sigma^2} \right\|^2 \right]$$

- Step2. Hessian으로 구조적 metric 구성

- 학습된 에너지 함수의 평균 Hessian으로 local metric을 생성

$$M = \frac{1}{N} \sum_{i=1}^N \nabla_x^2 E_{\theta}(x_i) + \rho I$$

- 고유값 분해 후 음수나 0에 가까운 값을 clipping하여 가장 가까운 안정적 SPD surrogate로 만들어서 metric에 사용

Methodology

- Step3. Structure-aware transport cost 정의

- 기존 WDRO 대신 Hessian 기반 Mahalanobis cost 사용

$$c_{M(s)}(\xi, \hat{\xi}_i) = (\xi - \hat{\xi}_i)^\top M(s)(\xi - \hat{\xi}_i)$$

- 따라서 ambiguity set은 다음과 같음

$$\mathcal{U}_\epsilon^{SA}(\hat{P}_N, s) = \{Q: W_{c_{M(s)}}(Q, \hat{P}_N) \leq \epsilon\}$$

- Step4. SA DRO portfolio

- learned ambiguity set 위의 worst-case distribution을 방어하기 위한 SA-DRO 문제를 해결
- Robustness budget + 경험적 평균 수익 최대화 + structure-aware penalty

$$\min_w \sup_{Q \in \mathcal{U}_\epsilon^{SA}(P_n, s)} \mathbb{E}_Q[-w^\top \xi] \Rightarrow \inf_{\lambda \geq 0} \left\{ \lambda \epsilon^2 + \frac{1}{N} \sum_{i=1}^N \sup_{\xi} \left[-w^\top \xi - \lambda (\xi - \hat{\xi}_i)^\top M(s)(\xi - \hat{\xi}_i) \right] \right\}$$

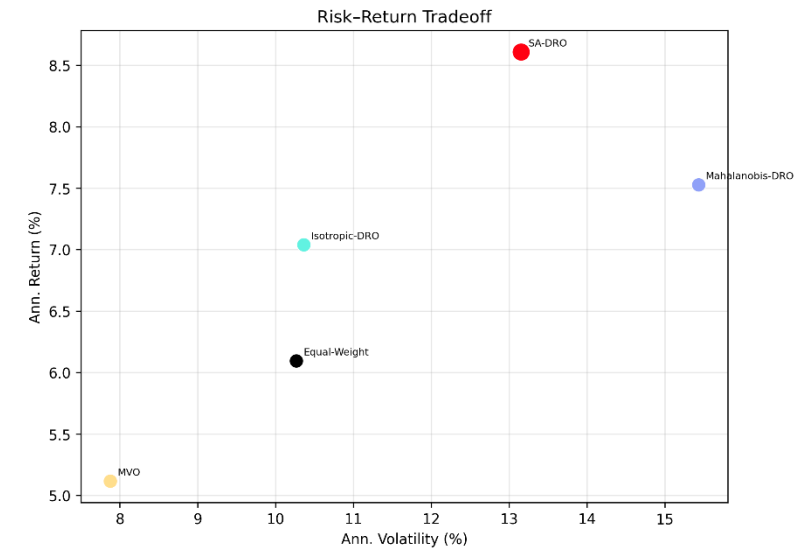
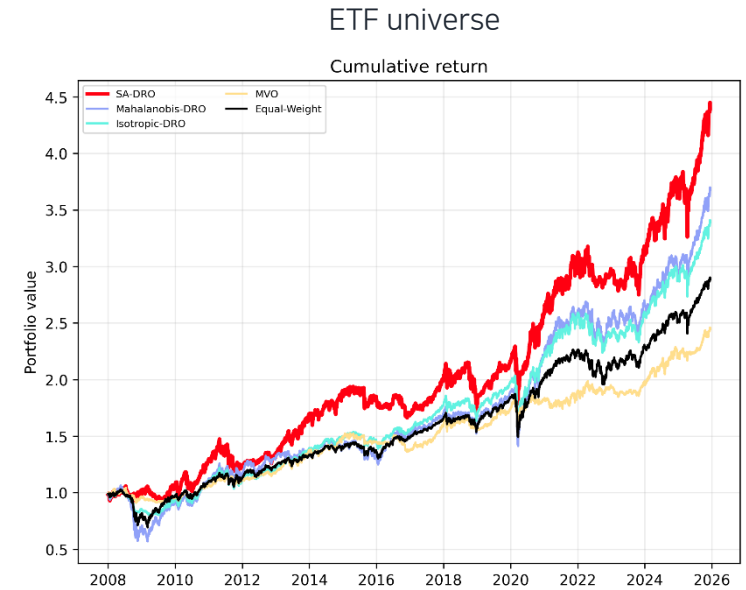
$$\sup_{\xi} \left[-w^\top \xi - \lambda (\xi - \hat{\xi}_i)^\top M(s)(\xi - \hat{\xi}_i) \right] = -w^\top \xi + \frac{1}{4\lambda} w^\top M(s)^{-1} w \Rightarrow -\hat{\mu}^\top w + \frac{1}{4\lambda} w^\top M(s)^{-1} w$$

$$\min_{w, \lambda \geq 0} \left\{ \lambda \epsilon^2 - \hat{\mu}^\top w + \frac{1}{4\lambda} w^\top M(s)^{-1} w \right\}$$

Results

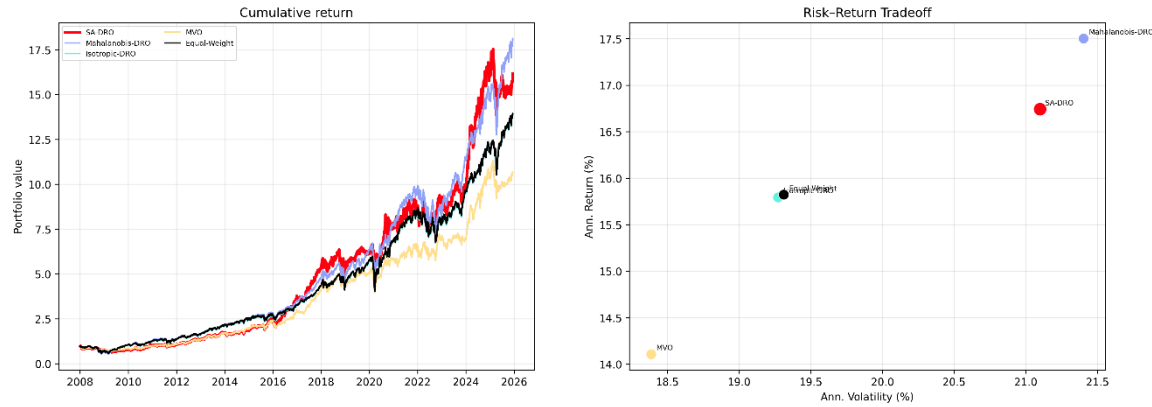
- Data
 - 3가지 투자 universe (ETF / Dow 30 / Nasdaq 100)
 - 2007.01.01~2025.12.31에 모두 존재하는 종목
 - 1년 (252 영업일)데이터 train -> 1달 (21영업일) 데이터 test
 - Sliding window 방식
 - 거래 수수료 : 10bp
- Benchmark
 - Isotropic DRO : Wasserstein 거리를 이용한 DRO [4]
 - Mahalanobis DRO : 공분산 행렬을 이용한 DRO
- Results

	Equal Weight	MVO	WDRO (Isotropic DRO)	Mahalanobis DRO	SA-DRO
Ann. Return	6.10%	5.12%	7.04%	7.53%	8.61%
Sharpe ratio	0.628	0.674	0.709	0.548	0.695
Sortino	0.974	1.012	1.089	0.852	1.070
MDD	-32.72%	-13.11%	-22.42%	-46.44%	-22.04%

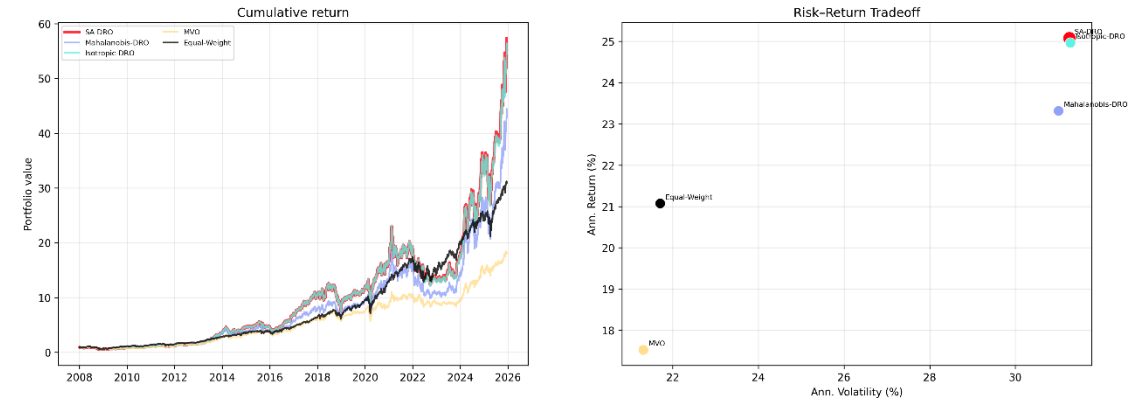


Results

Dow Jones 30 universe



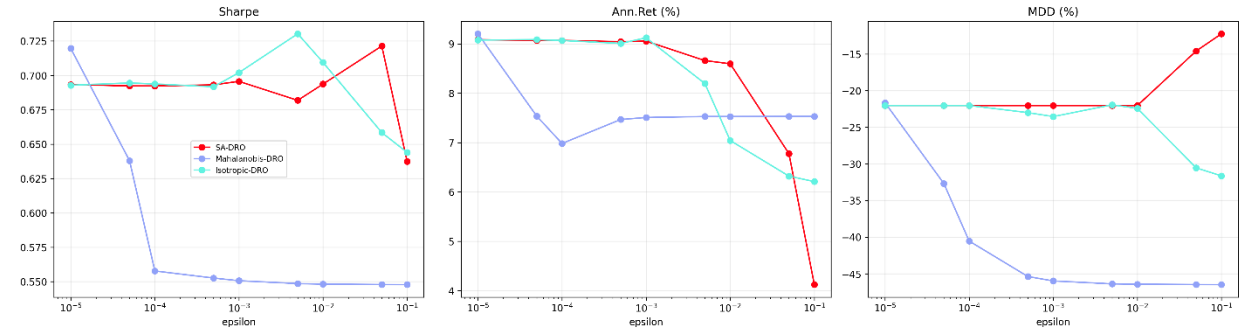
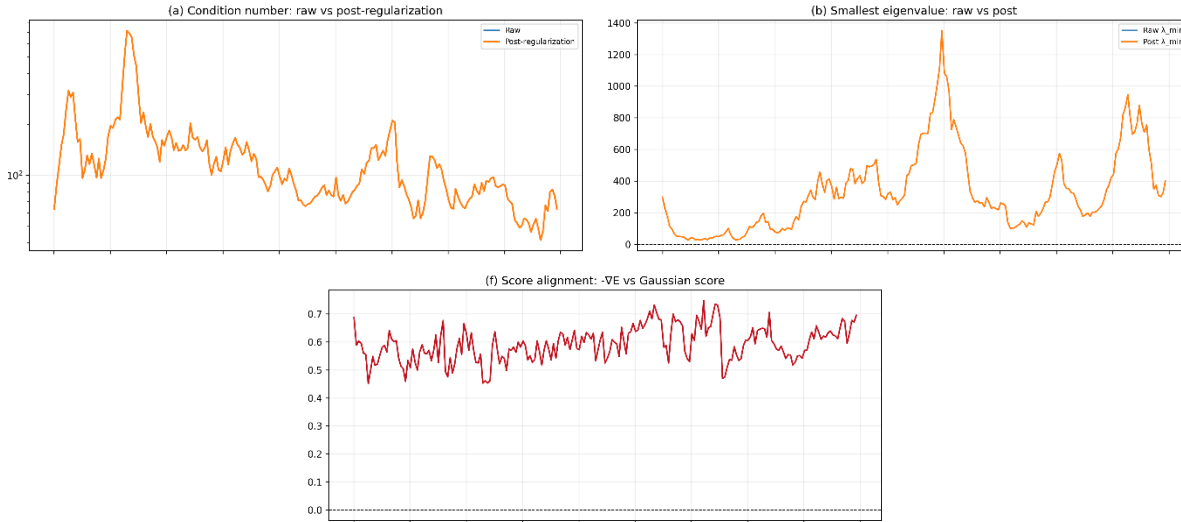
Nasdaq 100 universe



	Equal Weight	MVO	WDRO	Mahalanobis DRO	SA-DRO
Ann. Return	15.82%	14.11%	15.79%	17.50%	16.74%
Sharpe ratio	0.858	0.811	0.859	0.862	0.840
Sortino	1.360	1.318	1.359	1.372	1.388
MDD	-42.53%	-33.80%	-42.73%	-46.34%	-37.99%

	Equal Weight	MVO	WDRO	Mahalanobis DRO	SA-DRO
Ann. Return	21.07%	17.53%	24.96%	23.31%	25.08%
Sharpe ratio	0.991	0.866	0.870	0.833	0.874
Sortino	1.583	1.380	1.426	1.351	1.431
MDD	-49.29%	-43.42%	-51.67%	-52.67%	-51.67%

Results – ETF universe



- Hessian 행렬 진단

- Condition number: $\kappa(H) = \frac{\lambda_{max}(H)}{\lambda_{min}(H)}$
- Hessian 행렬이 anisotropic 구조
- 원본 Hessian은 SPD 행렬
- EBM score와 가우시안 조건하에서 score가 유사하게 정렬됨
 - EBM 모델이 데이터 중심의 geometry와 어긋나있지 않음

- ϵ 에 따른 성능 변화

- Sharpe 지수 측면에서는 SA-DRO가 전반적으로 안정적 / 최고점 기준으로는 WDRO
- Return 측면에서도 SA-DRO가 전반적으로 안정적 / 과도한 보수성은 좋지 않음
- MDD 측면에서는 ϵ 이 커질수록 SA-DRO는 성능 상승 / 나머지는 하락
- 최적의 ϵ 은 달라질 수 있고, SA-DRO는 이에 덜 민감한 편

conclusion

- Summary

- 기존 WDRO 는 isotropic ambiguity set을 사용하여 모든 방향의 분포 이동을 동일하게 취급
- 그러나 실제 금융 시장은 비등방적(asymmetric/anisotropic) 구조를 가지면, 위기 시 특정 방향으로 위험이 집중됨
- 본 연구는 EBM으로 return-space geometry를 학습하고, 그 Hessian curvature를 사용한 anisotropic ambiguity set을 구성
- SA-DRO는 기존 WDRO와 다른 데이터 기반의 robust penalty를 형성
- 실험적으로 일부 파라미터 구간에서 더 안정적인 sharpe/MDD를 보여줌

- Future works

- 데이터 생존편향 문제
- 학습된 Hessian이 위기 구간에서 실제로 잘 작동하는지 확인
- Epsilon 선택 시 단순 grid search가 아닌 coverage, realized tail loss 등의 기준으로 정교하게 calibration (or validation set 활용)
- 현재 구조에서는 포트폴리오 집중도가 높아질 수 있으므로 diversification penalty등도 고려

- Connection to Prior work

- 이전 연구 (ME-IRL)은 시장이 무엇을 선호하는지 학습하는 연구
- SA-DRO는 어떤 방향의 위험을 더 경계해야 하는지 학습하는 연구
- 두 방법 모두 금융 목표나 위험구조를 고정적으로 설계 하지 않고 데이터로부터 energy-based 표현을 통해 학습
- 선호학습 → 위험구조 학습 → 강건한 의사결정 단계로 연결하여 일관된 연구 목표 달성

References

1. Michaud, R. O. (1989). The Markowitz optimization enigma: Is 'optimized' optimal?. *Financial analysts journal*, 45(1), 31-42.
2. Mohajerin Esfahani, P., & Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1), 115-166.
3. Fonseca, D., & Junca, M. (2021). Wasserstein Distributionally Robust Optimization with Expected Value Constraints. *arXiv preprint arXiv:2111.04663*.
4. Blanchet, J., Chen, L., & Zhou, X. Y. (2022). Distributionally robust mean-variance portfolio selection with Wasserstein distances. *Management science*, 68(9), 6382-6410.
5. Costa, G., & Iyengar, G. N. (2023). Distributionally robust end-to-end portfolio construction. *Quantitative Finance*, 23(10), 1465-1482.
6. Finn, C., Christiano, P., Abbeel, P., & Levine, S. (2016). A connection between generative adversarial networks, inverse reinforcement learning, and energy-based models. *arXiv preprint arXiv:1611.03852*.
7. Girolami, M., & Calderhead, B. (2011). Riemann manifold langevin and hamiltonian monte carlo methods. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 73(2), 123-214.