Chapter 2 Exercise

1

Fitting the Marcenko-Pastur Distribution

- 2 Using a series of matrix of stock returns:
  - A. Compute the covariance matrix. What is the condition number of the correlation matrix?
  - B. Compute one hundred efficient frontiers by drawing one hundred alternative vectors of expected returns from a Normal distribution with mean 10% and standard deviation 10%.
  - C. Compute the variance of the errors against the mean efficient frontier.
- 3. Repeat Exercise 2, where this time you denoise the covariance matrix before computing the one hundred efficient frontiers.
  - A. What is the value of σ^2 implied by the Marcenko–Pastur distribution?
  - B. How many eigenvalues are associated with random components?
  - C. Is the variance of the errors significantly higher or lower? Why?
- 5. Repeat Exercise 3, where this time you also detone the covariance matrix before computing the one hundred efficient frontiers. Is the variance of the errors significantly higher or lower? Why?

1

Fitting the Marcenko-Pastur Distribution

- 3. Repeat Exercise 2, where this time you denoise the covariance matrix before computing the one hundred efficient frontiers.
  - A. What is the value of σ^2 implied by the Marcenko–Pastur distribution?
  - B. How many eigenvalues are associated with random components?
  - C. Is the variance of the errors significantly higher or lower? Why?
- 5. Repeat Exercise 3, where this time you also detone the covariance matrix before computing the one hundred efficient frontiers. Is the variance of the errors significantly higher or lower? Why?

- 2 Using a series of matrix of stock returns:
- a Compute the covariance matrix. What is the condition number of the
- correlation matrix?
- b Compute one hundred efficient frontiers by drawing one hundred alternative
- vectors of expected returns from a Normal distribution with mean
- 10% and standard deviation 10%.
- c Compute the variance of the errors against the mean efficient frontier.

1

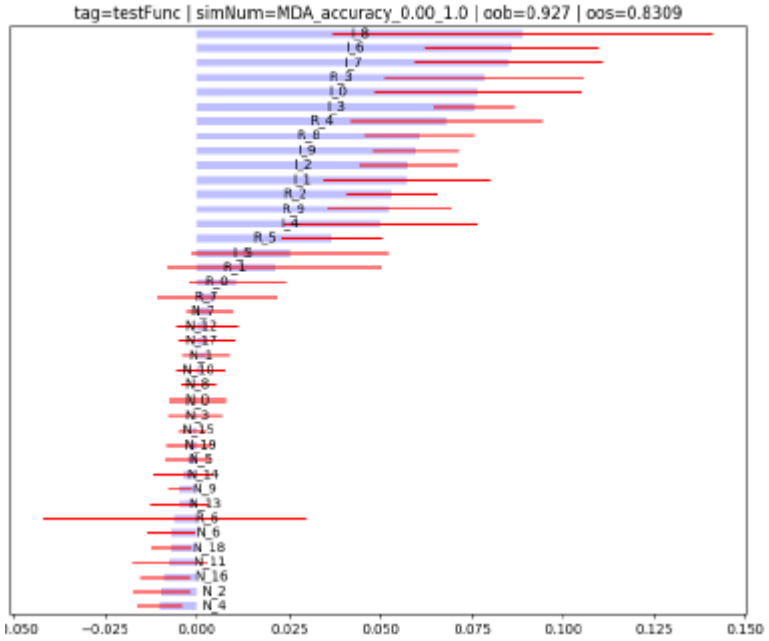
Fitting the Marcenko-Pastur Distribution

- **What are popular misconceptions about financial ML?**
  - **ML Is the Holy Grail versus ML Is Useless**
    - ML은 만능이 아니며, 무용한 것도 아님
    - 오히려 ML은 현대적 통계학에 가까움

1

Fitting the Marcenko-Pastur Distribution

- **What are popular misconceptions about financial ML?**
  - **ML Is a Black Box**
    - 오히려 traditional statistical methods보다 더 인사이트가 클 수 있음
      - PDP, ICE, ALE, Friedman's H-stat,MDI, MDA, global surrogate,
      - LIME, and Shapley values 등을 활용 가능
    - 사람들이 차의 내부 작동 방식에 대해서는 모르지만, 차를 운전할 수 있음
      - ML 모델도 내용을 잘 파악하고 있다면 원하는 실험을 진행하는데 유용한 툴(white box)가 될 수 있음



1

Fitting the Marcenko-Pastur Distribution

- **What are popular misconceptions about financial ML?**
  - **Finance Has Insufficient Data for ML**
    - Price prediction 관점에서는 많은 데이터가 필요한 것이 사실
    - 하지만 금융 영역이 price prediction만 있는 것은 아님
      - Risk analysis, portfolio construction, outlier detection, feature importance 등..
    - Deep hedging, credit rating, execution의 경우 데이터 수가 많음

1 Fitting the Marcenko-Pastur Distribution

- **What are popular misconceptions about financial ML?**
  - **The Signal-to-Noise Ratio Is Too Low in Finance**
    - 분명 금융이 다른 영역에 비해 SNR이 낮은 것은 사실
    - 하지만 이는 금융 영역에서는 ML을 다르게 사용해야 한다는 것을 의미하지, ML을 이용할 수 없다는 것을 말하지는 않음

1 Fitting the Marcenko-Pastur Distribution

- **What are popular misconceptions about financial ML?**
  - **The Risk of Overfitting Is Too High in Finance**
    - ML 알고리즘이 classical method보다 오히려 오버피팅 확률이 낮다
      - 반면 비전문가가 행할 경우 오버피팅 확률이 높을 수 있다


Thank you
YONSEI UNIVERSITY