WEEK 3 PROJECT

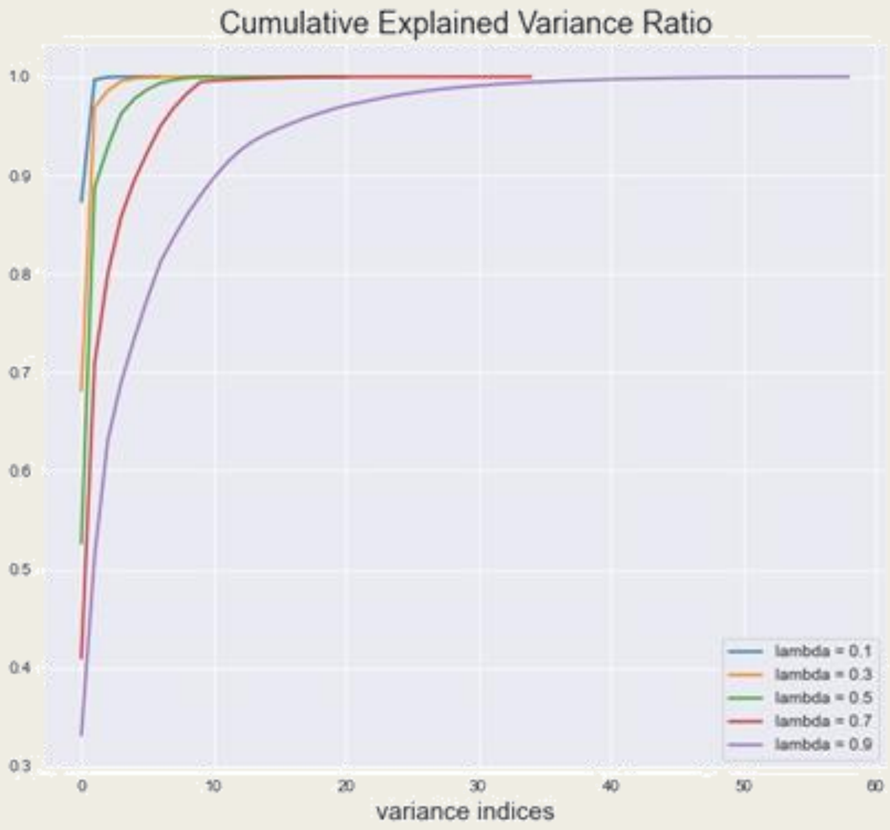
Peiyi Xu

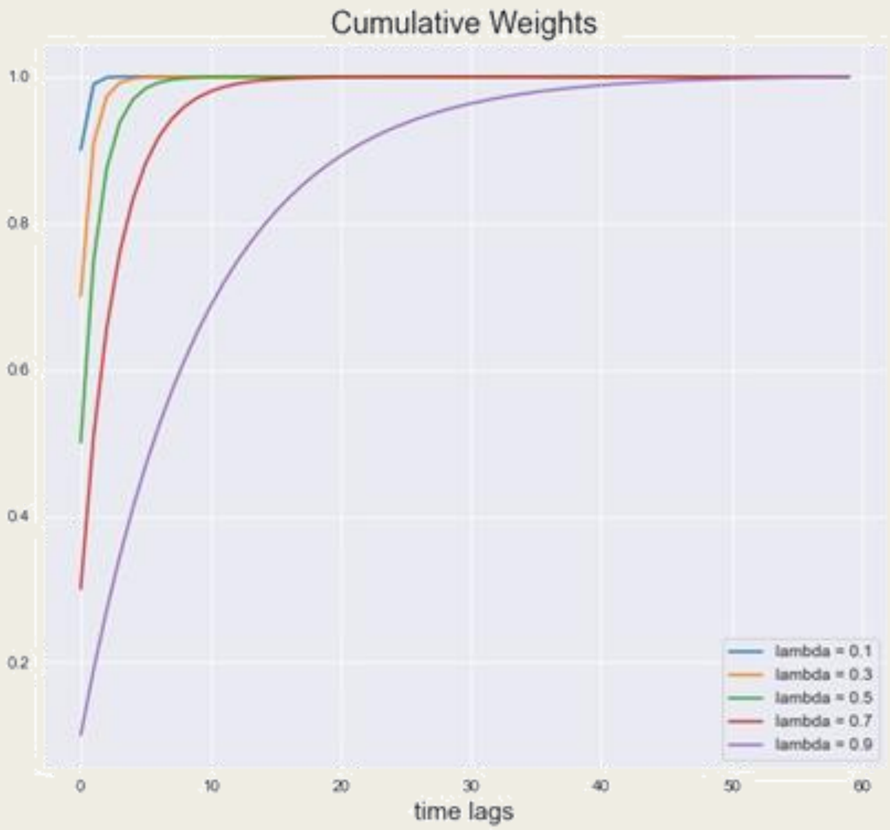
2022.1.28

Problem 1

- Use the stock returns in DailyReturn.csv for this problem. DailyReturn.csv contains returns for 100 large US stocks and as well as the ETF, SPY which tracks the S&P500.
- Create a routine for calculating an exponentially weighted covariance matrix. If you have a package that calculates it for you, verify that it calculates the values you expect. This means you still have to implement it.
- Vary $\lambda \in (0, 1)$. Use PCA and plot the cumulative variance explained by each eigenvalue for each λ chosen. What does this tell us about values of λ and the effect it has on the covariance matrix?

Answer 1





$$time\ lag = 1, \ldots, 60$$
$$\lambda = 0.1, 0.3, 0.5, 0.7, 0.9$$

Conclusions

- The less λ is, the greater weight we put on current information
- The greater λ is, the weights are more equally distributed on information of each time periods.
- When λ is closer to 0, we need less principal components to explain for the variance;
- When λ is closer to 1, we will need more.

Problem 2

- Copy the chol_psd(), and near_psd() functions from the course repository. Implement in your programming language of choice.
- Implement Higham's 2002 nearest psd correlation function.
- Generate a non-psd correlation matrix that is 500x500. Use near_psd() and Higham's method to fix the matrix. Confirm the matrix is now PSD.
- Compare the results of both using the Frobenius Norm. Compare the run time between the two. How does the run time of each function compare as N increases?
- Based on the above, discuss the pros and cons of each method and when you would use each.

Answer 2

When input size increases…

Rebonato & Jackel

- run times don’t increase much
- Frobenius Norm increased rapidly.
- fast but inaccurate

Higham

- run times increases at a nearly quadratic rate
- Frobenius Norm increased
- accurate but slow one.

Run Time (in seconds)
Frobenius Norm
n = 100
n = 300
n = 500
Rebonato and Jackel
Higham

Problem 3

- Using DailyReturn.csv. Implement a multivariate normal simulation that allows for simulation directly from a covariance matrix or using PCA with an optional parameter for % variance explained. Generate a correlation matrix and variance vector 2 ways: 1. Standard Pearson correlation/variance 2. Exponentially weighted $\lambda = 0.97$
- Combine these to form 4 different covariance matricesSimulate 25,000 draws from each covariance matrix using:
  1. Direct Simulation
  2. PCA with 100% explained.
  3. PCA with 75% explained.
  4. PCA with 50% explained.
- Calculate the covariance of the simulated values. Compare the simulated covariance to it's input matrix using the Frobenius Norm. Compare the run times for each simulation. What can we say about the trade offs between time to run and accuracy?

Answer 3

- Though we can be 10 times faster on generating random simulations using PCA with 75% and 50% variance explained, we end up having hundreds times larger Frobenius Norms.

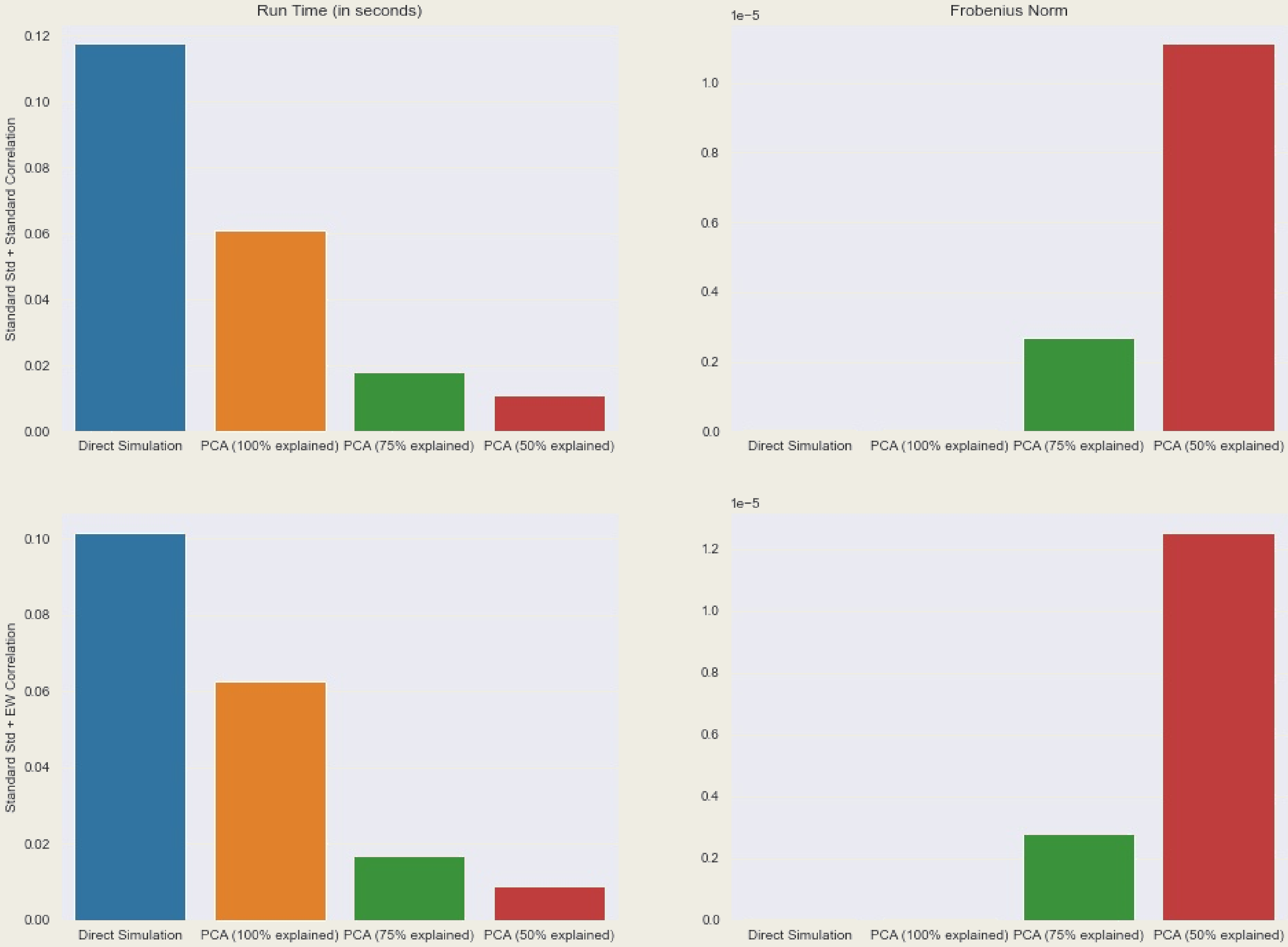
- We don’t necessarily gain improvements in efficiency in proportion to the sacrifice we made on accuracy.

Run Time

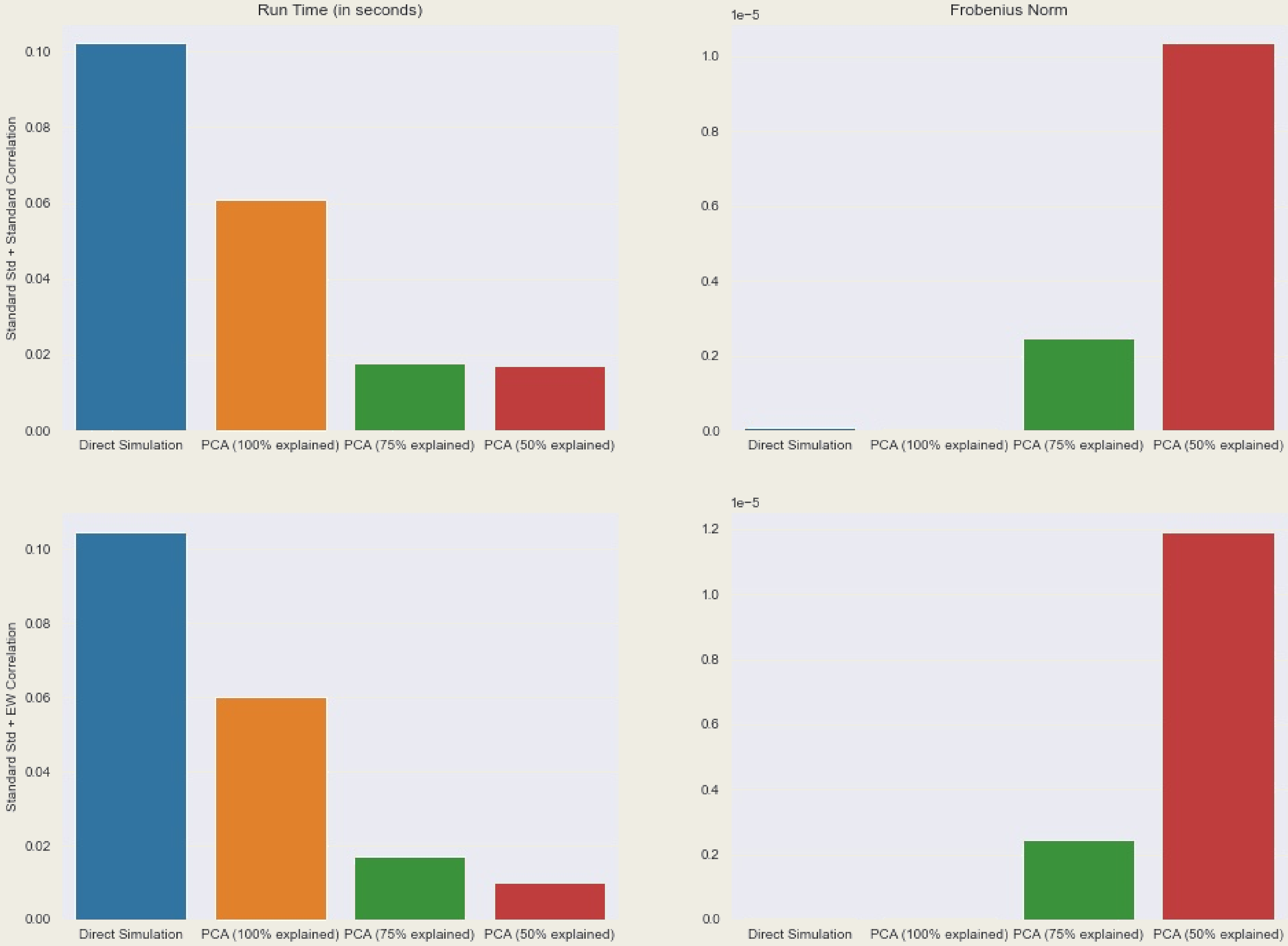
Run Time (seconds)	Direct Simulation	PCA 100% explained	PCA 75% explained	PCA 50% explained
std & correlation	0.126219	0.064827	0.018953	0.010971
std & EW correlation	0.126661	0.063829	0.017951	0.010972
EW std & correlation	0.129653	0.062834	0.017952	0.011969
EW std & EW correlation	0.149073	0.068400	0.017954	0.013962

Frobenius Norm

Frobenius Norm	Direct Simulation	PCA 100% explained	PCA 75% explained	PCA 50% explained
std & correlation	4.866301	6.633571	270.392342	1114.777441
std & EW correlation	4.488397	4.954941	283.530388	1254.964411
EW std & correlation	7.057699	4.196981	249.222515	1042.556673
EW std & EW correlation	4.547950	5.540411	245.092077	1191.059750



Run Time (in seconds)
Frobenius Norm
Standard Std + Standard Correlation
Standard Std + EW Correlation
1e−5
Direct Simulation
PCA (100% explained)
PCA (75% explained)
PCA (50% explained)


Run Time (in seconds)
Frobenius Norm
1e−5
Standard Std + Standard Correlation
Standard Std + EW Correlation
0.00
0.02
0.04
0.06
0.08
0.10
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Direct Simulation
PCA (100% explained)
PCA (75% explained)
PCA (50% explained)