1 Introduction

The covariance matrix is important in many areas, like constructing a Mean-Variance portfolio. High dimensional covariance matrix Σ_T contains huge estimation error given small data X_T. One potential method to solve this problem is shrinkage method[1]. One simple way to do this would be taking the convex combination between covariance matrix and a shrinkage target. This target could be identity matrix or other customer-tailered matrix. Upon this approach we can reduce the variance of sample covariance matrix because of introducing an biased but low-variant item. An alternative aspect thinking this shrinkage method would be a trade off between bias and variance.

2 Methodology

2.1 Shrinkage of the covariance matrix

Let $\mathbf{H}_t$ denotes covariance matrix that used to construct GMV portfolio.[2][1]

$$\mathbf{H}_t = \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t \tag{1}$$

where $\mathbf{D}_t$ denoting the diagonal matrix of volatilities and $\mathbf{R}_t$ is the correlation matrix. By shrinking the correlation matrix $\mathbf{R_t}$,

$$\tilde{\mathbf{R}}_t = (1-\alpha)\mathbf{R}_t + \alpha\mathring{\mathbf{R}}_t \tag{2}$$

We can shrink the covariance matrix $\mathbf{H_t}$ indirectly,

$$\tilde{\mathbf{H}}_t = \mathbf{D}_t\tilde{\mathbf{R}}_t\mathbf{D}_t = \mathbf{D}_t\left[(1-\alpha)\mathbf{R}_t + \alpha\mathring{\mathbf{R}}_t\right]\mathbf{D}_t = (1-\alpha)\mathbf{H}_t + \alpha\mathbf{D}_t\mathring{\mathbf{R}}_t\mathbf{D}_t \tag{3}$$

2.2 Shrank GMV portfolio

We can use shrank covariance matrix $\tilde{\mathbf{H}}_t$ to construct new shrank GMV portfolio.

$$\begin{aligned}\mathbf{x}_t(\alpha) &= \frac{\tilde{\mathbf{H}}_t^{-1}\mathbf{1}}{\mathbf{1}'\tilde{\mathbf{H}}_t^{-1}\mathbf{1}} \\ &= \frac{\mathbf{D}_t^{-1}\left[(1-\alpha)\mathbf{R}_t + \alpha\mathring{\mathbf{R}}_t\right]^{-1}\mathbf{D}_t^{-1}\mathbf{1}}{\mathbf{1}'\mathbf{D}_t^{-1}\left[(1-\alpha)\mathbf{R}_t + \alpha\mathring{\mathbf{R}}_t\right]^{-1}\mathbf{D}_t^{-1}\mathbf{1}}\end{aligned}$$

As shown above, $\mathbf{x}_t(\alpha)$ is a function of α, $\mathbf{R}_t$ and $\mathring{\mathbf{R}}_t$. Here, $\mathbf{R}_t$ and $\mathring{\mathbf{R}}_t$ is given from data and α is decision variable that we want to control to optimize our objective.

2.3 Reinforcement Learning

Reinforcement Learning(RL) can be used for this non-linear optimization problem finding the optimal α with changing time period. In particular, two kinds of RL method could be applied.One is Value-based methods like Q-learning, the other one is Policy-based methods likepolicy gradient. While leaving the discussion between these two kinds of methods, we firstly introduce the basic elements for RL problem.

- State space $S = \{r_{p,t}\}$
- Action space $A = [0, 1]$
- Reward $R_t = \frac{\mathbf{E}[r_p - r_f]}{\sigma[r_p - r_f]}$
- Objective function $\max_{\pi} E_{\pi}\left[R_1 + \gamma R_2 + \cdots + \gamma^{T-1} R_T\right]$

Where, $\{r_{p,t}\}$ is portfolio return. π is policy, which is discussed under each method below. More measurable variables could be added into state space to attain a more precise description of exposure to information environment. Also, reward at each time R_t could be log return or some risk-adjusted measurement.

The first method, Value-based methods like Q-learning[3], can be firstly applied for gaining some insight toward this problem while searching α^* in a discrete space. Taking SARSA[3] as an example, for each episode we go through the training data once. For each step in an episode the agent observes current state S_t and take action A_t according to current policy π. Then agent observes reward R_{t+1} and next time state S_{t+1}, and perform another action A_{t+1} again. After gather these information we are ready to update our Q function, the value for each state and action pair. The updating rule described as below.

$$Q\left(S_t, A_t\right) \leftarrow Q\left(S_t, A_t\right) + \alpha\left[R_{t+1} + \gamma Q\left(S_{t+1}, A_{t+1}\right) - Q\left(S_t, A_t\right)\right]$$

Where α is learning rate. We can see this as an approximation for value function

$$Q(s, a) = \mathbb{E}\left[R_{t+1} + \gamma Q\left(S_{t+1}, A_{t+1}\right) \mid S_t = s, A_t = a\right]$$

And our policy for value-based method is deterministic all the time.

$$\pi(a \mid s) = \begin{cases} \epsilon/m + 1 - \epsilon & \text{if } a = \underset{a' \in \mathcal{A}}{\operatorname{argmax}} Q(s, a') \\ \epsilon/m & \text{otherwise} \end{cases}$$

where ϵ is a predefined probability to choose an action at random. m is the number of action could be choosen randomly.

In summary, we obtain a huge table of Q value for each state and action pair (s, a). Each iteration we going through training data, we update this Q table until it doesn't change any more.

After the convergence, our desired policy π^* is straight forward to simply act greedily to choose the action A_t giving us highest Q value $Q(S_t, A_t)$ under each state S_t.

After gaining some insight from using value-based methods, policy-based methods like policy gradient can be used. And policy gradient method would perform better, since it works better on continuous space. Policy gradient method share a similar structure while don't bother with value function $Q(s,a)$ or $V(s)$. Instead we introduce a new variable θ named policy parameter, which governs the new policy function $\pi(a \mid s, \theta)$. A new performance function $J(\theta)$ is also introduced to measure agent's performance. We can choose our performance function as

$$J(\theta) = E_{\pi_\theta}\left[R_1 + \gamma R_2 + \cdots + \gamma^{T-1} R_T\right]$$

For implementing policy gradient we take REINFORCE algorithm as an example. In each episode we go through the training data once. For each step in an episode, the agent takes action A_t given current state S_t and policy $\pi(a \mid s, \theta)$, then observes reward R_{t+1} and next state S_{t+1}. This process is looped til the end of each episode. At the end of each episode, the agent update its policy by updating policy parameter θ_t

$$\theta_{t+1} = \theta_t + \alpha \widehat{\nabla J\left(\theta_t\right)}$$

where α is learning rate. We use $\widehat{\nabla J\left(\theta_t\right)}$the gradient of performance function with respect to policy parameter θ_t to update θ_t. We learn the value of θ which controls the policy giving us the optimal policy.

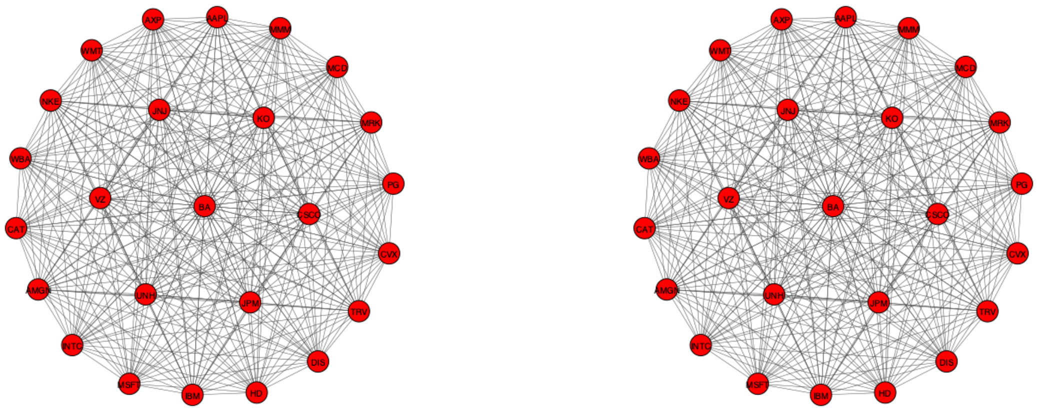
3 Empirical Results

3.1 Data description

We include two examples of correlation matrix and TBN for 1996 in the appendix.

3.1.1 Correlation-Based Networks

Given stock price time series, we can get stock relation matrix $\mathbf{R}_t$ for each year. An example of correlation-based network graph for 1996 and 2008 are shown below.



(a) year 1996

(b) year 2008

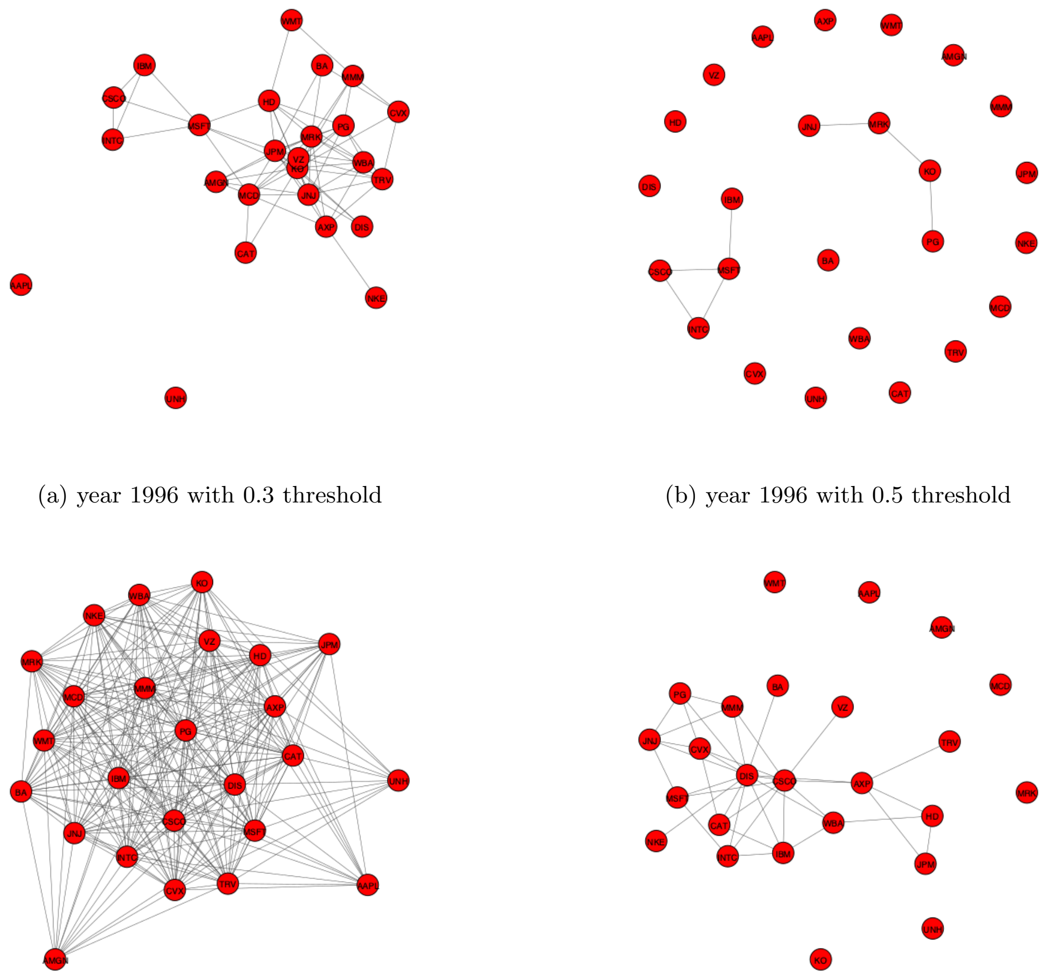
Figure 1: correlation-based network graph for 1996 and 2008

In graph 1a, all diagonal elements on its original adjacency matrix are non-zero. So it contains self loop for each node. We remove this loop in following graphs, making them simple graphs to concentrate more on the relationship between each other. The simple graph for 1996 and 2008 correlation-based network are shown above.

In graph 1b, we are plotting the adjacency matrix $\mathbf{R}_{2008}$ for 26 stocks return in 2008. Each vertex(node) on the graph represent a stock. The edge(line) linking each node is the correlation in $\mathbf{R}_{2008}$. Near node has higher correlation.

To look into these graphs, we remove all correlation below 50% and re-do these graphs. We find that the link is disappearing when correlation threshold is increasing. It appears for both years. On higher correlation threshold, only companies linked closely exist edge then. For example, in graph 2b, IBM, MSFT and INTC, which are all technology companies, still has link between each other under high correlation threshold.

Comparing 2008 and 1996 networks, while they appears similar under 0 threshold, 1996 graph has more pars layout under 0.5 threshold. This means the same group companies in 2008 has higher correlation than 1996.



(a) year 1996 with 0.3 threshold

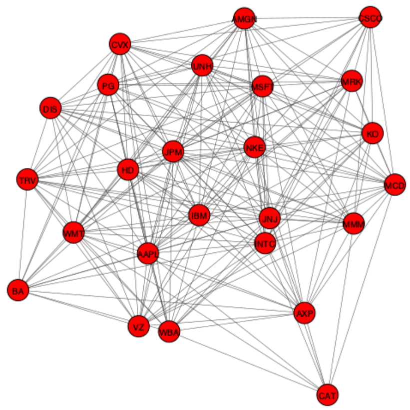
(b) year 1996 with 0.5 threshold

(c) year 2008 with 0.5 threshold

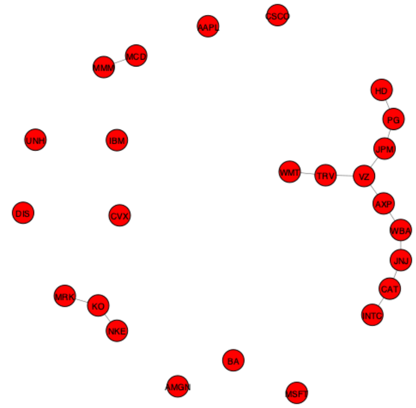
(d) year 2008 with 0.7 threshold

Figure 2: correlation-based network graph for 1996 and 2008 with parse adjacency matrix

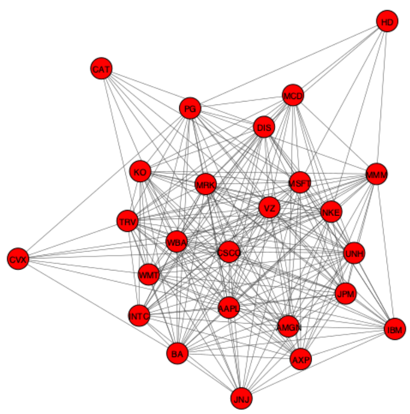
We re-do the similar things for TBN, which is provided by Hoberg-Phillips Advanced Data Options



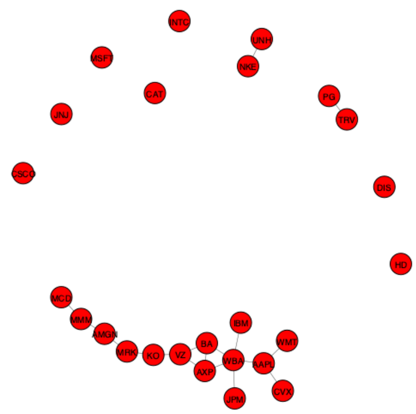
(a) year 1996 with 0 threshold



(b) year 1996 with 0.1 threshold



(c) year 2008 with 0 threshold



(d) year 2008 with 0.1 threshold

Figure 3: text-based network graph for 1996 and 2008 with various thresholds

Comparing TBN with correlation-based network, we find that TBN has fewer links between each node. And the links are also weaker than correlation-based network. Then we can do some linear combination between these two to form a new shrank correlation matrix $\tilde{\mathbf{R}}_t$

3.2 Back-testing

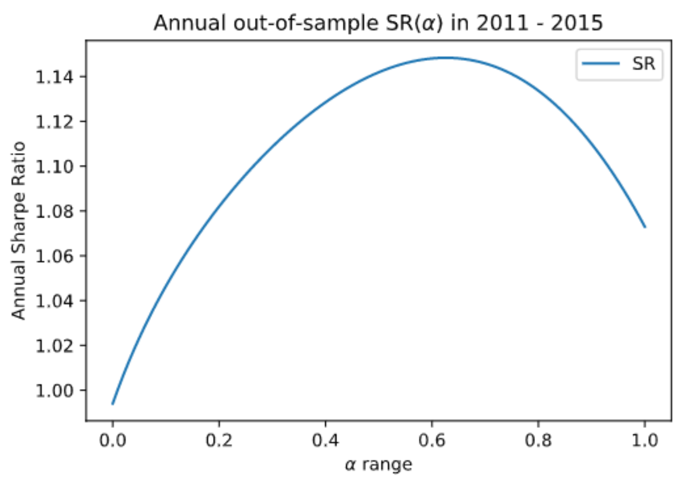
Following above we can construct GMVP $\mathbf{x}_t(\alpha)$ at the end of each year for a givce shrinkage intensity α. Hence we can compute out-of-sample GMVP return $r_{p,t+1}$ using following year's stock returns $\mathbf{Y}_{t+1}$. The concrete equation is described bellow,

$$r_{p,t+1}(\alpha) = \mathbf{Y}_{t+1}\mathbf{x}_t(\alpha) \tag{4}$$

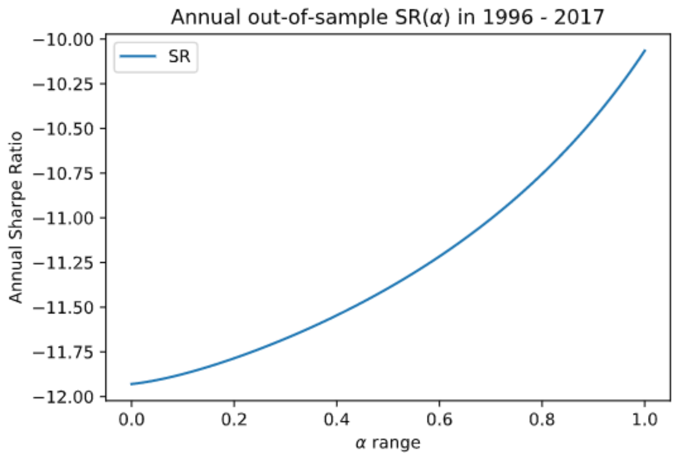
Further, we can compute Sharpe ratio $SR(\alpha)$ for different time series $r_{p,t+1}$ given different α.

$$SR = \frac{\mathbf{E}[r_p - r_f]}{\sigma[r_p - r_f]} \tag{5}$$

Given a concrete time series r_p, we output a Sharpe ratio as above. Here, the risk free rate is provided by Kenneth R. French - Data Library. We repeat this process to back test on selected time range. An example of back testing from 2011 to 2015 is shown below.



In the sample above, the optimal shrinkage intensity α^* exists in middle, which is about 0.6 giving us the highest annual Sharpe ratio. While it's not always the case like above sample. Actually, $SR(\alpha)$ is also time dependent. When we expand the time period from 1996 to 2017, optimal α^* lie on the boundary and $SR(\alpha)$ has different convex shape.



For future research, I think we need to keep an eye on time range t while optimize the shrinkage intensity α.

Appendix

Table 1: correlation matrix of stock return in 1996

	AAPL	AMGN	AXP	BA	CAT	CSCO	CVX	DIS	HD	IBM	INTC	JNJ	JPM	KO	MCD	MMM	MRK	MSFT	NKE	PG	TRV	UNH	VZ	WBA	WMT
AAPL	1.00	0.05	-0.00	0.06	0.14	0.22	0.04	0.01	0.14	0.12	0.24	0.08	0.08	0.14	0.13	0.20	0.03	0.15	0.03	0.07	0.01	0.16	0.07	0.03	0.07
AMGN	0.05	1.00	0.25	0.22	0.13	0.23	0.17	0.27	0.23	0.22	0.09	0.30	0.32	0.34	0.33	0.17	0.33	0.24	0.08	0.25	0.24	0.08	0.16	0.27	0.19
AXP	-0.00	0.25	1.00	0.24	0.28	0.25	0.19	0.29	0.28	0.20	0.13	0.28	0.44	0.38	0.37	0.20	0.34	0.29	0.31	0.29	0.31	0.20	0.30	0.31	0.20
BA	0.06	0.22	0.24	1.00	0.27	0.29	0.34	0.18	0.24	0.21	0.17	0.19	0.30	0.26	0.21	0.14	0.35	0.22	0.09	0.23	0.27	0.12	0.24	0.21	0.16
CAT	0.14	0.13	0.28	0.27	1.00	0.29	0.14	0.22	0.18	0.29	0.19	0.20	0.17	0.36	0.33	0.20	0.20	0.24	0.12	0.22	0.20	0.13	0.15	0.23	0.14
CSCO	0.22	0.23	0.25	0.29	0.29	1.00	0.17	0.21	0.30	0.41	0.50	0.21	0.29	0.27	0.23	0.21	0.22	0.56	0.21	0.15	0.11	0.26	0.13	0.21	0.18
CVX	0.04	0.17	0.19	0.34	0.14	0.17	1.00	0.17	0.18	0.20	0.08	0.21	0.27	0.31	0.26	0.32	0.30	0.16	0.18	0.27	0.34	0.11	0.13	0.27	0.19
DIS	0.01	0.27	0.29	0.18	0.22	0.21	0.17	1.00	0.29	0.16	0.18	0.35	0.31	0.36	0.23	0.29	0.29	0.25	0.24	0.22	0.18	0.21	0.20	0.28	0.16
HD	0.14	0.23	0.28	0.24	0.18	0.30	0.18	0.29	1.00	0.16	0.20	0.30	0.30	0.42	0.27	0.24	0.34	0.32	0.29	0.32	0.27	0.16	0.24	0.38	0.33
IBM	0.12	0.22	0.20	0.21	0.29	0.41	0.20	0.16	0.16	1.00	0.44	0.21	0.26	0.20	0.26	0.14	0.10	0.51	0.05	0.14	0.19	0.11	-0.02	0.20	-0.03
INTC	0.24	0.09	0.13	0.17	0.19	0.50	0.08	0.18	0.20	0.44	1.00	0.19	0.17	0.19	0.12	0.23	0.11	0.60	0.19	0.10	0.18	0.22	0.09	0.07	0.21
JNJ	0.08	0.30	0.28	0.19	0.20	0.21	0.21	0.35	0.30	0.21	0.19	1.00	0.45	0.46	0.37	0.26	0.50	0.25	0.14	0.47	0.31	0.15	0.29	0.31	0.14
JPM	0.08	0.32	0.44	0.30	0.17	0.29	0.27	0.31	0.30	0.26	0.17	0.45	1.00	0.50	0.40	0.26	0.45	0.34	0.18	0.36	0.33	0.20	0.30	0.32	0.23
KO	0.14	0.34	0.38	0.26	0.36	0.27	0.31	0.36	0.42	0.20	0.19	0.46	0.50	1.00	0.41	0.41	0.52	0.35	0.23	0.56	0.41	0.19	0.28	0.34	0.27
MCD	0.13	0.33	0.37	0.21	0.33	0.23	0.26	0.23	0.27	0.26	0.12	0.37	0.40	0.41	1.00	0.27	0.32	0.32	0.19	0.32	0.28	0.19	0.20	0.30	0.20
MMM	0.20	0.17	0.20	0.14	0.20	0.21	0.32	0.29	0.24	0.14	0.23	0.26	0.26	0.41	0.27	1.00	0.31	0.29	0.24	0.40	0.21	0.22	0.25	0.24	0.31
MRK	0.03	0.33	0.34	0.35	0.20	0.22	0.30	0.29	0.34	0.10	0.11	0.50	0.45	0.52	0.32	0.31	1.00	0.24	0.17	0.46	0.33	0.22	0.23	0.34	0.22
MSFT	0.15	0.24	0.29	0.22	0.24	0.56	0.16	0.25	0.32	0.51	0.60	0.25	0.34	0.35	0.32	0.29	0.24	1.00	0.25	0.20	0.26	0.18	0.15	0.23	0.21
NKE	0.03	0.08	0.31	0.09	0.12	0.21	0.18	0.24	0.29	0.05	0.19	0.14	0.18	0.23	0.19	0.24	0.17	0.25	1.00	0.09	0.11	0.21	0.18	0.13	0.16
PG	0.07	0.25	0.29	0.23	0.22	0.15	0.27	0.22	0.32	0.14	0.10	0.47	0.36	0.56	0.32	0.40	0.46	0.20	0.09	1.00	0.33	0.19	0.28	0.32	0.18
TRV	0.01	0.24	0.31	0.27	0.20	0.11	0.34	0.18	0.27	0.19	0.18	0.31	0.33	0.41	0.28	0.21	0.33	0.26	0.11	0.33	1.00	0.13	0.26	0.28	0.20
UNH	0.16	0.08	0.20	0.12	0.13	0.26	0.11	0.21	0.16	0.11	0.22	0.15	0.20	0.19	0.19	0.22	0.22	0.18	0.21	0.19	0.13	1.00	0.05	0.09	0.09
VZ	0.07	0.16	0.30	0.24	0.15	0.13	0.13	0.20	0.24	-0.02	0.09	0.29	0.30	0.28	0.20	0.25	0.23	0.15	0.18	0.28	0.26	0.05	1.00	0.30	0.22
WBA	0.03	0.27	0.31	0.21	0.23	0.21	0.27	0.28	0.38	0.20	0.07	0.31	0.32	0.34	0.30	0.24	0.34	0.23	0.13	0.32	0.28	0.09	0.30	1.00	0.21
WMT	0.07	0.19	0.20	0.16	0.14	0.18	0.19	0.16	0.33	-0.03	0.21	0.14	0.23	0.27	0.20	0.31	0.22	0.21	0.16	0.18	0.20	0.09	0.22	0.21	1.00

Table 2: text based network in 1996

	AAPL	AMGN	AXP	BA	CAT	CSCO	CVX	DIS	HD	IBM	INTC	JNJ	JPM	KO	MCD	MMM	MRK	MSFT	NKE	PG	TRV	UNH	VZ	WBA	WMT
AAPL	1.00	0.00	0.03	0.04	0.03	0.00	0.04	0.02	0.02	0.03	0.01	0.04	0.04	0.00	0.00	0.00	0.01	0.05	0.01	0.09	0.02	0.01	0.05	0.04	0.00
AMGN	0.00	1.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.02	0.02	0.00	0.08	0.01	0.07	0.03	0.05	0.00	0.01	0.00	0.00	0.00	0.00	0.02
AXP	0.03	0.00	1.00	0.01	0.01	0.00	0.00	0.00	0.01	0.00	0.03	0.00	0.08	0.00	0.01	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.13	0.06	0.01
BA	0.04	0.00	0.01	1.00	0.00	0.00	0.06	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.04	0.06	0.09
CAT	0.03	0.00	0.01	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00
CSCO	0.00	0.00	0.00	0.00	0.00	1.00	0.03	0.00	0.00	0.00	0.01	0.00	0.09	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
CVX	0.04	0.01	0.00	0.06	0.00	0.03	1.00	0.04	0.03	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.02	0.02	0.01	0.04	0.00	0.00	0.03	0.08
DIS	0.02	0.00	0.00	0.00	0.00	0.00	0.04	1.00	0.00	0.00	0.00	0.00	0.04	0.00	0.04	0.00	0.00	0.05	0.01	0.01	0.02	0.01	0.00	0.00	0.03
HD	0.00	0.02	0.01	0.09	0.00	0.00	0.08	0.03	1.00	0.03	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.14	0.02	0.00	0.03	0.00	0.00
IBM	0.02	0.01	0.01	0.01	0.00	0.00	0.03	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.03	0.01	0.00	0.05	0.00	0.00	0.05	0.00	0.08	0.02	0.02
INTC	0.03	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.04	0.00	1.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.01	0.08	0.01	0.00	0.00	0.03
JNJ	0.01	0.02	0.03	0.00	0.12	0.01	0.00	0.00	0.00	0.00	0.00	1.00	0.06	0.01	0.00	0.00	0.09	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00
JPM	0.04	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.00	1.00	0.00	0.00	0.01	0.02	0.05	0.00	0.02	0.00	0.01	0.13	0.07	0.02
KO	0.04	0.00	0.08	0.00	0.00	0.09	0.04	0.04	0.00	0.00	0.06	0.02	0.00	1.00	0.04	0.00	0.05	0.04	0.10	0.00	0.00	0.02	0.00	0.00	0.00
MCD	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	1.00	0.10	0.04	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MMM	0.00	0.01	0.01	0.00	0.02	0.01	0.00	0.04	0.03	0.02	0.00	0.00	0.04	0.01	0.00	1.00	0.00	0.03	0.03	0.00	0.06	0.04	0.00	0.00	0.00
MRK	0.00	0.07	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.10	0.00	0.00	1.00	0.01	0.00	0.00	0.00	0.02	0.00	0.00	0.00
MSFT	0.01	0.03	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.09	0.02	0.05	0.04	0.00	0.02	0.00	1.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
NKE	0.05	0.05	0.02	0.00	0.00	0.00	0.02	0.05	0.05	0.01	0.01	0.05	0.04	0.03	0.03	0.01	0.02	0.00	1.00	0.02	0.03	0.00	0.00	0.04	0.03
PG	0.01	0.00	0.00	0.00	0.00	0.01	0.02	0.01	0.00	0.00	0.00	0.00	0.10	0.00	0.03	0.00	0.01	0.02	0.00	1.00	0.06	0.07	0.00	0.00	0.00
TRV	0.09	0.01	0.00	0.07	0.00	0.00	0.01	0.01	0.00	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	1.00	0.00	0.02	0.03	0.14
UNH	0.02	0.00	0.00	0.00	0.00	0.00	0.04	0.02	0.05	0.08	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.03	0.06	0.00	0.00	1.00	0.00	0.00	0.02
VZ	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.01	0.02	0.00	0.04	0.02	0.00	0.00	0.07	0.00	0.15	0.00	1.00	0.00	0.00
WBA	0.05	0.00	0.13	0.04	0.00	0.00	0.00	0.00	0.08	0.00	0.01	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	1.00	0.03
WMT	0.04	0.00	0.06	0.06	0.02	0.00	0.03	0.00	0.02	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.03	0.00	0.00	0.05	0.00	1.00

References

[1] Olivier Ledoit and Michael Wolf. The power of (non-) linear shrinking: a review and guide to covariance matrix estimation. *University of Zurich, Department of Economics, Working Paper*, (323), 2020.

[2] Gerard Hoberg and Gordon Phillips. Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5):1423–1465, 2016.

[3] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.