Principal Component Analysis and Portfolio Optimization

Applied Finance Project, Ji Tan (ji_tan@mfe.berkeley.edu)

Supervisor: Professor Hayne Leland (leland@haas.berkeley.edu)

Abstract

Principle Component Analysis (PCA) is one of the common techniques used in Risk modeling, i.e. statistical factor models. When using PCA to estimate the covariance matrix, and applying it to portfolio optimization, we formally analyze its performance, and find positive results in terms of portfolio efficiency (Information Ratio) and transaction cost reduction. We also propose using PCA to manage beta against alpha, and show how to apply the idea within Black-Litterman framework. Finally, we invent the technique "Mean-Reverting PCA" to improve the stability of conventional PCA analysis.

Table of content

Introduction:

In this paper, we first address two interrelated problems in portfolio construction. People usually deal with them separately. We find principle component analysis (PCA) could address both problems at the same time.

The first problem: if we follow standard mean-variance optimization (MVO) of Markowitz (1952, 1959), we would very likely get extreme positions for many assets. Such result would not be practical for most portfolio managers. There is a trade-off between MV efficiency and acceptable weights. When dealing with such trade-off, people use various techniques. Some are purely mathematical, while others do have economical meanings, notably Black-Litterman (1993), which can regarded as blending stable equilibrium weights and unstable MVO weights.

The second problem: When a portfolio manager changes his expectations of asset returns over time, and that, in standard MVO, also leads to substantial rebalancing, incurring huge transaction cost. People usually add another term to the MV objective function, to penalize the big deviations from the original weights. This method is mathematically optimal, but not tractable. In other words, it is a black box.

Now we turn to PCA, which gives us a series of independent "factors" in the market. In approximation, we can maintain those more important factors and ignore the others. *A lower rank PCA approximation will give us very conservative weights, and simultaneously make the weights insensitive to the expected returns.* If we can still keep a good Sharpe Ratio, both problems are solved. And solving the second problem means we can easily achieve a higher after-transaction-cost Sharpe Ratio. This method is shown in Chapter 1 and 2, and empirical study in Chapter 5.

Sometimes by using PCA approximation, the Sharpe Ratio could be less appealing. Nonetheless, PCA separates major market factors from asset-class specific features. In that way, it reduces our investment decision to 2 independent sets of choices: one is how to allocate our exposure to the major market factors, which would give us a stable portfolio similar to what we get above; the other is how to allocate our exposure to asset-class specific features, which is alpha-like portfolio with less stability.

This idea is an extension of Leibowitz-Bova (2005), and is applicable to every aspect of portfolio management. As we can see, it draws similarity to Black-Litterman in terms of blending two things together, one stable and the other unstable. To give an example of this idea's application, we show how to combine it with Black-Litterman approach. This example is shown in Chapter 3, and empirical study in Chapter 6. In Chapter 4, we discuss the general properties of PCA as a starting point for subsequent empirical studies, and propose Mean-Reverting PCA, the detail of which is put in the appendix.

Chapter 1: PCA: transform the assets to orthogonal factors

1.1 Mean Variance Optimization and its difficulties

Denote the assets as X_i, i=0, 1,…, n. When i=0 it is risk free asset, while others are risky. For this one period problem, each asset X_i has unit price at the beginning. Our initial wealth is also 1. The position in each asset is w_i, so we can express our portfolio as: $w_0X_0 + w_1X_1 + \cdots + w_nX_n$.

Now we will group the risky assets as a vector X (not including X_0), its weights vector w; its expected return is μ, and the covariance matrix Σ. We will also denote risk free rate r_0, risk aversion parameter γ, and vector of 1s **1**. The standard MVO problem can be expressed as:

$$\max_{w_0,w}\{w_0r_0 + w'\mu - \gamma w'\Sigma w\}$$

$$\text{s.t.}\, w_0 + w'\mathbf{1} = 1$$

Subtracting the risk free rate, we can rewrite it as:

$$\max_{w}\{w'(\mu - r_0\mathbf{1}) - \gamma w'\Sigma w\}$$

Consequentially the optimal weights are

$$w = \frac{1}{2\gamma}\Sigma^{-1}(\mu - r_0\mathbf{1})$$

Consider Eigen Value Decomposition of the covariance matrix,

$$\Sigma = H'VH$$

H is orthogonal matrix, V is diagonal matrix, and $\Sigma^{-1} = H'V^{-1}H$

$$w = \frac{1}{2\gamma}H'V^{-1}H(\mu - r_0\mathbf{1})$$

If V has some small values on its diagonal, V^{-1} would be big, which may also cause extreme weights w. This is the first problem in the introduction: given a set of arbitrary expected returns, mean variance optimization tends to give extreme positions.

When we consider small changes, the relation is exactly the same because of linearity.

$$dw = \frac{1}{2\gamma}H'V^{-1}Hd\mu$$

This is the second problem in the introduction: a small change in the expected returns usually leads to big change in positions.

1.2 Asset transform (PCA)

We may rewrite the above equation as:

$$dHw = \frac{1}{2\gamma}V^{-1}dH\mu$$

$H\mu$ is the orthogonal transform of expect return vector. This inspires us to do an orthogonal transform on the risky assets $X \rightarrow HX$, so each New asset is a linear combination of the original assets, i.e. a portfolio. To facilitate the calculation, we make the weights in each portfolio sum to 1, so that its price is also 1.

Let H_i be the i^{th} row of H (Here we assume $H_i\mathbf{1} \neq 0$).

$$H_i H_j' = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{otherwise} \end{cases}$$

Denote

$$\frac{1}{H_i\mathbf{1}} \equiv A_i, A \equiv \text{diag}(A_1, \cdots, A_n)$$

We have the transform $\widetilde{X} \equiv AHX$, or

$$\widetilde{X}_i \equiv A_i H_i X = A_i \sum_{j=1}^{n} H_{i,j} X_j = \frac{\sum_{j=1}^{n} H_{i,j} X_j}{\sum_{j=1}^{n} H_{i,j}}$$

$\widetilde{X}$ can be regarded as independent "factors" in the market. Each factor can also be regarded as a *Portfolio* of original assets. This idea is also called Principal Component Analysis, or PCA.

1.3 Property of the factors

The new factors have the following mean and covariance.

$$\tilde{\mu} = AH\mu$$
$$\tilde{\Sigma} = AH\Sigma H'A' = A^2 V$$

$\tilde{\Sigma}$ is diagonal matrix diag{σ_i^2}. Denote V= $\text{diag}(v_1, \cdots, v_n)$, we have $\sigma_i^2 = v_i A_i^2$.

The optimal weight on these factors is

$$\widetilde{w} = \frac{1}{2\gamma}\tilde{\Sigma}^{-1}(\tilde{\mu} - r\mathbf{1}) = A^{-2}\left(\frac{1}{2\gamma}V^{-1}\right)(AH\mu - r_0\mathbf{1})$$

The corresponding weights in original assets would be $(AH)'\widetilde{w}$.

We find the following properties:

1. $(AH)'\widetilde{w}$ is the same as the original optimal weights w, i.e. the asset allocation is invariant with regard to the transform, and it is equivalent to work on either side and transform the result to the other. (See *proof 1*)
2. The factors have zero correlations with each other. The optimal weight of that particular factor is only determined by the mean and variance of itself. Any change in other factors would not affect this value.

$$\widetilde{w}_i = \frac{\tilde{\mu}_i - r_0}{2\gamma\sigma_i^2}$$

Chapter 2: PCA approach to improve weight stability

2.1 Weight Stabilization

If we discard some factors, and k factors remain, we will have the following optimal weights (When k=n, then it returns to the original problem)

$$\tilde{w}_{(k)} = \frac{1}{2\gamma}\tilde{\Sigma}_{(k)}^{-1}\left(\tilde{\mu}_{(k)} - r_0\mathbf{1}_{(\mathbf{k})}\right)$$

Here footnote (k) denotes the first k elements of the vector or that of the diagonal. According to the 2nd property, this would not change the optimal weights on these k factors.

$$\begin{pmatrix} \tilde{w}_{(k)} \\ 0 \end{pmatrix} = I(n,k)\tilde{w}$$

I(n,k) means identity matrix setting the last n-k diagonals to 0.

However, we may have little idea about these factors, and still want to work with the original assets. So we transform it back

$$w(k) = (AH)'\begin{pmatrix} \tilde{w}_{(k)} \\ 0 \end{pmatrix} = (AH)'I(n,k)\tilde{w}$$

The allocation to each factor can be regarded as a portfolio of original assets that is independent from the portfolio by other factors. The final portfolio is simply the sum of all these factor portfolios. Thus the following (See *proof 2*):

$$w(k) = \frac{1}{2\gamma}\left(\sum_{i=1}^{k}(v_i^{-1}H_i{}'H_i)\right)(\mu - r_0\mathbf{1})$$

$$\frac{dw(k)_q}{d\mu_j} = \frac{1}{2\gamma}\sum_{i=1}^{k} v_i^{-1}H_{i,q}H_{i,j}$$

If we discard the factors with small v_i, we would expect both numbers to be small. Therefore, both problems are solved.

2.2 Impact to Performance

The new optimal portfolio has weight w(k). In the space of new assets, we can express the Sharpe ratio of the portfolio by the Sharpe ratio of each factor. (See *proof 3*)

$$SR(k) = \sqrt{\sum_{i=1}^{k} SR_i^2}$$

As we are discarding some factors, is it possible for us to miss factors with high Sharpe Ratios? From above equation, we can see, it is all about whether the major PCA factors also have bigger Sharpe Ratios. It is usually true, but not always.

To be more specific, we find (See *proof 4*):

$$SR_i = \frac{H_i(\mu - r_0\mathbf{1})}{\sqrt{v_i}} = 2\gamma\sqrt{v_i}H_iw$$

Remember here w is the mean-variance optimal weights on original assets. This equation tells us that, if $H_i w$ are not very different from each other, the main driver of Sharpe ratio would be $\sqrt{v_i}$. Since we keep big v_i, we would also keep big Sharpe Ratios.

2.3 Adjust for transaction cost

You may have your own judgment in choosing the factors, while we can also do it a little systematically. Suppose there is no big Sharpe ratio decline, we would maximize the Sharpe ratio after adjusted by transaction cost.

Consider n different scenarios, in each scenario, one risky asset's expected return increases by 1%, and the others remain unchanged. This would lead to change in the optimal weights. The cost adjusted Sharpe Ratio would be:

$$\begin{aligned} SR_{ca}(k) &= \frac{\text{New Excess Return}(k) - \text{cost}}{\sqrt{\text{New Excess Variance}(k)}} \\ &= \frac{\sum_{i=1}^{k} \widetilde{w}(\text{new})_i (\tilde{\mu}(\text{new})_i - r_0) - \theta \sum_{i=1}^{n} |w(k,\text{new})_i - w(k,\text{old})_i|}{\sqrt{\sum_{i=1}^{k} \widetilde{w}(\text{new})_i^2 \sigma_i^2}} \end{aligned}$$

θ is unit transaction cost, and we assume to be 1%, which is a higher than typical stock cost but lower than bond cost. We would maximize this cost adjusted Sharpe Ratio with regards to k in each scenario and expect to find similar results.

Chapter 3: What if PCA hurts Sharpe Ratio badly?

3.1 Rationale

In PCA, we deleted many minor factors. It is however possible, that these factors turn out to have high Sharpe Ratio ($H_i w$ become the dominant term over $\sqrt{v_i}$), but we delete them. We must note, such case cannot be as common because their $\sqrt{v_i}$ are very small.

Assume it happens, we can however, look at this problem from a different angle. There are a few major factors, very stable, and explain a large amount of risk within each asset, whereas minor factors, much more idiosyncratic, are related to more specific features. We can hold the stable part as a core portfolio, and actively manage the less stable part as opportunities for alpha.

This idea is similar to Leibowitz-Bova (2005). They argue that US Stock is the main source of risk in the market; every other asset class has some loading on US Stock, and a remaining part is orthogonal to it. People can manage these remaining parts for alpha generating, and finally blend with US Stock to achieve a desired risk level.

This idea can be perfectly combined with Black-Litterman (BL) model, and make the model much more powerful. In original BL model, we give views on the relation among the different asset classes. In PCA, we know every asset is driven by some major factors and many minor factors. When we express a view, it involves these major factors as well. The problem is, in asset-class-specific views, we do not have strong views on these common major factors, but rather the remaining part that characterizes the asset-class-specific features. Therefore, by separating the major and minor factors, we can have much more economic sensible views!

It requires us to think in a different paradigm. We should no longer look at the overall performance of each asset class, but only the part that is not correlated with major factors. Each asset class has different beta on major factors, while that is easy to manage. It is easy to understand the major factors, and these betas are quite stable over time. The remaining part of the asset is much more meaningful, because that distinguish this asset from the others.

This idea is also related to the famous Arbitrage Pricing Theory (Stephen Ross, 1976), and we will show the relationship in the appendix.

2.2 Procedure

We assume:

$$\mu \sim N(\mu^*, \tau\Sigma)$$

where μ^* is the equilibrium expected returns, and τ is a small number. In factor space:

$$\tilde{\mu} \sim N(AH\mu^*, \tau A^2 V)$$

Say we have 2 major factors, we can express our view on them like:

$$\begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} \tilde{\mu}_1 \\ \tilde{\mu}_2 \end{bmatrix} \sim N\left(\begin{bmatrix} P_1 \\ P_2 \end{bmatrix}, \Omega_1 \right)$$

Since $X \equiv (AH)^{-1}\widetilde{X}$, write

$$X_i = \sum_{j=1}^{n} C_{i,j}\widetilde{X}_j = C_{i,1}\widetilde{X}_1 + C_{i,2}\widetilde{X}_2 + \sum_{j=3}^{n} C_{i,j}\widetilde{X}_j \equiv C_{i,1}\widetilde{X}_1 + C_{i,2}\widetilde{X}_2 + (1 - C_{i,1} - C_{i,2})Y_i$$

Then Y_i is the part of X_i that is orthogonal to the first 2 factors. We will get something like:

$$Y = B\begin{bmatrix}\widetilde{X}_3 \\ \vdots \\ \widetilde{X}_n\end{bmatrix}$$

So we can express our view on the residual Y as:

$$\begin{bmatrix} Q_{11} & \cdots & Q_{1n} \\ \vdots & \vdots & \vdots \\ Q_{k'1} & \cdots & Q_{k'n} \end{bmatrix} B \begin{bmatrix}\tilde{\mu}_3 \\ \vdots \\ \tilde{\mu}_n\end{bmatrix} \sim N\left(\begin{bmatrix} q_1 \\ \vdots \\ q_{k'}\end{bmatrix}, \Omega_2\right)$$

We can see the major factors are independent from minor factors, in both prior and views, so they can be solved separately. Finally when we get the posterior distribution, we just transform back to asset space.

Chapter 4: Data and initial analysis

4.1 Data description

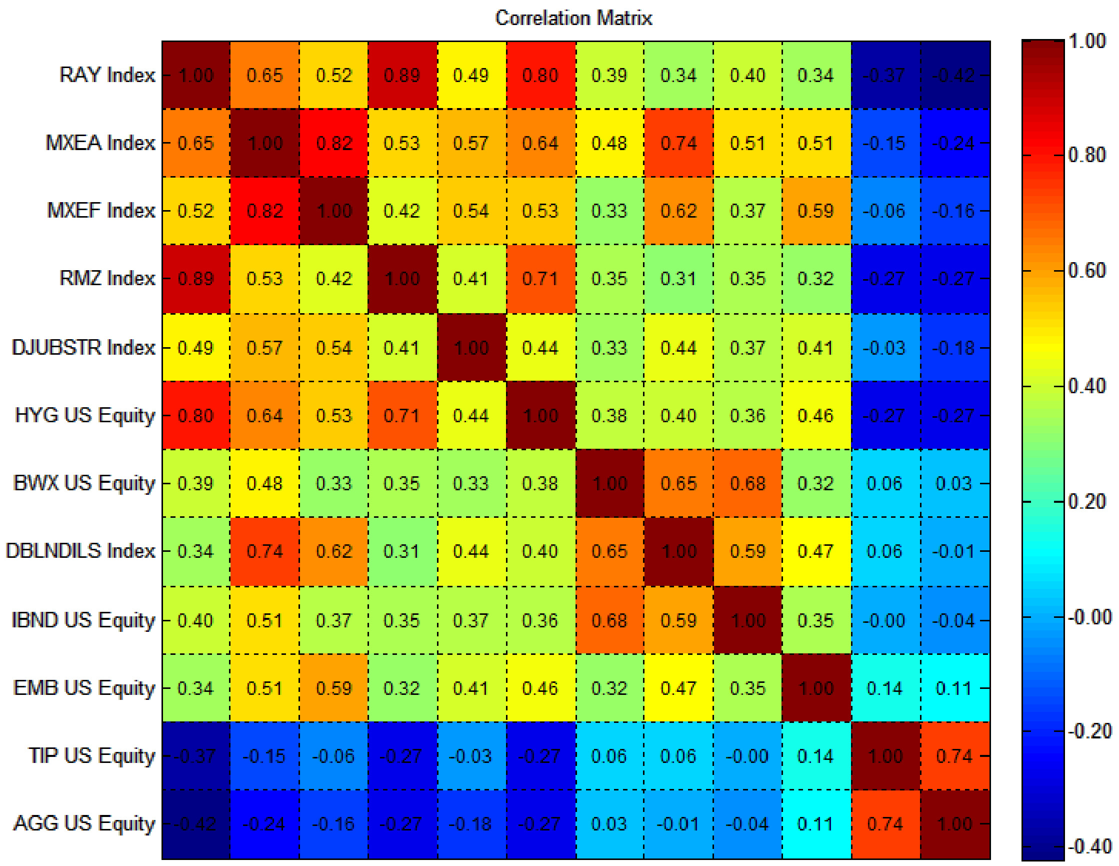
We selected 13 major asset classes in the world, in which there are 12 risky assets and 1 risk free asset. The data is downloaded from Bloomberg.

Major Asset Class	Asset Class	Bloomberg Ticker	Description	
Equity	US Stocks	RAY	Russell 3000	Group 1
	Foreign Developed Market Stocks	MXEA	MSCI EAFE	
	Emerging Market Stocks	MXEF	MSCI Emerging Market Index	
Alternatives	Real Estate	RMZ	MSCI US REIT Index	
	Commodity	DJUBSTR	DJ-UBS Commodity Index	
Bonds	US High Yield Bonds	HYG US	Ishares iBoxx high yield corporate bond fund	
	Foreign Government Bonds	BWX US Equity	SPDR Barclays Cap International Treasury bond ETF	Group 2
	Foreign Government Inflation linked Bonds	DBLNDILS Index	DB Global Govt ex-US Inflation Linked Bond	
	Foreign Corporate Bonds	IBND US Equity	SPDR Barclays Cap International Corporate bond ETF	
	Emerging Market Bonds	EMB US Equity	Ishares JPM USD Emerging Markets bond fund	
	TIPS	TIP US	Ishares Barclays Tips Bond	Group 3
	US Investment Grade Bonds	AGG US	Ishares Barclays Aggregate bond	
Risk Free	US Treasury	H15T1Y	Fed US H.15 T Note Treasury Constant Maturity 1 Year	

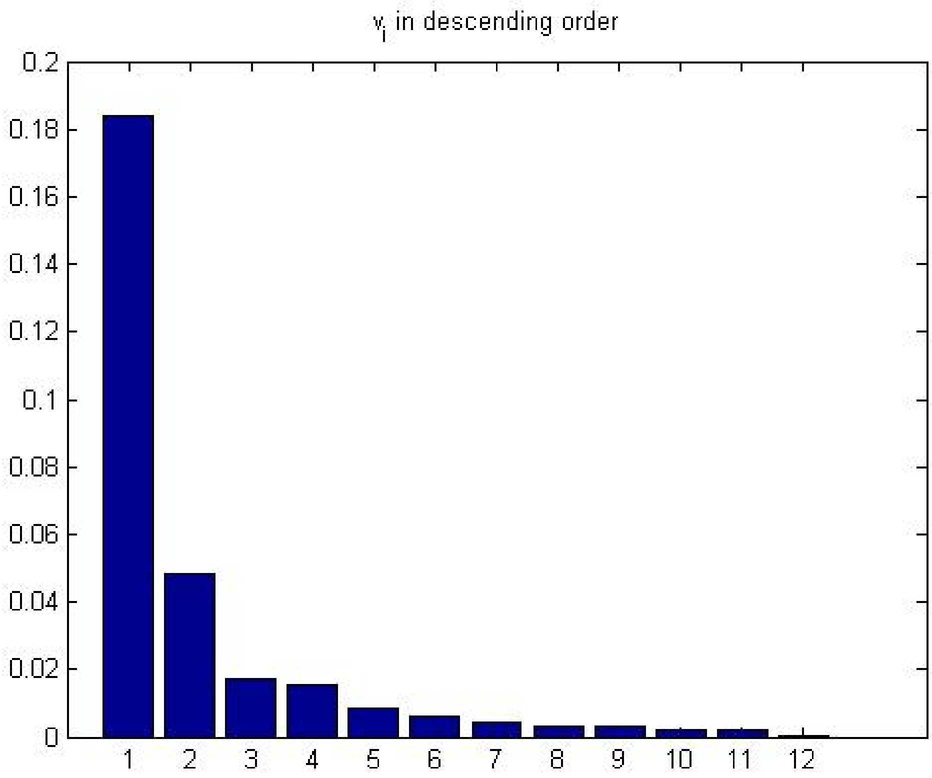
Some of these indexes were only initiated recently, and the most recent is IBND US Equity (Foreign Corporate Bonds), which starts from 20-May-2010. We will do the static analysis from 20-May-2010 to 24-Feb-2012, with 442 common observations of returns. Later we will backtest the result, but because of the data shortage, we will exclude IBND US Equity.

4.2 Covariance structure and PCA transform

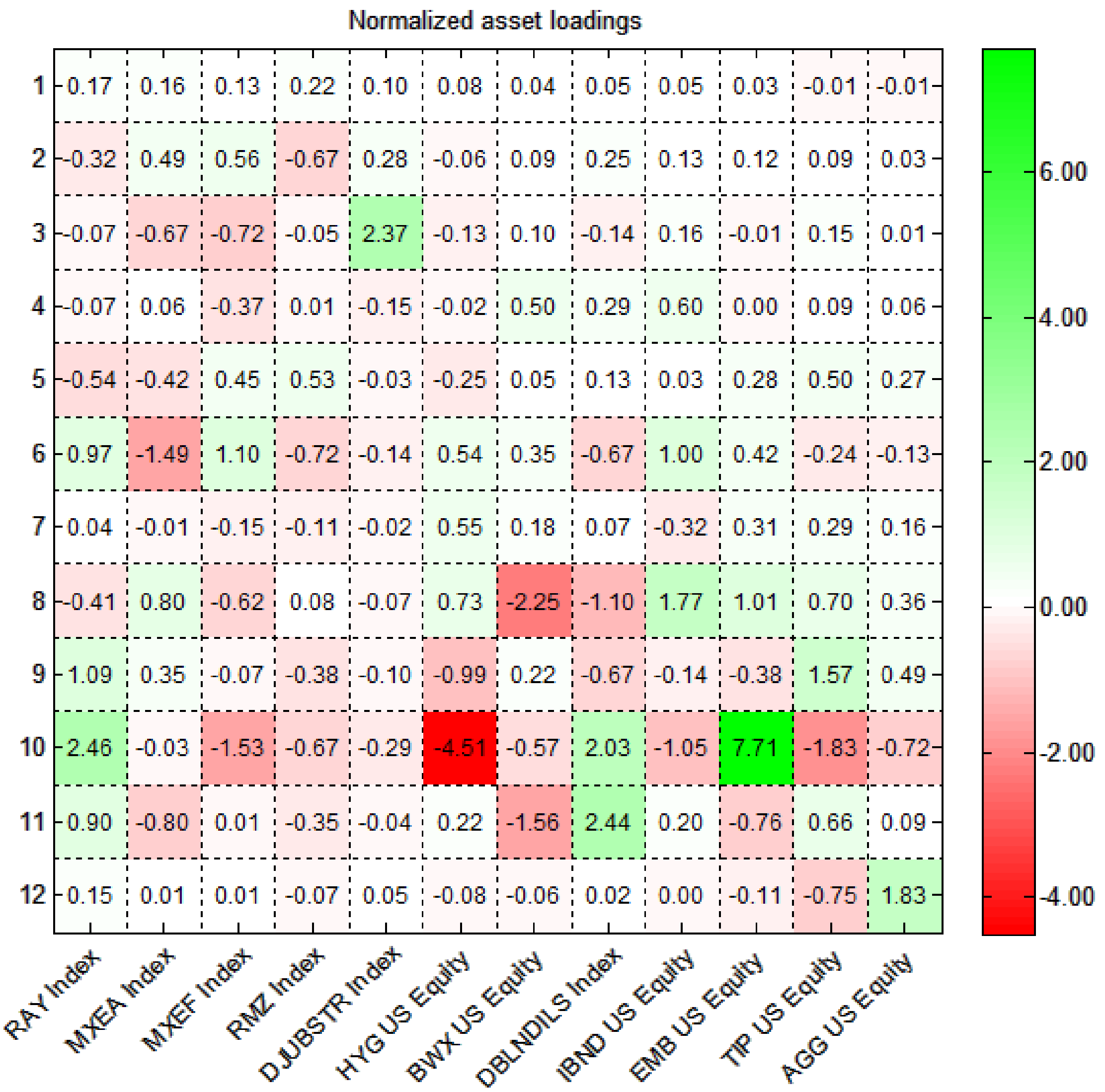
The following graph shows the heat map of correlation matrix. There is obvious clustering, and we can identify 3 economically meaningful groups (marked in the table). The first group is stocks, alternatives and high yield bonds; the second is non-US bonds; and the last is US bonds. There are also some correlations between the first and second. The last group is quite close to risk free asset.



Next, we decompose the covariance matrix to get our diagonal matrix V and transform matrix H. The next graph shows v_i.



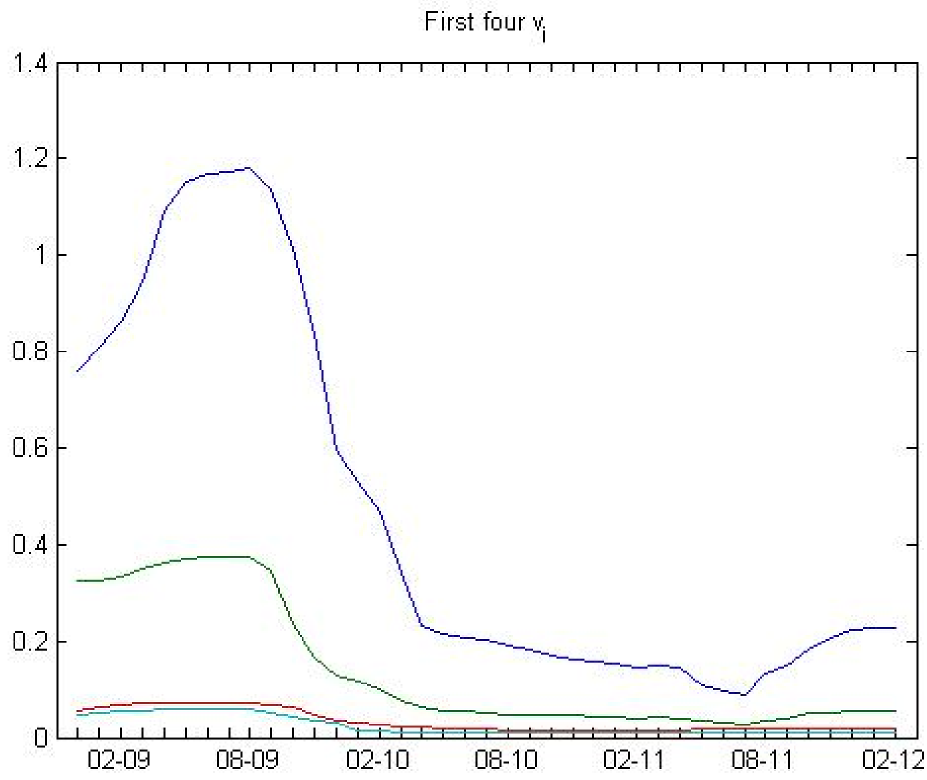
Let’s look at factors $\tilde{X} \equiv AHX$. In the following graph, we show the loadings given by AH. So the first row corresponds to the first factor. It has quite high loadings in Group 1, low loadings in Group 2, and almost nil in Group 3 (about 86%, 17%, -2%). Within each group, the loadings are quite uniform. We can call it “Stock” factor. The second row has quite high loadings in Group 2, small loadings in Group 1 and 3 (about 28%, 59%, 12%), and we can call it “(Foreign) Bond” factor. If we go further towards minor factors, it becomes more difficult for us to see the economical meaning of these factors, and the loadings become mixed.

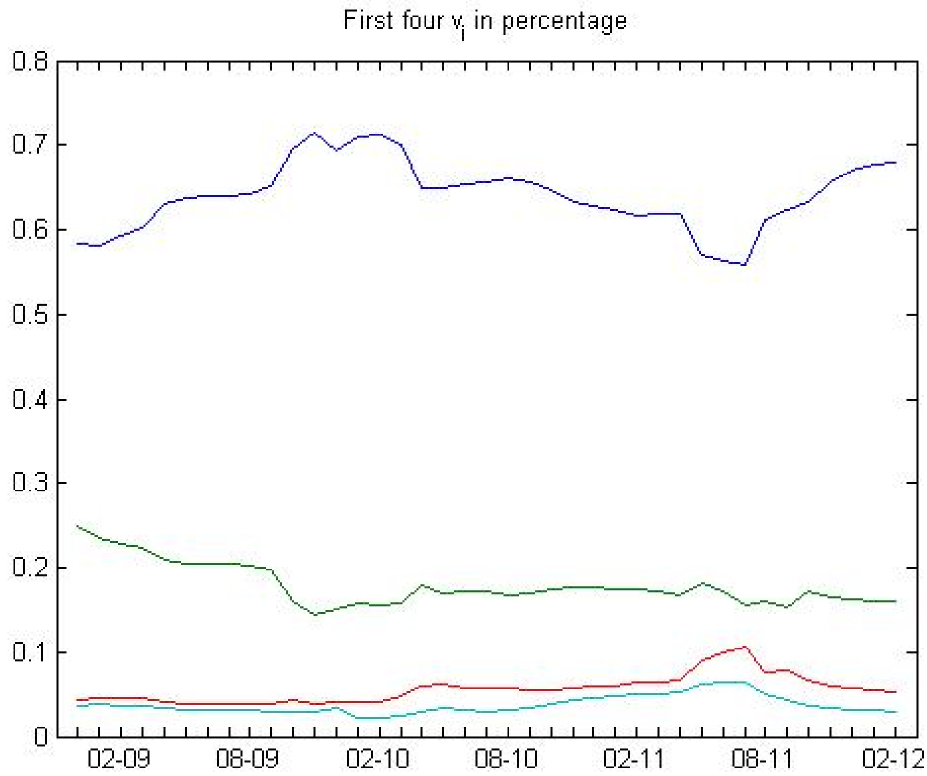


4.3 PCA factors in historical setting

Now we discard IBND US Equity, so our sample starts from 19-Dec-2007. From then, we use a yearly window (250 observations), and conduct above analysis monthly (every other 21 observations). There are 39 times of analysis in total. Later we will show that k=4 is the optimal number of factors from our assumption.

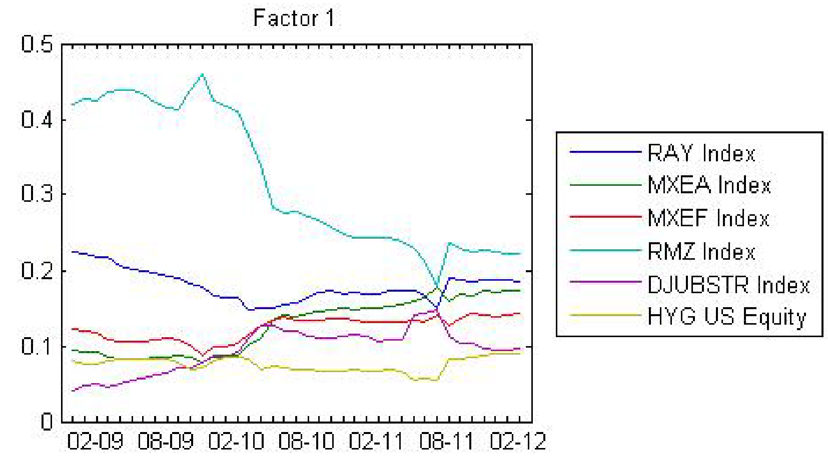
In the following 2 graphs, we show the corresponding v_i. As the market changes, v_i as a measure of volatility also change, but in percentage, v_i stays remarkably constant.

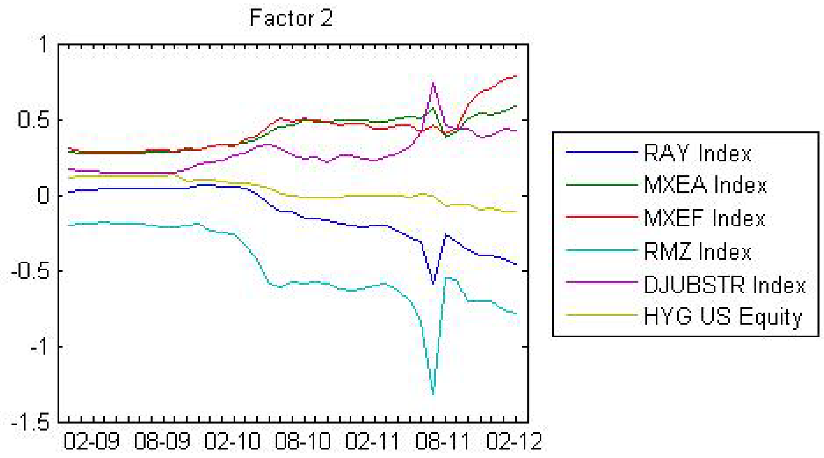


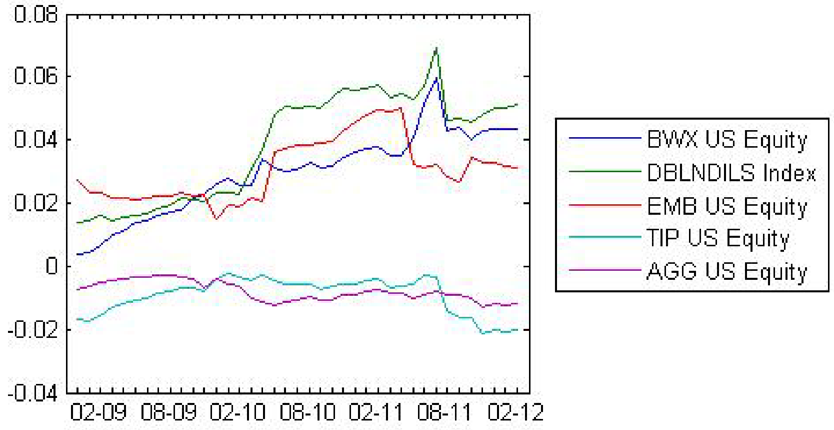


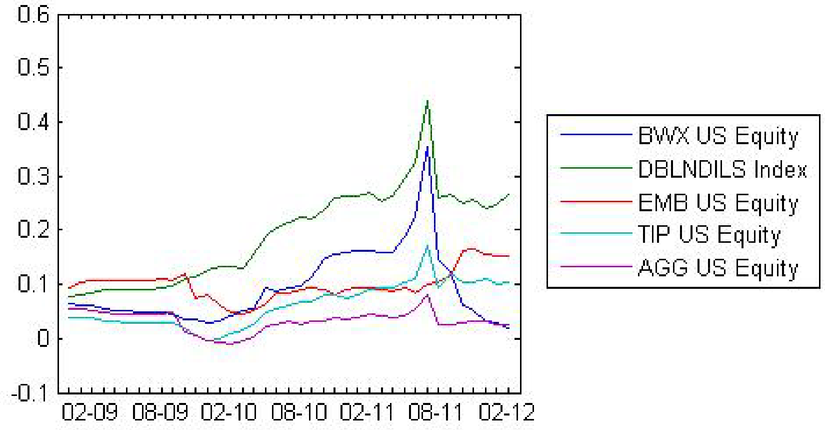
Now we give the asset loadings over time for the first 4 factors. To display clearly we show the assets in two groups(Group 1 vs Group 2&3). The first factor has quite stable loadings over time, except the Real Estate asset declines. If there is only one factor in the market, it would be the market itself, and the loadings would be proportional to beta of each asset. The decline of Real Estate can be regarded as the decline of its beta in the whole market. The second factor is less stable, especially in the latter half of 2011. As we know the second factor is "Foreign bond", and these volatile times correspond to the Europe bond crisis. And it is still volatile now. The 3rd and 4th factor are much more unstable, and they have irregular spikes. That means they are sometimes highly leveraged.

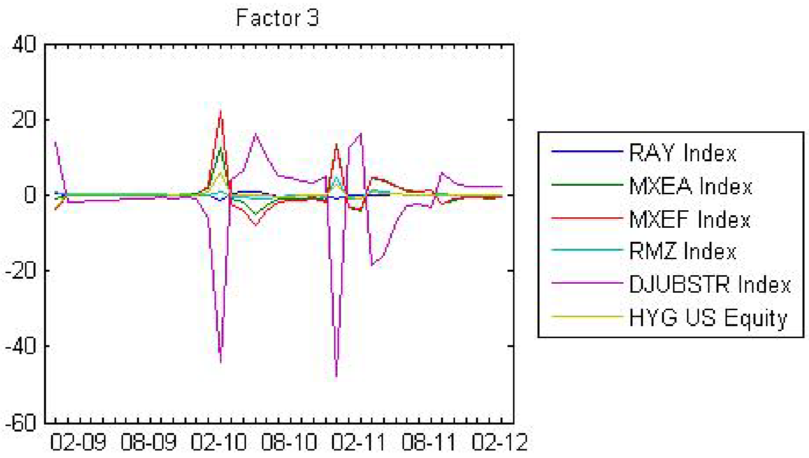
When we examine the loadings, we must note this transform is solely based on covariance matrix, which has estimation errors. Therefore, there is an extra-layer of instability in the result. In the appendix, we will show a method to filter the noises in the data so the patterns become more evident. We call this method Mean-Reverting PCA. For simplicity, we do not use the result of Mean-Reverting PCA in the following chapters.

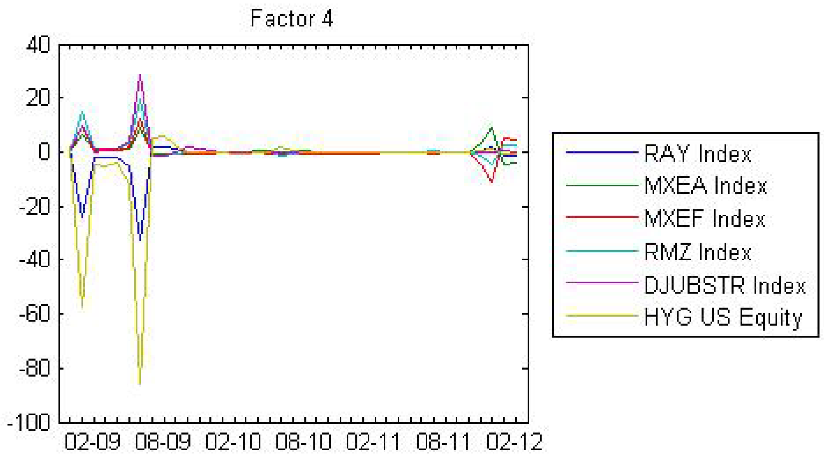

Factor 1
0.5
0.4
0.3
0.2
0.1
0
02-09 08-09 02-10 08-10 02-11 08-11 02-12
RAY Index
MXEA Index
MXEF Index
RMZ Index
DJUBSTR Index
HYG US Equity

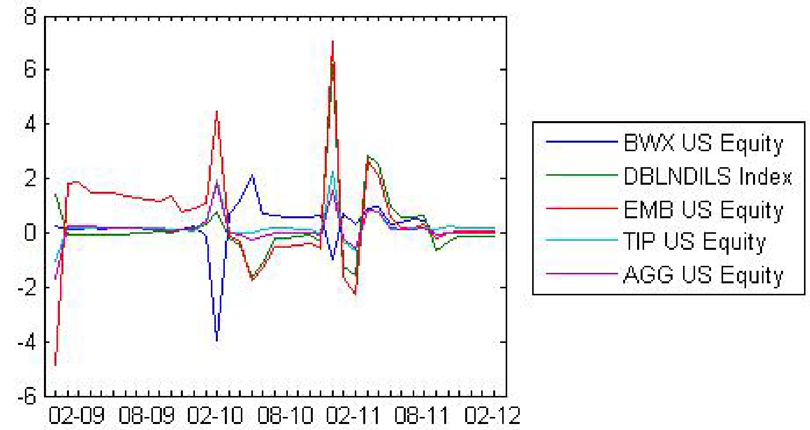

Factor 2
1
0.5
0
-0.5
-1
-1.5
02-09 08-09 02-10 08-10 02-11 08-11 02-12
RAY Index
MXEA Index
MXEF Index
RMZ Index
DJUBSTR Index
HYG US Equity

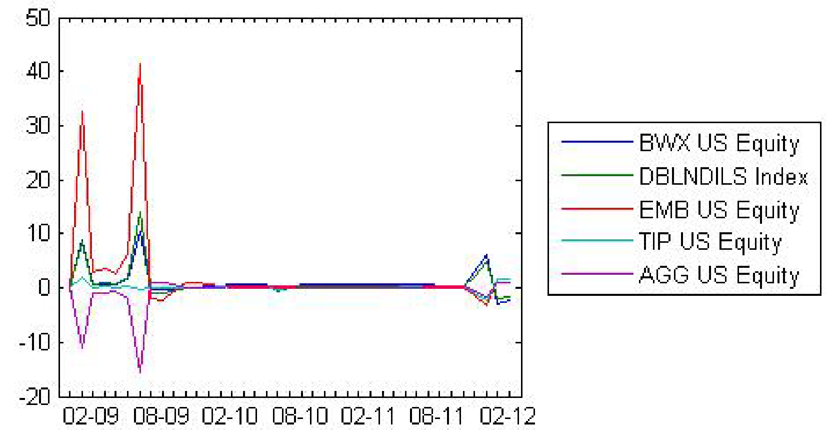

0.08
0.06
0.04
0.02
0
-0.02
-0.04
02-09 08-09 02-10 08-10 02-11 08-11 02-12
BWX US Equity
DBLNDILS Index
EMB US Equity
TIP US Equity
AGG US Equity


0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
02-09 08-09 02-10 08-10 02-11 08-11 02-12
BWX US Equity
DBLNDILS Index
EMB US Equity
TIP US Equity
AGG US Equity


Factor 3
40
20
0
-20
-40
-60
02-09
08-09
02-10
08-10
02-11
08-11
02-12
RAY Index
MXEA Index
MXEF Index
RMZ Index
DJUBSTR Index
HYG US Equity


Factor 4
40
20
0
-20
-40
-60
-80
-100
02-09
08-09
02-10
08-10
02-11
08-11
02-12
RAY Index
MXEA Index
MXEF Index
RMZ Index
DJUBSTR Index
HYG US Equity


8
6
4
2
0
-2
-4
-6
02-09
08-09
02-10
08-10
02-11
08-11
02-12
BWX US Equity
DBLNDILS Index
EMB US Equity
TIP US Equity
AGG US Equity


50
40
30
20
10
0
-10
-20
02-09
08-09
02-10
08-10
02-11
08-11
02-12
BWX US Equity
DBLNDILS Index
EMB US Equity
TIP US Equity
AGG US Equity

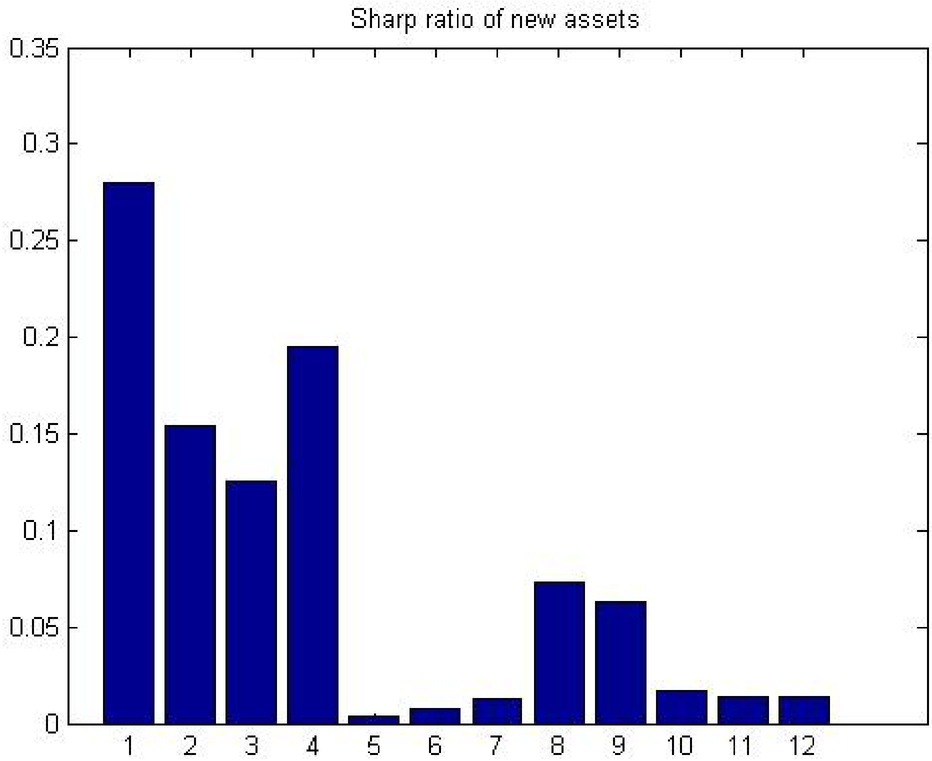
Chapter 5: Empirically sound case for PCA

Usually Sharpe Ratio is not substantially reduced. We give an example where the MVO yields some large positions.

The weights and the corresponding expected returns are shown in the 2nd and 3rd column below. Basing on the excess returns, we calculate the weights given by our PCA approach w(4) in the 5th column. The original weights in the 2nd column can be regarded as w(12).

Asset Class	Optimal weights	Expected excess returns	Portfolio Sensitivity to each asset's return	Optimal Weights by PCA (k=4)	Portfolio Sensitivity by PCA
US Stocks	5.90	4.40	45.99	-1.27	1.35
Foreign Developed Market Stocks	33.97	6.42	26.70	14.13	2.70
Emerging Market Stocks	-17.72	4.29	17.93	-10.86	4.58
Real Estate	-3.94	4.85	22.84	-1.28	2.50
Commodity	31.08	5.02	8.58	32.07	4.87
US High Yield Bonds	-7.08	2.03	18.87	0.55	0.29
Foreign Government Bonds	14.78	2.74	18.56	33.80	2.18
Foreign Government Inflation linked Bonds	-10.90	2.71	22.22	22.65	1.46
Foreign Corporate Bonds	61.51	3.92	11.89	41.52	3.10
Emerging Market Bonds	-1.77	1.29	15.43	4.09	0.11
TIPS	30.62	0.38	24.78	7.86	0.36
US Investment Grade Bonds	33.45	0.03	31.48	3.49	0.11
US Treasury	-69.89	0.00		-46.74	

Immediately we find w(4) is less extreme than w(12), with standard deviation of 17% and 24% respectively. The following graph shows the Sharpe ratio of each factor. We have not missed the big ones. Sharpe Ratio of w(4) is 0.3947, just slightly lower than that of w(12), which is 0.4074. Therefore, we have improved the situation in the first problem.

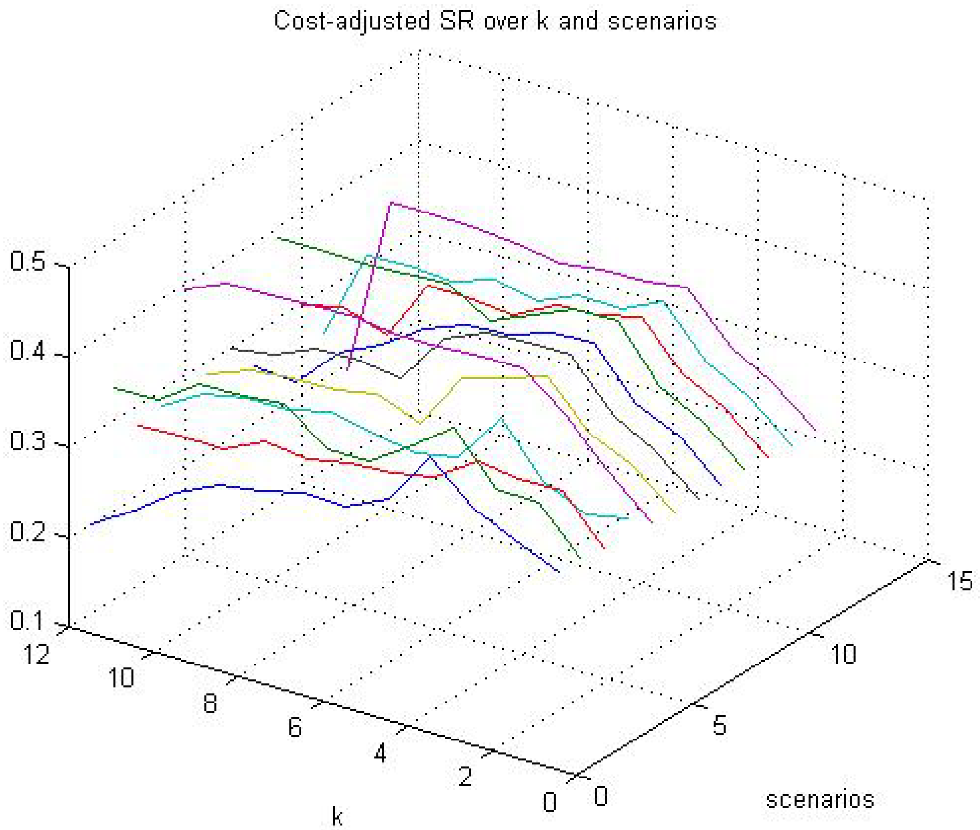


Next we analyze the portfolio sensitivity with regard to each asset's expected returns. For each risky asset, we increase its excess return by 10% of its standard deviation, and keep the other excess returns unaltered, we would derive two new set of weights $w(k)^+$ and $w(12)^+$. Finally we calculate the average absolute difference for the risky assets

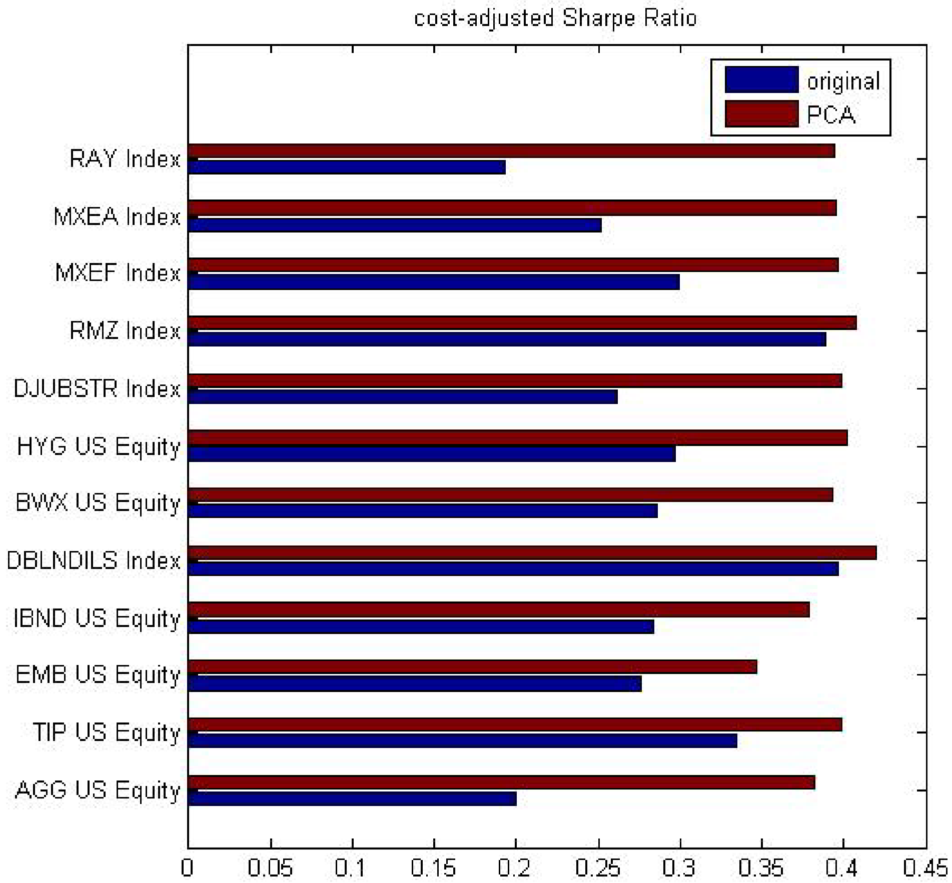
$$\frac{1}{12}\sum_{i=1}^{12}|w(12)_i - w(12)_i^+| \text{ and } \frac{1}{12}\sum_{i=1}^{12}|w(k)_i - w(k)_i^+|$$

and shown them in column 4 and 6. k is selected so that cost-adjusted Sharpe Ratio is optimized. We can see there is huge improvement in turns of transaction cost.

For the 12 different scenarios, the best k is always 4 except for 5th asset: commodity, where k=8.



Finally, we compare these cost-adjusted Sharpe Ratios with the original, and as we can see from the following graph the improvement on the second problem is obvious.



Chapter 6: Combine with Black-Litterman model

We assume the assets being equal weighted as the equilibrium, and calculate equilibrium excess return μ^*. We choose $\tau = 1/427$, the inverse of number of observations.

Next, we need to introduce our views. Just for illustration purpose, we base our view on observed returns calculated in the following table. We choose to have views on the 2 major factors, and 7 assets with big Minor-factor Sharpe-Ratio. For each one, the view is simply that: we expect the mean is its historically observed mean; the variance the corresponding observed variance multiplies 0.1. Basically τ gives the confidence on the prior; and 0.1 gives the confidence on the views. Thus we are leaning towards the prior, rather than the views.

As shown in Chapter 3, we can transform these views to the view of factor mean, and carry out a conventional BL analysis. The result is shown in the last column of the table. In our case, the equilibrium has low expected returns, while observed returns are high. These high returns would result crazy positions, while BL reconcile these two and give a less extreme result.

What our method adds is the understanding of these excess returns. For example, US Stock has high returns, but it is due to the major factors. When we subtract them, US Stock actually has a negative return. The same is to Real Estate. So we actually distinguished the major factors from the asset-class specific features, and that gives us more understanding of the past, as well as a clearer way to set views.

The result shows large increase in the positions of 4 bond assets with views: Foreign Government Inflation Linked Bonds, Emerging Market Bonds, TIPS, US Investment Grade Bonds. They also have the largest Sharpe Ratio for Y. The views on the 2 major factors do not have much impact on these bond assets, because the first factor is Stocks, and the second factor has relatively small change.

Asset Class	Equilibrium excess returns(%,μ^*)	Observed excess returns (%)	Observed Minor factor excess returns(%,Y)	Observed Minor factor Sharpe Ratio(Y)	Equilibrium weights (%)	BL final weights (%)
US Stocks	4.84	15.85	-5.05	-0.02	7.69	8.83
Foreign Developed Market Stocks	5.05	11.47	12.06	0.90	7.69	-2.40
Emerging Market Stocks	4.23	12.92	5.68	0.25	7.69	6.96
Real Estate	5.91	19.27	-53.77	-0.07	7.69	9.62
Commodity	3.21	10.88	2.79	0.02	7.69	7.41
US High Yield Bonds	2.25	13.45	12.08	1.00	7.69	19.96
Foreign Government Bonds	1.57	9.35	7.69	0.60	7.69	7.61
Foreign Government Inflation linked Bonds	1.82	12.72	15.17	1.14	7.69	20.92
Foreign Corporate Bonds	1.92	7.51	3.49	0.20	7.69	7.57
Emerging Market Bonds	1.11	11.37	11.41	1.47	7.69	26.89
TIPS	-0.18	9.72	10.48	1.74	7.69	27.24
US Investment Grade Bonds	-0.19	5.90	6.50	1.92	7.69	45.81
First Factor	4.24	14.04				
Second Factor	1.11	4.67				

Conclusion

We tackle two problems: irregular weights and big weights changes in MVO. PCA would improve both, while other methods are usually designed to solve only one problem, e.g. add a penalty term of transaction cost to solve the second one.

For most cases, PCA would maintain a good Sharpe Ratio. However, when the MVO weights are extreme, PCA may reduce Sharpe Ratio to some extent, and that is THE problem of trade-off between probable weights and MV efficiency. This problem is irreducible, and we must accept it, e.g. Black-Litterman does not tackle MV efficiency. Even in such difficult situation, PCA can still help us make investment decisions: we can make separate decisions in regard to major market factors and remaining asset-specific features.

There are still a lot of researches needed in studying the major market factors: e.g. their dynamics, changes, relation to each asset classes, and maybe relation with fundamental economic variables, etc.

References:

[1] Harry Markowitz (1952), Portfolio Selection, Journal of Finance, 7, pp. 77–91

[2] Harry Markowitz (1959), Portfolio Selection: Efficient Diversification of Investments, Cowles Foundation Monograph #16, John Wiley & Sons, New York

[3] Stephen Ross (1976). The arbitrage theory of capital asset pricing, Journal of Economic Theory, 13, pp. 341-60

[4] Black-Litterman (1993) Global Portfolio Optimization, Financial Analysts Journal, 48, pp. 28-43

[5] Leibowitz-Bova (2005) Allocation Betas, Financial Analysts Journal, 61, pp. 70-82

Appendix:

Proof 1.

$$(AH)'\widetilde{w} = H'A^{-1}\left(\frac{1}{2\gamma}V^{-1}\right)(AH\mu - r_0\mathbf{1})$$

$$= \frac{1}{2\gamma}(H'V^{-1}H\mu - H'A^{-1}V^{-1}r_0\mathbf{1})$$

$$= \frac{1}{2\gamma}\Sigma^{-1}\mu - \frac{r_0}{2\gamma}H'V^{-1}A^{-1}1$$

We also know that $(A^{-1}\mathbf{1})_i = A_i^{-1} = H_i\mathbf{1} = (H\mathbf{1})_i$, so $A^{-1}\mathbf{1} = H\mathbf{1}$

$$(AH)'\widetilde{w} = \frac{1}{2\gamma}\Sigma^{-1}\mu - \frac{r_0}{2\gamma}H'V^{-1}H\mathbf{1}$$

$$= \frac{1}{2\gamma}\Sigma^{-1}\mu - \frac{r_0}{2\gamma}\Sigma^{-1}\mathbf{1}$$

$$= \frac{1}{2\gamma}\Sigma^{-1}(\mu - r_0\mathbf{1})$$

$$= w$$

That means we have the same asset allocation despite the transform!

Proof 2.

$$w(k) = (AH)'I(n,k)\widetilde{w}$$

$$= H'AI(n,k)A^{-2}\left(\frac{1}{2\gamma}V^{-1}\right)(AH\mu - r_0\mathbf{1})$$

$$= H'I(n,k)A^{-1}\left(\frac{1}{2\gamma}V^{-1}\right)(AH\mu - r_0\mathbf{1})$$

$$= H'I(n,k)\left(\frac{1}{2\gamma}V^{-1}\right)H\mu - \frac{r_0}{2\gamma}H'I(n,k)V^{-1}A^{-1}\mathbf{1}$$

$$= H'I(n,k)\left(\frac{1}{2\gamma}V^{-1}\right)H(\mu - r_0\mathbf{1})$$

$$= \frac{1}{2\gamma}\left(\sum_{i=1}^{k}(v_i^{-1}H_i{}'H_i)\right)(\mu - r_0\mathbf{1})$$

Therefore,

$$\frac{dw(k)}{d\mu} = \frac{1}{2\gamma}\sum_{i=1}^{k}(v_i^{-1}H_i{}'H_i)$$

$$\frac{dw(k)_q}{d\mu_j} = \frac{1}{2\gamma}\sum_{i=1}^{k}(v_i^{-1}H_i{}'H_i)_{q,j} = \frac{1}{2\gamma}\sum_{i=1}^{k}v_i^{-1}H_{i,q}H_{i,j}$$

Proof 3.

In the new asset space, the weights are $\widetilde{w}_{(k)}$.

$$SR(k) = \frac{\text{Excess Return}(k)}{\sqrt{\text{Excess Variance}(k)}} = \frac{\sum_{i=1}^{k} \widetilde{w}_i(\tilde{\mu}_i - r_0)}{\sqrt{\sum_{i=1}^{k} \widetilde{w}_i^2 \sigma_i^2}}$$

We also know

$$\widetilde{w}_i = \frac{\tilde{\mu}_i - r_0}{2\gamma\sigma_i^2}$$

So

$$SR(k) = \frac{\sum_{i=1}^{k} \frac{\tilde{\mu}_i - r_0}{2\gamma\sigma_i^2}(\tilde{\mu}_i - r_0)}{\sqrt{\sum_{i=1}^{k} \left(\frac{\tilde{\mu}_i - r_0}{2\gamma\sigma_i^2}\right)^2 \sigma_i^2}}$$

$$= \frac{\frac{1}{2\gamma}\sum_{i=1}^{k} \left(\frac{\tilde{\mu}_i - r_0}{\sigma_i}\right)^2}{\frac{1}{2\gamma}\sqrt{\sum_{i=1}^{k} \left(\frac{\tilde{\mu}_i - r_0}{\sigma_i}\right)^2}}$$

$$= \sqrt{\sum_{i=1}^{k} \left(\frac{\tilde{\mu}_i - r_0}{\sigma_i}\right)^2}$$

$$= \sqrt{\sum_{i=1}^{k} SR_i^2}$$

Additionally, Excess Return(k) and Excess Variance(k) are just

$$\frac{1}{2\gamma}\sum_{i=1}^{k} \left(\frac{\tilde{\mu}_i - r_0}{\sigma_i}\right)^2 = \frac{1}{2\gamma} SR(k)^2 \text{ and } \frac{1}{4\gamma^2}\sum_{i=1}^{k} \left(\frac{\tilde{\mu}_i - r_0}{\sigma_i}\right)^2 = \frac{1}{4\gamma^2} SR(k)^2$$

So they are in line with Sharpe Ratio.

Proof 4.

Given equilibrium weights w, the expected return is

$$\mu - r_0\mathbf{1} = 2\gamma\Sigma w$$

So the Sharpe ratio of the new asset i would be:

$$SR_i = \frac{\tilde{\mu}_i - r_0}{\sigma_i}$$

$$= \frac{A_i H_i \mu - r_0}{A_i\sqrt{v_i}}$$

$$= \frac{A_i H_i (\mu - r_0\mathbf{1})}{A_i\sqrt{v_i}}$$

$$= \frac{H_i (\mu - r_0\mathbf{1})}{\sqrt{v_i}}$$

$$= \frac{H_i(2\gamma\Sigma w)}{\sqrt{v_i}}$$
$$= 2\gamma \frac{H_i \sum_{i=1}^{n}(v_i H_i' H_i)\, w}{\sqrt{v_i}}$$
$$= 2\gamma \frac{v_i H_i w}{\sqrt{v_i}}$$
$$= 2\gamma\sqrt{v_i} H_i w$$

Relation with APT

PCA has a close relative: Factor Analysis, which is equivalent to Arbitrage Pricing Theory (APT) here. It says the return of each asset can be expressed by a few independent factors, but also have an idiosyncratic risk left that is uncorrelated with each other.

$$\Sigma \approx H'\widehat{V}H + D$$

Here $\widehat{V}$ is the diagonal matrix, and some of its diagonals are 0 because we expect relatively few factors. D is a diagonal matrix of variances of the residual risk for each asset. Note it is possible to assume heteroskedasticity in D.

In optimization, here D is small, so $H'\widehat{V}H + D$ would still have small eigenvalues. It does not have the stable property of PCA. If instead, we use $H'\widehat{V}H$, that would not be materially different from our method, and because of the existence of D, it would be fuzzier.

In portfolio construction, we can always use a pure factor model like APT. The issue is, we do not need to go that far, because it is hard to be confident about the minor factors (in APT, even how to separate them from noises). As we show in Chapter 4, the minor factors are highly unstable and hard to identify.

That is why we want to go with both dimensions: on factor dimension, we take the 2 major factors, and stop there; on asset dimension, we take the remaining part of all assets. We would not say the remaining parts are pure noises, because they are not, and the market cannot be merely 2-dimensional.

Mean Reverting PCA

The asset loadings of each factor AH have noise, and we may want to filter them. There is more than one way to do it, depending on your particular purpose. For example, we can add a L1 norm penalty term, so that some of these loadings shrink to 0. Here we propose a procedure to make the loadings more stable *over time*.

At each evaluation time point, we calculated the covariance and PCA factors. For each factor $\widetilde{X}_{i,t} = A_{i,t}H_{i,t}X_t$, we want to find an "Approximate factor" $\widehat{X}_{i,t} = a_{i,t}X_t$, so that

1. It has very high correlation with the original factor
2. It is close to the average Approximate factor from the very beginning to that time point

The correlation is:

$$\mathrm{corr}(\widehat{X}_{i,t}, \widetilde{X}_{i,t}) = \frac{\mathrm{cov}(\widehat{X}_{i,t}, \widetilde{X}_{i,t})}{\sqrt{\mathrm{var}(\widehat{X}_{i,t})}\sqrt{\mathrm{var}(, \widetilde{X}_{i,t})}} = \frac{a_{i,t}\Sigma_t H'_{i,t} A_{i,t}}{\sqrt{a_{i,t}\Sigma_t a'_{i,t}}\sigma_{i,t}} = \frac{a_{i,t}\Sigma_t H'_{i,t}}{\sqrt{(a_{i,t}\Sigma_t a'_{i,t}) v_{i,t}}}$$

We solve this problem stepwise:

$$\widehat{X}_{i,1} = \widetilde{X}_{i,1}$$

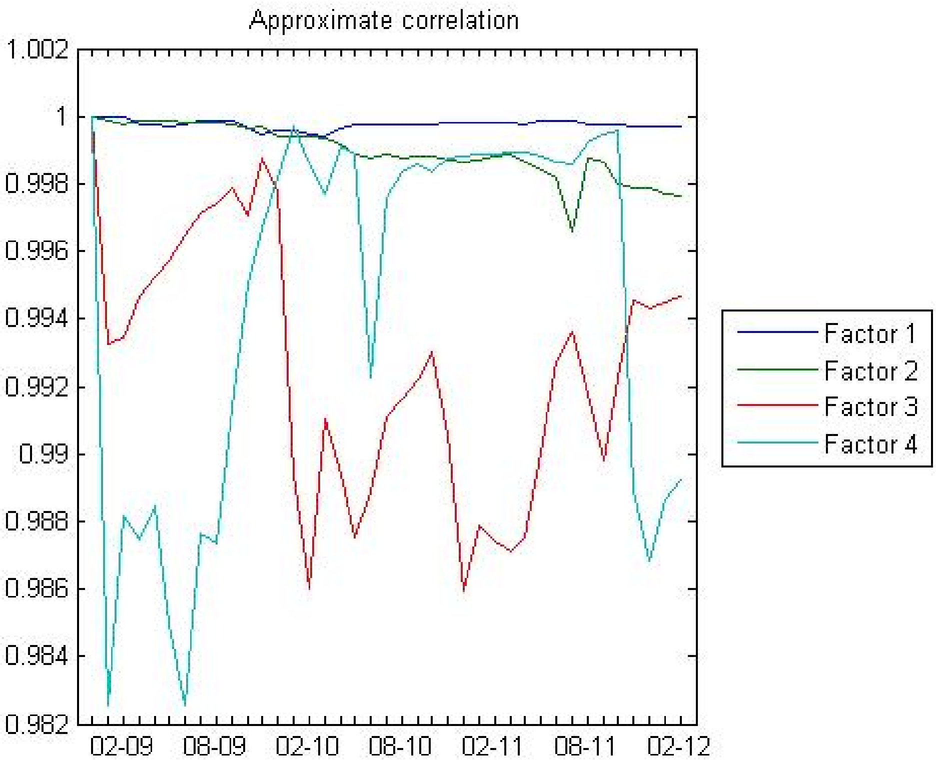
When t>1,

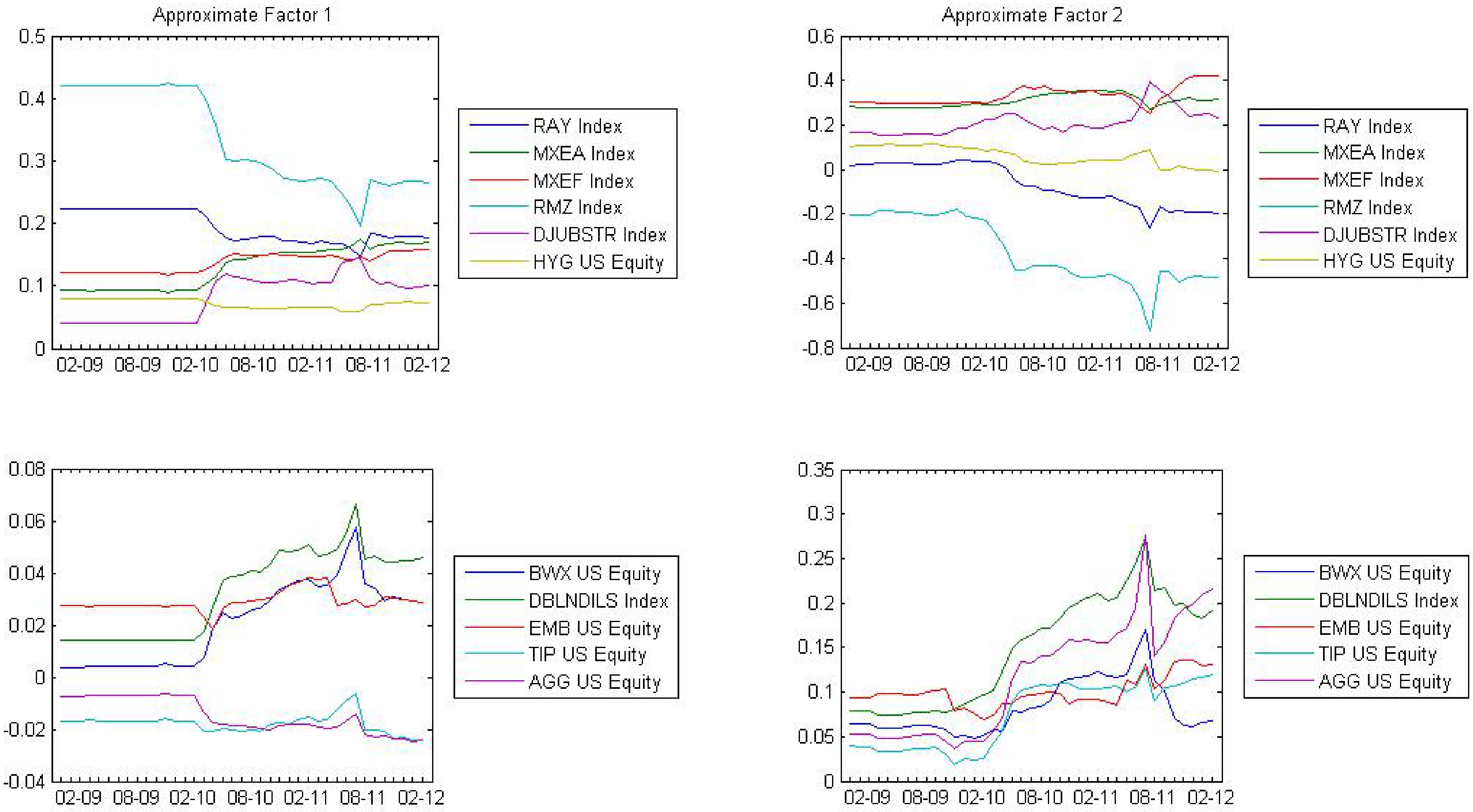
$$\max_{a_{i,t}} \left\{ \frac{a_{i,t}\Sigma_t H'_{i,t}}{\sqrt{(a_{i,t}\Sigma_t a'_{i,t}) v_{i,t}}} - \lambda \left\| a_{i,t} - \frac{1}{t-1}\sum_{s=1}^{t-1} a_{i,s} \right\|_2 \right\}$$

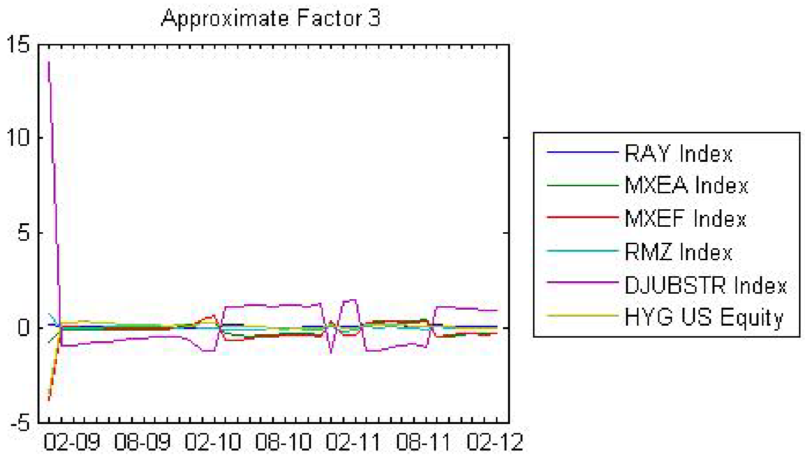
$$\text{s.t. } a_{i,t} \times \mathbf{1} = 1$$

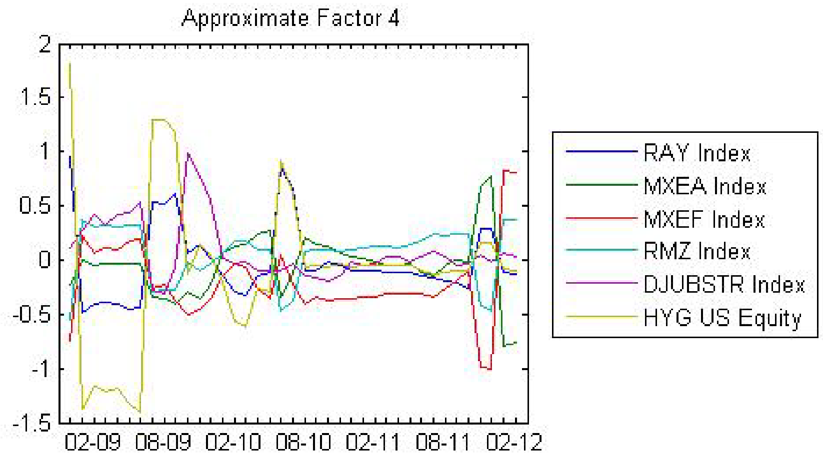
We apply this method to the data in Chapter 4, choose λ as 0.2. When we increase penalty, we would get more stable loadings, while it would be less close to the original factors.

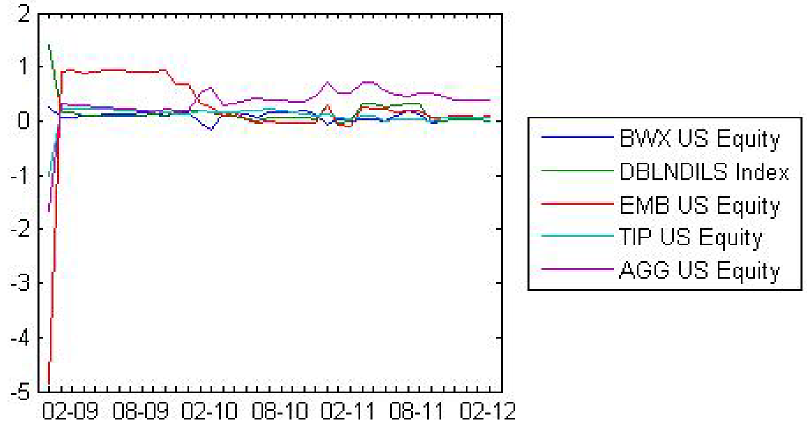
The correlation graph shows that the Approximation factors are highly correlated with the original factors, with correlation above 99.8% for the first 2 factors. For them, these are not materially different from the original.

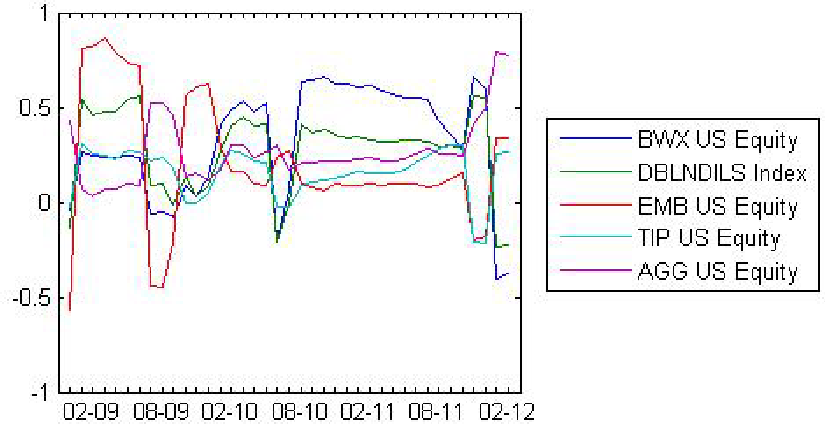



Approximate Factor 1
RAY Index
MXEA Index
MXEF Index
RMZ Index
DJUBSTR Index
HYG US Equity
02-09 08-09 02-10 08-10 02-11 08-11 02-12
Approximate Factor 2
RAY Index
MXEA Index
MXEF Index
RMZ Index
DJUBSTR Index
HYG US Equity
02-09 08-09 02-10 08-10 02-11 08-11 02-12
BWX US Equity
DBLNDILS Index
EMB US Equity
TIP US Equity
AGG US Equity
02-09 08-09 02-10 08-10 02-11 08-11 02-12
BWX US Equity
DBLNDILS Index
EMB US Equity
TIP US Equity
AGG US Equity
02-09 08-09 02-10 08-10 02-11 08-11 02-12


Approximate Factor 3
15
10
5
0
-5
02-09 08-09 02-10 08-10 02-11 08-11 02-12
RAY Index
MXEA Index
MXEF Index
RMZ Index
DJUBSTR Index
HYG US Equity


Approximate Factor 4
2
1.5
1
0.5
0
-0.5
-1
-1.5
02-09 08-09 02-10 08-10 02-11 08-11 02-12
RAY Index
MXEA Index
MXEF Index
RMZ Index
DJUBSTR Index
HYG US Equity


2
1
0
-1
-2
-3
-4
-5
02-09 08-09 02-10 08-10 02-11 08-11 02-12
BWX US Equity
DBLNDILS Index
EMB US Equity
TIP US Equity
AGG US Equity


1
0.5
0
-0.5
-1
02-09 08-09 02-10 08-10 02-11 08-11 02-12
BWX US Equity
DBLNDILS Index
EMB US Equity
TIP US Equity
AGG US Equity