

USING ORTHOGONAL GARCH TO FORECAST COVARIANCE MATRIX
OF STOCK RETURNS

A Thesis

Presented to

The Faculty of the Department

of Economics

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Master of Applied Economics

By

Jingjing Bai

August, 2011

USING ORTHOGONAL GARCH TO FORECAST COVARIANCE MATRIX
OF STOCK RETURNS

An Abstract of a Thesis

Presented to

The Faculty of the Department

Of Economics

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Master of Applied Economics

By

Jingjing Bai

August, 2011

ABSTRACT

The motivation of this paper is to study the estimation problems in large dimension systems in quantitative finance. The paper firstly presents principal component analysis to obtain the most important information in the data. Then, the orthogonal GARCH model introduced by Alexander and Chibumba (1997) and Alexander (2000) is provided to forecast five energy stocks' monthly volatilities and correlations. I show that as long as the stocks are already highly correlated with one another, the orthogonal GARCH approach will reduce computational complexity, control the amount of 'noise', and produce volatilities and correlations for all the assets. All the computation procedures were accomplished in Microsoft Excel. Forecasting of volatility and correlation of stock returns is significant in the analysis of option pricing, portfolio optimization and value-at-risk models.

ACKNOWLEDGEMENTS

I would like to express my gratitude to all those who give me the possibility to complete the thesis. First of all, I want to thank my parents, Xuexia Wang and Jie Bai. They bore me, raised me, supported me, taught me and loved me. Thank you for all your care and love! I would furthermore thank my my grandparents Shuping Bai and Huiwen Zhao, my aunt Yonghong Bai, my aunt Yanfang Wang, Caixia Wang, my cousin Yang Fan and entire family for financial support and providing a loving environment for me. I also would like to give my enormous thanks to my aunt Jinhua Bai and uncle Shijiang Lu. They suggested me to go to University of Houston and made great effort to help me and take care of me in Houston.

I deeply thank Dr. Brett Jiu who not only served as my thesis advisor but also encouraged and helped me to achieve the academic success.

I wish to thank Dr. Rebecca Thornton and Dr. Elaine Liu for their help, thank my classmate Jaweria Seth for useful discussion in the whole academic year, and thank Department of Economics for offering a great master program in Applied Economics! My senior Long Wen, Shu Zhang, Fei Lu, Bin Gao and Sa Bui, deserved special mention for kind advice.

Especially, I would like to give my special thanks to my boyfriend Zhun Zhao, who looked closely at the final version of the thesis for English style and grammar, correcting both and offering suggestions for improvement, and it is your sweet love that enables me to successfully complete graduate school.

TABLE OF CONTENTS

1. Introduction	1
2. Data	3
3. Methodologies	3
3.1 Principal component analysis.....	3
3.2 Generation of positive semi-definite covariance matrices	6
3.3 Implementing GARCH Model in Excel.....	9
4. Results	10
4.1 Eigenvalue Analysis	10
4.2 Using GARCH (1,1) model to forecast stock returns.....	13
4.3 Energy stocks' Volatility and Correlation Comparison.....	13
5. Conclusions	16
6. References	17

Using Orthogonal GARCH to Forecast Covariance

Matrix of Stock Returns

1 Introduction

It is well known that the covariance matrix of all the assets in a portfolio is highly associated with the portfolio's risk. Therefore, the creation of large, positive semi-definite covariance matrices is important in the field of modern risk management and factor modeling. Some large covariance matrices of the returns to many risk factors include foreign exchange rates, interest rates, money market rates, equity indices, and other key commodities. However, it is a challenge to process the large dimensions in the covariance matrix with numerous data in the system. Simple methods such as Exponentially Weighted Moving Average (EWMA, hereafter) are commonly used but they lack a solid statistical basis. Financial researchers are always looking for ways to reduce the dimensionality of large datasets while retaining most information in the data.

This paper will implement the approach of orthogonal GARCH, which in turn incorporates principal component analysis (PCA hereafter), as a solution to the dimensionality problem. The objective of the paper is to demonstrate how principal component analysis works to obtain the orthogonal factors and makes the process of generating large covariance matrices easier. The generalized autoregressive conditional heteroskedasticity (GARCH) model will be implemented in Excel to forecast volatilities and correlations.

Many research studies have been done on GARCH models. Engle (1982) and Bollerslev (1986) mentioned that the univariate GARCH could be successfully used for short term volatility forecasting in financial markets. In addition, the GARCH model is widely acknowledged to be able to generate more realistic long-term forecasts than exponentially weighted moving averages. Duan *et al.* (1995, 1996) used a new test of volatility forecasting model in their pricing and hedging options. An evolution of the GARCH model was introduced by Engle and Rosenberg in 1995, leading to an obvious advantage of Black-Scholes methods. However, the actual implementation of multivariate GARCH models is not as easy as generalizing the univariate models to multivariate parameterizations (Engle and Kroner, 1995). The reason is that the computation of large covariance matrices becomes more and more complicated as the dimensions increase.

Alexander and Chibumba (1997) first introduced the orthogonal GARCH model for generating large GARCH covariance forecast and Alexander (2000) further developed the orthogonal GARCH model. As Alexander (2001) pointed out, the advantages of the orthogonal method for generating covariance matrices are obvious. It outperformed EWMA model in generating estimates for volatilities and correlations of variables in the system. In the computing process, we only need to impose a few constraints on the movements in volatility and correlations. It reduces the computational burden since it is based on only the first few components of all the risk factors. It converts the multivariate problem to a univariate one and allows us to implement GARCH directly; therefore, the correlation estimates become more stable and less affected by the noise in the data.

The structure of my paper is as follows: Section 2 describes the two datasets used in the paper. Section 3 explains the methodology of principal component analysis. Section 4

illustrates how to generate a covariance matrix and implement the GARCH forecasting model. Section 5 shows the results and discusses the performance of the orthogonal GARCH model. Section 6 concludes the paper.

2 Data

In order to compare the efficiency of principle component analysis, two different datasets are selected in this paper. First, five sectors are selected and one stock in each Select Sector SPDRs funds was picked out (Select Sector SPDRs are unique ETFs that divide the S&P 500 into nine sectors): Walt-Disney from the consumer discretionary fund, Exxon Mobil from the energy fund, Bruker Corp from the healthcare industry fund, Bank of America from the financial fund, and Microsoft Corp from the technology fund. The second dataset comprises of the following five top energy companies: Exxon Mobil, Shell, Chevron, BP, and ConocoPhillips. These five energy companies' stocks are expected to be more stable and highly correlated with one another because the energy industry is more significant affected by the macro-economy than some of the other industries. In both datasets, the sample period extends from April 2001 to April 2011; since monthly data is used, each stock returns dataset is a 120×5 matrix in the analysis.

3 Methodologies

3.1 Principal component analysis

In mathematics, principal component analysis is defined as a procedure that uses an orthogonal transformation to extract the most important information of a set of possibly correlated variables into a set of uncorrelated variables. The new orthogonal variables are

called principal components (PCs hereafter), and the number of PCs is less than or equal to the number of original variables. For example, if K is the number of variables in the original system and M is the number of principal components used to explain the system, M is expected to be much less than K because one would wish to exclude the “noise” from the data and simplify the calculation. Meanwhile, the number of principal components used in the analysis will determine the accuracy of calculations because PCA indicates how much of the total variation in the original data is explained by each principal component (Alexander 2002). In general, the first principal component should accounts for the largest possible variance, and each following component has the highest possible variance under the constraints of being orthogonal to the prior components.

The material presented in this section follows Alexander (2001).

We define Y as a 120×5 matrix because both datasets used in the paper have 120 observations on 5 stock returns .Since the results of PCA will be different if the scale of the data changes, the first step of principal component analysis is to normalize the data into a $T \times K$ matrix X , where each column of X is $X_i = (Y_i - \mu_i) / \sigma_i$ where μ_i is the mean of original stock returns Y_i and σ_i is the standard deviation. This equation states that X is standardized to have a mean of 0 and a variance of 1, and X represents the same variables as Y . By multiplying transpose(X) and X , we get $X'X$, which is a 5×5 matrix. At this time, we could observe that all diagonal number is 119 (Table 1a). The reason is simple: the mean and standard deviations were sample statistics; therefore we lost one degree of freedom. (All calculations were done using Excel’s built-in statistics functions.)

Table 1a: X'X Matrix

119.00	88.52	81.62	97.90	92.86
88.52	119.00	88.76	87.49	81.03
81.62	88.76	119.00	74.59	73.66
97.90	87.49	74.59	119.00	85.65
92.86	81.03	73.66	85.65	119.00

The next step is to divide X'X by 119 (i.e., X'X/(T-1)(Table 1b). Now we have derived the unconditional sample correlation matrix of the stock returns.

Table 1b: X'X/119 Matrix

1.0000	0.7439	0.6859	0.8227	0.7804
0.7439	1.0000	0.7459	0.7352	0.6809
0.6859	0.7459	1.0000	0.6268	0.6190
0.8227	0.7352	0.6268	1.0000	0.7198
0.7804	0.6809	0.6190	0.7198	1.0000

Eigenvalues serve a very important role in principal component analysis. Each eigenvalue gives the percentage of risk due to each PC. The eigenvalue λ of a square matrix A is defined as a number that satisfies $Ax=\lambda x$, where x is a non-zero vector, In this case, x is called an eigenvector (corresponding to λ), and the pair (λ, x) is called an eigenpair for A.

Jackson and Staunton (2001) provide an easy way to calculate eigenvalues and eigenvectors in Excel VBA (Visual Basic for Applications). This method seeks to “nudge the input matrix towards diagonal form by a sequence of transformations or matrix rotations”. Table 2 gives the eigenvalues and their corresponding eigenvectors for the X'X/119 matrix based on their method. After sorting the eigenvalues from largest to smallest, we will be able to produce the principal components from the most important component to the least important component.

Table 2: Eigenvalues and Corresponding Eigenvectors

Eigenvalues				
3.8681	0.4491	0.2935	0.2292	0.1600
Eigenvectors				
0.4678	-0.2629	-0.0820	-0.3338	-0.7707
0.4517	0.3317	-0.2883	0.7638	-0.1391
0.4231	0.7305	0.2499	-0.4418	0.1724
0.4527	-0.3399	-0.5660	-0.2332	0.5520
0.4395	-0.4143	0.7262	0.2358	0.2287

The proportion of the total variation in X that is explained by the m -th PC is each eigenvalue divided by the sum of all eigenvalues. (Alexander 2000). In this paper, only the first three principal components will be used.

Let's define W as the matrix of eigenvectors of $X'X/T$. W is also called the matrix of 'factor weights'. The principal components of Y are then given by the $T \times K$ matrix $P=XW$.

Now we calculate the variance of each principle component in order to generate matrix $D=\text{diag}(V(P_1), \dots V(P_m))$, which is the diagonal matrix of each variance as shown in Table 3.

Table 3: Diagonal Matrix of PCs' Variance

3.8681	0.0000	0.0000	0.0000	0.0000
0.0000	0.4491	0.0000	0.0000	0.0000
0.0000	0.0000	0.2935	0.0000	0.0000
0.0000	0.0000	0.0000	0.2292	0.0000
0.0000	0.0000	0.0000	0.0000	0.1600

It's worth noting that the principal component variances are exactly the same as the eigenvalues. We can now re-write the normalized returns matrix X as follows:

$$X_i = w_{i1}P_1 + w_{i2}P_2 + \dots + w_{ik}P_k$$

3.2 Generation of positive semi-definite covariance matrices

$$V \approx ADA'$$

where $A=(w_{ij}^*)$ is the $K \times M$ matrix of normalized factor weights, A equals matrix W multiple standard deviation. The accuracy of V , the conditional covariance matrix of the original returns, is determined by how many components are selected to represent the system.

Since we only take the first three PCs into analysis, A is a 5×3 matrix as in Table 4:

Table 4: A Matrix

0.0277	-0.0175	-0.0063
0.0267	0.0221	-0.0222
0.0251	0.0487	0.0193
0.0268	-0.0226	-0.0436
0.0260	-0.0276	0.0559

Our task of forecasting the original covariance matrix V has been reduced to finding the diagonal covariance matrix of the principal components, D . Principal component analysis has allowed us to decrease the number of factors from the original K to a smaller number M , or, in our case, from 5 to 3. While our case does not benefit much from this reduction in dimensionality, the savings in computational time in a larger dataset can be significant. For instance, if we have 1000 stocks in the datasets, then the number for forecasting items is $(1+1000) \times 1000 / 2 = 500,500$. It is a very involved and time-consuming task to forecast half a million covariance numbers. However, if we use PCA, we could reduce the number of factors from 1000 to a much smaller number M , and A becomes a $1000 \times M$ matrix and D is now a $M \times M$ matrix. Let's say M is 10, then the number of forecasts we need is only $(1+10) \times 10 / 2 = 55$, just 0.01% of the original number of forecasts needed! From this relatively tiny covariance matrix D , we can apply ADA' to get back to V , the covariance matrix of the original 1,000 stocks is a 1000×1000 matrix.

Table 5 is the D matrix forecasting matrices. Alexander (2000) found that although D is positive-definite as it is a diagonal matrix with positive elements, ADA' cannot be guaranteed positive-definite when $M < K$. If it is only positive semi-definite, some weights in X

could be given zero portfolio variance. The only way to guarantee exact positive definiteness is to take all K principal components in the model, in which case there is no error term, too.

Table 5: Forecast D Matrix

Dt121		
2.6923	0.0000	0.0000
0.0000	0.5411	0.0000
0.0000	0.0000	0.2395
Dt122		
2.5744	0.0000	0.0000
0.0000	0.5411	0.0000
0.0000	0.0000	0.2295
Dt123		
2.4615	0.0000	0.0000
0.0000	0.5149	0.0000
0.0000	0.0000	0.2199
Dt124		
2.3537	0.0000	0.0000
0.0000	0.4900	0.0000
0.0000	0.0000	0.2107
Dt 125		
2.2505	0.0000	0.0000
0.0000	0.4663	0.0000
0.0000	0.0000	0.2019

Table 6 show the V matrix using the five energy stocks data.

Table 6: V Matrix

V121				
0.0022	0.0018	0.0014	0.0023	0.0021
0.0018	0.0023	0.0023	0.0019	0.0012
0.0014	0.0023	0.0031	0.0010	0.0013
0.0023	0.0019	0.0010	0.0027	0.0016
0.0021	0.0012	0.0013	0.0016	0.0030
V122				
0.0022	0.0017	0.0013	0.0022	0.0020
0.0017	0.0022	0.0022	0.0018	0.0012
0.0013	0.0022	0.0030	0.0009	0.0012
0.0022	0.0018	0.0009	0.0026	0.0016
0.0020	0.0012	0.0012	0.0016	0.0029
V123				
0.0021	0.0017	0.0012	0.0021	0.0019
0.0017	0.0021	0.0021	0.0017	0.0011
0.0012	0.0021	0.0028	0.0009	0.0011
0.0021	0.0017	0.0009	0.0025	0.0015
0.0019	0.0011	0.0011	0.0015	0.0027
V124				
0.0020	0.0016	0.0012	0.0020	0.0019
0.0016	0.0020	0.0020	0.0016	0.0011
0.0012	0.0020	0.0027	0.0009	0.0011
0.0020	0.0016	0.0009	0.0023	0.0014
0.0019	0.0011	0.0011	0.0014	0.0026
V125				
0.0019	0.0015	0.0011	0.0019	0.0018
0.0015	0.0019	0.0019	0.0016	0.0010
0.0011	0.0019	0.0026	0.0008	0.0011
0.0019	0.0016	0.0008	0.0022	0.0014
0.0018	0.0010	0.0011	0.0014	0.0025

3.3 Implementing GARCH Model in Excel

The GARCH model can be implemented in Excel via the maximum likelihood method. We run a GARCH (1, 1) model for each of the first three principal components (PCs) of the

original normalized returns. The unconditional variance is the variance of the first PC, while the forecast variance is calculated according to the GARCH (1, 1) model equation. For instance, the corresponding formula for second variance is

$$G7+ABS(G3)*C4+ABS(G4)*(B4^2),$$

where G7 is the constant coefficient ω ; G3 is the autoregressive coefficient β ; G4 is error coefficient α ; C4 is the first variance; and B4 is the error term. I calculate the log likelihood value of each observation and then use the Excel data analysis tool Solver to maximize the sum of log likelihood by changing constant term ω , error term α , and autoregressive term β , subject to the constraint that $\alpha + \beta < 1$.

4 Results

4.1 Eigenvalue Analysis

Since the orthogonal GARCH model works ideally for highly correlated system, the energy stocks will be used in the following forecasting procedure. The monthly returns from April 2001 to April 2011 are shown in Figure 1. The eigenvalue analysis and correlation comparison are shown in Table 7 and Table 8.

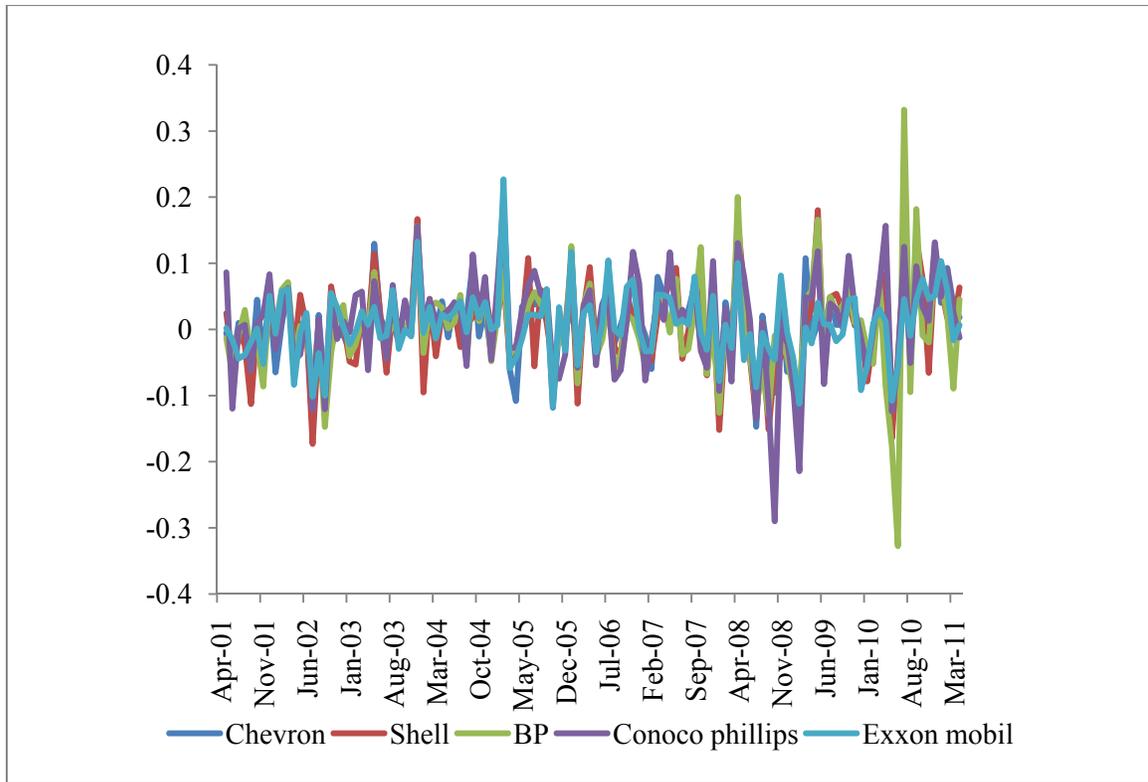


Figure 1: Energy stocks monthly returns

The five energy stocks monthly returns have similar trends indicated that they are highly correlated with each other.

Table 7: Eigenvalue Analysis

Component	Energy stocks		General stocks	
	Eigenvalue	R ²	Eigenvalue	R ²
P1	3.8681	77.36%	1.8343	36.69%
P2	0.4491	86.34%	1.0368	57.42%
P3	0.2935	92.22%	0.8380	74.18%
P4	0.2292	96.80%	0.7205	88.59%

In Table 7, for the general stocks, it is observed that 1.83 is the largest eigenvalue and the sum of eigenvalue is 5. So the proportion of the system variance that the first PC explains is only 1.83/5, or 33.6%. Following the same pattern calculation, the second largest eigenvalue (1.04), explains another 1.04/5, which is 20% and the third largest eigenvalue, 0.84, represents 16.8% of total variation. Thus the first three principal components together explain

74.1% of the total variation. The PCA will work better for the energy stocks because the first PC alone explains 77.4% and the first three PCs together explain over 92.2 % of the total variation. The reason for the energy stock' better performance is that the monthly returns of the energy stocks are highly correlated with one another while those of the five general stocks are not. PCA works best when the system is highly correlated.

Table 8: Correlation Comparison: Energy Stocks and General Stocks

	Chevron	Shell	BP	Conoco phillips	Exxon mobil
Chevron	100.00%				
Shell	74.39%	100.00%			
BP	68.59%	74.59%	100.00%		
Conoco phillips	82.27%	73.52%	62.68%	100.00%	
Exxon mobil	78.04%	68.09%	61.90%	71.98%	100.00%

	BOA	Exxon Mobil	Bruker	Microsoft	Walt-Disney
BOA	100.00%				
Exxon Mobil	2.48%	100.00%			
Bruker	-7.91%	13.37%	100.00%		
Microsoft	0.10%	33.21%	32.26%	100.00%	
Walt-Disney	2.21%	22.93%	23.94%	37.90%	100.00%

The correlation comparison is illustrated in Table 8. The five big energy companies' monthly return correlations are in a range of 61.9% to 82.27 %. For example, the correlation between Conoco Phillips and Exxon Mobil is 61.9%; the correlation between Conoco Phillips and Chevron is 82.27%. The other correlations are within this range. However, the correlation of five general stocks is randomly distributed. For instance, the highest correlation is 37.9% between Microsoft and Walt-Disney; and there are some really small correlations such as 0.1%, 2.21%, and even a negative correlation -7.91% between Bank of America and Bruker. Therefore, the five general stock's return data is not suited for our next forecasting procedure.

4.2 Using GARCH (1,1) model to forecast stock returns

The GARCH (1, 1) model defines the conditional variance at time t as:

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

Where α is the coefficient of volatility and last month's unexpected market return, and the coefficient β measures the persistence of volatility. The constraints is $\alpha + \beta < 1$. Standard deviation is a good measure of variation for a normal distribution. Risk is variation around the expected or mean value.

Table 9: Garch(1,1) Models of the First Three Principal Components

	1st PC	2nd PC	3rd PC
β	0.9562	0.9515	0.9582
α	0.0237	0.0331	0.0247

Since each β is over 0.95, all the three principal components have high persistence but low market reactions.

4.3 Energy stocks' Volatility and Correlation Comparison

Table 10: Energy Stocks' Volatility and Correlation Comparison

2PCs				
V121				
0.0472	0.8074	0.5464	0.9963	0.9860
0.8074	0.0468	0.9353	0.7539	0.6976
0.5464	0.9353	0.0545	0.4726	0.3989
0.9963	0.7539	0.4726	0.0470	0.9967
0.9860	0.6976	0.3989	0.9967	0.0473
3 PCs				
V121				
0.0473	0.7996	0.5261	0.9324	0.8189
0.7996	0.0481	0.8591	0.7623	0.4746
0.5261	0.8591	0.0553	0.3538	0.4254
0.9324	0.7623	0.3538	0.0517	0.5783
0.8189	0.4746	0.4254	0.5783	0.0546
4PCs				
V121				
0.0489	0.5463	0.5587	0.9314	0.7449
0.5463	0.0555	0.5748	0.5701	0.4836
0.5587	0.5748	0.0576	0.3804	0.3605
0.9314	0.5701	0.3804	0.0523	0.5387

Table 10 shows how closely the volatilities that are obtained using the 3 PCs compare with 2 PCs and 4 PCs. It is very easy to calculate 2 PCs and 4 PCs in Excel, just by taking the first 2×2 submatrix of the $D(t)$'s and 5×2 sub-matrix of the original A matrix. For the 4 PC case, just add another spreadsheet and implement the fourth principal component vector to the GARCH model. For instance, V121 (i.e., the T+1 covariance matrix, since our T=120) will be 0.0472, 0.0473 and 0.0489 for the first volatility. The values are very close. Results by using 2 PCs compared with 3 PCs have 15 similar correlations, indicating that three PCs even will not make a big difference if data is highly correlated. While for results by using 3 PCs vs. 4 PCs, they have 12 similar correlations. In this case, 4 PCs explains 3 more correlations. Thus, we decided to use 3 PCs. The forecasting results are listed in Table 11.

Table 11: 3PCs Volatility and Correlation Matrix

V121				
0.0473	0.7996	0.5261	0.9324	0.8189
0.7996	0.0481	0.8591	0.7623	0.4746
0.5261	0.8591	0.0553	0.3538	0.4254
0.9324	0.7623	0.3538	0.0517	0.5783
0.8189	0.4746	0.4254	0.5783	0.0546
V122				
0.0464	0.7919	0.5122	0.9324	0.8191
0.7919	0.0471	0.8586	0.7535	0.4658
0.5122	0.8586	0.0546	0.3399	0.4094
0.9324	0.7535	0.3399	0.0506	0.5800
0.8191	0.4658	0.4094	0.5800	0.0536
V123				
0.0453	0.7927	0.5137	0.9323	0.8187
0.7927	0.0461	0.8585	0.7544	0.4664
0.5137	0.8585	0.0534	0.3411	0.4112
0.9323	0.7544	0.3411	0.0495	0.5790
0.8187	0.4664	0.4112	0.5790	0.0524
V124				
0.0443	0.7935	0.5151	0.9322	0.8184
0.7935	0.0450	0.8584	0.7554	0.4669
0.5151	0.8584	0.0521	0.3424	0.4130
0.9322	0.7554	0.3424	0.0484	0.5781
0.8184	0.4669	0.4130	0.5781	0.0512
V125				
0.0433	0.7943	0.5166	0.9320	0.8180
0.7943	0.0440	0.8583	0.7564	0.4674
0.5166	0.8583	0.0509	0.3436	0.4148
0.9320	0.7564	0.3436	0.0473	0.5771
0.8180	0.4674	0.4148	0.5771	0.0501

Table 11 shows five forecasting covariance matrices from V121 (i.e., V at time T+1, or the forecast covariance matrix one month into the future) to V125 (five months into the future)

on the five energy stocks using the first three principal components of their monthly returns. The diagonal values in each matrix stand for volatilities (standard deviations) of the five stocks; the off-diagonal values are the correlations of the stocks returns. However, there is a limitation of the GARCH (1, 1) model. Take the first volatility of each matrix as an example, the values 0.0473, 0.0464, 0.0453, 0.0443 and 0.0433 are very close to each other. Moreover, volatilities became smaller and smaller in the forecasting results. Since one constraint in Garch Model is $\alpha + \beta < 1$, β has to be smaller than 1. Therefore the forecasting results will tend to 0 as we forecast further and further into the future; this is a well-known problem with ARCH and GARCH models in general. What mitigates this problem in practice is people usually do not forecast far into the future, so the orthogonal GARCH model can be used to forecast a few periods and, as shown in Alexander's papers, it performs just as well, yet at a small fraction of the computational resources, as other multivariate models, as long as the time series in the system are already highly correlated and a judicious number of first principal components are chosen.

5 Conclusions

To summarize, the orthogonal GARCH model relies on using principal component analysis to extract the factors that explain the variation in the original time series, and then using the principal components' own covariance matrix to arrive back at the original time series' covariance matrix. It can be used to solve practical problems in finance as it reduces the dimensions of the covariance matrix that needs to be forecasted. The example of five energy stocks' monthly returns data demonstrates that orthogonal GARCH is applicable in high dimensions with highly correlated data. Since Excel is one of the most convenient and widely

used data analysis tools, an advantage of this method is that the entire model, including both the principal component analysis and the GRACH forecasting, can be easily implemented in Excel. This method overcomes computational complexity and avoids the need to impose parameter constraints. In addition, it could reduce the ‘noise’ and produce volatilities and correlations for all the variances in the system. It also provides a direct and simply way to observe the contribution of each key factor to the risk. Consequently, financial professionals could optimally structure their positions to minimize the risk.

6 References

- Alexander, C.O.(2000). A Primer on the Orthogonal GARCH Model. ISMA Centre, University of Reading, Working paper.
- Alexander, C.O. (2001). Orthogonal GARCH. In C.O. Alexander(ed.), *Mastering Risk*, Volume 2, pp.21-38. London: Financial Times-prentice Hall.
- Alexander, C.O.(2002). Principal component models for generating large GARCH covariance matrices. *Economic Notes*, 31(2), 337-359.
- Alexander, C.O., and Chibumba, A.M.(1997). Multivariate Orthogonal Factor GARCH. University of Sussex, Discussion Papers in Mathematics.
- Bollerslev, T(1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31,307-327.
- Duan, J.(1995). The GARCH option pricing model, *Mathematical Finance*, 5:1, pp.13-22.

Duan, J.(1996) .Cracking the smile. *Risk*, 9, pp55-59.

Engle, R.F(1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50, 987-1008.

Engle, R.F and J. Rosenberg (1995). GARCH gamma. *Journal of Derivatives*, 2, pp47-59.

Engle, R.F. and K.F. Kroner (1995) .Multivariate Simultaneous Generalized ARCH. *Econometric Theory*, 11, 122-150.

Jackson and Staunton (2001). *Advanced Modelling in Finance using Excel and VBA*. Hoboken, NJ: John Wiley and Sons Ltd.