

Principal Component Models for Generating Large GARCH Covariance Matrices

Professor Carol Alexander, Chair of Risk Management, ISMA Centre, University of Reading, UK.

1. Introduction

The univariate generalised autoregressive conditional heteroscedasticity (GARCH) models that were introduced by Engle (1982) and Bollerslev (1986) have been very successful for short and medium term volatility forecasting in financial markets. Alternative univariate GARCH models, for use in different financial markets, have been developed in an academic literature that spans two decades. Bollerslev, Engle and Nelson (1994) provide a good review of most of the earlier literature on GARCH model and see Alexander (2001a) for a review of the more recent work. Many of these are being successfully applied to generating convergent term structure volatility forecasts, and in stochastic volatility models for option pricing and hedging (see Duan 1995, 1996; Engle and Rosenberg, 1995).

Multivariate GARCH models have been much less successful. It is straightforward to generalize univariate GARCH models to multivariate parameterizations (see for example, Engle and Kroner, 1993; Engle and Mezrich, 1996). However the implementation of these models in more than a few dimensions is extremely difficult: because the model has very many parameters, the likelihood function becomes very flat, and consequently the optimization of the likelihood becomes practicably impossible. There is simply no way that full multivariate GARCH models can be used to estimate directly the very large covariance matrices that are required to net all the risks in a large trading book.

This paper describes the principal component GARCH or 'orthogonal GARCH' (O-GARCH) model for generating large GARCH covariance matrices that was first introduced in Alexander and Chibumba (1996) and subsequently developed in Alexander (2000, 2001b).¹ The method is computationally very simple, as it is based on the univariate GARCH volatilities of the first few principal components of a system of risk factors. Empirical results on term structure forecasts for covariance matrices of commodity futures, interest rates and equities are used to illustrate the strengths and weaknesses of the O-GARCH model. The principal component model may also be used to generate large positive semi-definite exponentially weighted moving average (EWMA) covariance matrices where the smoothing constant is not the same for all risk factors, as it is in the RiskMetrics data (JPMorgan, 1996).

The outline of the paper is as follows: the next section puts the principal component model in context by reviewing the RiskMetrics methodology and surveying some recent

¹ Ding (1994) was the first to suggest the use of principal components with GARCH models, but with the full number of principal components. As we shall see, the empirical success of the O-GARCH model relies on the use of only a few principal components to represent the system.

developments in multivariate GARCH models. The O-GARCH and O-EWMA models are described in section 3 with particular reference to the correlation constraints that have to be satisfied in order that very large covariance matrices will be positive semi-definite. The implementation of O-GARCH is discussed, with reference to several empirical examples, in section 4. Section 5 describes the advantages and limitations of the principal component method for generating covariance matrices and section 6 concludes.

2. A Survey of Methods for Generating Large Covariance Matrices

The RiskMetrics data that are downloadable daily from the internet at www.riskmetrics.com consist of three large covariance matrices of the returns to many risk factors: major foreign exchange rates, money market rates, equity indices, interest rates and some key commodities. There are three covariance matrices: a 1-day matrix, a 25-day (1-month) matrix and a 'regulatory' matrix. Each matrix is based on a different weighted average model for volatility and correlation.

An EWMA model on squared and cross-products returns is used for the 1-day matrix, with the same smoothing constant $\lambda = 0.94$ for all variances and covariances in the matrix. EWMA's of the squares and cross products of returns are a standard method for generating covariance matrices, but a substantial limitation is that the covariance matrix is only guaranteed to be positive semi-definite if the same smoothing constant is used for all the data.² A consequence of this is that the reaction of volatility to market events and the persistence in volatility must be assumed to be the same in all the assets or risk factors that are represented in the covariance matrix.

A major advantage of the orthogonal EWMA (O-EWMA) model described in this paper is that it allows positive semi-definite exponentially weighted moving average methods to be used without the unrealistic constraint that the smoothing constant is the same for all risk factors. The exponentially weighted moving average variance of each principal component would normally be applied with a different smoothing constant. A consequence of this is that the reaction and persistence in the variance of any particular asset or risk factor will depend on the factor weights in the principal component representation and these factor weights are determined by the correlation in the system. Thus the market reaction and volatility persistence of a given asset or risk factor will not be the same at the other assets or risk factors in the system, but instead it will be related to its correlation with the other variables.

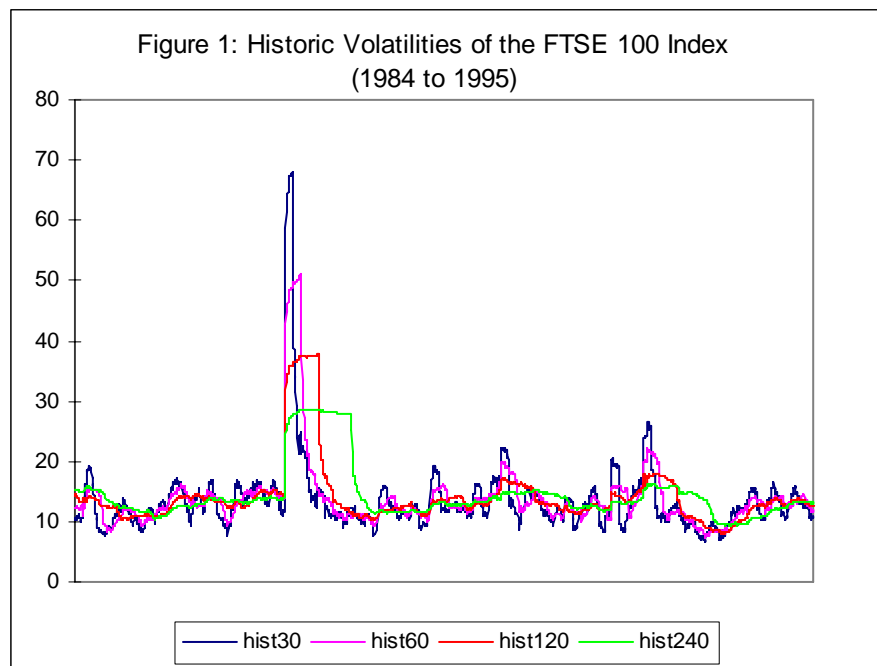
The 25-day matrix employs a similar EWMA methodology but with a smoothing constant $\lambda = 0.97$, and multiplying the resulting variance and covariance estimates by 25. This matrix not only suffers from the constraint that the volatility-correlation reaction-persistence is the same for all risk factors, it makes the additional assumption that volatilities and correlations

² See the RiskMetrics Technical Document, 3rd Edition, 1996 (www.riskmetrics.com)

Principal Component Models for Generating Large Covariance Matrices

are constant for 25 days. The constant volatility and correlation assumption is a major disadvantage of all EWMA models; however GARCH models do not make this assumption – they generate convergent term structures for volatility and correlation.

The RiskMetrics 'regulatory' matrix is so called because it complies with the quantitative standards for internal models for market risk requirements that were set out in the 1996 Amendment to the 1988 Basle Accord (see *Basle Committee on Banking Supervision, 1995*). To be specific, these quantitative requirements include the fact that at least 250 trading days of historical data should be used to construct the Value-at-Risk of a bank's positions. The two RiskMetrics EWMA matrices do not use 250 trading days of historical data, but the 'regulatory' matrix is constructed from equally weighted moving averages of squares and cross-products of returns over the past 250 trading days. It needs no constraints to be positive definite, but nevertheless the equally weighted methodology is well-known to have a major limitations. Figure 1 illustrates that one large return will continue to keep the volatility estimates based on equally weighted average of n squared returns high for exactly n trading days, and then exactly n trading days after this major market event the equally weighted volatility estimate will jump down again as abruptly as it jumped up. By the same token, the covariance estimates that are based on equally weighted average of n cross-products of returns will appear too stable for exactly n days following a single event that affects both markets. These 'ghost features' in the regulatory covariance matrix that will certainly follow large market movements are going to bias estimates of portfolio risk for a whole year, but the direction of the bias is not always easy to determine as it depends on the portfolio construction.



One of the great disadvantages of covariance matrices that are generated using moving average methodologies is that they assume that volatility and correlation is constant over time; a significant advantage of GARCH covariance matrices is that they do *not* assume that volatility and correlation is constant. Instead, the GARCH forecasts of these parameters will converge to their long term average values as the forecast horizon increases. We now turn attention to some recent developments in research into GARCH covariance matrices.

The conditional variance of an n -dimensional multivariate process is a time-series of $n \times n$ covariance matrices, one matrix for each point in time, denoted \mathbf{H}_t . Each of the $n(n+1)/2$ distinct elements of these covariance matrices has its own GARCH model, so even if there are just three parameters in each conditional variance or conditional covariance equation there will be $3n(n+1)/2$ parameters in total, plus the n parameters, one in each conditional mean equation. Thus the three-variate GARCH model has a minimum of 21 parameters. This is already a large number, but just think of a 10-dimensional system, with at least 175 parameters to estimate! It is therefore not surprising that estimation of multivariate GARCH models can pose problems. However, there are some approximations that allow multivariate GARCH covariance matrices \mathbf{H}_t to be generated by univariate GARCH models alone; these models are now described.

The idea of using factor models with GARCH goes back to Engle, Ng and Rothschild (1990) who use the capital asset pricing model to show how the volatilities and correlations between individual equities can be generated from the univariate GARCH variance of the market risk factor. In the CAPM individual asset returns are related to market returns X_t by the regression equation

$$r_{it} = \alpha_i + \beta_i X_t + \varepsilon_{it} \quad \text{for } i = 1, 2, \dots, n$$

Simultaneous estimation of the n linear regression equations will give factor sensitivities β_i and specific components ε_{it} . Denote by $\sigma_{i,t}$ the conditional standard deviation of asset i and by $\sigma_{ij,t}$ the conditional covariance between assets i and j . Assuming no conditional correlation between the market and specific components, taking variances and covariances of the CAPM equations yields

$$\begin{aligned} \sigma_{i,t}^2 &= \beta_i^2 \sigma_{X,t}^2 + \sigma_{\varepsilon_i,t}^2 \\ \sigma_{ij,t} &= \beta_i \beta_j \sigma_{X,t}^2 + \sigma_{\varepsilon_i \varepsilon_j,t} \end{aligned}$$

Thus all the GARCH variances and covariances of the assets in a portfolio are obtained from the GARCH variance of the market risk factor, and the GARCH variances and covariances of the stock-specific components. From a computational point of view there is much to be said for ignoring the covariance between specific components, the second term of the covariance equation. Then the above gives the individual asset conditional variances and covariances in terms of univariate GARCH models only. The limitation of this model is that it is not straightforward to generalize this framework to the case where the assets have more than one risk factor: firstly, generalized least squares would have to be used for the factor

sensitivity and residual estimation; and secondly, multivariate models for the covariances between risk factors will need to be employed.

Another model, introduced by Bollerslev (1990), approximates the time-varying covariance matrix as a product of time-varying volatilities and a correlation matrix that does not vary over time. This is the 'constant GARCH correlation model', defined by:

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{C} \mathbf{D}_t,$$

where \mathbf{D}_t is a diagonal matrix of time-varying GARCH volatilities, and \mathbf{C} is the constant correlation matrix. Individual return data are used to estimate univariate GARCH volatilities and the correlation matrix is estimated by taking equally weighted moving averages over the whole data period.

This model has been generalized by Engle (2000) to the case where the correlation matrix is time-varying. To avoid the need for a multivariate GARCH parameterization and to keep the model as simple as possible, this 'dynamic conditional correlation model' uses a GARCH(1,1) model with the same parameters for all the elements of the correlation matrix. A small limitation of this model is the same as one of the limitations of the RiskMetrics data: the same reaction and persistence parameters are used for all risk factor correlations. However, the model does not assume constant volatility and correlation and it does not require that the volatility reaction and persistence is the same for all risk factors. Therefore it represents a substantial improvement on the RiskMetrics methodology.

3. The Principal Component Covariance Matrix Model: O-GARCH and O-EWMA

Suppose historical data with T observations on k correlated asset or risk factor returns is summarized in a $T \times k$ matrix \mathbf{Y} . Principal component analysis (PCA) will give up to k uncorrelated returns, called the principal components of \mathbf{Y} , each component being a simple linear combination of the original returns as in (1) below. The weights in these linear combinations are determined by the eigenvectors the correlation matrix of \mathbf{Y} , and the eigenvalues of this matrix are the variances of the principal components. The principal components are ordered according to the size of eigenvalue so that the first principal component, the one corresponding to the largest eigenvalue (i.e. the one with the largest variance) explains most of the variation; in a highly correlated system the largest eigenvalue will be much larger than the rest and only the first few eigenvalues will be significantly different from zero. Thus only a few principal components are required to represent the original variables to a fairly high degree of accuracy.

To understand the mathematics suppose we normalize the \mathbf{Y} data into a $T \times k$ matrix \mathbf{X} where each column is standardized to have mean zero and variance one: if the i th risk factor or asset return in the system is y_i , then the normalized variables are $x_i = (y_i - \mu_i)/\sigma_i$ where μ_i and σ_i are the mean and standard deviation of y_i for $i = 1, \dots, k$. Now let \mathbf{W} be the matrix of eigenvectors of $\mathbf{X}'\mathbf{X}/T$, and $\mathbf{\Lambda}$ be the associated diagonal matrix of eigenvalues, ordered

according to decreasing magnitude of eigenvalue. The principal components of \mathbf{Y} are given by the $T \times k$ matrix

$$\mathbf{P} = \mathbf{X}\mathbf{W} \quad (1)$$

It is easy to show that such a linear transformation of the original risk factor returns produces transformed risk factor returns \mathbf{P} that are orthogonal and have variances equal to the eigenvalues in Λ . Since \mathbf{W} is an orthogonal matrix (1) is equivalent to $\mathbf{X} = \mathbf{P}\mathbf{W}'$, that is

$$x_i = w_{i1} \mathbf{p}_1 + w_{i2} \mathbf{p}_2 + \dots + w_{ik} \mathbf{p}_k$$

or

$$y_i = \mu_i + \omega_{i1}^* \mathbf{p}_1 + \omega_{i2}^* \mathbf{p}_2 + \dots + \omega_{ir}^* \mathbf{p}_r + \varepsilon_i \quad (2)$$

where $\omega_{ij}^* = w_{ij}\sigma_i$ and the error term in (2) picks up the approximation from using only the first r of the k principal components. These r principal components are the 'key' risk factors of the system; it is important to choose only a few of these, as the empirical results in this paper will show. For example, in a term structure of eleven interest rates, and in a term structure of twelve commodity futures, only two principal components will be used to generate the covariance matrices.

These m principal components are orthogonal so their covariance matrix a diagonal matrix \mathbf{D} . Taking variances of (2) gives

$$\mathbf{V} = \mathbf{A}\mathbf{D}\mathbf{A}' + \mathbf{V}_\varepsilon$$

where $\mathbf{A} = (\omega_{ij}^*)$ is the $k \times m$ matrix of normalized factor weights, $\mathbf{D} = \text{diag}(V(\mathbf{p}_1), \dots, V(\mathbf{p}_r))$ is the covariance matrix of the principal components and \mathbf{V}_ε is the covariance matrix of the errors. Ignoring \mathbf{V}_ε gives the approximation that forms the basis of the principal component model for large covariance matrices:

$$\mathbf{V} \approx \mathbf{A}\mathbf{D}\mathbf{A}' \quad (3)$$

Much computational efficiency is gained by calculating only r variances instead of the $k(k+1)/2$ variances and covariances of the original system. Note also that \mathbf{V} will always be positive semi-definite: to see this write

$$\mathbf{x}'\mathbf{A}\mathbf{D}\mathbf{A}'\mathbf{x} = \mathbf{y}'\mathbf{D}\mathbf{y}$$

where $\mathbf{A}'\mathbf{x} = \mathbf{y}$. Since \mathbf{y} can be zero for some non-zero \mathbf{x} , $\mathbf{x}'\mathbf{A}\mathbf{D}\mathbf{A}'\mathbf{x}$ will not be strictly positive for all non-zero \mathbf{x} , it may be zero, and so $\mathbf{A}\mathbf{D}\mathbf{A}'$ is positive semi-definite.

Now let us examine the covariance matrix \mathbf{D} more closely. The principal components are only unconditionally uncorrelated, so \mathbf{D} is only truly diagonal if its elements are the 'historical' variances and covariances of the principal components. However the O-EWMA model is based on EWMA variances in \mathbf{D} with different smoothing constants for each principal component variance, and it assumes the off-diagonal elements of \mathbf{D} are zero. Similarly the O-GARCH model is based on GARCH variances in \mathbf{D} with different GARCH parameters for each principal component and it again assumes the off-diagonal elements of

\mathbf{D} are zero. A very strong assumption of the basic O-EWMA and O-GARCH models is that zero conditional correlations are assumed. However, recently this assumption has been relaxed by van der Weide (2001) who develops a generalization of the O-GARCH model called 'generalized orthogonal GARCH' (GO-GARCH).

In order to generate a large covariance matrix covering all the risk factors such as equity market indices, exchange rates, commodities, government bond and money market rates, that are relevant to a banks operations, the risk factors must first be grouped into reasonably highly correlated categories. These categories will normally reflect geographic locations and instrument types. Care needs to be taken with the initial calibration, in the choice of correlated categories, the number of principal components used in each category, and the time period used to estimate them. But once calibrated the principal component model may be run very quickly and efficiently on a daily basis.

Principal component analysis is used to extract the key risk factors from each category and the orthogonal GARCH model or the orthogonal EWMA model is applied to generate the covariance matrix for each category. Then, in the second stage, the factor weights from the principal component analysis are used to 'splice' together a large covariance matrix for the original system. The method is explained for just two categories; the generalization to any number of categories is straightforward. Suppose there are m variables in the first category and n variables in the second category. It is not the dimensions that matter. What does matter is that each category of risk factors is suitably co-dependent, so that it justifies the categorization as a separate and correlated category.

First compute the principal components of each category, $\mathbf{P} = (P_1, \dots, P_r)$, and separately $\mathbf{Q} = (Q_1, \dots, Q_s)$ where r and s are number of principal components that are used in the representation of each category. Generally r will be much less than m and s will be much less than n . Denote by \mathbf{A} ($m \times r$) and \mathbf{B} ($n \times s$) the normalized factor weights matrices obtained in the PCA of the first and second categories, respectively. Then the 'within-category' covariances are given by $\mathbf{A}\mathbf{D}_1\mathbf{A}'$ and $\mathbf{B}\mathbf{D}_2\mathbf{B}'$, respectively where \mathbf{D}_1 and \mathbf{D}_2 are the diagonal matrices of the univariate GARCH or EWMA variances of the principal components of each system.

Denote by \mathbf{C} the $r \times s$ matrix of covariances of principal components across the two systems, that is, $\mathbf{C} = \text{cov}(P_i, Q_j)$. This 'cross-category' covariance matrix is computed using orthogonal EWMA or orthogonal GARCH a second time, now on a system of the $r + s$ principal components $P_1, \dots, P_r, Q_1, \dots, Q_s$. The cross-category covariances of the original system will then be given by \mathbf{ACB}' and the full covariance matrix of the original system is:

$$\begin{pmatrix} \mathbf{AD}_1\mathbf{A}' & \mathbf{ACB}' \\ (\mathbf{ACB})' & \mathbf{BD}_2\mathbf{B}' \end{pmatrix} \quad (4)$$

Since \mathbf{D}_1 and \mathbf{D}_2 are diagonal matrices with positive elements (the variances of the principal components) the within-factor covariance matrices $\mathbf{A}\mathbf{D}_1\mathbf{A}'$ and $\mathbf{B}\mathbf{D}_2\mathbf{B}'$ will always be positive semi-definite. But it is not always possible to guarantee positive semi-definiteness of the full covariance matrix of the original system, unless the following constraint holds:

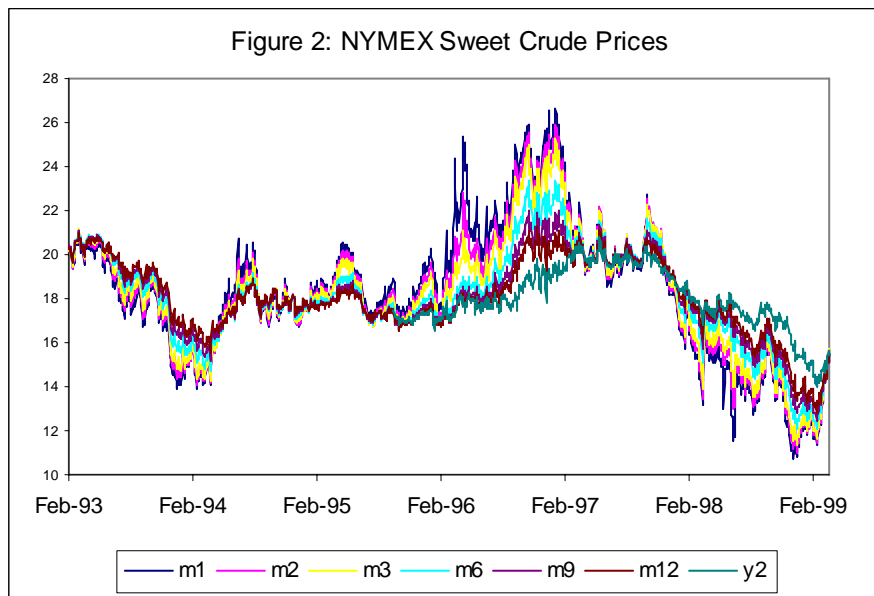
$$|\text{corr}(P_i, Q_j)| \leq (rs)^{-1/2} \quad (5)$$

The proof of this is given in the appendix. For example, if three principal components are used to represent each category, the large covariance matrix will be positive semi-definite if each O-GARCH (or O-EWMA) correlation never exceeds 1/3 in absolute value. But clearly, this constraint is more likely to be violated as more principal components are used in the model, and this is another reason why only few principal components should be used.

For the generation of the very large covariance matrices that are required for asset management and risk measurement, the implication of (5) is that the cross-factor correlations of the principal components cannot always be freely estimated. In fact, if the resulting covariance matrix does turn out to be non positive semi-definite, the constraint (5) will need to be imposed.

4. Implementation of the O-GARCH Model

First consider an O-GARCH(1,1) model for WTI crude oil futures from 1 month to 12 months, sampled daily between 4th February 1993 and 24th March 1999.³ The 1, 2, 3, 6, 9 and 12-month maturity futures prices are shown in figure 2 and the results of a principal component analysis on daily returns are given in table 1.



³ See Alexander (1999) for a full discussion of these data and of correlations in energy markets in general. Many thanks to Enron for providing these data.

Principal Component Models for Generating Large Covariance Matrices

This system is so highly correlated that over 99% of its variation may be explained by just two principal components and the first principal component alone explains almost 96% of the variation over the period.

Table 1a: Eigenvalue Analysis

Component	Eigenvalue	Cumulative R ²
P1	11.51	0.9592
P2	0.397	0.9923
P3	0.069	0.9981

Table 1b: Factor Weights

	P1	P2	P3
1mth	0.89609	0.40495	0.18027
2mth	0.96522	0.24255	-0.063052
3mth	0.98275	0.15984	-0.085002
4mth	0.99252	0.087091	-0.080116
5mth	0.99676	0.026339	-0.065143
6mth	0.99783	-0.020895	-0.046369
7mth	0.99702	-0.062206	-0.023588
8mth	0.99451	-0.098582	0.000183
9mth	0.99061	-0.13183	0.020876
10mth	0.98567	-0.16123	0.040270
11mth	0.97699	-0.19269	0.064930
12mth	0.97241	-0.21399	0.075176

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

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (5)$$

where $\omega > 0$, $\alpha, \beta \geq 0$. This simple GARCH model effectively captures volatility clustering and provides convergent term structure forecasts to the long-term average level of volatility $100\sqrt{250\omega/(1-\alpha-\beta)}$. The coefficient α measures the intensity of reaction of volatility to yesterday's unexpected market return ε_{t-1}^2 , and the coefficient β measures the persistence in volatility.

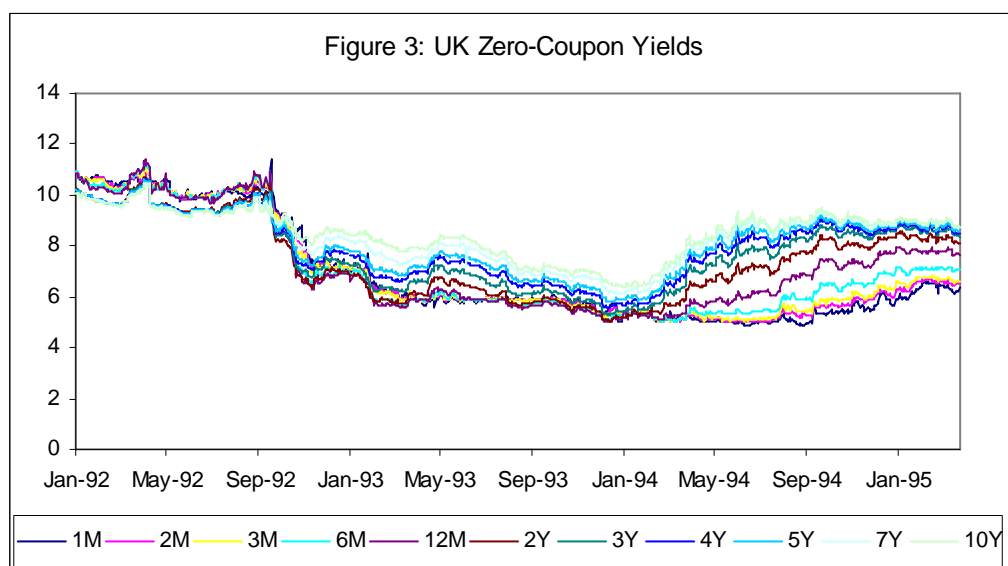
Table 2: GARCH(1,1) models of the first two principal components

	<i>1st Principal Component</i>		<i>2nd Principal Component</i>	
	<i>Coefficient</i>	<i>t-stat</i>	<i>Coefficient</i>	<i>t-stat</i>
constant	.650847E-02	.304468	.122938E-02	.066431
ω	.644458E-02	3.16614	.110818	7.34255
α	.037769	8.46392	.224810	9.64432
β	.957769	169.198	.665654	21.5793

Applying (5) to the first two principal components of these data gives the parameter estimates reported in table 2. Note that the first component has low market reaction but high persistence, and the opposite is true for the second component. All the variation in correlations comes from the second principal component because with only one component all variables would be perfectly correlated. The second component here has a 'spiky' volatility, and this gives rise to orthogonal GARCH correlations that also have only temporary deviations from normal levels. Thus the orthogonal GARCH model is capturing the true nature of crude oil futures markets where price decoupling occurs for only very short periods of time.

Note that only two univariate GARCH(1,1) variances, of the trend and tilt principal components need to be estimated and the entire 12x12 covariance matrix (containing 78 volatilities and correlations) of the original system is simply a transformation of these two variances, as defined in (3) above. Since the system is so highly correlated there is negligible loss of precision – the 12 volatility estimates are very similar to those obtained by applying 12 separate univariate GARCH(1, 1) models and, if it were possible to estimate a well-defined 12-dimensional multivariate GARCH model, the correlation estimates would also be very similar.

An advantage of the orthogonal method is that volatilities and correlations of all variables in the system can be estimated even when data on some variables are sparse and unreliable. For an example, consider daily zero coupon yield data in the UK with 11 different maturities between 1st January 1992 to 24th March 1995, shown in figure 3. It is not an easy task to estimate even univariate GARCH models on these data directly because the yields may remain relatively fixed for a number of days. Particularly on the more illiquid maturities, there is insufficient conditional heteroscedasticity for univariate GARCH models to converge well. So an 11-dimensional multivariate GARCH model is completely out of the question.



Principal Component Models for Generating Large Covariance Matrices

Again two principal components were used in the orthogonal GARCH model. This time, the PCA results reported in table 3 shows that these two components only account for 72% of the total variation in returns to zero coupon bonds over the period:

Table 3a: Eigenvalue Analysis

Component	Eigenvalue	Cumulative R ²
P1	5.9284117	0.53894652
P2	1.9899323	0.71984946
P3	0.97903180	0.80885235

Table 3b: Factor Weights

	P1	P2	P3
1mth	0.50916	0.60370	0.12757
2mth	0.63635	0.62136	-0.048183
3mth	0.68721	0.57266	-0.10112
6mth	0.67638	0.47617	-0.10112
12mth	0.83575	0.088099	-0.019350
2yr	0.88733	-0.21379	0.033486
3yr	0.87788	-0.30805	-0.033217
4yr	0.89648	-0.36430	0.054061
5yr	0.79420	-0.37981	0.14267
7yr	0.78346	-0.47448	0.069182
10yr	0.17250	-0.18508	-0.95497

The GARCH (1,1) parameter estimates of the principal components are given in table 4. This time both components have fairly persistent volatilities, and both are less reactive than the volatility models reported in table 2.

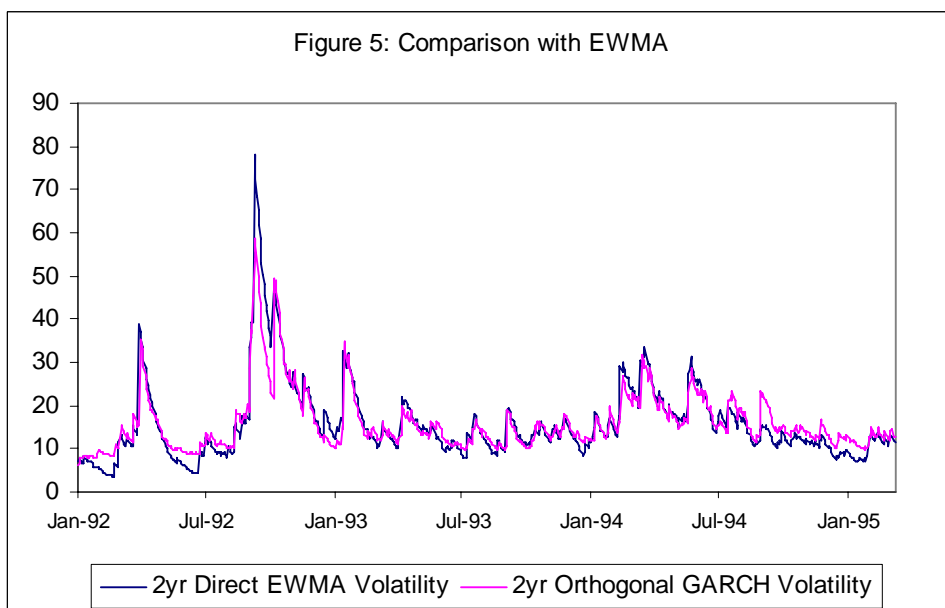
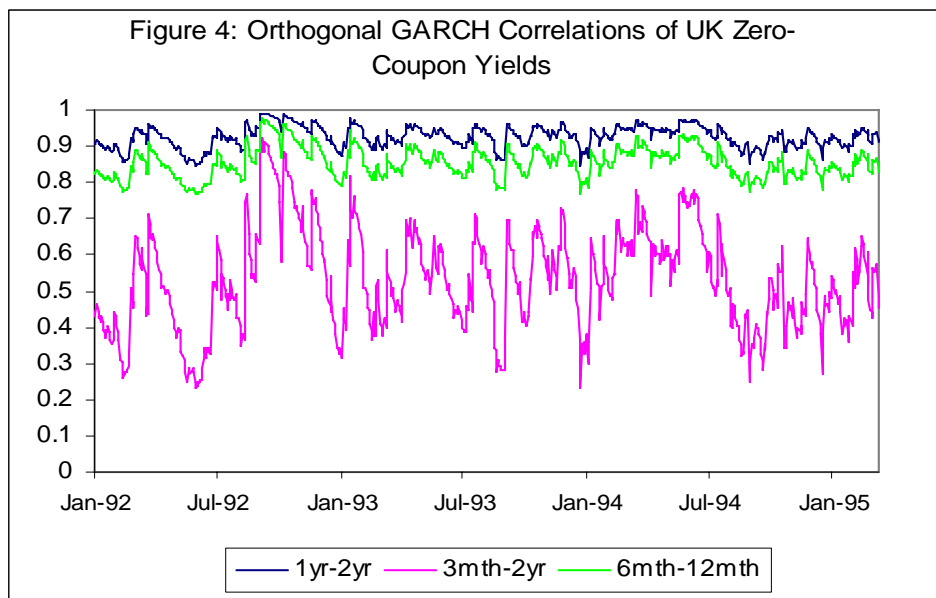
Table 4: GARCH(1,1) models of the first two principal components

	<i>1st Principal Component</i>		<i>2nd Principal Component</i>	
	<i>Coefficient</i>	<i>t-stat</i>	<i>Coefficient</i>	<i>t-stat</i>
constant	.769758E-02	.249734	.033682	1.09064
ω	.024124	4.50366	.046368	6.46634
α	.124735	6.46634	.061022	9.64432
β	.866025	135.440	.895787	50.8779

Not only does the orthogonal method provide a way of estimating GARCH volatilities and generating volatility term structures that may be difficult to obtain by direct univariate GARCH estimation; it also produces GARCH correlations that would be very difficult indeed to estimate using direct multivariate GARCH. Figure 4 shows some of the orthogonal GARCH correlations for the UK zero coupon yields. All these are obtained from just two

principal components – the rest of the variation in the system is ascribed to 'noise' and is not included in the model. Thus these correlation estimates are more stable over time than the estimates that would be obtained by direct application of a bivariate GARCH model, or indeed a EWMA model.

Figure 5 compares the volatility of the 2yr rate obtained using the orthogonal GARCH model to that obtained by direct estimation of exponentially weighted moving averages with the same smoothing constant (0.95) for all variables. This type of comparison is an important part of the calibration of the model. The fit is good but may have been even closer if the system had not included the 10yr rate. This is because the 10yr rate has a very low correlation with the rest of the system, as reflected by its factor weight on the 1st principal component in table 3b.



Principal Component Models for Generating Large Covariance Matrices

Turning now to equity markets, table 5 gives the results of PCA on daily return data from 1st April 1993 to 31st December 1996 on four European equity indices: France (CAC40), Germany (DAX30), Holland (AEX), and the UK (FTSE100):

Table 5a: Eigenvalue analysis

Component	Eigenvalue	Cumulative R ²
P_1	2.686141	0.671535
P_2	0.596853	0.820749
P_3	0.382549	0.916386
P_4	0.334456	1

Table 5b: Factor weights

	P_1	P_2	P_3	P_4
AEX	0.866	0.068	0.224	0.441
CAC	0.834	-0.238	-0.496	0.036
DAX	0.755	0.615	-0.027	-0.226
FTSE	0.818	-0.397	0.294	-0.296

The weights on the first principal component are comparable and quite high. Since this is the trend component the indices are, on the whole, moving together: indeed the eigenvalue analysis in table 5a indicates that common movements in the trend explain 67% of the total variation over the 4-year period. All four principal components are used in the orthogonal GARCH model, and table 6 gives the univariate GARCH(1,1) parameter estimates of each component:

Table 6: GARCH(1,1) models of the principal components

	1 st PC		2 nd PC		3 rd PC		4 th PC	
	Coefficien t	<i>t</i> -stat	Coefficien t	<i>t</i> -stat	Coefficien t	<i>t</i> -stat	Coefficien t	<i>t</i> -stat
Constan t	0.00613	0.19446	0.00262	0.09008	-0.0801	-0.26523	0.00267	0.087511
ω	0.032609	1.90651	0.066555	3.10594	0.089961	2.12915	0.203359	1.80057
α	0.033085	2.69647	0.086002	4.57763	0.067098	2.92511	0.070417	2.00423
β	0.934716	35.9676	0.846648	25.9852	0.841618	14.4038	0.726134	5.22888

Figure 6 shows just one of the correlations from an orthogonal GARCH(1,1) model of the system compared with those obtained from two different direct parameterizations of a multivariate GARCH(1,1) model: (a) the Vech model, and (b) the BEKK model. The four-dimensional diagonal vech model has 10 equations, each with three parameters. The 30 parameter estimates and their *t*-statistics (in italics) are reported in table 7. The four-

Principal Component Models for Generating Large Covariance Matrices

dimensional BEKK model has 42 parameters and the estimates of the matrices **A**, **B**, and **C** are given in table 8.

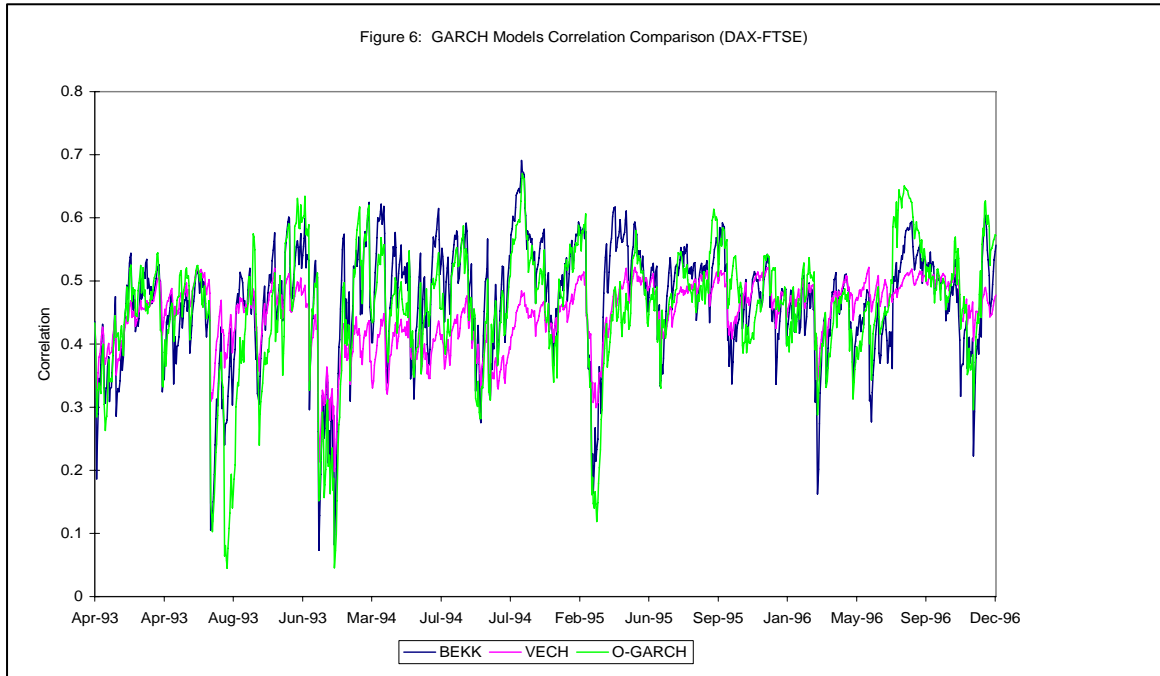


Table 7: Diagonal vech parameter estimates

Variance equations				Covariance equations					
AEX	CAC	DAX	FTSE	AEX-CAC	AEX-DAX	AEX-FTSE	CAC-DAX	CAC-FTSE	DAX-FTSE
ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
5.8×10^{-6}	3.4×10^{-6}	5.0×10^{-6}	1.8×10^{-6}	1.9×10^{-6}	9.3×10^{-6}	1.8×10^{-6}	8.6×10^{-6}	3.0×10^{-6}	1.6×10^{-6}
3.11	2.19	1.71	2.11	2.45	1.63	2.25	3.24	1.58	2.28
α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
0.054900	0.028889	0.028264	0.024601	0.021826	0.031806	0.028069	0.059739	0.022341	0.028377
4.14	3.15	2.28	2.68	3.75	2.40	3.68	3.56	2.18	2.88
β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
0.82976	0.89061	0.83654	0.91227	0.95802	0.74503	0.926414	0.829753	0.861387	0.934363
18.23	21.17	9.58	25.2	79.49	5.20	36.36	17.93	10.67	38.74

Table 8: BEKK parameter estimates

	AEX	CAC	DAX	FTSE
A	0.00160	0	0	0
	0.00008	-0.00176	0	0
	0.00094	0.00197	-0.00087	0
	0.00142	-0.00003	-0.00051	2.5×10^{-6}
B	0.22394	-0.04156	0.019373	0.04785
	-0.07147	0.18757	-0.05247	0.031895
	-0.06286	-0.04764	0.29719	0.07003
	-0.016277	-0.027589	-0.017405	0.178563
C	0.951805	0.027231	-0.050236	0.026130
	0.033141	0.9615723	0.023822	0.013623
	0.067985	0.053024	0.844291	0.005211
	0.022278	0.029257	-0.014482	0.948453

5. Summary of the Advantages and Limitations of the Principal Component Method

The empirical examples in the previous section have been chosen to highlight some of the great advantages of the O-GARCH and O-EWMA models: Large positive semi-definite covariance matrices can be generated in a computationally efficient manner that have:

1. Different degrees of market reaction and volatility persistence for different assets or risk factors;
2. Correlations that are more stable over time than those produced by other time-varying models;
3. Good results for illiquid markets, where lack of data would otherwise preclude the efficient forecasting of volatility and correlation;

and, for the O-GARCH model only:

4. Convergent volatility and correlation term structure forecasts.

The main limitation of the basic principal component model that is described here is the assumption that within-category conditional correlations are zero. That is, that within a particular asset category the covariance matrix is given by

$$\mathbf{V} = \mathbf{A}\mathbf{D}\mathbf{A}'$$

Where \mathbf{A} is a normalized matrix of eigenvectors of the correlation matrix of the original returns system and \mathbf{D} is a *diagonal* matrix of GARCH (or EWMA) variances of the first few principal components of the system.

To overcome this limitation a generalization of the O-GARCH model has been proposed by van der Weide (2001). In the O-GARCH model the univariate GARCH models are applied to uncorrelated principal components \mathbf{P} given by (1). In the generalized orthogonal GARCH (GO-GARCH) model of van der Weide the univariate GARCH models are applied to transformed variables $\mathbf{P}^* = \mathbf{P}\mathbf{U}$ where \mathbf{U} is an orthonormal matrix that is estimated using *conditional* information from the observed data.

6. Conclusion

Until recently the very large covariance matrices that are required for asset management and for firm-wide Value-at-Risk models could only be generated by imposing some very strong constraints on the movements in volatility and correlations. For example if exponentially weighted moving averages are used, the same smoothing constant had to be used for every risk factor, as in the RiskMetrics data. Without this assumption the covariance matrix would not be positive semi-definite. This is already a very strong assumption, but to make matters even worse, the assumption of constant volatility and correlation over the next 10 days was also made. This was because the only practicable way to generate large covariance matrices with mean-reverting term structures is to use a multivariate GARCH model and very little progress had been made towards developing empirically practical methods for generating large GARCH covariance matrices – until very recently.

This paper reviews the O-GARCH and O-EWMA models for generating covariance matrices that were first introduced in Alexander and Chibumba (1996). Large covariance matrices are based on the GARCH (or EWMA) volatilities of only a few, uncorrelated key market risk factors. Because only the most important risk factors are used in the model – the rest of the variation is ascribed to 'noise' – correlation estimates become more stable. The method produces positive semi-definite covariance matrices with, in the case of the O-GARCH model, mean-reverting term structure forecasts. The method is computationally efficient: it allows a great reduction in the dimension of the set of risk factors and moreover these 'key' risk factors are orthogonal. The model also quantifies how much risk is associated with each risk factor, so that managers can focus their attention on the most important sources of risk.

An important advantage of the model from the VaR modelling perspective is that the O-GARCH and O-EWMA models will conform to the BIS regulatory requirements on historic data if least one year of data is used in the principal component analysis. Moreover the O-GARCH model will provide properly convergent 10-day covariance matrices and does not assume that 10-day covariance matrices are simply 10 times the 1-day matrix, which is a substantial disadvantage of any EWMA or 'historical' covariance matrix.

Periods of sparse trading on some (but not all) assets do not present a problem either, because their current volatilities and correlations will be inferred (via the principal components) from their historic relationship with the other variables in the system. It is not only in illiquid markets that the principal component model can be used to overcome data problems. For example, when a new stock comes on the market, lack of historical data can

present enormous difficulties, not just for risk management, where a minimum standard of one year of historical data is often imposed, but it will also be very difficult to price long term derivatives if the firm wishes to raise capital in this way. However it is possible to obtain volatility forecasts using O-GARCH or O-EWMA if the price of the stock is related to the prices of several other stocks for which there is a reasonable amount of historic data.

Acknowledgements

Many thanks to Aubrey Chibumba for generating some of the results in section 4 as part of his MPhil thesis and to Walter Ledermann for useful discussions on bilinear forms.

References

- Alexander, C.O. (1999) "Correlation and cointegration in energy markets" in *Managing Energy Price Risk* (2nd Edition) RISK Publications pp291-304
- Alexander, C.O. (2000) 'Orthogonal methods for generating large positive semi-definite covariance matrices'. ISMA Centre Discussion Papers in Finance 2000-06, available from www.ssrn.com or www.ismacentre.rdg.ac.uk
- Alexander, C.O. (2001a) "Market Models: A Practitioners Guide to Financial Data Analysis" Wileys
- Alexander, C.O. (2001b) 'Orthogonal GARCH' in C.O. Alexander (ed.), *Mastering Risk Volume 2*. Financial Times – Prentice Hall.
- Alexander, C.O. and Chibumba, A. (1996) 'Multivariate orthogonal factor GARCH'. University of Sussex Discussion Papers in Mathematics (now available from www.ismacentre.rdg.ac.uk)
- Basle Committee on Banking Supervision (1995) "An internal model based approach to market risk capital requirements" available from www.bis.org
- Bollerslev, T. (1986) "Generalized autoregressive conditional heteroskedasticity". *Journal of Econometrics* 31 pp 307-327.
- Bollerslev, T., R.F. Engle and D.B. Nelson (1994) "ARCH models" in *Handbook of Econometrics* vol. 4, ed. R.F. Engle and D.L. McFadden, North-Holland
- Ding, Z. (1994) 'Time series analysis of speculative returns'. PhD thesis, University of California, San Diego.
- Duan, J-C (1995) "The GARCH option pricing model" *Mathematical Finance* 5:1 pp13-32.
- Duan, J-C (1996) "Cracking the smile" *RISK Magazine* 9:12 pp55- 59
- Engle, R.F. (1982) "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation". *Econometrica* 50:4 pp987-1007.
- Engle, R.F. (2000) "Dynamic conditional correlation - A simple class of multivariate GARCH models" pdf version available from <http://weber.ucsd.edu/~mbacci/engle/cv.html>
- Engle, R.F., and J. Mezrich (1996) "GARCH for groups" *RISK Magazine* 9:8 pp36-40
- Engle, R.F., Ng, V. and Rothschild, M. (1990) "Asset pricing with a factor ARCH covariance structure: Empirical estimates for treasury bills" *Journal of Econometrics* 45 pp213-238
- Engle, R.F. and K.F. Kroner (1993) "Multivariate simultaneous generalized ARCH" *Econometric Theory* 11 pp22-150
- Engle, R.F. and J. Rosenberg (1995) "GARCH gamma" *Journal of Derivatives* 2 pp47-59

Appendix

With the notation defined at the end of section 3, let

$$\mathbf{H} = \begin{pmatrix} \mathbf{A}\mathbf{D}_1\mathbf{A}' & \mathbf{A}\mathbf{C}\mathbf{B}' \\ (\mathbf{A}\mathbf{C}\mathbf{B}')' & \mathbf{B}\mathbf{D}_2\mathbf{B}' \end{pmatrix}$$

Since

$$\mathbf{H} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{D}_1 & \mathbf{C} \\ \mathbf{C}' & \mathbf{D}_2 \end{pmatrix} \begin{pmatrix} \mathbf{A}' & \mathbf{0} \\ \mathbf{0} & \mathbf{B}' \end{pmatrix}$$

clearly \mathbf{H} is positive semi-definite if and only if $\begin{pmatrix} \mathbf{D}_1 & \mathbf{C} \\ \mathbf{C}' & \mathbf{D}_2 \end{pmatrix}$ is positive semi-definite.

Denote by $\alpha_1 \dots \alpha_r$ the r elements of the diagonal matrix \mathbf{D}_1 and by $\beta_1 \dots \beta_s$ the s elements of the diagonal matrix \mathbf{D}_2 . These are the eigenvalues, the variances of the principal components; that is, $\alpha_i = V(P_i)$ and $\beta_j = V(Q_j)$. Denoting $\text{corr}(P_i, Q_j)$ by ρ_{ij} we write the ij th element of \mathbf{C} as

$$c_{ij} = \text{cov}(P_i, Q_j) = \rho_{ij} (\alpha_i \beta_j)^{1/2}.$$

This appendix derives the following sufficient condition for \mathbf{H} to be positive semi-definite:

$$\rho_{ij}^2 \leq (rs)^{-1}$$

Consider the quadratic form

$$\begin{aligned} f(\mathbf{u}, \mathbf{v}) &= (\mathbf{u}', \mathbf{v}') \begin{pmatrix} \mathbf{D}_1 & \mathbf{C} \\ \mathbf{C}' & \mathbf{D}_2 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \mathbf{u}'\mathbf{D}_1\mathbf{u} + 2\mathbf{u}'\mathbf{C}\mathbf{v} + \mathbf{v}'\mathbf{D}_2\mathbf{v} \\ &= \sum_{i=1}^r \alpha_i u_i^2 + 2 \sum_{i=1}^r \sum_{j=1}^s c_{ij} u_i v_j + \sum_{j=1}^s \beta_j v_j^2 \\ &= \sum_{i=1}^r \sum_{j=1}^s [(\alpha_i / r) u_i^2 + 2c_{ij} u_i v_j + (\beta_j / s) v_j^2] \end{aligned}$$

Now $f(\mathbf{u}, \mathbf{v})$ is positive semi-definite if each binary form

$$(\alpha_i / r) u_i^2 + 2c_{ij} u_i v_j + (\beta_j / s) v_j^2 \geq 0$$

A binary form $au^2 + buv + cv^2$ is positive semi-definite if $a > 0$, $c > 0$ and $b^2 \leq 4ac$. By definition $\alpha_1 \dots \alpha_r > 0$ and $\beta_1 \dots \beta_s > 0$, so $f(\mathbf{u}, \mathbf{v})$ is positive semi-definite if for each i and j

$$c_{ij}^2 \leq \alpha_i \beta_j / rs$$

But $c_{ij}^2 = \rho_{ij}^2 (\alpha_i \beta_j)$, so the result follows.