

## 2. Multivariate GARCH models

Consider a  $k$  component multivariate return series  $\mathbf{r}_t = (r_{1t}, \dots, r_{kt})'$ , where the prime denotes transpose. As in the univariate case, let

$$\mathbf{r}_t = \boldsymbol{\mu}_t + \mathbf{u}_t, \quad (12)$$

where  $\boldsymbol{\mu}_t = E[\mathbf{r}_t | \mathcal{F}_{t-1}]$  is the conditional expectation of  $\mathbf{r}_t$  given the past information  $\mathcal{F}_{t-1}$ . We assume that  $\boldsymbol{\mu}_t$  has a vector AR representation.

The conditional covariance matrix of  $\mathbf{u}_t$

$$\text{cov}[\mathbf{u}_t | \mathcal{F}_{t-1}] = \boldsymbol{\Sigma}_t \quad (13)$$

is a  $k \times k$  matrix, assumed positive definite.

There are many ways to generalize the univariate models to multivariate. The dimensionality, however, rapidly grows to unmanageable magnitudes because there are  $k(k+1)/2$  variance and covariance (or correlation) parameters in  $\boldsymbol{\Sigma}_t$  to model.

## 2.1 Simple bivariate GARCH(1,1) model

Consider the bivariate covariance matrix

$$\boldsymbol{\Sigma}_t = \begin{pmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{pmatrix} \quad (14)$$

The *vech* operator stacks all columns of a matrix into a column vector. In the covariance matrix we account for only the distinct elements, such that

$$\text{vech}(\boldsymbol{\Sigma}_t) = \begin{pmatrix} \sigma_{1,t}^2 \\ \sigma_{21,t} \\ \sigma_{2,t}^2 \end{pmatrix}, \quad (15)$$

where  $\sigma_{i,t}^2 = \sigma_{ii,t}$ ,  $i = 1, 2$ .

The simplest GARCH(1,1) parameterization in a bivariate GARCH is

$$\begin{aligned} \sigma_{1,t}^2 &= \omega_1 + \alpha_{11}u_{1,t-1}^2 + \beta_{11}\sigma_{1,t-1}^2 \\ \sigma_{2,t}^2 &= \omega_2 + \alpha_{22}u_{2,t-1}^2 + \beta_{22}\sigma_{2,t-1}^2 \\ \sigma_{12,t} &= \omega_{12} + \alpha_{12}u_{1,t-1}u_{2,t-1} + \beta_{12}\sigma_{12,t-1} \end{aligned} \quad (16)$$

Estimation of these models must be done with some programmable packages, like EViews ([www.eviews.com](http://www.eviews.com)), RATS ([www.estima.com](http://www.estima.com)) or GAUSS. Assuming conditional normality of the  $\mathbf{u}_t$ , the estimation can be accomplished with the ML method.

Let  $\boldsymbol{\theta}$  denote the vector of all the estimated parameters. Then the likelihood function is of the form (5) with

$$\ell(\boldsymbol{\theta}) = -\frac{k}{2} \log(2\pi) - \frac{1}{2} \log(|\boldsymbol{\Sigma}_t|) - \frac{1}{2} \mathbf{u}_t' \boldsymbol{\Sigma}_t^{-1} \mathbf{u}_t, \quad (17)$$

where  $\mathbf{u}_t = \mathbf{y}_t - \boldsymbol{\mu}_t$  with  $\boldsymbol{\mu}_t$  modeled (usually) with suitable vector ARMA model.

**Example.** Consider the forest and metal industry with the covariance matrix specification (16) and

$$\begin{aligned} y_{1,t} &= \phi_{10} + \phi_{11}y_{1,t-1} + u_{1,t} \\ y_{2,t} &= \phi_{20} + \phi_{22}y_{2,t-1} + u_{2,t} \end{aligned} \quad (18)$$

Below is a RATS code to run the estimation

```
* Read data
Open Data \Users\Seppo\TeX\Teaching\mvs\data\hex40.xls
Cal(Irregular)
Comp Start = 1
Comp End = 2253
all end
Data(org=obs,format=Excel)

* Define log return series
set r_fo = 100*log(forest/forest{1})
set r_me = 100*log(metal/metal{1})

* Initialize residual series
set u_fo = 0.0
set u_me = 0.0

* Initialize conditional variance and covariance series
set s_ff = 0.0
set s_mm = 0.0
set s_fm = 0.0

* define model parameters
nonlin p_of p_if p_om p_im $
      w_f a_f b_f $
      w_m a_m b_m $
      w_fm a_fm b_fm

* residual series
frml fu_fo = r_fo - p_of - p_if*r_fo{1}
frml fu_me = r_me - p_om - p_im*r_me{1}

* Simple bivariate GARCH specifications
frml fs_ff = u_fo(T) = fu_fo(T), w_f + a_f*u_fo{1}**2 + b_f*s_ff{1}
frml fs_mm = u_me(T) = fu_me(T), w_m + a_m*u_me{1}**2 + b_m*s_mm{1}
frml fs_fm = w_fm + a_fm*u_fo{1}*u_me{1} + b_fm*s_fm{1}
```

```

* Define covariance matrix and normal random vector
dec symm sigma
dec vect uvect

* Bivariate GARCH
frml glogl = s_ff(T) = fs_ff, s_mm(T) = fs_mm, s_fm(T) = fs_fm, $
sigma = ||s_ff(T)|s_fm(T),s_mm(T)||, $
uvect = ||u_fo(T),u_me||, $
%logdensity(sigma,uvect)

* Initialize AR parameter values
linreg(noprint) r_fo 2 end u_fo
# constant r_fo{1}

comp p_0f = %beta(1), p_1f = %beta(2), w_f = %seesq

linreg(noprint) r_me 2 end u_me
# constant r_me{1}

comp p_0m = %beta(1), p_1m = %beta(2), w_m = %seesq

set s_ff = u_fo**2
set s_mm = u_me**2
set s_fm = u_fo*u_me
comp a_f = b_f = 0.05
comp a_m = b_m = 0.05
comp w_fm = 0.20*sqrt(w_f*w_m)
comp a_fm = b_fm = 0.0

* Run ML estimation
maximize(method=bhhh,recursive) glogl 4 end

```

28

The results are:

```

MAXIMIZE - Estimation by BHHH
Convergence in 33 Iterations.
Final criterion was 0.0000024 < 0.0000100
Usable Observations 2250
Function Value -3535.89034261

```

Variable	Coeff	Std Error	T-Stat	Signif
1. P_OF	0.0855002050	0.0340704678	2.50951	0.01208985
2. P_1F	0.1159464672	0.0196845133	5.89024	0.00000000
3. P_OM	0.0553909885	0.0273991444	2.02163	0.04321438
4. P_1M	0.0944441986	0.0181307258	5.20907	0.00000019
5. W_F	0.3934345803	0.0362057391	10.86664	0.00000000
6. A_F	0.0922674237	0.0076690906	12.03108	0.00000000
7. B_F	0.7719222596	0.0177172980	43.56885	0.00000000
8. W_M	0.0225641119	0.0036558059	6.17213	0.00000000
9. A_M	0.0307861692	0.0033573738	9.16972	0.00000000
10. B_M	0.9563418700	0.0042109260	227.10964	0.00000000
11. W_FM	0.0756784673	0.0155469342	4.86774	0.00000113
12. A_FM	0.0291452752	0.0055289646	5.27138	0.00000014
13. B_FM	0.9074262457	0.0167555881	54.15663	0.00000000

The estimated coefficient seem reasonable. Nevertheless, the model has inherently severe shortcomings. First, covariance does not include conditional variances, and thus does not (explicitly) model the observed fact that correlation tends to increase as variability increases.

29

### Example. (Continued)

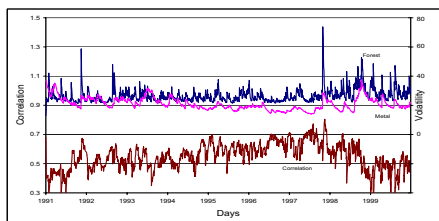


Figure. Conditional correlation and variance processes of the simple bivariate GARCH given in (16)

The (estimated) conditional correlation process depicted in the Figure is defined as

$$\hat{\rho}_t = \frac{\hat{\sigma}_{12,t}}{\hat{\sigma}_{1,t} \hat{\sigma}_{2,t}}. \quad (19)$$

As seen in the Figure above, the observed relationship between volatility and correlation is, however, captured in some places by the simple model, although the variances are not explicitly included into the covariance equation.

30

Second, additional constraints should be imposed on the coefficients to ensure positive definiteness of the covariance matrix.

Third, and perhaps the most serious drawback is that the structure is not invariant with respect to linear combinations. This is not a problem with returns, but for example in exchange rate markets it makes difference in which currency the variables are denominated (e.g. Euro, Dollar, UK pound, Yen, if one is interested in the relationship of these currencies).

Fourth, the model is not invariant under portfolio aggregation (this a drawback for most ARCH models). This means that aggregating individual stocks to portfolios does not preserve the volatility structure (as an exercise consider a portfolio of the two stocks with weights  $w_1$  and  $w_2$ ).

31

Before going to the more advanced models, let us briefly look at an even simpler model, Exponentially Weighted Moving Average (EWMA), which sometimes preferred by practitioners [see e.g., RiskMetrics™ (www.riskmetrics.com)]

The EWMA is a special case of the simple GARCH in (16)

$$\begin{aligned}\sigma_{1,t}^2 &= (1-\lambda)u_{1,t-1}^2 + \lambda\sigma_{1,t-1}^2 \\ \sigma_{2,t}^2 &= (1-\lambda)u_{2,t-1}^2 + \lambda\sigma_{2,t-1}^2 \\ \sigma_{12,t} &= (1-\lambda)u_{1,t-1}u_{2,t-1} + \lambda\sigma_{12,t-1},\end{aligned}\quad (20)$$

where the  $\lambda$ -parameter, called *persistence* is defined by the user. A smaller  $\lambda$  implies a higher reaction of the volatility to the market information in yesterday's return. The range of  $\lambda$  is usually between 0.75 (highly reactive) and 0.98 (very persistent but not highly reactive).

An  $n$ -period moving average of a time series  $x_t$  is

$$\hat{x}_t(n) = \frac{x_{t-1} + \lambda x_{t-2} + \lambda^2 x_{t-3} + \dots + \lambda^{n-1} x_{t-n}}{1 + \lambda + \lambda^2 + \dots + \lambda^{n-1}}. \quad (21)$$

Because  $0 < \lambda < 1$ ,  $n \rightarrow \infty$  the denominator converges to  $1 - \lambda$ . Thus

$$\lim_{n \rightarrow \infty} \hat{x}_t(n) = (1 - \lambda) \sum_{i=1}^{\infty} \lambda^{i-1} x_{t-i}.$$

So for volatility and correlation one first calculates the exponentially weighted variance and covariance estimates

$$\hat{\sigma}_t^2 = (1 - \lambda) \sum_{i=1}^{\infty} \lambda^{i-1} r_{t-i}^2, \quad (22)$$

and

$$\hat{\sigma}_{12,t} = (1 - \lambda) \sum_{i=1}^{\infty} \lambda^{i-1} r_{1,t-1} r_{2,t-i}. \quad (23)$$

Note. It is standard to use in above formulas daily returns,  $r_t$ , not the residuals, and not deviations from the mean.

Rewriting (21) and (22) in a recursive form gives expressions (20)

Example. In the graphs below are volatility estimate series,  $\hat{\sigma}_t = \sqrt{(1-\lambda)r_{t-1}^2 + \lambda\hat{\sigma}_{t-1}^2}$ , and correlation estimate series,  $\hat{\rho}_t = \hat{\sigma}_{12,t} / \hat{\sigma}_{1t}\hat{\sigma}_{2t}$  with  $\lambda = 0.84$  and  $\lambda = 0.96$ .

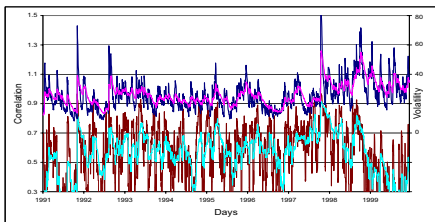


Figure. EWMA series for forest return volatility and correlation between forest and metal returns with  $\lambda = 0.84$  (blue/brown) and  $\lambda = 0.96$  (pink/cyan).

As indicated by the figure, the effect of  $\lambda$  on EWMA volatility and correlation can be quite substantial.

Remark. EWMA is a special case of the IGARCH, Integrated GARCH, which is a GARCH(1,1) model with  $\alpha + \beta = 1$

$$\sigma_t^2 = \omega + (1 - \lambda)u_{t-1}^2 + \lambda\sigma_{t-1}^2. \quad (24)$$

Note that in these models the unconditional variance is not defined (goes to infinity).

BEKK model

Generally an  $n$ -dimensional vec-model can be given as

$$\text{vech}(\Sigma_t) = \mathbf{W} + \mathbf{A} \text{vech}(\mathbf{u}_{t-1}\mathbf{u}'_{t-1}) + \mathbf{B} \text{vech}(\Sigma_{t-1}), \quad (25)$$

which becomes (16) if the coefficient matrices are defined as diagonal matrices.

There are altogether

$$\frac{n(n+1)}{2} + 2 \times \left( \frac{n(n+1)}{2} \right)^2$$

plus the parameters of the mean equation to be estimated. Thus the model can be used in the general form only in the case of, say, a two or three variables.

Even in the diagonal case it is hard to ensure the positive definiteness.

Different restrictions may lead substantial differences in the estimated model. This is why the vech model should be employed with caution.

In order to reduce the number of parameters and guarantee positive definiteness several alternatives have been suggested.

A general parameterization that involves the minimum number of parameters, while imposing no gross equation restrictions, and ensuring positive definiteness is the BEKK model (Baba, Engle, Kraft and Kroner, see Engle and Kroner 1995, *Econometric Theory* **11**, 122–150). The model for  $n$  series is of the form

$$\Sigma_t = \mathbf{C}'\mathbf{C} + \mathbf{A}'\mathbf{u}_{t-1}\mathbf{u}'_{t-1}\mathbf{A} + \mathbf{B}'\Sigma_{t-1}\mathbf{B}, \quad (26)$$

where  $\mathbf{C}$  an  $n \times n$  triangular matrix,  $\mathbf{B}$  and  $\mathbf{A}$  are  $n \times n$  matrices. It is clear the (26) is positive definite under fairly general assumptions.

Example. A bivariate case

$$\begin{aligned} \Sigma_t = & \mathbf{C}'\mathbf{C} \\ & + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}' \begin{pmatrix} u_{1,t-1}^2 & u_{1,t-1}u_{2,t-1} \\ u_{2,t-1}u_{1,t-1} & u_{2,t-1}^2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \\ & + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}' \Sigma_{t-1} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}. \end{aligned}$$

or, suppressing the time subscripts and the GARCH terms

$$\begin{aligned} \sigma_1^2 &= c_{11}^* + a_{11}^2 u_1^2 + 2a_{11}a_{21}u_1u_2 + a_{21}^2 u_2^2 \\ \sigma_{12} &= c_{12}^* + a_{11}b_{12}u_1^2 + (a_{21}b_{12} + a_{11}a_{22})u_1u_2 + a_{21}a_{22}u_2^2 \\ \sigma_2^2 &= c_{22}^* + a_{12}^2 u_1^2 + 2a_{12}a_{22}u_1u_2 + a_{22}^2 u_2^2, \end{aligned}$$

where  $c_{ij}^*$  are the relevant elements of the  $\mathbf{C}^* = \mathbf{C}'\mathbf{C}$  matrix.

In the general vech model, excluding the constant terms, there would be 18 parameters to be estimated, compared to the BEKK which has "only" 8 parameters to be estimated. Generally the number of parameters are of order  $n^2$  in a system of  $n$  variables.

As indicated by the example the formulas become highly non-linear, and there are usually convergence problems in the estimation procedures even in this bivariate case. This was also the case with our forest-metal data. Even SAS procedure VARMAX which includes a multivariate GARCH option did not find the solution.

Below is an example from RATS manual.

Several simplifications of the general BEKK model have been suggested to alleviate the computational problems.

#### (a) The Scalar BEKK

In the scalar BEKK the parameters in  $\mathbf{A}$  are the same (all series react similarly the market information). Similarly if the persistence in the correlations and variances are assumed to be identical, the parameters in  $\mathbf{B}$  are the same.

#### (b) The diagonal BEKK

Matrices  $\mathbf{A}$  and  $\mathbf{B}$  are assumed to be diagonal.

Example. The BEKK model for monthly SP500 and SP Mid Capital stock index (monthly observation from Jan, 1986 to Dec 1996). A version of RATS code can be found at [www.estima.com](http://www.estima.com) >  
Proc/Examples > ARCH/GARCH models > garchmv.prg

Modifying the program and running different  $\Sigma_t$  specifications yield the following criterion function values

GARCH Model	AIC	BIC
Simple vec	639.97	682.60
Full vec	641.49	737.85
Constant corr	636.18	671.05
Scalar BEKK	766.67	797.67
Diagonal BEKK	641.67	688.17
BEKK	641.49	695.74

The constant correlation model (considered in more detail below) seems to fit best according to both criterion functions.

Covariance matrices based on univariate GARCH

Even in the diagonal vech there are  $3n(n+1)/2$  parameters to be estimated. Thus for example in a moderate portfolio problem with 10 securities there would be 165 parameters to be estimated! In the diagonal BEKK there are  $n(n+1)/2 + 2n$  parameters to be estimated. So even in that case with 10 variables there would be 75 parameters to be estimated.

Thus computational problems become soon overwhelming even in relatively low dimensional systems. This has cast severe doubt to the practical usefulness of full multivariate GARCH modeling. However, there are some approximations that allow generate multivariate GARCH matrices by univariate GARCH.

(a) Constant correlation model

Assuming the correlation matrix  $\mathbf{R}$  time invariant, then we can write

$$\Sigma_t = \mathbf{D}_t \mathbf{R} \mathbf{D}_t, \quad (27)$$

where  $\mathbf{D}_t$  is a diagonal matrix of GARCH volatilities (standard deviations).

Thus the time varying covariances in  $\Sigma_t$  are of the form

$$\sigma_{ij,t} = \rho_{ij} \sigma_{it} \sigma_{jt}, \quad (28)$$

where  $\rho_{ij}$  is the correlation of the series  $i$  and  $j$ .

Example. A bivariate constant correlation model is defined as

$$\begin{aligned} \sigma_{1,t}^2 &= \alpha_{10} + \alpha_{11} u_{1,t-1}^2 + \beta_{11} \sigma_{1,t-1}^2 \\ \sigma_{2,t}^2 &= \alpha_{20} + \alpha_{22} u_{2,t-1}^2 + \beta_{22} \sigma_{2,t-1}^2 \\ \sigma_{12,t} &= \rho \sigma_{1,t} \sigma_{2,t}, \end{aligned} \quad (29)$$

where  $\rho$  is the contemporaneous correlation of the return series.

With the above specification and constant mean, i.e.,  $r_{i,t} = \mu_i + u_{i,t}$ ,  $i = 1, 2$ , the following estimation results are obtained for the data in the previous example.

```
MAXIMIZE - Estimation by BFGS
Convergence in 24 Iterations.
Final criterion was 0.0000051 < 0.0000100
Monthly Data From 1986:02 To 1996:12
Usable Observations 131
Function Value -386.41471767
```

Parameter	Coeff	Std Error	T-Stat	Signif
mu1	1.61439553	0.30186031	5.34815	0.00000009
mu2	1.59240710	0.34384462	4.63118	0.00000364
alpha10	17.03517135	2.62326668	6.49388	0.00000000
rho	0.92883173	0.01186534	78.28110	0.00000000
alpha20	19.22156624	3.41553898	5.62768	0.00000002
alpha11	-0.18301176	0.07652723	-2.39146	0.01678155
alpha22	-0.08607381	0.07809343	-1.10219	0.27037897
beta11	0.29150416	0.11587040	2.51578	0.01187701
beta22	0.24541543	0.14689110	1.67073	0.09477494

```
AIC = 636.18
BIC = 671.05
```

It seems that there is quite little evidence of GARCH in the volatility.

(b) Factor ARCH

In order to reduce the dimensionality factor analysis approach may prove useful.

Consider for simplicity a single factor model,

$$r_{i,t} = \mu + \gamma_i f_t + u_{i,t}, \quad (30)$$

where  $r_{i,t}$  is the return of an asset  $i$ , and  $f_t$  is the common factor for all assets (may be observable or unobservable). The underlying assumptions are that

$$\begin{aligned} \text{cov}[f_t, u_{i,t}] &= 0, \forall i, \\ E[u_{i,t} | \mathbf{f}_{t-1}, \mathbf{u}_{t-1}] &= 0, \forall i, \\ \text{cov}[\mathbf{u}_t | \mathbf{f}_{t-1}, \mathbf{u}_{t-1}] &= \Omega, \\ E[f_t | \mathbf{f}_{t-1}, \mathbf{u}_{t-1}] &= 0, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 f_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned} \quad (31)$$

where  $\sigma_t^2 = \text{var}[f_t | \mathbf{f}_{t-1}, \mathbf{u}_{t-1}]$ ,  $\Omega$  is the time invariant contemporaneous covariance matrix of  $\mathbf{u}_t = (u_{1,t}, \dots, u_{n,t})'$ , the vector of the residual terms, and  $\mathbf{f}_{t-1}$  and  $\mathbf{u}_{t-1}$  denote the historical values of the factor and residuals up to time point  $t - 1$ .

The common factor may be observable or latent. All the time varying volatility in the returns are governed by the volatility process of the single common factor.

The conditional covariance matrix of the asset returns becomes then

$$\Sigma_t = \text{cov}_{t-1}[\mathbf{r}_t] = \mathbf{g}\mathbf{g}'\sigma_t^2 + \Omega, \quad (32)$$

where  $\mathbf{g} = (\gamma_1, \dots, \gamma_n)'$ .

Engle, et al.\* consider a more general factor ARCH and give an application to model bond yield volatilities.

\*Engle, R.F., V. Ng, and M. Rothschild (1990). Asset pricing with a factor-ARCH covariance structure. *Journal of Econometrics*, 45, 213–237.

(c) Orthogonal models

In the orthogonal methods Principal Component Analysis (PCA) is utilized to orthogonalize the original returns.

Let

$$r_i = \mu_i + w_{i1}p_1 + \dots + w_{in}p_n, \quad (33)$$

where  $r_i$  is the return of the share  $i = 1, \dots, n$ ,  $p_j$  is the  $j$ th zero mean principal component,  $j = 1, \dots, n$ . If only, say,  $m < n$  first principal components are utilized, then

$$r_i = \mu_i + w_{i1}p_1 + \dots + w_{im}p_m + \epsilon_i, \quad (34)$$

where the residual  $\epsilon_i$  includes the discarded components.

Remark. PCA solution is usually calculated from the correlation matrix (standardized solution), in which case the weights must be scaled, such that in place of  $w_{ij}$  must be used  $w_{ij}^* \sigma_i$ , where  $w_{ij}^*$  is the weight of the standardized solution. Note that in the general case  $w_{ij} \neq w_{ij}^* \sigma_i$ !

In terms of (34) the covariance matrix of the returns becomes

$$\Sigma = \mathbf{W}\mathbf{D}\mathbf{W}' + \mathbf{V}_\epsilon, \quad (35)$$

where  $\mathbf{W} = (w_{ij})$  is the matrix of weights, and  $\mathbf{D} = \text{diag}(\text{var}(p_1), \dots, \text{var}(p_m))$  is the diagonal matrix of variances of principal components, and  $\mathbf{V}_\epsilon$  is the covariance matrix of residuals. In PCA  $\mathbf{V}_\epsilon$  is ignored and the approximation

$$\Sigma = \mathbf{W}\mathbf{D}\mathbf{W}' \quad (36)$$

is used.

Remark. Formula (34) defines essentially a Factor Analysis (FA) model with the difference that the residual terms in (35) are correlated.

### Orthogonal EWMA

In the basic EWMA model given by equations (22) and (23) the  $\lambda$  parameter is assumed the same in all equations. This guarantees that the covariance matrix is positive semidefinite (Exercise: prove it).

48

Utilizing PCA allows use of different weights in the PCA variances, and the matrix  $\Sigma$  defined via (36) will still be p.s.d. (Note that it is not p.d. if  $m < n$ .)

Remark. The advantage of PCA to generate risk covariance matrix is the reduction in equations. In stead of generating  $n(n+1)/2$  variance and covariance equations, only  $m$  variance equations are needed.

Usually the criterion of the successfulness of the PCA approach is how well it can reproduce the direct EWMA series.

Example. Consider daily returns from Nordic stock exchanges of Copenhagen (Den), Helsinki (Fin), Oslo (Nor) and Stockholm (Swe). The sample period is January 2, 1990 to February 4, 2000.

Below are contemporaneous correlations of the daily returns

49

Contemporaneous daily return correlations, means and standard deviations of Nordic stock exchanges [Jan 2, 1990 to Feb 4, 2000].

	DEN	FIN	NOR	SWE
DEN	1.000			
FIN	0.420	1.000		
NOR	0.486	0.452	1.000	
SWE	0.486	0.550	0.580	1.000
Mean	0.0363	0.0923	0.0379	0.0619
Std. Dev.	0.9625	1.4275	1.1457	1.1866
Obs.	2546	2546	2546	2546

Eigen values and eigen vectors computed from the contemporaneous covariance matrix

	Comp 1	Comp 2	Comp 3	Comp 4
Eigenvalue	3.619747	0.956054	0.562894	0.543779
Variance Prop.	0.637002	0.168246	0.099058	0.095694
Cumulative Prop.	0.637002	0.805248	0.904306	1.000000
Eigenvectors:				
Variable	Vector 1	Vector 2	Vector 3	Vector 4
DEN	0.334915	0.270409	0.606230	0.668727
FIN	0.640295	-0.749719	0.114759	-0.121550
NOR	0.459146	0.536906	0.240941	-0.665480
SWE	0.516758	0.276646	-0.749175	0.308488

50

Defining the EWMA variances  $\widehat{\text{var}}_t(p_i)$  for the principal components  $p_i$  as

$$\widehat{\text{var}}_t(p_i) = (1 - \lambda)p_{i,t-1}^2 + \lambda\widehat{\text{var}}_{t-1}(p_i) \quad (37)$$

we get the conditional variances

$$\hat{\sigma}_{i,t}^2 = \sum_{k=1}^4 \hat{w}_{ik}^2 \widehat{\text{var}}_t(p_k), \quad (38)$$

covariances

$$\hat{\sigma}_{ij,t} = \sum_{k=1}^4 \hat{w}_{ik} \hat{w}_{jk} \widehat{\text{var}}_t(p_k), \quad (39)$$

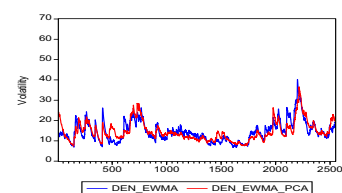
and correlations

$$\hat{\rho}_{ij,t} = \frac{\hat{\sigma}_{ij,t}}{\hat{\sigma}_{i,t} \hat{\sigma}_{j,t}}, \quad (40)$$

$i, j = 1, \dots, 4$ .

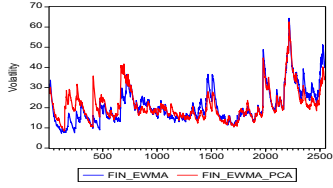
Below are graphs of EWMA and PCA-EWMA (annualized) volatilities:

Denmark.

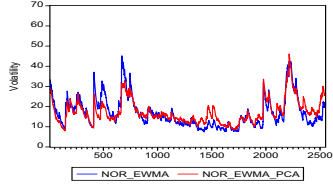


51

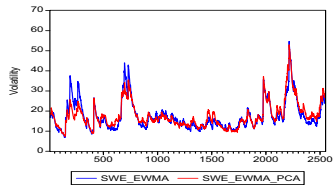
Finland



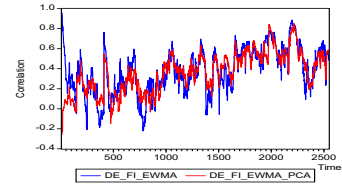
Norway



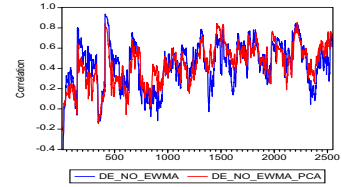
Sweden



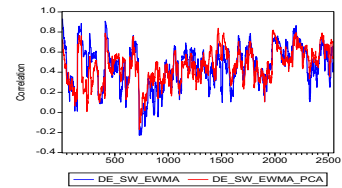
EWMA and PCA-EWMA correlations:  
Denmark–Finland



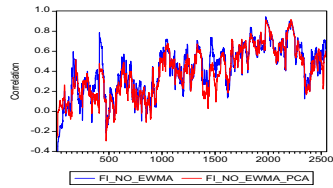
Denmark–Norway



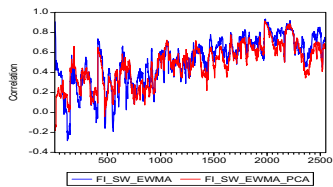
Denmark–Sweden



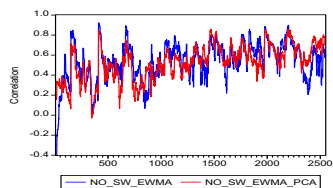
Finland–Norway



Finland–Sweden



Norway–Sweden



From the above graphs we observe that the four EWMA-PCA volatilities pretty well can reproduce the 10 EWMA volatilities and correlations.

The real computational strength of the PCA-EWMA comes with highly correlated series, where only few principal components are needed. For example bond yields of different maturities are this kind of data.

### Orthogonal GARCH

As was found earlier it is extremely difficult to use multivariate GARCH to generate covariance matrices even in the case of low dimension.

PCA-GARCH is obtained by replacing above EWMA principal component variances by GARCH variances.

Note further that PCA is just a special case of multi-factor model with orthogonal factors.

Let  $D_t$  denote the diagonal matrix of  $m \leq n$  time varying principal component variances, then the time varying covariance matrix  $V_t$  of the original variables is approximated by

$$V_t = W D_t W', \quad (41)$$

where  $W$  is, as earlier, the  $n \times m$  matrix of eigenvectors (principal component weights). Model (41) is called *orthogonal* GARCH.

**Remark.** The representation (41) always p.s.d.

**Remark.** The principal component are only unconditionally uncorrelated (orthogonal), so the assumed conditional orthogonality is just an approximation.

**Example.** Below are GARCH(1,1) models for the principal components of the Nordic indices.

Dependent Variable: P1				
Method: ML - ARCH (Marquardt)				
Date: 02/25/03				
Sample(adjusted): 2 2547				
Included observations: 2546 after adjusting endpoints				
Convergence achieved after 13 iterations				
Variance backcast: ON				
	Coefficient	Std. Error	z-Statistic	Prob.
Variance Equation				
C	0.191986	0.019683	9.753948	0.0000
ARCH(1)	0.113826	0.011293	10.07962	0.0000
GARCH(1)	0.628716	0.016554	37.98138	0.0000
R-squared	0.000000	Mean dependent var	1.98E-16	
Adjusted R-squared	-0.007786	S.D. dependent var	1.902937	
S.E. of regression	1.903685	Akaike info criterion	3.892484	
Sum squared resid	9215.875	Schwarz criterion	3.899368	
Log likelihood	-4992.133	Durbin-Watson stat	1.699886	

56

Dependent Variable: F2				
Method: ML - ARCH (Marquardt)				
Date: 02/25/03				
Sample(adjusted): 2 2547				
Included observations: 2546 after adjusting endpoints				
Convergence achieved after 16 iterations				
Variance backcast: ON				
	Coefficient	Std. Error	z-Statistic	Prob.
Variance Equation				
C	0.022161	0.003264	6.788933	0.0000
ARCH(1)	0.101581	0.008900	10.26179	0.0000
GARCH(1)	0.878827	0.010100	87.01079	0.0000
R-squared	0.000000	Mean dependent var	-7.95E-16	
Adjusted R-squared	-0.000786	S.D. dependent var	0.977972	
S.E. of regression	0.978357	Akaike info criterion	2.625073	
Sum squared resid	2434.114	Schwarz criterion	2.631957	
Log likelihood	-3338.718	Durbin-Watson stat	1.764620	

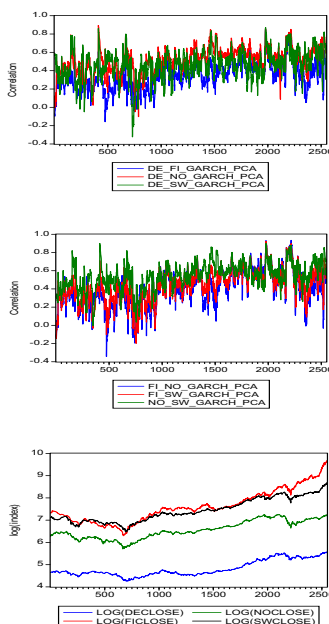
Dependent Variable: F3				
Method: ML - ARCH (Marquardt)				
Date: 02/25/03				
Sample(adjusted): 2 2547				
Included observations: 2546 after adjusting endpoints				
Convergence achieved after 13 iterations				
Variance backcast: ON				
	Coefficient	Std. Error	z-Statistic	Prob.
Variance Equation				
C	0.015461	0.003174	4.870846	0.0000
ARCH(1)	0.050795	0.008898	5.641993	0.0000
GARCH(1)	0.887398	0.011787	75.28359	0.0000
R-squared	0.000000	Mean dependent var	8.77E-16	
Adjusted R-squared	-0.000786	S.D. dependent var	0.750410	
S.E. of regression	0.750705	Akaike info criterion	2.106615	
Sum squared resid	1433.128	Schwarz criterion	2.113499	
Log likelihood	-2678.721	Durbin-Watson stat	1.769628	

Dependent Variable: F4				
Method: ML - ARCH (Marquardt)				
Date: 02/25/03				
Sample(adjusted): 2 2547				
Included observations: 2546 after adjusting endpoints				
Convergence achieved after 13 iterations				
Variance backcast: ON				
	Coefficient	Std. Error	z-Statistic	Prob.
Variance Equation				
C	0.064904	0.001477	3.319334	0.0009
ARCH(1)	0.059132	0.006038	9.793347	0.0000
GARCH(1)	0.933042	0.006701	139.2406	0.0000
R-squared	0.000000	Mean dependent var	-6.62E-16	
Adjusted R-squared	-0.000786	S.D. dependent var	0.737559	
S.E. of regression	0.737848	Akaike info criterion	2.088950	
Sum squared resid	1384.461	Schwarz criterion	2.106534	
Log likelihood	-2668.963	Durbin-Watson stat	1.805472	

57

The conditional variances, covariances and correlations again estimated as in (38)–(40).

Below are graphs of the resulted conditional correlations.



58

It would be important here to compare the estimation results with BEKK and vech model to see how the relatively simple orthogonal GARCH compares with them. We however skip the comparison and look at some other approaches to analyze large scale systems.

Before that let us summarize the orthogonal approach. The idea is simply that:

- (i) Find a transformation  $W$  such that  $y_t = W r_t$ , and  $\text{cov}[y_t] = D$  is *diagonal*.
- (ii) Estimate GARCH for the components of the  $y_t$  vector and compile the conditional variances to the diagonal matrix  $D_t$ .
- (iii) Under the (strong) assumption  $\text{cov}_{t-1}[y_t] = D_t$ , we get

$$\Sigma_t = \text{cov}_{t-1}[r_t] = W^{-1} D_t W^{-1}. \quad (42)$$

59