CHAPTER 10

Multivariate Volatility Models and Their Applications

In this chapter, we generalize the univariate volatility models of Chapter 3 to the multivariate case and discuss some simple methods for modeling the dynamic relationships between volatility processes of multiple asset returns. By multivariate volatility, we mean the conditional covariance matrix of multiple asset returns. Multivariate volatilities have many important financial applications. They play an important role in portfolio selection and asset allocation, and they can be used to compute the value at risk of a financial position consisting of multiple assets.

Consider a multivariate return series $\{\boldsymbol{r}_t\}$. We adopt the same approach as the univariate case by rewriting the series as

$$\boldsymbol{r}_t = \boldsymbol{\mu}_t + \boldsymbol{a}_t,$$

where $\boldsymbol{\mu}_t = E(\boldsymbol{r}_t|\boldsymbol{F}_{t-1})$ is the conditional expectation of $\boldsymbol{r}_t$ given the past information $\boldsymbol{F}_{t-1}$, and $\boldsymbol{a}_t = (a_{1t}, \cdots, a_{kt})'$ is the shock, or innovation, of the series at time t. In addition, we assume that $\boldsymbol{r}_t$ follows a multivariate time series model of Chapter 8 so that $\boldsymbol{\mu}_t$ is the 1-step-ahead prediction of the model. For most return series, it suffices to employ a simple vector ARMA structure with exogenous variables for $\boldsymbol{\mu}_t$—that is,

$$\boldsymbol{\mu}_t = \boldsymbol{\Upsilon}\boldsymbol{x}_t + \sum_{i=1}^{p}\boldsymbol{\Phi}_i\boldsymbol{r}_{t-i} - \sum_{i=1}^{q}\boldsymbol{\Theta}_i\boldsymbol{a}_{t-i}, \tag{10.1}$$

where $\boldsymbol{x}_t$ denotes an m-dimensional vector of exogenous (or explanatory) variables with $x_{1t} = 1$, $\boldsymbol{\Upsilon}$ is a $k \times m$ matrix, and p and q are nonnegative integers. We refer to Eq. (10.1) as the mean equation of $\boldsymbol{r}_t$.

Analysis of Financial Time Series, Third Edition, By Ruey S. Tsay

The conditional covariance matrix of $\boldsymbol{a}_t$ given $\boldsymbol{F}_{t-1}$ is a $k \times k$ positive-definite matrix $\boldsymbol{\Sigma}_t$ defined by $\boldsymbol{\Sigma}_t = \text{Cov}(\boldsymbol{a}_t|\boldsymbol{F}_{t-1})$. Multivariate volatility modeling is concerned with the time evolution of $\boldsymbol{\Sigma}_t$. We refer to a model for the $\{\boldsymbol{\Sigma}_t\}$ process as a volatility model for the return series $\boldsymbol{r}_t$.

There are many ways to generalize univariate volatility models to the multivariate case, but the curse of dimensionality quickly becomes a major obstacle in applications because there are $k(k+1)/2$ quantities in $\boldsymbol{\Sigma}_t$ for a k-dimensional return series. To illustrate, there are 15 conditional variances and covariances in $\boldsymbol{\Sigma}_t$ for a five-dimensional return series. The goal of this chapter is to introduce some relatively simple multivariate volatility models that are useful, yet remain manageable in real application. In particular, we discuss some models that allow for time-varying correlation coefficients between asset returns. Time-varying correlations are useful in finance. For example, they can be used to estimate the time-varying beta of the market model for a return series.

We begin by using an exponentially weighted approach to estimate the covariance matrix in Section 10.1. This estimated covariance matrix can serve as a benchmark for multivariate volatility estimation. Section 10.2 discusses some generalizations of univariate GARCH models that are available in the literature. We then introduce two methods to reparameterize $\boldsymbol{\Sigma}_t$ for volatility modeling in Section 10.3. The reparameterization based on the Cholesky decomposition is found to be useful. We study some volatility models for bivariate returns in Section 10.4, using the GARCH model as an example. In this particular case, the volatility model can be bivariate or three dimensional. Section 10.5 is concerned with volatility models for higher dimensional returns and Section 10.6 addresses the issue of dimension reduction. We demonstrate some applications of multivariate volatility models in Section 10.7. Finally, Section 10.8 gives a multivariate Student-t distribution useful for volatility modeling.

10.1 EXPONENTIALLY WEIGHTED ESTIMATE

Given the innovations $F_{t-1} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{t-1}\}$, the (unconditional) covariance matrix of the innovation can be estimated by

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{t-1}\sum_{j=1}^{t-1} \boldsymbol{a}_j \boldsymbol{a}_j',$$

where it is understood that the mean of $\boldsymbol{a}_j$ is zero. This estimate assigns equal weight $1/(t-1)$ to each term in the summation. To allow for a time-varying covariance matrix and to emphasize that recent innovations are more relevant, one can use the idea of exponential smoothing and estimate the covariance matrix of $\boldsymbol{a}_t$ by

$$\widehat{\boldsymbol{\Sigma}}_t = \frac{1-\lambda}{1-\lambda^{t-1}}\sum_{j=1}^{t-1} \lambda^{j-1} \boldsymbol{a}_{t-j}\boldsymbol{a}_{t-j}', \tag{10.2}$$

where $0 < \lambda < 1$ and the weights $(1-\lambda)\lambda^{j-1}/(1-\lambda^{t-1})$ sum to one. For a sufficiently large t such that $\lambda^{t-1} \approx 0$, the prior equation can be rewritten as

$$\widehat{\boldsymbol{\Sigma}}_t = (1-\lambda)\boldsymbol{a}_{t-1}\boldsymbol{a}'_{t-1} + \lambda\widehat{\boldsymbol{\Sigma}}_{t-1}.$$

Therefore, the covariance estimate in Eq. (10.2) is referred to as the exponentially weighted moving-average (EWMA) estimate of the covariance matrix.

Suppose that the return data are $\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_T\}$. For a given λ and initial estimate $\widehat{\boldsymbol{\Sigma}}_1$, $\widehat{\boldsymbol{\Sigma}}_t$ can be computed recursively. If one assumes that $\boldsymbol{a}_t = \boldsymbol{r}_t - \boldsymbol{\mu}_t$ follows a multivariate normal distribution with mean zero and covariance matrix $\boldsymbol{\Sigma}_t$, where $\boldsymbol{\mu}_t$ is a function of parameter $\boldsymbol{\Theta}$, then λ and $\boldsymbol{\Theta}$ can be estimated jointly by the maximum-likelihood method because the log-likelihood function of the data is

$$\ln L(\boldsymbol{\Theta}, \lambda) \propto -\frac{1}{2}\sum_{t=1}^{T}|\boldsymbol{\Sigma}_t| - \frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{r}_t - \boldsymbol{\mu}_t)'\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{r}_t - \boldsymbol{\mu}_t),$$

which can be evaluated recursively by substituting $\widehat{\boldsymbol{\Sigma}}_t$ for $\boldsymbol{\Sigma}_t$.

Example 10.1. To illustrate, consider the daily log returns of the Hang Seng index of Hong Kong and the Nikkei 225 index of Japan from January 4, 2006, to December 30, 2008, for 713 observations. The indexes were obtained from Yahoo Finance. For simplicity, we only employ data when both markets were open to calculate the log returns, which are in percentages. Figure 10.1 shows the time plots of the two index returns. The effect of recent global financial crisis is clearly seen from the plots. Let r_{1t} and r_{2t} be the log returns of the Hong Kong and Japanese markets, respectively. If univariate GARCH models are entertained, we obtain the models

$$\begin{aligned} r_{1t} &= 0.109 + a_{1t}, \qquad a_{1t} = \sigma_{1t}\epsilon_{1t}, \\ \sigma_{1t}^2 &= 0.038 + 0.143a_{1,t-1}^2 + 0.855\sigma_{1,t-1}^2, \qquad (10.3) \\ r_{2t} &= 0.003 + a_{2t}, \qquad a_{2t} = \sigma_{2t}\epsilon_{2t}, \\ \sigma_{2t}^2 &= 0.044 + 0.127a_{2,t-1}^2 + 0.861\sigma_{2,t-1}^2, \qquad (10.4) \end{aligned}$$

where all of the parameter estimates are significant at the 5% level except for the constant term of the mean equation for the Nikkei 225 index returns. The Ljung–Box statistics of the standardized residuals and their squared series of the two univariate models fail to indicate any model inadequacy. The two volatility equations are close to an IGARCH(1,1) model. This is reasonable because of the increased volatility caused by the subprime financial crisis. Figure 10.2 shows the estimated volatilities of the two univariate GARCH(1,1) models. Indeed, the volatility series confirm that both markets were more volatile than usual in 2008.

Turn to bivariate modeling. We apply the EWMA approach to obtain volatility estimates, using the command `mgarch` in S-Plus FinMetrics:

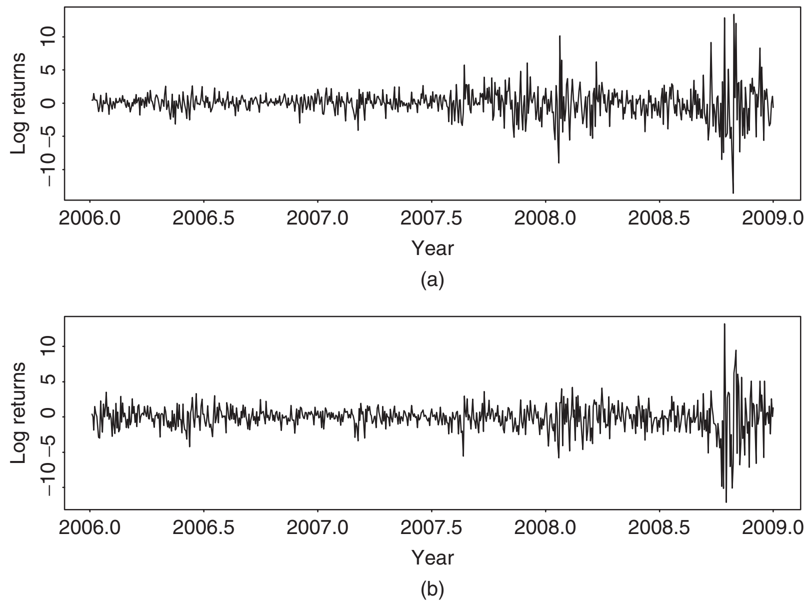


Figure 10.1 Time plots of daily log returns in percentages of stock market indexes for Hong Kong and Japan from January 4, 2006, to December 30, 2008: (a) Hong Kong market and (b) Japanese market.

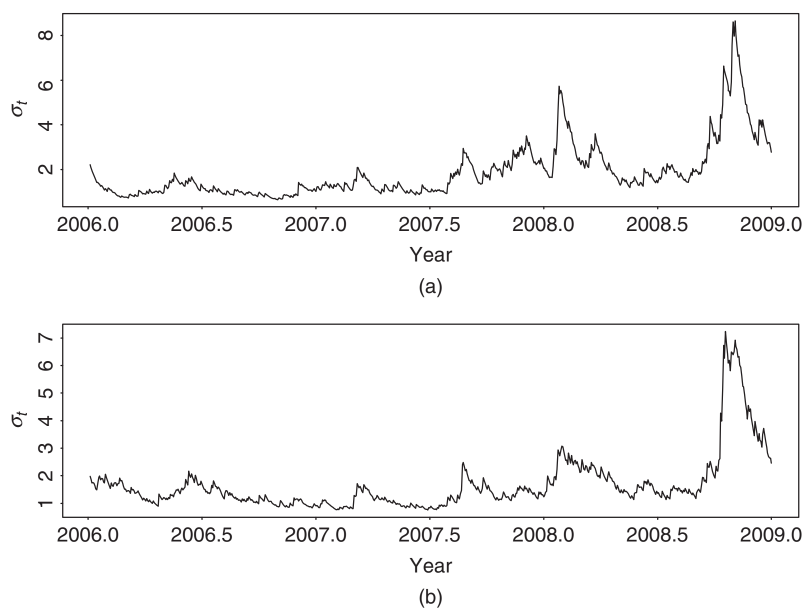


Figure 10.2 Estimated volatilities (standard error) for daily log returns in percentages of stock market indexes for Hong Kong and Japan from January 4, 2006, to December 30, 2008: (a) Hong Kong market and (b) Japanese market. Univariate models are used.

```
> m3=mgarch(formula.mean=~arma(0,0),formula.var=~ewma1,
   series=rtn,trace=F)
> summary(m3)
Call:
mgarch(formula.mean =~arma(0,0), formula.var=~ewma1,
   series=rtn,trace = F)
Mean Equation: structure(.Data = ~arma(0,0), class="formula")
Conditional Var. Eq.: structure(.Data=~ewma1,class="formula")

Conditional Distribution:  gaussian
--------------------------------------------------------------
Estimated Coefficients:
--------------------------------------------------------------
           Value Std.Error  t value Pr(>|t|)
 C(1)  0.082425  0.030900   2.6675 0.007816
 C(2) -0.006849  0.030093  -0.2276 0.820020
ALPHA  0.069492  0.004945  14.0517 0.000000
```

The estimate of λ is $1 - \hat{\alpha} = 1 - 0.0695 \approx 0.9305$, which is in the typical range commonly seen in practice. Figure 10.3 shows the estimated volatility series by the

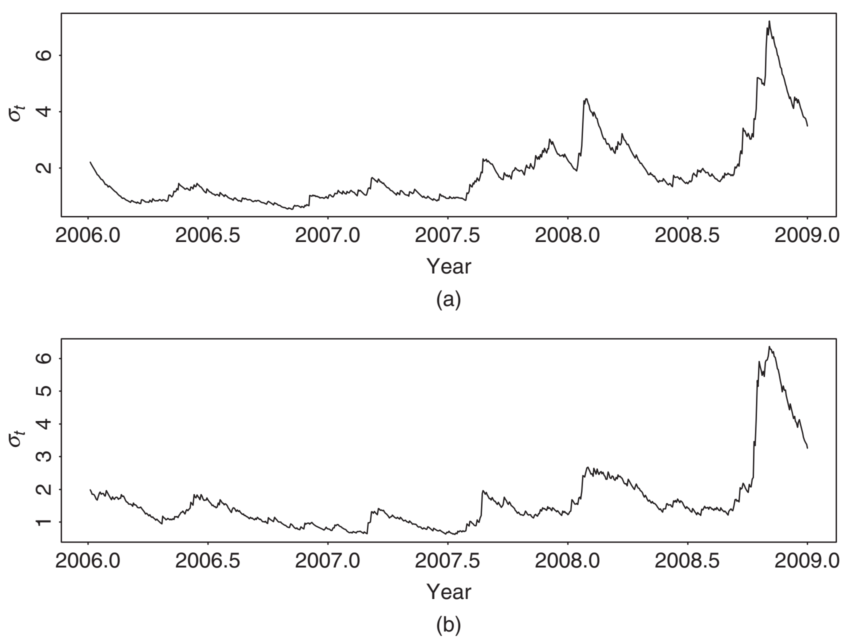


Figure 10.3 Estimated volatilities (standard error) for daily log returns in percentages of stock market indices for Hong Kong and Japan from January 4, 2006, to December 30, 2008: (a) Hong Kong market and (b) Japanese market. Exponentially weighted moving-average approach is used.

EWMA approach. Compared with those in Figure 10.2, the EWMA approach produces smoother volatility series, even though the two plots show similar volatility patterns.

10.2 SOME MULTIVARIATE GARCH MODELS

Many authors have generalized univariate volatility models to the multivariate case. In this section, we discuss some of the generalizations. For more details, readers are referred to the survey article of Bauwens, Laurent, and Rombouts (2004).

10.2.1 Diagonal Vectorization (VEC) Model

Bollerslev, Engle, and Wooldridge (1988) generalize the exponentially weighted moving-average approach to propose the model

$$\boldsymbol{\Sigma}_t = \boldsymbol{A}_0 + \sum_{i=1}^{m} \boldsymbol{A}_i \odot (\boldsymbol{a}_{t-i}\boldsymbol{a}'_{t-i}) + \sum_{j=1}^{s} \boldsymbol{B}_j \odot \boldsymbol{\Sigma}_{t-j}, \tag{10.5}$$

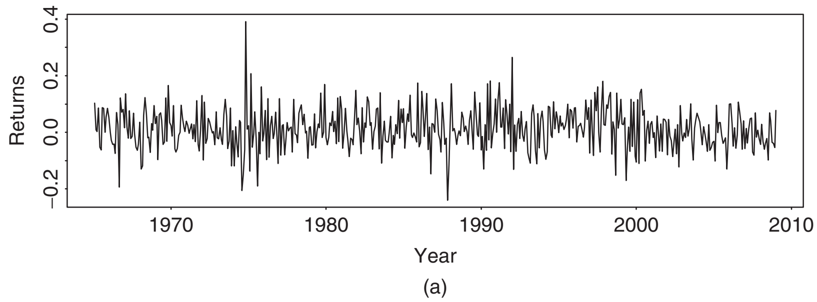
where m and s are nonnegative integers, $\boldsymbol{A}_i$ and $\boldsymbol{B}_j$ are symmetric matrices, and $\odot$ denotes the Hadamard product, that is, element-by-element multiplication. This is referred to as the diagonal VEC(m, s) model or DVEC(m, s) model. To appreciate the model, consider the bivariate DVEC(1,1) case satisfying

$$\begin{aligned}
\begin{bmatrix} \sigma_{11,t} & \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix} = {} & \begin{bmatrix} A_{11,0} & \\ A_{21,0} & A_{22,0} \end{bmatrix} \\
& + \begin{bmatrix} A_{11,1} & \\ A_{21,1} & A_{22,1} \end{bmatrix} \odot \begin{bmatrix} a^2_{1,t-1} & \\ a_{1,t-1}a_{2,t-1} & a^2_{2,t-1} \end{bmatrix} \\
& + \begin{bmatrix} B_{11,1} & \\ B_{21,1} & B_{22,1} \end{bmatrix} \odot \begin{bmatrix} \sigma_{11,t-1} & \\ \sigma_{21,t-1} & \sigma_{22,t-1} \end{bmatrix},
\end{aligned}$$

where only the lower triangular part of the model is given. Specifically, the model is

$$\begin{aligned}
\sigma_{11,t} &= A_{11,0} + A_{11,1}a^2_{1,t-1} + B_{11,1}\sigma_{11,t-1}, \\
\sigma_{21,t} &= A_{21,0} + A_{21,1}a_{1,t-1}a_{2,t-1} + B_{21,1}\sigma_{21,t-1}, \\
\sigma_{22,t} &= A_{22,0} + A_{22,1}a^2_{2,t-1} + B_{22,1}\sigma_{22,t-1},
\end{aligned}$$

where each element of $\boldsymbol{\Sigma}_t$ depends only on its own past value and the corresponding product term in $\boldsymbol{a}_{t-1}\boldsymbol{a}'_{t-1}$. That is, each element of a DVEC model follows a GARCH(1,1)-type model. The model is, therefore, simple. However, it may not produce a positive-definite covariance matrix. Furthermore, the model does not allow for dynamic dependence between volatility series.



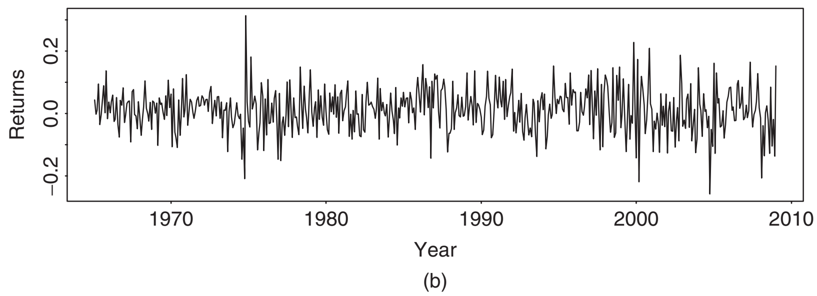


Figure 10.4 Time plot of monthly simple returns, including dividends, for Pfizer and Merck stocks from January 1965 to December 2008: (a) Pfizer stock and (b) Merck stock.

Example 10.2. For illustration, consider the monthly simple returns, including dividends, of two U.S. major drug companies from January 1965 to December 2008 for 528 observations. Let r_{1t} and r_{2t} be the monthly returns of Pfizer and Merck stock, respectively. The bivariate return series $\boldsymbol{r}_t = (r_{1t}, r_{2t})'$, shown in Figure 10.4, has no significant serial correlations with $Q(10)$ being 10.48(0.40) and 11.42(0.33), respectively, for the two series. Therefore, the mean equation of $\boldsymbol{r}_t$ consists of a constant term only. We fit a DVEC(1,1) model to the series using the command `mgarch` in FinMetrics of S-Plus:

```
> rtn=cbind(pfe,mrk)  % Output edited.
> mdvec=mgarch(rtn~1,~dvec(1,1))
> summary(mdvec)
Call:
mgarch(formula.mean=rtn ~ 1, formula.var= ~ dvec(1, 1))
Mean Equation: structure(.Data =rtn ~ 1, class="formula")
Conditional Var. Eq.: structure(.Data=~dvec(1,1),
   class="formula")
Conditional Distribution:  gaussian
-------------------------------------------------------------
Estimated Coefficients:
-------------------------------------------------------------
                   Value Std.Error t value  Pr(>|t|)
          C(1) 1.350e-02 3.149e-03   4.285 2.174e-05
```

```
          C(2) 1.313e-02 3.043e-03   4.314 1.921e-05
       A(1, 1) 7.544e-04 3.939e-04   1.916 5.597e-02
       A(2, 1) 7.543e-05 3.468e-05   2.175 3.010e-02
       A(2, 2) 7.941e-05 3.871e-05   2.051 4.072e-02
  ARCH(1; 1, 1) 7.078e-02 2.757e-02   2.568 1.051e-02
  ARCH(1; 2, 1) 2.513e-02 8.492e-03   2.960 3.220e-03
  ARCH(1; 2, 2) 4.095e-02 1.213e-02   3.375 7.939e-04
 GARCH(1; 1, 1) 7.858e-01 9.055e-02   8.677 0.000e+00
 GARCH(1; 2, 1) 9.499e-01 1.671e-02  56.831 0.000e+00
 GARCH(1; 2, 2) 9.454e-01 1.469e-02  64.358 0.000e+00
-------------------------------------------------------------
Ljung-Box test for standardized residuals:
-------------------------------------------------------------
    Statistic P-value Chi^2-d.f.
pfe     9.531  0.6570         12
mrk    12.349  0.4181         12

Ljung-Box test for squared standardized residuals:
-------------------------------------------------------------
    Statistic P-value Chi^2-d.f.
pfe    22.077 0.03666         12
mrk     6.437 0.89246         12
> names(mdvec)
 [1] "residuals"    "sigma.t"       "df.residual"  "coef"
 [5] "model"        "cond.dist"     "likelihood"   "opt.index"
 [9] "cov"          "std.residuals" "R.t"          "S.t"
[13] "prediction"   "call"          "series"
```

From the output, all parameter estimates, but A(1,1), are significant at the 5% level, and the fitted volatility model is

$$
\begin{aligned}
\sigma_{11,t} &= 0.00075 + 0.071a_{1,t-1}^2 + 0.786\sigma_{11,t-1},\\
\sigma_{21,t} &= 0.00008 + 0.025a_{1,t-1}a_{2,t-1} + 0.950\sigma_{21,t-1},\\
\sigma_{22,t} &= 0.00008 + 0.041a_{2,t-1}^2 + 0.945\sigma_{22,t-1}.
\end{aligned}
$$

The output also provides some model checking statistics for individual stock returns. For instance, the Ljung–Box statistics for the standardized residual series and its squared series of Pfizer stock returns give $Q(12) = 9.53(0.66)$ and $Q(12) = 12.35(0.42)$, respectively, where the number in parentheses denotes the p value. Thus, checking the fitted model individually, one cannot reject the DVEC(1,1) model. A more informative model-checking approach is to apply the multivarite Q statistics to the bivariate standardized residual series and its squared process. Details are omitted. Interested readers are referred to Li (2004). Figure 10.5 shows the fitted volatility and correlation series. These series are stored in “sigma.t” and “R.t”, respectively. The correlations range from 0.37 to 0.83.

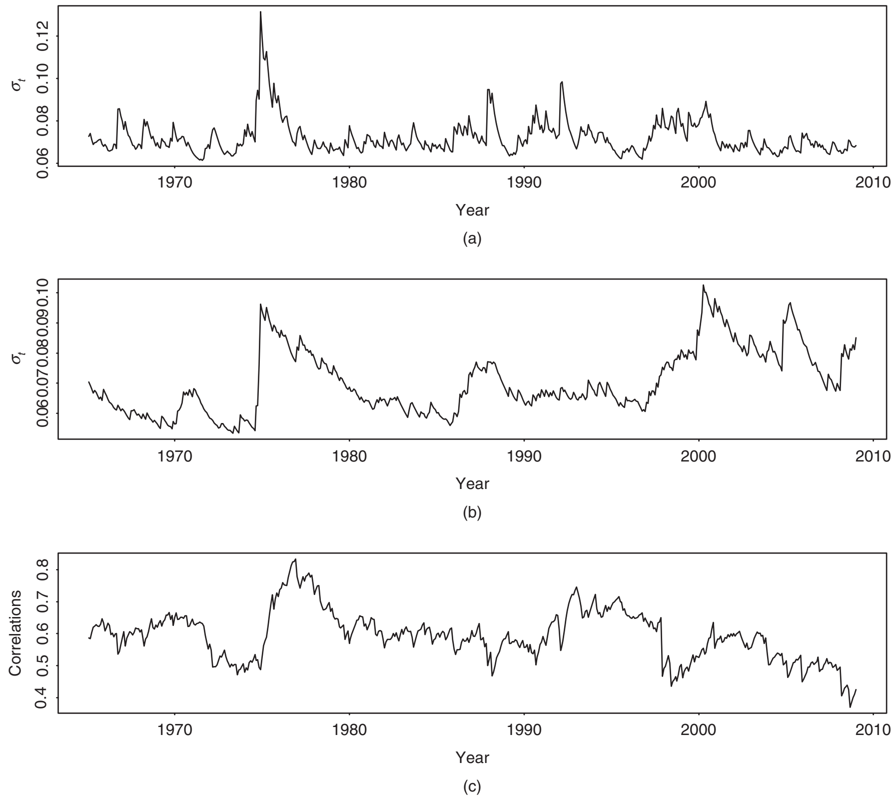


Figure 10.5 Estimated volatilities (standard error) and time-varying correlations of DVEC(1,1) model for monthly simple returns of two major drug companies from January 1965 to December 2008: (a) Pfizer stock volatility, (b) Merck stock volatility, and (c) time-varying correlations.

10.2.2 BEKK Model

To guarantee the positive-definite constraint, Engle and Kroner (1995) propose the Baba-Engle-Kraft-Kroner (BEKK) model,

$$\boldsymbol{\Sigma}_t = \boldsymbol{A}\boldsymbol{A}' + \sum_{i=1}^{m} \boldsymbol{A}_i(\boldsymbol{a}_{t-i}\boldsymbol{a}'_{t-i})\boldsymbol{A}'_i + \sum_{j=1}^{s} \boldsymbol{B}_j \boldsymbol{\Sigma}_{t-j} \boldsymbol{B}'_j, \tag{10.6}$$

where $\boldsymbol{A}$ is a lower triangular matrix and $\boldsymbol{A}_i$ and $\boldsymbol{B}_j$ are $k \times k$ matrices. Based on the symmetric parameterization of the model, $\boldsymbol{\Sigma}_t$ is almost surely positive definite provided that $\boldsymbol{A}\boldsymbol{A}'$ is positive definite. This model also allows for dynamic dependence between the volatility series. On the other hand, the model has several disadvantages. First, the parameters in $\boldsymbol{A}_i$ and $\boldsymbol{B}_j$ do not have direct interpretations concerning lagged values of volatilities or shocks. Second, the number of parameters employed is $k^2(m+s) + k(k+1)/2$, which increases rapidly with m and s.

Limited experience shows that many of the estimated parameters are statistically insignificant, introducing additional complications in modeling.

Example 10.3. To illustrate, we consider the monthly simple returns of Pfizer and Merck stocks of Example 10.2 and employ a BEKK(1,1) model. Again, S-Plus is used to perform the estimation:

```
> mbekk=mgarch(rtn~1,~bekk(1,1))
> summary(mbekk)
Call:
mgarch(formula.mean = rtn ~ 1, formula.var = ~ bekk(1, 1))
Mean Equation: structure(.Data = rtn ~ 1, class = "formula")
Conditional Var. Eq.: structure(.Data=~bekk(1,1),
   class="formula")
Conditional Distribution:  gaussian
--------------------------------------------------------------
Estimated Coefficients:
--------------------------------------------------------------
                    Value  Std.Error    t value  Pr(>|t|)
          C(1)  1.329e-02  0.003247  4.094e+00 4.907e-05
          C(2)  1.269e-02  0.003095  4.100e+00 4.792e-05
       A(1, 1)  2.505e-02  0.008382  2.988e+00 2.938e-03
       A(2, 1)  1.349e-02  0.004979  2.710e+00 6.946e-03
       A(2, 2)  3.272e-06  8.453262  3.870e-07 1.000e+00
 ARCH(1; 1, 1)  2.129e-01  0.084340  2.524e+00 1.190e-02
 ARCH(1; 2, 1)  9.963e-02  0.072156  1.381e+00 1.680e-01
 ARCH(1; 1, 2)  6.336e-02  0.076065  8.330e-01 4.052e-01
 ARCH(1; 2, 2)  1.824e-01  0.062133  2.936e+00 3.467e-03
GARCH(1; 1, 1)  9.090e-01  0.063239  1.437e+01 0.000e+00
GARCH(1; 2, 1) -5.888e-02  0.047766 -1.233e+00 2.182e-01
GARCH(1; 1, 2) -8.231e-03  0.031512 -2.612e-01 7.940e-01
GARCH(1; 2, 2)  9.824e-01  0.022587  4.349e+01 0.000e+00
--------------------------------------------------------------
Ljung-Box test for standardized residuals:
--------------------------------------------------------------
    Statistic P-value Chi^2-d.f.
pfe    9.465  0.6628         12
mrk   11.591  0.4791         12

Ljung-Box test for squared standardized residuals:
--------------------------------------------------------------
    Statistic P-value Chi^2-d.f.
pfe     21.55 0.04291         12
mrk      9.19 0.68664         12
```

Model-checking statistics based on the individual residual series and provided by S-Plus fail to suggest any model inadequacy of the fitted BEKK(1,1) model. Figure 10.6 shows the fitted volatilities and the time-varying correlations of the

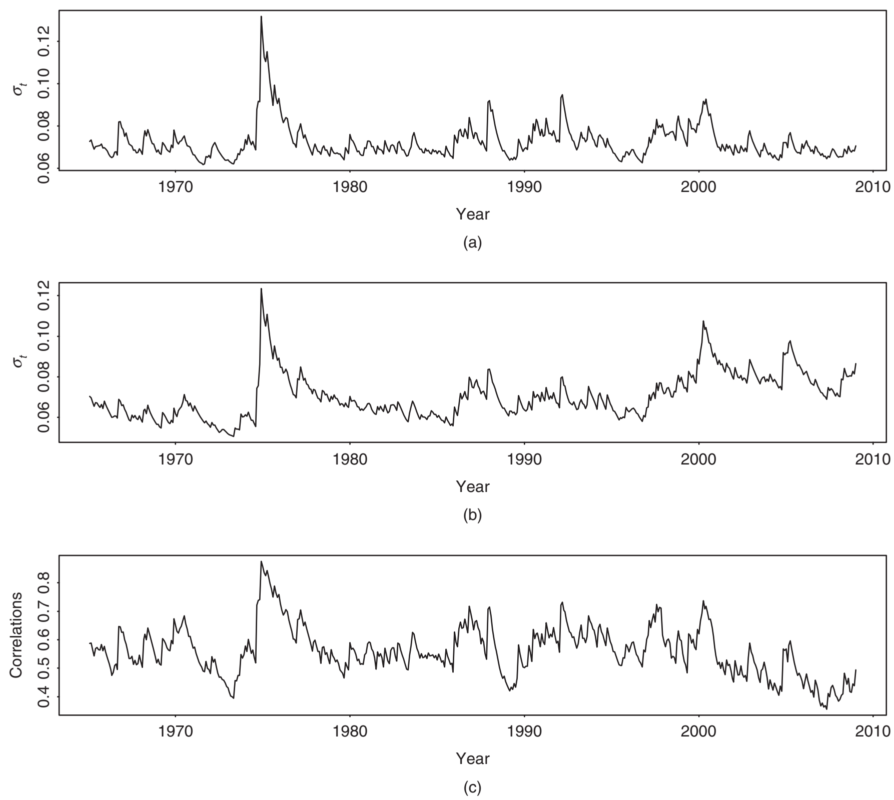


Figure 10.6 Estimated volatilities (standard error) and time-varying correlations of BEKK(1,1) model for monthly simple returns of two major drug companies from January 1965 to December 2008: (a) Pfizer stock volatility, (b) Merck stock volatility, and (c) time-varying correlations.

BEKK(1,1) model. Compared with Figure 10.5, there are some differences between the two fitted volatility models. For instance, the time-varying correlations of the BEKK(1,1) model appear to be more volatile.

The volatility equation of the fitted BEKK(1,1) model is

$$\begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} 0.025 & 0 \\ 0.013 & 3\times 10^{-6} \end{bmatrix} \begin{bmatrix} 0.025 & 0.013 \\ 0 & 3\times 10^{-6} \end{bmatrix}$$
$$+ \begin{bmatrix} 0.213 & 0.063 \\ 0.100 & 0.182 \end{bmatrix} \begin{bmatrix} a_{1,t-1}^2 & a_{1,t-1}a_{2,t-1} \\ a_{2,t-1}a_{1,t-1} & a_{2,t-1}^2 \end{bmatrix}$$
$$\begin{bmatrix} 0.213 & 0.100 \\ 0.063 & 0.182 \end{bmatrix} + \begin{bmatrix} 0.901 & -0.008 \\ -0.059 & 0.982 \end{bmatrix}$$
$$\begin{bmatrix} \sigma_{11,t-1} & \sigma_{12,t-1} \\ \sigma_{21,t-1} & \sigma_{22,t-1} \end{bmatrix} \begin{bmatrix} 0.901 & -0.059 \\ -0.008 & 0.982 \end{bmatrix},$$

where three estimates are insignificant at the 5% level. In general, the BEKK model tends to contain some insignificant parameter estimates, and one needs to perform matrix multiplication to decipher the fitted model.

10.3 REPARAMETERIZATION

A useful step in multivariate volatility modeling is to reparameterize $\boldsymbol{\Sigma}_t$ by making use of its symmetric property. We consider two reparameterizations.

10.3.1 Use of Correlations

The first reparameterization of $\boldsymbol{\Sigma}_t$ is to use the conditional correlation coefficients and variances of $\boldsymbol{a}_t$. Specifically, we write $\boldsymbol{\Sigma}_t$ as

$$\boldsymbol{\Sigma}_t \equiv [\sigma_{ij,t}] = \boldsymbol{D}_t \boldsymbol{\rho}_t \boldsymbol{D}_t, \tag{10.7}$$

where $\boldsymbol{\rho}_t$ is the conditional correlation matrix of $\boldsymbol{a}_t$, and $\boldsymbol{D}_t$ is a $k \times k$ diagonal matrix consisting of the conditional standard deviations of elements of $\boldsymbol{a}_t$ (i.e., $\boldsymbol{D}_t = \text{diag}\{\sqrt{\sigma_{11,t}}, \ldots, \sqrt{\sigma_{kk,t}}\}$).

Because $\boldsymbol{\rho}_t$ is symmetric with unit diagonal elements, the time evolution of $\boldsymbol{\Sigma}_t$ is governed by that of the conditional variances $\sigma_{ii,t}$ and the elements $\rho_{ij,t}$ of $\boldsymbol{\rho}_t$, where $j < i$ and $1 \le i \le k$. Therefore, to model the volatility of $\boldsymbol{a}_t$, it suffices to consider the conditional variances and correlation coefficients of a_{it}. Define the $k(k+1)/2$-dimensional vector

$$\boldsymbol{\Xi}_t = (\sigma_{11,t}, \ldots, \sigma_{kk,t}, \boldsymbol{\varrho}_t')', \tag{10.8}$$

where $\boldsymbol{\varrho}_t$ is a $k(k-1)/2$-dimensional vector obtained by stacking columns of the correlation matrix $\boldsymbol{\rho}_t$, but using only elements below the main diagonal. Specifically, for a k-dimensional return series,

$$\boldsymbol{\varrho}_t = (\rho_{21,t}, \ldots, \rho_{k1,t} | \rho_{32,t}, \ldots, \rho_{k2,t} | \cdots | \rho_{k,k-1,t})'.$$

To illustrate, for $k = 2$, we have $\boldsymbol{\varrho}_t = \rho_{21,t}$ and

$$\boldsymbol{\Xi}_t = (\sigma_{11,t}, \sigma_{22,t}, \rho_{21,t})', \tag{10.9}$$

which is a three-dimensional vector, and for $k = 3$, we have $\boldsymbol{\varrho}_t = (\rho_{21,t}, \rho_{31,t}, \rho_{32,t})'$ and

$$\boldsymbol{\Xi}_t = (\sigma_{11,t}, \sigma_{22,t}, \sigma_{33,t}, \rho_{21,t}, \rho_{31,t}, \rho_{32,t})', \tag{10.10}$$

which is a six-dimensional random vector.

If $\boldsymbol{a}_t$ is a bivariate normal random variable, then $\boldsymbol{\Xi}_t$ is given in Eq. (10.9) and the conditional density function of $\boldsymbol{a}_t$ given $\boldsymbol{F}_{t-1}$ is

$$f(a_{1t}, a_{2t}|\boldsymbol{\Xi}_t) = \frac{1}{2\pi\sqrt{\sigma_{11,t}\sigma_{22,t}(1-\rho_{21,t}^2)}} \exp\left[-\frac{Q(a_{1t}, a_{2t}, \boldsymbol{\Xi}_t)}{2(1-\rho_{21,t}^2)}\right],$$

where

$$Q(a_{1t}, a_{2t}, \boldsymbol{\Xi}_t) = \frac{a_{1t}^2}{\sigma_{11,t}} + \frac{a_{2t}^2}{\sigma_{22,t}} - \frac{2\rho_{21,t}a_{1t}a_{2t}}{\sqrt{\sigma_{11,t}\sigma_{22,t}}}.$$

The log probability density function of $\boldsymbol{a}_t$ relevant to the maximum-likelihood estimation is

$$\ell(a_{1t}, a_{2t}, \boldsymbol{\Xi}_t) = -\frac{1}{2}\left\{\ln[\sigma_{11,t}\sigma_{22,t}(1-\rho_{21,t}^2)] + \frac{1}{1-\rho_{21,t}^2}\left(\frac{a_{1t}^2}{\sigma_{11,t}} + \frac{a_{2t}^2}{\sigma_{22,t}} - \frac{2\rho_{21,t}a_{1t}a_{2t}}{\sqrt{\sigma_{11,t}\sigma_{22,t}}}\right)\right\}. \quad (10.11)$$

This reparameterization is useful because it models covariances and correlations directly. Yet the approach has several weaknesses. First, the likelihood function becomes complicated when $k \geq 3$. Second, the approach requires a constrained maximization in estimation to ensure the positive definiteness of $\boldsymbol{\Sigma}_t$. The constraint becomes complicated when k is large.

10.3.2 Cholesky Decomposition

The second reparameterization of $\boldsymbol{\Sigma}_t$ is to use the Cholesky decomposition; see Appendix A of Chapter 8. This approach has some advantages in estimation as it requires no parameter constraints for the positive definiteness of $\boldsymbol{\Sigma}_t$; see Pourahmadi (1999). In addition, the reparameterization is an orthogonal transformation so that the resulting likelihood function is extremely simple. Details of the transformation are given next.

Because $\boldsymbol{\Sigma}_t$ is positive definite, there exist a lower triangular matrix $\boldsymbol{L}_t$ with unit diagonal elements and a diagonal matrix $\boldsymbol{G}_t$ with positive diagonal elements such that

$$\boldsymbol{\Sigma}_t = \boldsymbol{L}_t\boldsymbol{G}_t\boldsymbol{L}_t'. \quad (10.12)$$

This is the well-known Cholesky decomposition of $\boldsymbol{\Sigma}_t$. A feature of the decomposition is that the lower off-diagonal elements of $\boldsymbol{L}_t$ and the diagonal elements of $\boldsymbol{G}_t$

have nice interpretations. We demonstrate the decomposition by studying carefully the bivariate and three-dimensional cases. For the bivariate case, we have

$$\boldsymbol{\Sigma}_t = \begin{bmatrix} \sigma_{11,t} & \sigma_{21,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix}, \qquad \boldsymbol{L}_t = \begin{bmatrix} 1 & 0 \\ q_{21,t} & 1 \end{bmatrix}, \qquad \boldsymbol{G}_t = \begin{bmatrix} g_{11,t} & 0 \\ 0 & g_{22,t} \end{bmatrix},$$

where $g_{ii,t} > 0$ for $i = 1$ and 2. Using Eq. (10.12), we have

$$\boldsymbol{\Sigma}_t = \begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{12,t} & \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} g_{11,t} & q_{21,t}g_{11,t} \\ q_{21,t}g_{11,t} & g_{22,t} + q_{21,t}^2 g_{11,t} \end{bmatrix}.$$

Equating elements of the prior matrix equation, we obtain

$$\sigma_{11,t} = g_{11,t}, \qquad \sigma_{21,t} = q_{21,t}g_{11,t}, \qquad \sigma_{22,t} = g_{22,t} + q_{21,t}^2 g_{11,t}. \tag{10.13}$$

Solving the prior equations, we have

$$g_{11,t} = \sigma_{11,t}, \qquad q_{21,t} = \frac{\sigma_{21,t}}{\sigma_{11,t}}, \qquad g_{22,t} = \sigma_{22,t} - \frac{\sigma_{21,t}^2}{\sigma_{11,t}}. \tag{10.14}$$

However, consider the simple linear regression

$$a_{2t} = \beta a_{1t} + b_{2t}, \tag{10.15}$$

where b_{2t} denotes the error term. From the well-known least-squares theory, we have

$$\beta = \frac{\text{Cov}(a_{1t}, a_{2t})}{\text{Var}(a_{1t})} = \frac{\sigma_{21,t}}{\sigma_{11,t}},$$

$$\text{Var}(b_{2t}) = \text{Var}(a_{2t}) - \beta^2\text{Var}(a_{1t}) = \sigma_{22,t} - \frac{\sigma_{21,t}^2}{\sigma_{11,t}}.$$

Furthermore, the error term b_{2t} is uncorrelated with the regressor a_{1t}. Consequently, using Eq. (10.14), we obtain

$$g_{11,t} = \sigma_{11,t}, \qquad q_{21,t} = \beta, \qquad g_{22,t} = \text{Var}(b_{2t}), \qquad b_{2t} \perp a_{1t},$$

where $\perp$ denotes no correlation. In summary, the Cholesky decomposition of the 2×2 matrix $\boldsymbol{\Sigma}_t$ amounts to performing an orthogonal transformation from $\boldsymbol{a}_t$ to $\boldsymbol{b}_t = (b_{1t}, b_{2t})'$ such that

$$b_{1t} = a_{1t} \qquad \text{and} \qquad b_{2t} = a_{2t} - q_{21,t}a_{1t},$$

where $q_{21,t} = \beta$ is obtained by the linear regression (10.15) and $\text{Cov}(\boldsymbol{b}_t)$ is a diagonal matrix with diagonal elements $g_{ii,t}$. The transformed quantities $q_{21,t}$ and $g_{ii,t}$ can be interpreted as follows:

1. The first diagonal element of $\boldsymbol{G}_t$ is simply the variance of a_{1t}.
2. The second diagonal element of $\boldsymbol{G}_t$ is the residual variance of the simple linear regression in Eq. (10.15).
3. The element $q_{21,t}$ of the lower triangular matrix $\boldsymbol{L}_t$ is the coefficient β of the regression in Eq. (10.15).

The prior properties continue to hold for the higher dimensional case. For example, consider the three-dimensional case in which

$$\boldsymbol{L}_t = \begin{bmatrix} 1 & 0 & 0 \\ q_{21,t} & 1 & 0 \\ q_{31,t} & q_{32,t} & 1 \end{bmatrix}, \qquad \boldsymbol{G}_t = \begin{bmatrix} g_{11,t} & 0 & 0 \\ 0 & g_{22,t} & 0 \\ 0 & 0 & g_{3,t} \end{bmatrix}.$$

From the decomposition in Eq. (10.12), we have

$$\begin{bmatrix} \sigma_{11,t} & \sigma_{21,t} & \sigma_{31,t} \\ \sigma_{21,t} & \sigma_{22,t} & \sigma_{32,t} \\ \sigma_{31,t} & \sigma_{32,t} & \sigma_{33,t} \end{bmatrix}$$
$$= \begin{bmatrix} g_{11,t} & q_{21,t}g_{11,t} & q_{31,t}g_{11,t} \\ q_{21,t}g_{11,t} & q_{21,t}^2 g_{11,t} + g_{22,t} & q_{31,t}q_{21,t}g_{11,t} + q_{32,t}g_{22,t} \\ q_{31,t}g_{11,t} & q_{31,t}q_{21,t}g_{11,t} + q_{32,t}g_{22,t} & q_{31,t}^2 g_{11,t} + q_{32,t}^2 g_{22,t} + g_{33,t} \end{bmatrix}.$$

Equating elements of the prior matrix equation, we obtain

$$\begin{aligned} \sigma_{11,t} &= g_{11,t}, & \sigma_{21,t} &= q_{21,t}g_{11,t}, \\ \sigma_{22,t} &= q_{21,t}^2 g_{11,t} + g_{22,t}, & \sigma_{31,t} &= q_{31,t}g_{11,t}, \\ \sigma_{32,t} &= q_{31,t}q_{21,t}g_{11,t} + q_{32,t}g_{22,t}, & \sigma_{33,t} &= q_{31,t}^2 g_{11,t} + q_{32,t}^2 g_{22,t} + g_{33,t} \end{aligned}$$

or, equivalently,

$$g_{11,t} = \sigma_{11,t}, \qquad q_{21,t} = \frac{\sigma_{21,t}}{\sigma_{11,t}}, \qquad g_{22,t} = \sigma_{22,t} - q_{21,t}^2 g_{11,t},$$
$$q_{31,t} = \frac{\sigma_{31,t}}{\sigma_{11,t}}, \qquad q_{32,t} = \frac{1}{g_{22,t}}\left(\sigma_{32,t} - \frac{\sigma_{31,t}}{\sigma_{\ 21,t}}\sigma_{11,t}\right),$$
$$g_{33,t} = \sigma_{33,t} - q_{31,t}^2 g_{11,t} - q_{32,t}^2 g_{22,t}.$$

These quantities look complicated, but they are simply the coefficients and residual variances of the orthogonal transformation

$$\begin{aligned} b_{1t} &= a_{1t}, \\ b_{2t} &= a_{2t} - \beta_{21}b_{1t}, \\ b_{3t} &= a_{3t} - \beta_{31}b_{1t} - \beta_{32}b_{2t}, \end{aligned}$$

where β_{ij} are the coefficients of least-squares regressions

$$a_{2t} = \beta_{21}b_{1t} + b_{2t},$$
$$a_{3t} = \beta_{31}b_{1t} + \beta_{32}b_{2t} + b_{3t}.$$

In other words, we have $q_{ij,t} = \beta_{ij}$, $g_{ii,t} = Var(b_{it})$ and $b_{it} \perp b_{jt}$ for $i \neq j$.

Based on the prior discussion, using Cholesky decomposition amounts to doing an orthogonal transformation from $\boldsymbol{a}_t$ to $\boldsymbol{b}_t$, where $b_{1t} = a_{1t}$, and b_{it}, for $1 < i \leq k$, is defined recursively by the least-squares regression

$$a_{it} = q_{i1,t}b_{1t} + q_{i2,t}b_{2t} + \cdots + q_{i(i-1),t}b_{(i-1)t} + b_{it}, \tag{10.16}$$

where $q_{ij,t}$ is the (i, j)th element of the lower triangular matrix $\boldsymbol{L}_t$ for $1 \leq j < i$. We can write this transformation as

$$\boldsymbol{b}_t = \boldsymbol{L}_t^{-1}\boldsymbol{a}_t, \qquad \text{or} \qquad \boldsymbol{a}_t = \boldsymbol{L}_t\boldsymbol{b}_t, \tag{10.17}$$

where, as mentioned before, $\boldsymbol{L}_t^{-1}$ is also a lower triangular matrix with unit diagonal elements. The covariance matrix of $\boldsymbol{b}_t$ is the diagonal matrix $\boldsymbol{G}_t$ of the Cholesky decomposition because

$$\mathrm{Cov}(\boldsymbol{b}_t) = \boldsymbol{L}_t^{-1}\boldsymbol{\Sigma}_t(\boldsymbol{L}_t^{-1})' = \boldsymbol{G}_t.$$

The parameter vector relevant to volatility modeling under such a transformation becomes

$$\boldsymbol{\Xi}_t = (g_{11,t}, \ldots, g_{kk,t}, q_{21,t}, q_{31,t}, q_{32,t}, \ldots, q_{k1,t}, \ldots, q_{k(k-1),t})', \tag{10.18}$$

which is also a $k(k+1)/2$-dimensional vector.

The previous orthogonal transformation also dramatically simplifies the likelihood function of the data. Using the fact that $|\boldsymbol{L}_t| = 1$, we have

$$|\boldsymbol{\Sigma}_t| = |\boldsymbol{L}_t\boldsymbol{G}_t\boldsymbol{L}_t'| = |\boldsymbol{G}_t| = \prod_{i=1}^{k} g_{ii,t}. \tag{10.19}$$

If the conditional distribution of $\boldsymbol{a}_t$ given the past information is multivariate normal $N(\boldsymbol{0}, \boldsymbol{\Sigma}_t)$, then the conditional distribution of the transformed series $\boldsymbol{b}_t$ is multivariate normal $N(\boldsymbol{0}, \boldsymbol{G}_t)$, and the log-likelihood function of the data becomes extremely simple. Indeed, we have the log probability density of $\boldsymbol{a}_t$ as

$$\ell(\boldsymbol{a}_t, \boldsymbol{\Sigma}_t) = \ell(\boldsymbol{b}_t, \boldsymbol{\Xi}_t) = -\frac{1}{2}\sum_{i=1}^{k}\left[\ln(g_{ii,t}) + \frac{b_{it}^2}{g_{ii,t}}\right], \tag{10.20}$$

where for simplicity the constant term is omitted and $g_{ii,t}$ is the variance of b_{it}.

Using the Cholesky decomposition to reparameterize $\mathbf{\Sigma}_t$ has several advantages. First, from Eq. (10.19), $\mathbf{\Sigma}_t$ is positive definite if $g_{ii,t} > 0$ for all i. Consequently, the positive-definite constraint of $\mathbf{\Sigma}_t$ can easily be achieved by modeling $\ln(g_{ii,t})$ instead of $g_{ii,t}$. Second, elements of the parameter vector $\mathbf{\Xi}_t$ in Eq. (10.18) have nice interpretations. They are the coefficients and residual variances of multiple linear regressions that orthogonalize the shocks to the returns. Third, the correlation coefficient between a_{1t} and a_{2t} is

$$\rho_{21,t} = \frac{\sigma_{21,t}}{\sqrt{\sigma_{11,t}\sigma_{22,t}}} = q_{21,t} \times \frac{\sqrt{\sigma_{11,t}}}{\sqrt{\sigma_{22,t}}},$$

which is time varying if $q_{21,t} \neq 0$. In particular, if $q_{21,t} = c \neq 0$, then $\rho_{21,t} = c\sqrt{\sigma_{11,t}}/\sqrt{\sigma_{22,t}}$, which continues to be time-varying provided that the variance ratio $\sigma_{11,t}/\sigma_{22,t}$ is not a constant. This time-varying property applies to other correlation coefficients when the dimension of $\boldsymbol{r}_t$ is greater than 2 and is a major difference between the two approaches for reparameterizing $\mathbf{\Sigma}_t$.

Using Eq. (10.16) and the orthogonality among the transformed shocks b_{it}, we obtain

$$\sigma_{ii,t} = \mathrm{Var}(a_{it}|F_{t-1}) = \sum_{v=1}^{i} q_{iv,t}^2 g_{vv,t}, \qquad i = 1, \ldots, k,$$

$$\sigma_{ij,t} = \mathrm{Cov}(a_{it}, a_{jt}|F_{t-1}) = \sum_{v=1}^{j} q_{iv,t} q_{jv,t} g_{vv,t}, \qquad j < i, \qquad i = 2, \ldots, k,$$

where $q_{vv,t} = 1$ for $v = 1, \ldots, k$. These equations show the parameterization of $\mathbf{\Sigma}_t$ under the Cholesky decomposition.

10.4 GARCH MODELS FOR BIVARIATE RETURNS

Since the same techniques can be used to generalize many univariate volatility models to the multivariate case, we focus our discussion on the multivariate GARCH model. Other multivariate volatility models can also be used.

For a k-dimensional return series $\boldsymbol{r}_t$, a multivariate GARCH model uses "exact equations" to describe the evolution of the $k(k+1)/2$-dimensional vector $\mathbf{\Xi}_t$ over time. By exact equation, we mean that the equation does not contain any stochastic shock. However, the exact equation may become complicated even in the simplest case of $k = 2$ for which $\mathbf{\Xi}_t$ is three dimensional. To keep the model simple, some restrictions are often imposed on the equations.

10.4.1 Constant-Correlation Models

To keep the number of volatility equations low, Bollerslev (1990) considers the special case in which the correlation coefficient $\rho_{21,t} = \rho_{21}$ is time invariant, where

$|\rho_{21}| < 1$. Under such an assumption, ρ_{21} is a constant parameter and the volatility model consists of two equations for $\boldsymbol{\Xi}_t^*$, which is defined as $\boldsymbol{\Xi}_t^* = (\sigma_{11,t}, \sigma_{22,t})'$. A GARCH(1,1) model for $\boldsymbol{\Xi}_t^*$ becomes

$$\boldsymbol{\Xi}_t^* = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 \boldsymbol{a}_{t-1}^2 + \boldsymbol{\beta}_1 \boldsymbol{\Xi}_{t-1}^*, \tag{10.21}$$

where $\boldsymbol{a}_{t-1}^2 = (a_{1,t-1}^2, a_{2,t-1}^2)'$, $\boldsymbol{\alpha}_0$ is a two-dimensional positive vector, and $\boldsymbol{\alpha}_1$ and $\boldsymbol{\beta}_1$ are 2×2 nonnegative definite matrices. More specifically, the model can be expressed in detail as

$$\begin{bmatrix} \sigma_{11,t} \\ \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} \alpha_{10} \\ \alpha_{20} \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} a_{1,t-1}^2 \\ a_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} \sigma_{11,t-1} \\ \sigma_{22,t-1} \end{bmatrix}, \tag{10.22}$$

where $\alpha_{i0} > 0$ for $i = 1$ and 2. Defining $\boldsymbol{\eta}_t = \boldsymbol{a}_t^2 - \boldsymbol{\Xi}_t^*$, we can rewrite the prior model as

$$\boldsymbol{a}_t^2 = \boldsymbol{\alpha}_0 + (\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1)\boldsymbol{a}_{t-1}^2 + \boldsymbol{\eta}_t - \boldsymbol{\beta}_1 \boldsymbol{\eta}_{t-1},$$

which is a bivariate ARMA(1,1) model for the $\boldsymbol{a}_t^2$ process. This result is a direct generalization of the univariate GARCH(1,1) model of Chapter 3. Consequently, some properties of model (10.22) are readily available from those of the bivariate ARMA(1,1) model of Chapter 8. In particular, we have the following results:

1. If all of the eigenvalues of $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1$ are positive, but less than 1, then the bivariate ARMA(1,1) model for $\boldsymbol{a}_t^2$ is weakly stationary and, hence, $E(\boldsymbol{a}_t^2)$ exists. This implies that the shock process $\boldsymbol{a}_t$ of the returns has a positive-definite unconditional covariance matrix. The unconditional variances of the elements of $\boldsymbol{a}_t$ are $(\sigma_1^2, \sigma_2^2)' = (\boldsymbol{I} - \boldsymbol{\alpha}_1 - \boldsymbol{\beta}_1)^{-1}\boldsymbol{\phi}_0$, and the unconditional covariance between a_{1t} and a_{2t} is $\rho_{21}\sigma_1\sigma_2$.
2. If $\alpha_{12} = \beta_{12} = 0$, then the volatility of a_{1t} does not depend on the past volatility of a_{2t}. Similarly, if $\alpha_{21} = \beta_{21} = 0$, then the volatility of a_{2t} does not depend on the past volatility of a_{1t}.
3. If both $\boldsymbol{\alpha}_1$ and $\boldsymbol{\beta}_1$ are diagonal, then the model reduces to two univariate GARCH(1,1) models. In this case, the two volatility processes are not dynamically related.
4. Volatility forecasts of the model can be obtained by using forecasting methods similar to those of a vector ARMA(1,1) model; see the univariate case in Chapter 3. The 1-step-ahead volatility forecast at the forecast origin h is

   $$\boldsymbol{\Xi}_h^*(1) = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 \boldsymbol{a}_h^2 + \boldsymbol{\beta}_1 \boldsymbol{\Xi}_h^*.$$

   For the ℓ-step-ahead forecast, we have

   $$\boldsymbol{\Xi}_h^*(\ell) = \boldsymbol{\alpha}_0 + (\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1)\boldsymbol{\Xi}_h^*(\ell - 1), \qquad \ell > 1.$$

These forecasts are for the marginal volatilities of a_{it}. The ℓ-step-ahead forecast of the covariance between a_{1t} and a_{2t} is $\hat{\rho}_{21}[\sigma_{11,h}(\ell)\sigma_{22,h}(\ell)]^{0.5}$, where $\hat{\rho}_{21}$ is the estimate of ρ_{21} and $\sigma_{ii,h}(\ell)$ are the elements of $\Xi_h^*(\ell)$.

Example 10.4. Again, consider the daily log returns of Hong Kong and Japanese markets of Example 10.1. Using a bivariate GARCH model, we obtain a constant correlation model that fits the data reasonably well. The mean equations of the bivariate model are

$$r_{1t} = 0.101 + a_{1t},$$
$$r_{2t} = 0.002 + a_{2t},$$

where the standard errors of the two estimates are 0.050 and 0.048, respectively. The volatility equations are

$$\begin{bmatrix} \sigma_{11,t} \\ \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} \begin{matrix} 0.079 \\ (0.019) \end{matrix} \\ \begin{matrix} 0.054 \\ (0.019) \end{matrix} \end{bmatrix} + \begin{bmatrix} \begin{matrix} 0.145 \\ (0.022) \end{matrix} & \cdot \\ & \begin{matrix} 0.105 \\ (0.014) \end{matrix} \end{bmatrix} \begin{bmatrix} a_{1,t-1}^2 \\ a_{2,t-1}^2 \end{bmatrix}$$

$$+ \begin{bmatrix} \begin{matrix} 0.833 \\ (0.023) \end{matrix} & \\ & \begin{matrix} 0.875 \\ (0.020) \end{matrix} \end{bmatrix} \begin{bmatrix} \sigma_{11,t-1} \\ \sigma_{22,t-1} \end{bmatrix}, \qquad (10.23)$$

where the numbers in parentheses are standard errors. The estimated constant correlation between the two returns is 0.668.

Let $\tilde{\boldsymbol{a}}_t = (\tilde{a}_{1t}, \tilde{a}_{2t})'$ be the standardized residuals, where $\tilde{a}_{it} = a_{it}/\sqrt{\sigma_{ii,t}}$. The Ljung–Box statistics of $\tilde{\boldsymbol{a}}_t$ give $Q_2(4) = 17.29(0.37)$ and $Q_2(12) = 48.21(0.46)$, where the number in parentheses denotes the p value. Here the p values are based on chi-squared distributions with 16 and 48 degrees of freedom, respectively. The Q statistics of individual series $\tilde{a}_{it}$ shown in S-Plus output also fail to indicate any model inadequancy. Consequently, the constant correlation model in Eq. (10.23) fits the data reasonably well. Figure 10.7 shows the fitted volatility processes of model (10.23), which can be compared with those of Example 10.1.

The model in Eq. (10.23) shows two uncoupled volatility equations, indicating that the volatilities of the two markets are not dynamically related, but they are contemporaneously correlated. We refer to the model as a bivariate *diagonal constant-correlation* model. In practice, this type of models might not be suitable because there exists the possibility of dynamic dependence in volatility among markets, that is, the spillover effect in volatility. Finally, the constant-correlation model can easily be estimated using S-Plus:

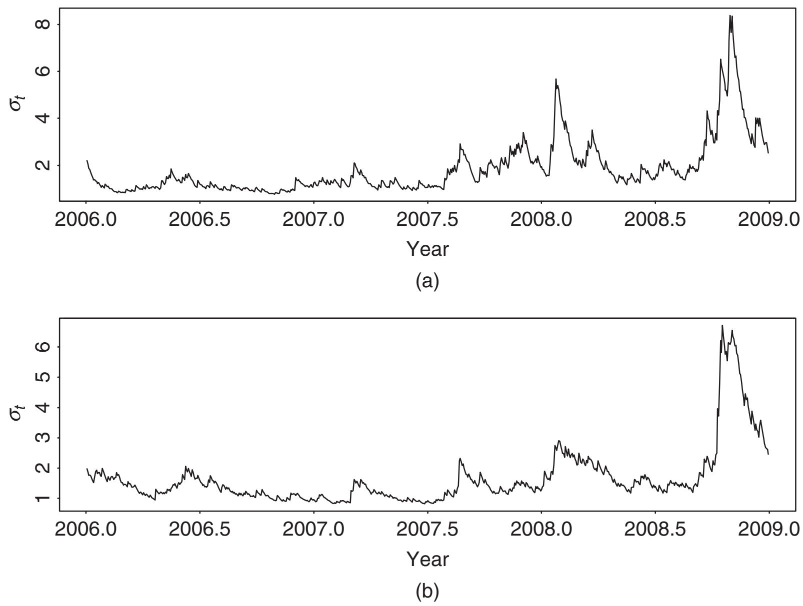


Figure 10.7 Estimated volatilities for daily log returns in percentages of stock market indexes for Hong Kong and Japan from January 4, 2006, to December 30, 2008: (a) Hong Kong market and (b) Japanese market. Model used is Eq. (10.23).

```
> mccc = mgarch(rtn~1,~ccc(1,1),trace=F)
> summary(mccc)
```

Example 10.5. As a second illustration, consider the monthly log returns, in percentages, of IBM stock and the S&P 500 index from January 1926 to December 1999 used in Chapter 8. Let r_{1t} and r_{2t} be the monthly log returns for IBM stock and the S&P 500 index, respectively. If a constant-correlation GARCH(1,1) model is entertained, we obtain the mean equations

$$\begin{aligned} r_{1t} &= 1.351 + 0.072 r_{1,t-1} + 0.055 r_{1,t-2} - 0.119 r_{2,t-2} + a_{1t}, \\ r_{2t} &= 0.703 + a_{2t}, \end{aligned}$$

where standard errors of the parameters in the first equation are 0.225, 0.029, 0.034, and 0.044, respectively, and the standard error of the parameter in the second equation is 0.155. The volatility equations are

$$\begin{aligned} \begin{bmatrix} \sigma_{11,t} \\ \sigma_{22,t} \end{bmatrix} &= \begin{bmatrix} 2.98 \\ (0.59) \\ 2.09 \\ (0.47) \end{bmatrix} + \begin{bmatrix} 0.079 & \\ (0.013) & \\ 0.042 & 0.045 \\ (0.009) & (0.010) \end{bmatrix} \begin{bmatrix} a^2_{1,t-1} \\ a^2_{2,t-1} \end{bmatrix} \\ &\quad + \begin{bmatrix} 0.873 & -0.031 \\ (0.020) & (0.009) \\ -0.066 & 0.913 \\ (0.015) & (0.014) \end{bmatrix} \begin{bmatrix} \sigma_{11,t-1} \\ \sigma_{22,t-1} \end{bmatrix}, \end{aligned} \tag{10.24}$$

where the numbers in parentheses are standard errors. The constant correlation coefficient is 0.614 with standard error 0.020. Using the standardized residuals, we obtain the Ljung–Box statistics $Q_2(4) = 16.77(0.21)$ and $Q_2(8) = 32.40(0.30)$, where the p values shown in parentheses are obtained from chi-squared distributions with 13 and 29 degrees of freedom, respectively. Here the degrees of freedom have been adjusted because the mean equations contain three lagged predictors. For the squared standardized residuals, we have $Q_2^*(4) = 18.00(0.16)$ and $Q_2^*(8) = 39.09(0.10)$. Therefore, at the 5% significance level, the standardized residuals $\tilde{\boldsymbol{a}}_t$ have no serial correlations or conditional heteroscedasticities. This bivariate GARCH(1,1) model shows a feedback relationship between the volatilities of the two monthly log returns.

10.4.2 Time-Varying Correlation Models

A major drawback of the constant-correlation volatility models is that the correlation coefficient tends to change over time in a real application. Consider the monthly log returns of IBM stock and the S&P 500 index used in Example 10.5. It is hard to justify that the S&P 500 index return, which is a weighted average, can maintain a constant-correlation coefficient with IBM return over the past 70 years. Figure 10.8 shows the sample correlation coefficient between the two monthly log return series using a moving window of 120 observations (i.e., 10 years). The correlation changes over time and appears to be decreasing in recent years. The decreasing trend in correlation is not surprising because the ranking of IBM market capitalization among large U.S. industrial companies has changed in recent years. A Lagrange multiplier statistic was proposed recently by Tse (2000) to test constant-correlation coefficients in a multivariate GARCH model.

A simple way to relax the constant-correlation constraint within the GARCH framework is to specify an exact equation for the conditional correlation coefficient. This can be done by two methods using the two reparameterizations of $\boldsymbol{\Sigma}_t$ discussed in Section 10.3. First, we use the correlation coefficient directly. Because the correlation coefficient between the returns of IBM stock and S&P 500 index is positive and must be in the interval [0, 1], we employ the equation

$$\rho_{21,t} = \frac{\exp(q_t)}{1 + \exp(q_t)}, \tag{10.25}$$

where

$$q_t = \varpi_0 + \varpi_1 \rho_{21,t-1} + \varpi_2 \frac{a_{1,t-1}a_{2,t-1}}{\sqrt{\sigma_{11,t-1}\sigma_{22,t-1}}},$$

where $\sigma_{ii,t-1}$ is the conditional variance of the shock $a_{i,t-1}$. We refer to this equation as a GARCH(1,1) model for the correlation coefficient because it uses the lag-1 cross correlation and the lag-1 cross product of the two shocks. If $\varpi_1 = \varpi_2 = 0$, then model (10.25) reduces to the case of constant correlation.

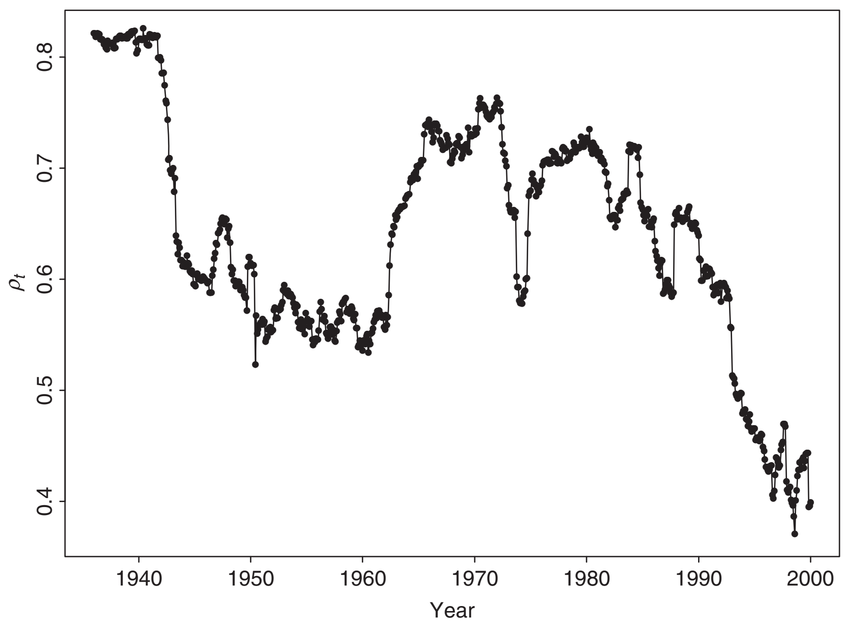


Figure 10.8 Sample correlation coefficient between monthly log returns of IBM stock and S&P 500 index. Correlation is computed by a moving window of 120 observations. Sample period is from January 1926 to December 1999.

In summary, a time-varying correlation bivariate GARCH(1,1) model consists of two sets of equations. The first set of equations consists of a bivariate GARCH(1,1) model for the conditional variances, and the second set of equation is a GARCH(1,1) model for the correlation in Eq. (10.25). In practice, a negative sign can be added to Eq. (10.25) if the correlation coefficient is negative. In general, when the sign of correlation is unknown, we can use the Fisher transformation for correlation

$$q_t = \ln\left(\frac{1+\rho_{21,t}}{1-\rho_{21,t}}\right) \qquad \text{or} \qquad \rho_{21,t} = \frac{\exp(q_t)-1}{\exp(q_t)+1}$$

and employ a GARCH model for q_t to model the time-varying correlation between two returns.

Example 10.5 (Continued). Augmenting Eq. (10.25) to the GARCH(1,1) model in Eq. (10.24) for the monthly log returns of IBM stock and the S&P 500 index and performing a joint estimation, we obtain the following model for the two series:

$$\begin{aligned} r_{1t} &= 1.318 + 0.076 r_{1,t-1} - 0.068 r_{2,t-2} + a_{1t}, \\ r_{2t} &= 0.673 + a_{2t}, \end{aligned}$$

where standard errors of the three parameters in the first equation are 0.215, 0.026, and 0.034, respectively, and standard error of the parameter in the second equation is 0.151. The volatility equations are

$$\begin{bmatrix} \sigma_{11,t} \\ \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} 2.80 \\ (0.58) \\ 1.71 \\ (0.40) \end{bmatrix} + \begin{bmatrix} 0.084 & \\ (0.013) & \\ 0.037 & 0.054 \\ (0.009) & (0.010) \end{bmatrix} \begin{bmatrix} a^2_{1,t-1} \\ a^2_{2,t-1} \end{bmatrix}$$

$$+ \begin{bmatrix} 0.864 & -0.020 \\ (0.021) & (0.009) \\ -0.058 & 0.914 \\ (0.014) & (0.013) \end{bmatrix} \begin{bmatrix} \sigma_{11,t-1} \\ \sigma_{22,t-1} \end{bmatrix}, \qquad (10.26)$$

where, as before, standard errors are in parentheses. The conditional correlation equation is

$$\rho_t = \frac{\exp(q_t)}{1+\exp(q_t)}, \qquad q_t = -2.024 + 3.983\rho_{t-1} + 0.088\frac{a_{1,t-1}a_{2,t-1}}{\sqrt{\sigma_{11,t-1}\sigma_{22,t-1}}}, \qquad (10.27)$$

where standard errors of the estimates are 0.050, 0.090, and 0.019, respectively. The parameters of the prior correlation equation are highly significant. Applying the Ljung–Box statistics to the standardized residuals $\tilde{\boldsymbol{a}}_t$, we have $Q_2(4) = 20.57(0.11)$ and $Q_2(8) = 36.08(0.21)$. For the squared standardized residuals, we have $Q_2^*(4) = 16.69(0.27)$ and $Q_2^*(8) = 36.71(0.19)$. Therefore, the standardized residuals of the model have no significant serial correlations or conditional heteroscedasticities.

It is interesting to compare this time-varying correlation GARCH(1,1) model with the constant-correlation GARCH(1,1) model in Eq. (10.24). First, the mean and volatility equations of the two models are close. Second, Figure 10.9 shows the fitted conditional correlation coefficient between the monthly log returns of IBM stock and the S&P 500 index based on model (10.27). The plot shows that the correlation coefficient fluctuated over time and became smaller in recent years. This latter characteristic is in agreement with that of Figure 10.8. Third, the average of the fitted correlation coefficients is 0.612, which is essentially the estimate 0.614 of the constant-correlation model in Eq. (10.24). Fourth, using the sample variances of r_{it} as the starting values for the conditional variances and the observations from $t = 4$ to $t = 888$, the maximized log-likelihood function is -3691.21 for the constant-correlation GARCH(1,1) model and -3679.64 for the time-varying correlation GARCH(1,1) model. Thus, the time-varying correlation model shows some significant improvement over the constant-correlation model. Finally, consider the 1-step-ahead volatility forecasts of the two models at the forecast origin $h = 888$. For the constant-correlation model in Eq. (10.24), we have $a_{1,888} = 3.075$,

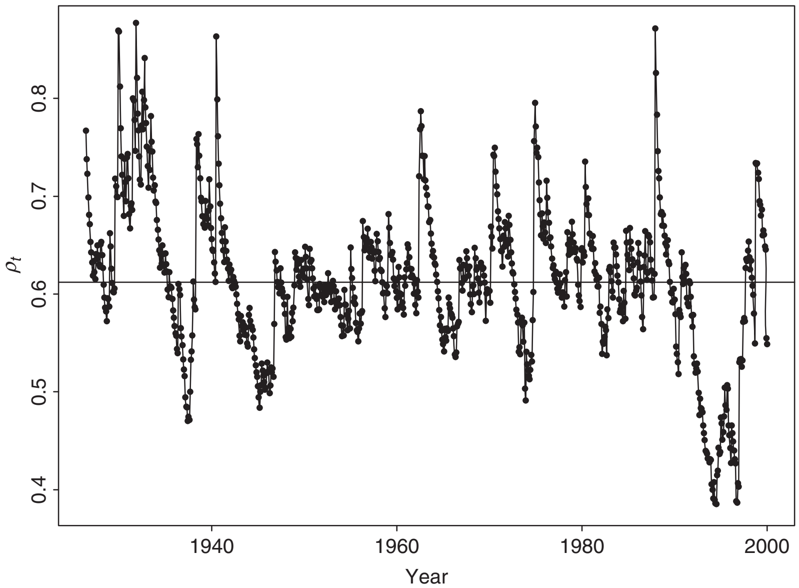


Figure 10.9 Fitted conditional correlation coefficient between monthly log returns of IBM stock and S&P 500 index using time-varying correlation GARCH(1,1) model of Example 10.5. Horizontal line denotes average of 0.612 of correlation coefficients.

$a_{2,888} = 4.931$, $\sigma_{11,888} = 77.91$, and $\sigma_{22,888} = 21.19$. Therefore, the 1-step-ahead forecast for the conditional covariance matrix is

$$\widehat{\boldsymbol{\Sigma}}_{888}(1) = \begin{bmatrix} 71.09 & 21.83 \\ 21.83 & 17.79 \end{bmatrix},$$

where the covariance is obtained by using the constant-correlation coefficient 0.614. For the time-varying correlation model in Eqs. (10.26) and (10.27), we have $a_{1,888} = 3.287$, $a_{2,888} = 4.950$, $\sigma_{11,888} = 83.35$, $\sigma_{22,888} = 28.56$, and $\rho_{888} = 0.546$. The 1-step-ahead forecast for the covariance matrix is

$$\widehat{\boldsymbol{\Sigma}}_{888}(1) = \begin{bmatrix} 75.15 & 23.48 \\ 23.48 & 24.70 \end{bmatrix},$$

where the forecast of the correlation coefficient is 0.545.

In the second method, we use the Cholesky decomposition of $\boldsymbol{\Sigma}_t$ to model time-varying correlations. For the bivariate case, the parameter vector is $\boldsymbol{\Xi}_t = (g_{11,t}, g_{22,t}, q_{21,t})'$; see Eq. (10.18). A simple GARCH(1,1) type model for $\boldsymbol{a}_t$ is

$$\begin{aligned} g_{11,t} &= \alpha_{10} + \alpha_{11} b_{1,t-1}^2 + \beta_{11} g_{11,t-1}, \\ q_{21,t} &= \gamma_0 + \gamma_1 q_{21,t-1} + \gamma_2 a_{2,t-1}, \\ g_{22,t} &= \alpha_{20} + \alpha_{21} b_{1,t-1}^2 + \alpha_{22} b_{2,t-1}^2 + \beta_{21} g_{11,t-1} + \beta_{22} g_{22,t-1}, \end{aligned} \tag{10.28}$$

where $b_{1t} = a_{1t}$ and $b_{2t} = a_{2t} - q_{21,t}a_{1t}$. Thus, b_{1t} assumes a univariate GARCH(1,1) model, b_{2t} uses a bivariate GARCH(1,1) model, and $q_{21,t}$ is autocorrelated and uses $a_{2,t-1}$ as an additional explanatory variable. The probability density function relevant to maximum-likelihood estimation is given in Eq. (10.20) with $k = 2$.

Example 10.5 (Continued). Again we use the monthly log returns of IBM stock and the S&P 500 index to demonstrate the volatility model in Eq. (10.28). Using the same specification as before, we obtain the fitted mean equations as

$$\begin{aligned} r_{1t} &= 1.364 + 0.075r_{1,t-1} - 0.058r_{2,t-2} + a_{1t}, \\ r_{2t} &= 0.643 + a_{2t}, \end{aligned}$$

where standard errors of the parameters in the first equation are 0.219, 0.027, and 0.032, respectively, and the standard error of the parameter in the second equation is 0.154. These two mean equations are close to what we obtained before. The fitted volatility model is

$$\begin{aligned} g_{11,t} &= 3.714 + 0.113b_{1,t-1}^2 + 0.804g_{11,t-1}, \\ q_{21,t} &= 0.0029 + 0.9915q_{21,t-1} - 0.0041a_{2,t-1}, \\ g_{22,t} &= 1.023 + 0.021b_{1,t-1}^2 + 0.052b_{2,t-1}^2 - 0.040g_{11,t-1} + 0.937g_{22,t-1}, \end{aligned} \tag{10.29}$$

where $b_{1t} = a_{1t}$, and $b_{2t} = a_{2t} - q_{21,t}b_{1t}$. Standard errors of the parameters in the equation of $g_{11,t}$ are 1.033, 0.022, and 0.037, respectively; those of the parameters in the equation of $q_{21,t}$ are 0.001, 0.002, and 0.0004; and those of the parameters in the equation of $g_{22,t}$ are 0.344, 0.007, 0.013, and 0.015, respectively. All estimates are statistically significant at the 1% level.

The conditional covariance matrix $\boldsymbol{\Sigma}_t$ can be obtained from model (10.29) by using the Cholesky decomposition in Eq. (10.12). For the bivariate case, the relationship is given specifically in Eq. (10.13). Consequently, we obtain the time-varying correlation coefficient as

$$\rho_t = \frac{\sigma_{21,t}}{\sqrt{\sigma_{11,t}\sigma_{22,t}}} = \frac{q_{21,t}\sqrt{g_{11,t}}}{\sqrt{g_{22,t} + q_{21,t}^2 g_{11,t}}}. \tag{10.30}$$

Using the fitted values of $\sigma_{11,t}$ and $\sigma_{22,t}$, we can compute the standardized residuals to perform model checking. The Ljung–Box statistics for the standardized residuals of model (10.29) give $Q_2(4) = 19.77(0.14)$ and $Q_2(8) = 34.22(0.27)$. For the squared standardized residuals, we have $Q_2^*(4) = 15.34(0.36)$ and $Q_2^*(8) = 31.87(0.37)$. Thus, the fitted model is adequate in describing the conditional mean and volatility. The model shows a strong dynamic dependence in the correlation; see the coefficient 0.9915 in Eq. (10.29).

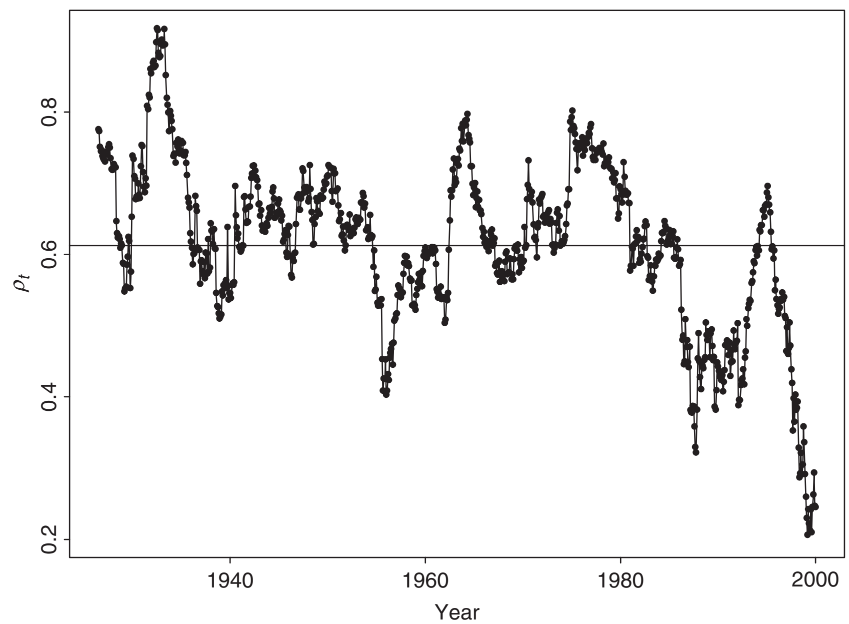


Figure 10.10 Fitted conditional correlation coefficient between monthly log returns of IBM stock and S&P 500 index using time-varying correlation GARCH(1,1) model of Example 10.5 with Cholesky decomposition. Horizontal line denotes average of 0.612 of the estimated coefficients.

Figure 10.10 shows the fitted time-varying correlation coefficient in Eq. (10.30). It shows a smoother correlation pattern than that of Figure 10.9 and confirms the decreasing trend of the correlation coefficient. In particular, the fitted correlation coefficients in recent years are smaller than those of the other models. The two time-varying correlation models for the monthly log returns of IBM stock and the S&P 500 index have comparable maximized-likelihood functions of about -3672, indicating the fits are similar. However, the approach based on the Cholesky decomposition may have some advantages. First, it does not require any parameter constraint in estimation to ensure the positive definiteness of $\boldsymbol{\Sigma}_t$. If one also uses log transformation for $g_{ii,t}$, then no constraints are needed for the entire volatility model. Second, the log-likelihood function becomes simple under the transformation. Third, the time-varying parameters $q_{ij,t}$ and $g_{ii,t}$ have nice interpretations. However, the transformation makes inference a bit more complicated because the fitted model may depend on the ordering of elements in $\boldsymbol{a}_t$; recall that a_{1t} is not transformed. In theory, the ordering of elements in $\boldsymbol{a}_t$ should have no impact on volatility.

Finally, the 1-step-ahead forecast of the conditional covariance matrix at the forecast origin $t = 888$ for the new time-varying correlation model is

$$\widehat{\boldsymbol{\Sigma}}_{888}(1) = \begin{bmatrix} 73.45 & 7.34 \\ 7.34 & 17.87 \end{bmatrix}.$$

The correlation coefficient of the prior forecast is 0.203, which is substantially smaller than those of the previous two models. However, forecasts of the conditional variances are similar as before.

10.4.3 Dynamic Correlation Models

Using the parameterization in Eq. (10.7), several authors have proposed parsimonious models for $\boldsymbol{\rho}_t$ to describe the time-varying correlations. We refer to those models as the dynamic conditional correlation (DCC) models.

For k-dimensional returns, Tse and Tsui (2002) assume that the conditional correlation matrix $\boldsymbol{\rho}_t$ follows the model

$$\boldsymbol{\rho}_t = (1 - \theta_1 - \theta_2)\boldsymbol{\rho} + \theta_1 \boldsymbol{\rho}_{t-1} + \theta_2 \boldsymbol{\psi}_{t-1},$$

where θ_1 and θ_2 are scalar parameters, $\boldsymbol{\rho}$ is a $k \times k$ positive-definite matrix with unit diagonal elements, and $\boldsymbol{\psi}_{t-1}$ is the $k \times k$ sample correlation matrix using shocks from $t - m, \ldots, t - 1$ for a prespecified m. Typically, one assumes that $0 \leq \theta_i < 1$ and $\theta_1 + \theta_2 < 1$ so that the resulting correlation matrix $\boldsymbol{\rho}_t$ is positive definite for all t. For a given $\boldsymbol{\rho}$, the model is parsimonious. In applications, the choice of $\boldsymbol{\rho}$ and m deserves a careful investigation. One possibility is to let $\boldsymbol{\rho}$ be the sample correlation matrix of the returns. The correlation equation then only employs two parameters.

Engle (2002) proposes the model

$$\boldsymbol{\rho}_t = \boldsymbol{J}_t \boldsymbol{Q}_t \boldsymbol{J}_t,$$

where $\boldsymbol{Q}_t = (q_{ij,t})_{k \times k}$ is a positive-definite matrix, $\boldsymbol{J}_t = \text{diag}\{q_{11,t}^{-1/2}, \ldots, q_{kk,t}^{-1/2}\}$, and $\boldsymbol{Q}_t$ satisfies

$$\boldsymbol{Q}_t = (1 - \theta_1 - \theta_2)\bar{\boldsymbol{Q}} + \theta_1 \boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}'_{t-1} + \theta_2 \boldsymbol{Q}_{t-1},$$

where $\boldsymbol{\epsilon}_t$ is the standardized innovation vector with elements $\epsilon_{it} = a_{it}/\sqrt{\sigma_{ii,t}}$, $\bar{\boldsymbol{Q}}$ is the unconditional covariance matrix of $\boldsymbol{\epsilon}_t$, and θ_1 and θ_2 are nonnegative scalar parameters satisfying $0 < \theta_1 + \theta_2 < 1$. The $\boldsymbol{J}_t$ matrix is a normalization matrix to guarantee that $\boldsymbol{R}_t$ is a correlation matrix.

An obvious drawback of the prior two models is that θ_1 and θ_2 are scalar so that all the conditional correlations have the same dynamics. This might be hard to justify in real applications, especially when the dimension k is large.

Tsay (2006) extends the previous DCC models in two ways. First, the standardized innovations are assumed to follow a multivariate Student-t distribution of Eq. (10.42). Second, the marginal volatility models have leverage effects. Specifically, the volatility equation for $\boldsymbol{r}_t$ is

$$\boldsymbol{D}_t^2 = \boldsymbol{\Lambda}_0 + \boldsymbol{\Lambda}_1 \boldsymbol{D}_{t-1}^2 + \boldsymbol{\Lambda}_2 \boldsymbol{A}_{t-1}^2 + \boldsymbol{\Lambda}_3 \boldsymbol{L}_{t-1}^2, \tag{10.31}$$

where $\boldsymbol{D}_t$ is the diagonal matrix of volatilities as defined in Eq. (10.7), $\boldsymbol{A}_j = \text{diag}\{a_{1j}, \ldots, a_{kj}\}$, $\boldsymbol{\Lambda}_i = \text{diag}\{\ell_{1i}, \ldots, \ell_{ki}\}$ are $k \times k$ diagonal matrices of parameters and $\boldsymbol{L}_{t-1} = \text{diag}\{L_{1,t-1}, \ldots, L_{k,t-1}\}$ is also a $k \times k$ diagonal matrix with diagonal elements

$$L_{i,t-1} = \begin{cases} a_{i,t-1} & \text{if } a_{i,t-1} < 0, \\ 0 & \text{otherwise.} \end{cases}$$

In Eq. (10.31), the parameters ℓ_{ij} satisfy $0 \leq \sum_{j=1}^{3} \ell_{ij} < 1$, $\ell_{i0} > 0$ for $i = 1, \ldots, k$, and $\ell_{ji} \geq 0$ for all positive i and j. The constraint ensures that the volatilities exist. Of course, if $\boldsymbol{\Lambda}_3 = \boldsymbol{0}$, then there is no leverage effect.

The correlation equation is

$$\boldsymbol{\rho}_t = (1 - \theta_1 - \theta_2)\hat{\boldsymbol{\rho}} + \theta_1 \boldsymbol{\psi}_{t-1} + \theta_2 \boldsymbol{\rho}_{t-1}, \tag{10.32}$$

where $\hat{\boldsymbol{\rho}}$ is the sample correlation matrix of the returns and $0 \leq \theta_1 + \theta_2 < 1$ with $\theta_i \geq 0$ for $i = 1, 2$.

Example 10.6. To illustrate the DCC model, we consider the daily exchange rates between U.S. dollar versus European euro and Japanese yen and the stock prices of IBM and Dell from January 1999 to December 2004. The exchange rates are the noon spot rate obtained from the Federal Reserve Bank of St. Louis and the stock returns are from the Center for Research in Security Prices (CRSP). We compute the simple returns of the exchange rates and remove returns for those days when one of the markets was not open. This results in a four-dimensional return series with 1496 observations. The return vector is $\boldsymbol{r}_t = (r_{1t}, r_{2t}, r_{3t}, r_{4t})'$ with r_{1t} and r_{2t} being the returns of euro and yen exchange rate, respectively, and r_{3t} and r_{4t} are the returns of IBM and Dell stock, respectively. All returns are in percentages. Figure 10.11 shows the time plot of the return series. From the plot, equity returns have higher variability than the exchange rate returns, and the variability of equity returns appears to be decreasing in latter years. Table 10.1 provides some descriptive statistics of the return series. As expected, the means of the returns are essentially zero and all four series have heavy tails with positive excess kurtosis.

The equity returns have some serial correlations, but the magnitude is small. If multivariate Ljung–Box statistics are used, we have $Q(3) = 59.12$ with a p value of 0.13 and $Q(5) = 106.44$ with a p value of 0.03. For simplicity, we use the sample mean as the mean equation and apply the proposed multivariate volatility model to the mean-corrected data. In estimation, we start with a general model, but add some equality constraints as some estimates appear to be close to each other. The results are given in Table 10.2 along with the value of likelihood function evaluated at the estimates.

For each estimated multivariate volatility model in Table 10.2, we compute the standardized residuals as

$$\hat{\boldsymbol{\epsilon}}_t = \widehat{\boldsymbol{\Sigma}}_t^{-1/2} \boldsymbol{a}_t,$$

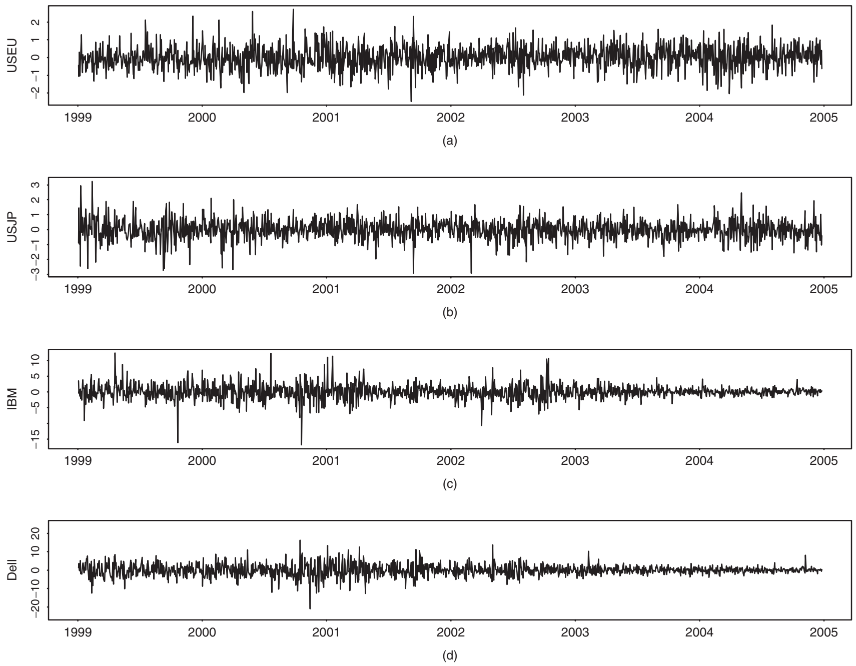


Figure 10.11 Time plots of daily simple return series from January 1999 to December 2004: (a) dollar–euro exchange rate, (b) dollar–yen exchange rate, (c) IBM stock, and (d) Dell stock.

TABLE 10.1 Descriptive Statistics of Daily Returns of Example 10.6.[a]

Asset	USEU	JPUS	IBM	DELL
Mean	0.0091	−0.0059	0.0066	0.0028
Standard error	0.6469	0.6626	5.4280	10.1954
Skewness	0.0342	−0.1674	−0.0530	−0.0383
Excess kurtosis	2.7090	2.0332	6.2164	3.3054
Box–Ljung $Q(12)$	12.5	6.4	24.1	24.1

[a]The returns are in percentages, and the sample period is from January 1999 to December 2004 for 1496 observations.

where $\widehat{\boldsymbol{\Sigma}}_t^{1/2}$ is the symmetric square root matrix of the estimated volatility matrix $\widehat{\boldsymbol{\Sigma}}_t$. We apply the multivariate Ljung–Box statistics to the standardized residuals $\hat{\boldsymbol{\epsilon}}_t$ and its squared process $\hat{\boldsymbol{\epsilon}}_t^2$ of a fitted model to check model adequacy. For the full model in Table 10.2(a), we have $Q(10) = 167.79(0.32)$ and $Q(10) = 110.19(1.00)$ for $\hat{\boldsymbol{\epsilon}}_t$ and $\hat{\boldsymbol{\epsilon}}_t^2$, respectively, where the number in parentheses denotes p value. Clearly, the model adequately describes the first two moments of the return series. For the model in Table 10.2(b), we have $Q(10) = 168.59(0.31)$ and $Q(10) = 109.93(1.00)$. For the final restricted model in Table 10.2(c), we obtain $Q(10) =$

TABLE 10.2 Estimation Results of Multivariate Volatility Models for Example 10.6[a]

(a) Full Model Estimation with $L_{\max} = -9175.80$

$\mathbf{\Lambda}_0$	$\mathbf{\Lambda}_1$	$\mathbf{\Lambda}_2$	$(v, \theta_1, \theta_2)'$
0.0041(0.0033)	0.9701(0.0114)	0.0214(0.0075)	7.8729(0.4693)
0.0088(0.0038)	0.9515(0.0126)	0.0281(0.0084)	0.9808(0.0029)
0.0071(0.0053)	0.9636(0.0092)	0.0326(0.0087)	0.0137(0.0025)
0.0150(0.0136)	0.9531(0.0155)	0.0461(0.0164)	

(b) Restricted Model with $L_{\max} = -9176.62$

$\mathbf{\Lambda}_0$	$\mathbf{\Lambda}_1 = \lambda \times I$	$\mathbf{\Lambda}_2$	$(v, \theta_1, \theta_2)'$
0.0066(0.0028)	0.9606(0.0068)	0.0255(0.0068)	7.8772(0.7144)
0.0066(0.0023)		0.0240(0.0059)	0.9809(0.0042)
0.0080(0.0052)		0.0355(0.0068)	0.0137(0.0025)
0.0108(0.0086)		0.0385(0.0073)	

(c) Final Restricted Model with $L_{\max} = -9177.44$

$\mathbf{\Lambda}_0(\lambda_1, \lambda_1, \lambda_3, \lambda_4)$	$\mathbf{\Lambda}_1 = \lambda \times \boldsymbol{I}$	$\mathbf{\Lambda}_2(b_1, b_1, b_2, b_2)$	$(v, \theta_1, \theta_2)'$
0.0067(0.0021)	0.9603(0.0063)	0.0248(0.0048)	7.9180(0.6952)
0.0067(0.0021)		0.0248(0.0048)	0.9809(0.0042)
0.0061(0.0044)		0.0372(0.0061)	0.0137(0.0028)
0.0148(0.0084)		0.0372(0.0061)	

(d) Model with Leverage Effects, $L_{\max} = -9169.04$

$\mathbf{\Lambda}_0(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$	$\mathbf{\Lambda}_1 = \lambda \times \boldsymbol{I}$	$\mathbf{\Lambda}_2(b_1, b_2, b_3, b_4)$	$(v, \theta_1, \theta_2)'$
0.0064(0.0027)	0.9600(0.0065)	0.0254(0.0063)	8.4527(0.7556)
0.0066(0.0023)		0.0236(0.0054)	0.9810(0.0044)
0.0128(0.0055)		0.0241(0.0056)	0.0132(0.0027)
0.0210(0.0099)		0.0286(0.0062)	

[a] $L_{\max}$ denotes the value of likelihood function evaluated at the estimates, v is the degrees of freedom of the multivariate Student-t distribution, and the numbers in parentheses are asymptotic standard errors.

168.50(0.31) and $Q(10) = 111.75(1.00)$. Again, the restricted models are capable of describing the mean and volatility of the return series.

From Table 10.2, we make the following observations. First, using the likelihood ratio test, we cannot reject the final restricted model compared with the full model. This results in a very parsimonious model consisting of only 9 parameters for the time-varying correlations of the four-dimensional return series. Second, for the two stock return series, the constant terms in $\mathbf{\Lambda}_0$ are not significantly different from zero, and the sum of GARCH parameters is $0.0372 + 0.9603 = 0.9975$, which is very close to unity. Consequently, the volatility series of the two equity returns exhibit IGARCH behavior. On the other hand, the volatility series of the two exchange rate returns appear to have a nonzero constant term and high persistence in GARCH

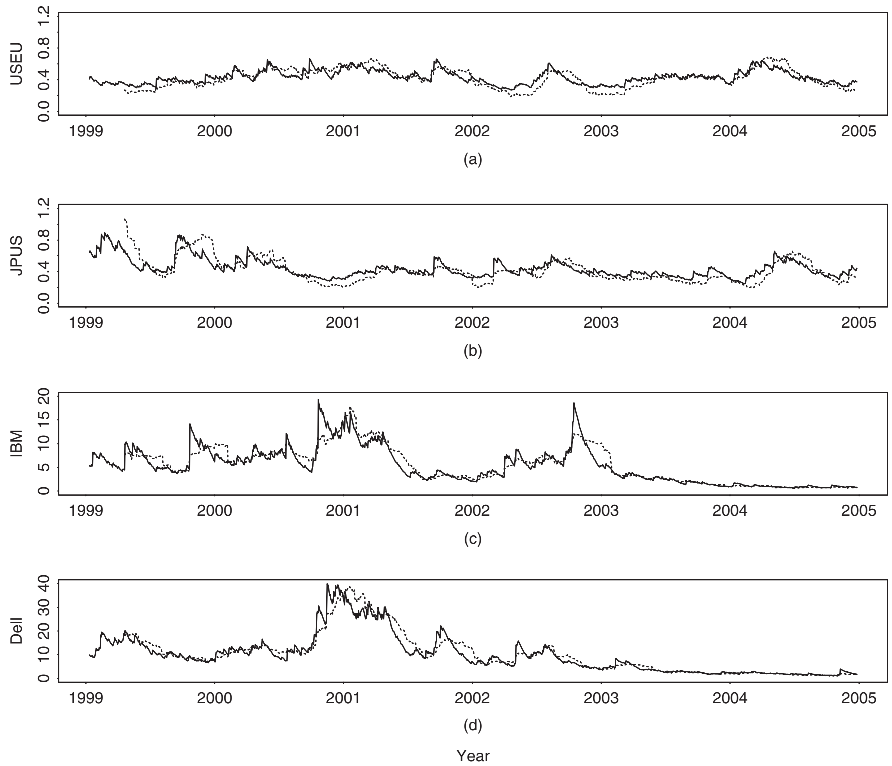


Figure 10.12 Time plots of estimated volatility series of four asset returns. Solid line is from proposed model and dashed line is from a rolling estimation with window size 69: (a) dollar–euro exchange rate, (b) dollar–yen exchange rate, (c) IBM stock, and (d) Dell stock.

parameters. Third, to better understand the efficacy of the proposed model, we compare the results of the final restricted model with those of rolling estimates. The rolling estimates of covariance matrix are obtained using a moving window of size 69, which is the approximate number of trading days in a quarter. Figure 10.12 shows the time plot of estimated volatility. The solid line is the volatility obtained by the proposed model and the dashed line is for volatility of the rolling estimation. The overall pattern seems similar, but, as expected, the rolling estimates respond more slowly than the proposed model to large innovations. This is shown by the faster rise and decay of the volatility obtained by the proposed model. Figure 10.13 shows the time-varying correlations of the four asset returns. The solid line denotes correlations obtained by the final restricted model of Table 10.2, whereas the dashed line is for rolling estimation. The correlations of the proposed model seem to be smoother.

Table 10.2(d) gives the results of a fitted integrated GARCH-type model with leverage effects. The leverage effects are statistically significant for equity returns only and are in the form of an IGARCH model. Specifically, the $\boldsymbol{\Lambda}_3$ matrix of the

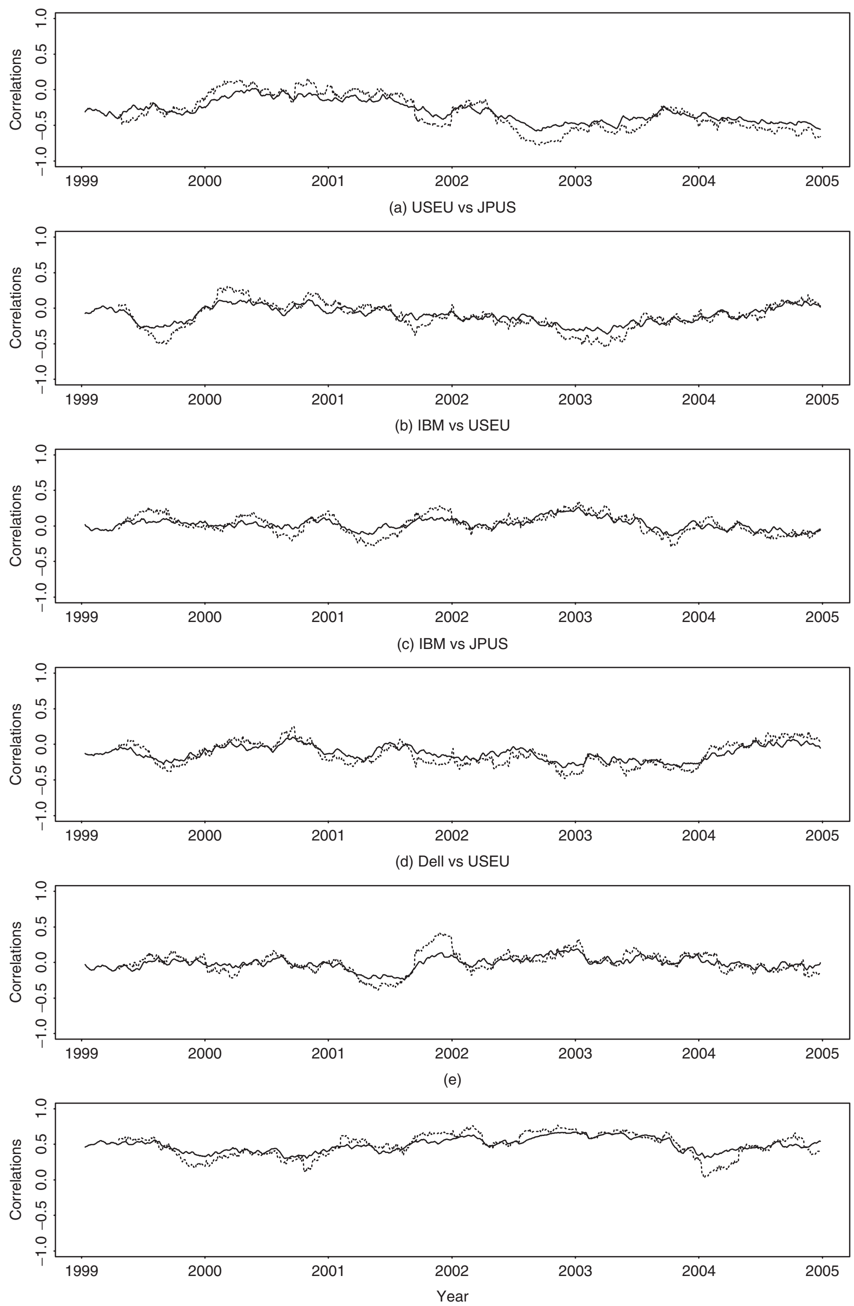


Figure 10.13 Time plots of time-varying correlations between percentage simple returns of four assets from January 1999 to December 2004. Solid line is from the proposed model, whereas dashed line is from a rolling estimation with window size 69.

correlation equation in Eq. (10.31) is

$$\begin{aligned}\mathbf{\Lambda}_3 &= \text{diag}\,\{0, 0, (1-0.96-0.0241), (1-0.96-0.0286)\}\\ &= \text{diag}\{0, 0, 0.0159, 0.0114\}.\end{aligned}$$

Although the magnitudes of the leverage parameters are small, they are statistically significant. This is shown by the likelihood ratio test. Specifically, comparing the fitted models in Table 10.2(b) and (d), the likelihood ratio statistic is 15.16, which has a p value of 0.0005 based on the chi-squared distribution with 2 degrees of freedom.

10.5 HIGHER DIMENSIONAL VOLATILITY MODELS

In this section, we make use of the sequential nature of Cholesky decomposition to suggest a strategy for building a high-dimensional volatility model. Again write the vector return series as $\boldsymbol{r}_t = \boldsymbol{\mu}_t + \boldsymbol{a}_t$. The mean equations for $\boldsymbol{r}_t$ can be specified by using the methods of Chapter 8. A simple vector AR model is often sufficient. Here we focus on building a volatility model using the shock process $\boldsymbol{a}_t$.

Based on the discussion of Cholesky decomposition in Section 10.3, the orthogonal transformation from a_{it} to b_{it} only involves b_{jt} for $j < i$. In addition, the time-varying volatility models built in Section 10.4 appear to be nested in the sense that the model for $g_{ii,t}$ depends only on quantities related to b_{jt} for $j < i$. Consequently, we consider the following sequential procedure to build a multivariate volatility model:

1. Select a market index or a stock return that is of major interest. Build a univariate volatility model for the selected return series.
2. Augment a second return series to the system, perform the orthogonal transformation on the shock process of this new return series, and build a bivariate volatility model for the system. The parameter estimates of the univariate model in step 1 can be used as the starting values in bivariate estimation.
3. Augment a third return series to the system, perform the orthogonal transformation on this newly added shock process, and build a three-dimensional volatility model. Again parameter estimates of the bivariate model can be used as the starting values in the three-dimensional estimation.
4. Continue the augmentation until a joint volatility model is built for all the return series of interest.

Finally, model checking should be performed in each step to ensure the adequacy of the fitted model. Experience shows that this sequential procedure can simplify substantially the complexity involved in building a high-dimensional volatility model. In particular, it can markedly reduce the computing time in estimation.

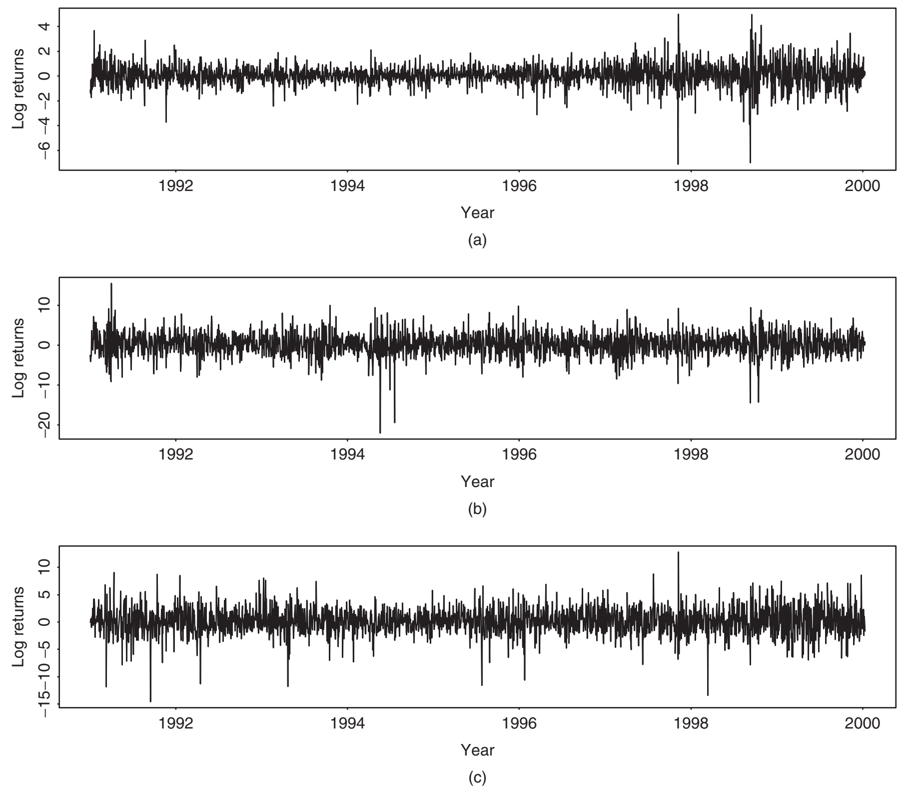


Figure 10.14 Time plots of daily log returns in percentages of (a) S&P 500 index and stocks of (b) Cisco Systems and (c) Intel Corporation from January 2, 1991, to December 31, 1999.

Example 10.7. We demonstrate the proposed sequential procedure by building a volatility model for the daily log returns of the S&P 500 index and the stocks of Cisco Systems and Intel Corporation. The data span is from January 2, 1991, to December 31, 1999, with 2275 observations. The log returns are in percentages and shown in Figure 10.14. Components of the return series are ordered as $\boldsymbol{r}_t = (\text{SP5}_t, \text{CSCO}_t, \text{INTC}_t)'$. The sample means, standard errors, and correlation matrix of the data are

$$\widehat{\boldsymbol{\mu}} = \begin{bmatrix} 0.066 \\ 0.257 \\ 0.156 \end{bmatrix}, \quad \begin{bmatrix} \hat{\sigma}_1 \\ \hat{\sigma}_2 \\ \hat{\sigma}_3 \end{bmatrix} = \begin{bmatrix} 0.875 \\ 2.853 \\ 2.464 \end{bmatrix}, \quad \widehat{\boldsymbol{\rho}} = \begin{bmatrix} 1.00 & 0.52 & 0.50 \\ 0.52 & 1.00 & 0.47 \\ 0.50 & 0.47 & 1.00 \end{bmatrix}.$$

Using the Ljung–Box statistics to detect any serial dependence in the return series, we obtain $Q_3(1) = 26.20$, $Q_3(4) = 79.73$, and $Q_3(8) = 123.68$. These test statistics are highly significant with p values close to zero as compared with

TABLE 10.3 Sample Cross-Correlation Matrices of Daily Log Returns of S&P 500 Index and Stocks of Cisco Systems and Intel Corporation from January 2, 1991, to December 31, 1999

Lag					
1	2	3	4	5	6
. . .	. . .	− . .	. . .	− . .	. . .
. . .	. − .	. . .	. . .	− . .	. − .
− . .	. . .	. . .	. . .	− . .	. . .

chi-squared distributions with degrees of freedom 9, 36, and 72, respectively. There is indeed some serial dependence in the data. Table 10.3 gives the first five lags of sample cross-correlation matrices shown in the simplified notation of Chapter 8. An examination of the table shows that (a) the daily log returns of the S&P 500 index does not depend on the past returns of Cisco or Intel, (b) the log return of Cisco stock has some serial correlations and depends on the past returns of the S&P 500 index (see lags 2 and 5), and (c) the log return of Intel stock depends on the past returns of the S&P 500 index (see lags 1 and 5). These observations are similar to those between the returns of IBM stock and the S&P 500 index analyzed in Chapter 8. They suggest that returns of individual large-cap companies tend to be affected by the past behavior of the market. However, the market return is not significantly affected by the past returns of individual companies.

Turning to volatility modeling and following the suggested procedure, we start with the log returns of the S&P 500 index and obtain the model

$$\begin{aligned} r_{1t} &= 0.078 + 0.042 r_{1,t-1} - 0.062 r_{1,t-3} - 0.048 r_{1,t-4} - 0.052 r_{1,t-5} + a_{1t}, \\ \sigma_{11,t} &= 0.013 + 0.092 a^2_{1,t-1} + 0.894 \sigma_{11,t-1}, \end{aligned} \tag{10.33}$$

where standard errors of the parameters in the mean equation are 0.016, 0.023, 0.020, 0.022, and 0.020, respectively, and those of the parameters in the volatility equation are 0.002, 0.006, and 0.007, respectively. Univariate Ljung–Box statistics of the standardized residuals and their squared series fail to detect any remaining serial correlation or conditional heteroscedasticity in the data. Indeed, we have $Q(10) = 7.38(0.69)$ for the standardized residuals and $Q(10) = 3.14(0.98)$ for the squared series.

Augmenting the daily log returns of Cisco stock to the system, we build a bivariate model with mean equations given by

$$\begin{aligned} r_{1t} &= 0.065 - 0.046 r_{1,t-3} + a_{1t}, \\ r_{2t} &= 0.325 + 0.195 r_{1,t-2} - 0.091 r_{2,t-2} + a_{2t}, \end{aligned} \tag{10.34}$$

where all of the estimates are statistically significant at the 1% level. Using the notation of Cholesky decomposition, we obtain the volatility equations as

$$\begin{aligned}
g_{11,t} &= 0.006 + 0.051 b_{1,t-1}^2 + 0.943 g_{11,t-1}, \\
q_{21,t} &= 0.331 + 0.790 q_{21,t-1} - 0.041 a_{2,t-1}, \\
g_{22,t} &= 0.177 + 0.082 b_{2,t-1}^2 + 0.890 g_{22,t-1},
\end{aligned} \tag{10.35}$$

where $b_{1t} = a_{1t}$, $b_{2t} = a_{2t} - q_{21,t} b_{1t}$, standard errors of the parameters in the equation of $g_{11,t}$ are 0.001, 0.005, and 0.006, those of the parameters in the equation of $q_{21,t}$ are 0.156, 0.099, and 0.011, and those of the parameters in the equation of $g_{22,t}$ are 0.029, 0.008, and 0.011, respectively. The bivariate Ljung–Box statistics of the standardized residuals fail to detect any remaining serial dependence or conditional heteroscedasticity. The bivariate model is adequate. Comparing with Eq. (10.33), we see that the difference between the marginal and univariate models of r_{1t} is small.

The next and final step is to augment the daily log returns of Intel stock to the system. The mean equations become

$$\begin{aligned}
r_{1t} &= 0.065 - 0.043 r_{1,t-3} + a_{1t}, \\
r_{2t} &= 0.326 + 0.201 r_{1,t-2} - 0.089 r_{2,t-1} + a_{2t}, \\
r_{3t} &= 0.192 - 0.264 r_{1,t-1} + 0.059 r_{3,t-1} + a_{3t},
\end{aligned} \tag{10.36}$$

where standard errors of the parameters in the first equation are 0.016 and 0.017, those of the parameters in the second equation are 0.052, 0.059, and 0.021, and those of the parameters in the third equation are 0.050, 0.057, and 0.022, respectively. All estimates are statistically significant at about the 1% level. As expected, the mean equations for r_{1t} and r_{2t} are essentially the same as those in the bivariate case.

The three-dimensional time-varying volatility model becomes a bit more complicated, but it remains manageable as

$$\begin{aligned}
g_{11,t} &= 0.006 + 0.050 b_{1,t-1}^2 + 0.943 g_{11,t-1}, \\
q_{21,t} &= 0.277 + 0.824 q_{21,t-1} - 0.035 a_{2,t-1}, \\
g_{22,t} &= 0.178 + 0.082 b_{2,t-1}^2 + 0.889 g_{22,t-1}, \\
q_{31,t} &= 0.039 + 0.973 q_{31,t-1} + 0.010 a_{3,t-1}, \\
q_{32,t} &= 0.006 + 0.981 q_{32,t-1} + 0.004 a_{2,t-1}, \\
g_{33,t} &= 1.188 + 0.053 b_{3,t-1}^2 + 0.687 g_{33,t-1} - 0.019 g_{22,t-1},
\end{aligned} \tag{10.37}$$

where $b_{1t} = a_{1t}$, $b_{2t} = a_{2t} - q_{21,t} b_{1t}$, $b_{3t} = a_{3t} - q_{31,t} b_{1t} - q_{32,t} b_{2t}$, and standard errors of the parameters are given in Table 10.4. Except for the constant term of the $q_{32,t}$ equation, all estimates are significant at the 5% level. Let

TABLE 10.4 Standard Errors of Parameter Estimates of Three-Dimensional Volatility Model for Daily Log Returns in Percentages of S&P 500 Index and Stocks of Cisco Systems and Intel Corporation from January 2, 1991, to December 31, 1999[a]

Equation	Standard Error				Equation	Standard Error		
$g_{11,t}$	0.001	0.005	0.006		$q_{21,t}$	0.135	0.086	0.010
$g_{22,t}$	0.029	0.009	0.011		$q_{31,t}$	0.017	0.012	0.004
$g_{33,t}$	0.407	0.015	0.100	0.008	$q_{32,t}$	0.004	0.013	0.001

[a]The ordering of the parameter is the same as appears in Eq. (10.37).

$\tilde{\boldsymbol{a}}_t = (a_{1t}/\hat{\sigma}_{1t}, a_{2t}/\hat{\sigma}_{2t}, a_{3t}/\hat{\sigma}_{3t})'$ be the standardized residual series, where $\hat{\sigma}_{it} = \sqrt{\hat{\sigma}_{ii,t}}$ is the fitted conditional standard error of the ith return. The Ljung–Box statistics of $\tilde{\boldsymbol{a}}_t$ give $Q_3(4) = 34.48(0.31)$ and $Q_3(8) = 60.42(0.70)$, where the degrees of freedom of the chi-squared distributions are 31 and 67, respectively, after adjusting for the number of parameters used in the mean equations. For the squared standardized residual series $\tilde{\boldsymbol{a}}_t^2$, we have $Q_3^*(4) = 28.71(0.58)$ and $Q_3^*(8) = 52.00(0.91)$. Therefore, the fitted model appears to be adequate in modeling the conditional means and volatilities.

The three-dimensional volatility model in Eq. (10.37) shows some interesting features. First, it is essentially a time-varying correlation GARCH(1,1) model because only lag-1 variables are used in the equations. Second, the volatility of the daily log returns of the S&P 500 index does not depend on the past volatilities of Cisco or Intel stock returns. Third, by taking the inverse transformation of the Cholesky decomposition, the volatilities of daily log returns of Cisco and Intel stocks depend on the past volatility of the market return; see the relationships between elements of $\boldsymbol{\Sigma}_t$, $\boldsymbol{L}_t$, and $\boldsymbol{G}_t$ given in Section 10.3. Fourth, the correlation quantities $q_{ij,t}$ have high persistence with large AR(1) coefficients.

Figure 10.15 shows the fitted volatility processes of the model (i.e., $\hat{\sigma}_{ii,t}$) for the data. The volatility of the index return is much smaller than those of the two individual stock returns. The plots also show that the volatility of the index return has increased in recent years, but this is not the case for the return of Cisco Systems. Figure 10.16 shows the time-varying correlation coefficients between the three return series. Of particular interest is to compare Figures 10.15 and 10.16. They show that the correlation coefficient between two return series increases when the returns are volatile. This is in agreement with the empirical study of relationships between international stock market indexes for which the correlation between two markets tends to increase during a financial crisis.

The volatility model in Eq. (10.37) consists of two sets of equations. The first set of equations describes the time evolution of conditional variances (i.e., $g_{ii,t}$), and the second set of equations deals with correlation coefficients (i.e., $q_{ij,t}$ with $i > j$). For this particular data set, an AR(1) model might be sufficient for the correlation equations. Similarly, a simple AR model might also be sufficient for the conditional variances. Define $\boldsymbol{v}_t = (v_{11,t}, v_{22,t}, v_{33,t})'$, where $v_{ii,t} = \ln(g_{ii,t})$, and $\boldsymbol{q}_t = (q_{21,t}, q_{31,t}, q_{32,t})'$. The previous discussion suggests that we can use the

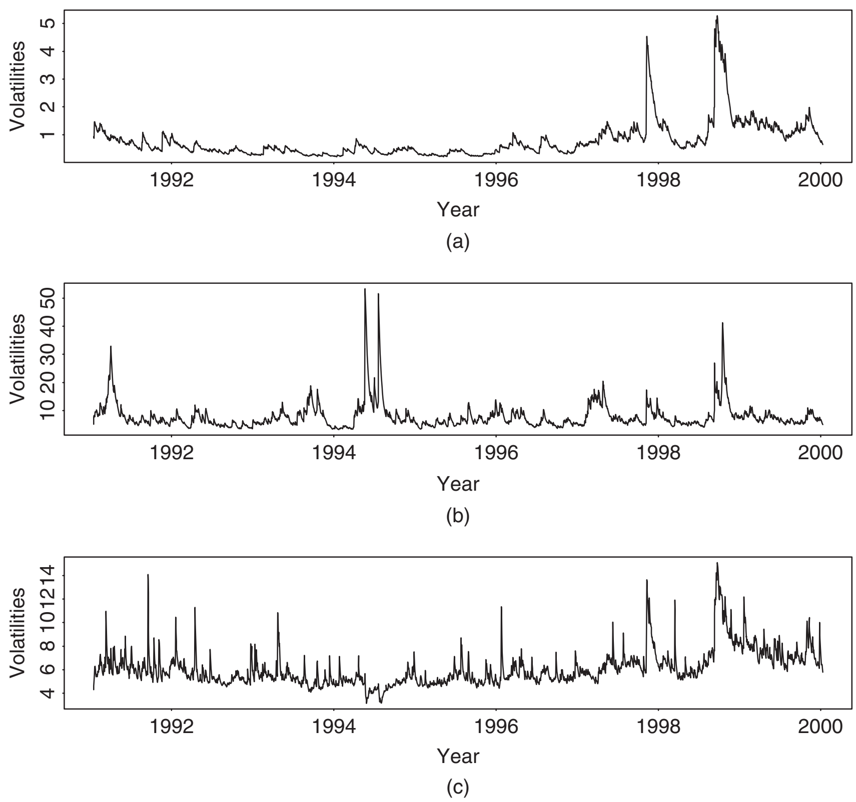


Figure 10.15 Time plots of fitted volatilities for daily log returns, in percentages, of (a) S&P 500 index and stocks of (b) Cisco Systems and (c) Intel Corporation from January 2, 1991, to December 31, 1999.

simple lag-1 models

$$\boldsymbol{v}_t = \boldsymbol{c}_1 + \boldsymbol{\beta}_1 \boldsymbol{v}_{t-1}, \qquad \boldsymbol{q}_t = \boldsymbol{c}_2 + \boldsymbol{\beta}_2 \boldsymbol{q}_{t-1}$$

as exact functions to model the volatility of asset returns, where $\boldsymbol{c}_i$ are constant vectors and $\boldsymbol{\beta}_i$ are 3×3 real-valued matrices. If a noise term is also included in the above equations, then the models become

$$\boldsymbol{v}_t = \boldsymbol{c}_1 + \boldsymbol{\beta}_1 \boldsymbol{v}_{t-1} + \boldsymbol{e}_{1t}, \qquad \boldsymbol{q}_t = \boldsymbol{c}_2 + \boldsymbol{\beta}_2 \boldsymbol{q}_{t-1} + \boldsymbol{e}_{2t},$$

where $\boldsymbol{e}_{it}$ are random shocks with mean zero and a positive-definite covariance matrix, and we have a simple multivariate stochastic volatility model. In a recent manuscript, Chib, Nardari, and Shephard (1999) use Markov chain Monte Carlo (MCMC) methods to study high-dimensional stochastic volatility models. The model considered there allows for time-varying correlations, but in a relatively restrictive manner. Additional references of multivariate volatility model include

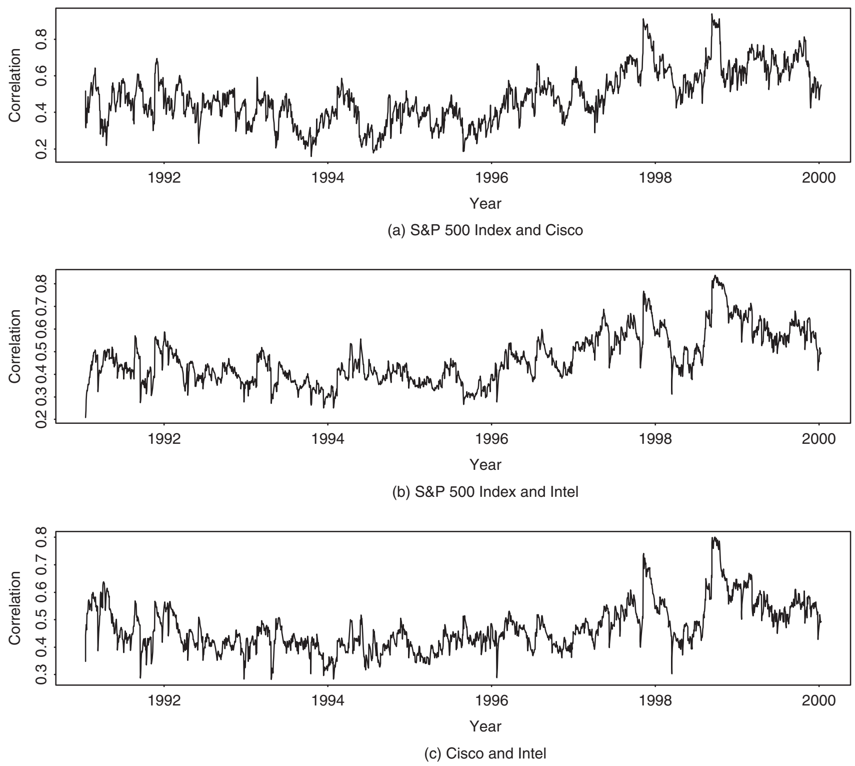


Figure 10.16 Time plots of fitted time-varying correlation coefficients between daily log returns of S&P 500 index and stocks of Cisco Systems and Intel Corporation from January 2, 1991, to December 31, 1999.

Harvey, Ruiz, and Shephard (1994). We discuss MCMC methods in volatility modeling in Chapter 12.

10.6 FACTOR–VOLATILITY MODELS

Another approach to simplifying the dynamic structure of a multivariate volatility process is to use factor models. In practice, the "common factors" can be determined a priori by substantive matter or empirical methods. As an illustration, we use the factor analysis of Chapter 8 to discuss factor–volatility models. Because volatility models are concerned with the evolution over time of the conditional covariance matrix of $\boldsymbol{a}_t$, where $\boldsymbol{a}_t = \boldsymbol{r}_t - \boldsymbol{\mu}_t$, a simple way to identify the "common factors" in volatility is to perform a principal component analysis (PCA) on $\boldsymbol{a}_t$; see the PCA of Chapter 8. Building a factor–volatility model thus involves a three-step procedure:

- Select the first few principal components that explain a high percentage of variability in $\boldsymbol{a}_t$.

- Build a volatility model for the selected principal components.
- Relate the volatility of each a_{it} series to the volatilities of the selected principal components.

The objective of such a procedure is to reduce the dimension but maintain an accurate approximation of the multivariate volatility.

Example 10.8. Consider again the monthly log returns, in percentages, of IBM stock and the S&P 500 index of Example 10.5. Using the bivariate AR(3) model of Example 8.4, we obtain an innovational series $\boldsymbol{a}_t$. Performing a PCA on $\boldsymbol{a}_t$ based on its covariance matrix, we obtained eigenvalues 63.373 and 13.489. The first eigenvalue explains 82.2% of the generalized variance of $\boldsymbol{a}_t$. Therefore, we may choose the first principal component $x_t = 0.797a_{1t} + 0.604a_{2t}$ as the common factor. Alternatively, as shown by the model in Example 8.4, the serial dependence in $\boldsymbol{r}_t$ is weak and, hence, one can perform the PCA on $\boldsymbol{r}_t$ directly. For this particular instance, the two eigenvalues of the sample covariance matrix of $\boldsymbol{r}_t$ are 63.625 and 13.513, which are essentially the same as those based on $\boldsymbol{a}_t$. The first principal component explains approximately 82.5% of the generalized variance of $\boldsymbol{r}_t$, and the corresponding common factor is $x_t = 0.796r_{1t} + 0.605r_{2t}$. Consequently, for the two monthly log return series considered, the effect of the conditional mean equations on PCA is negligible.

Based on the prior discussion and for simplicity, we use $x_t = 0.796r_{1t} + 0.605r_{2t}$ as a common factor for the two monthly return series. Figure 10.17(a) shows the time plot of this common factor. If univariate Gaussian GARCH models are entertained, we obtain the following model for x_t:

$$\begin{aligned} x_t &= 1.317 + 0.096x_{t-1} + a_t, \qquad a_t = \sigma_t\epsilon_t, \\ \sigma_t^2 &= 3.834 + 0.110a_{t-1}^2 + 0.825\sigma_{t-1}^2. \end{aligned} \tag{10.38}$$

All parameter estimates of the previous model are highly significant at the 1% level, and the Ljung–Box statistics of the standardized residuals and their squared series fail to detect any model inadequacy. Figure 10.17(b) shows the fitted volatility of x_t [i.e., the sample σ_t^2 series in Eq. (10.38)].

Using σ_t^2 of model (10.38) as a common volatility factor, we obtain the following model for the original monthly log returns. The mean equations are

$$\begin{aligned} r_{1t} &= 1.140 + 0.079r_{1,t-1} + 0.067r_{1,t-2} - 0.122r_{2,t-2} + a_{1t}, \\ r_{2t} &= 0.537 + a_{2t}, \end{aligned}$$

where standard errors of the parameters in the first equation are 0.211, 0.030, 0.031, and 0.043, respectively, and standard error of the parameter in the second equation

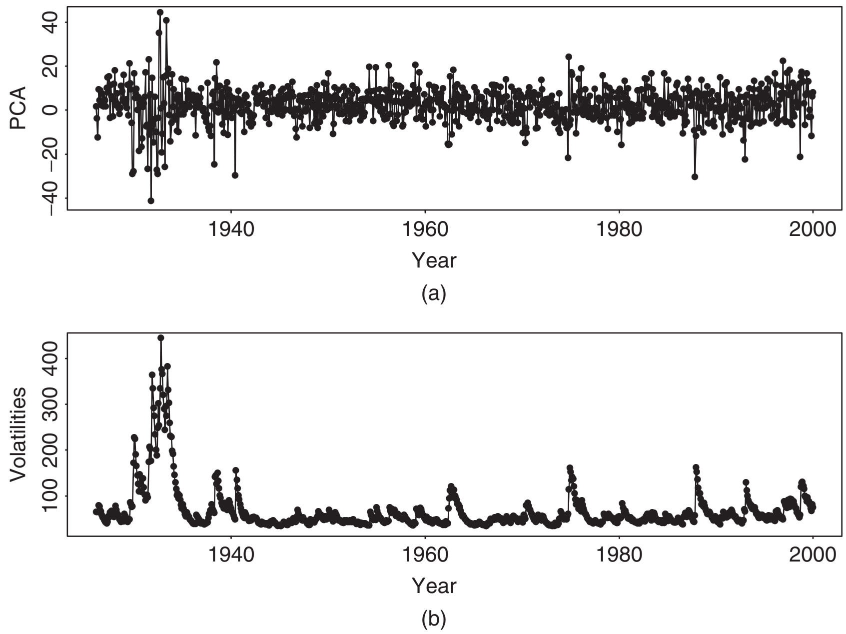


Figure 10.17 (a) Time plot of first principal component of monthly log returns of IBM stock and S&P 500 index. (b) Fitted volatility process based on a GARCH(1,1) model.

is 0.165. The conditional variance equation is

$$\begin{bmatrix} \sigma_{11,t} \\ \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} 19.08 \\ (3.70) \\ -5.62 \\ (2.36) \end{bmatrix} + \begin{bmatrix} 0.098 & \\ (0.044) & \\ & \\ & \end{bmatrix} \begin{bmatrix} a_{1,t-1}^2 \\ a_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} 0.333 \\ (0.076) \\ 0.596 \\ (0.050) \end{bmatrix} \sigma_t^2, \tag{10.39}$$

where, as before, standard errors are in parentheses, and σ_t^2 is obtained from model (10.38). The conditional correlation equation is

$$\begin{aligned} \rho_t &= \frac{\exp(q_t)}{1+\exp(q_t)}, \\ q_t &= -2.098 + 4.120\rho_{t-1} + 0.078\frac{a_{1,t-1}a_{2,t-1}}{\sqrt{\sigma_{11,t-1}\sigma_{22,t-1}}}, \end{aligned} \tag{10.40}$$

where standard errors of the three parameters are 0.025, 0.038, and 0.015, respectively. Defining the standardized residuals as before, we obtain $Q_2(4) = 15.37(0.29)$ and $Q_2(8) = 34.24(0.23)$, where the number in parentheses denotes the p value. Therefore, the standardized residuals have no serial correlations. Yet we have $Q_2^*(4) = 20.25(0.09)$ and $Q_2^*(8) = 61.95(0.0004)$ for the squared standardized residuals. The volatility model in Eq. (10.39) does not adequately handle the conditional heteroscedasticity of the data especially at higher lags. This is not surprising as the single common factor only explains about 82.5% of the generalized variance of the data.

Comparing the factor model in Eqs. (10.39) and (10.40) with the time-varying correlation model in Eqs. (10.26) and (10.27), we see that (a) the correlation equations of the two models are essentially the same, (b) as expected the factor model uses fewer parameters in the volatility equation, and (c) the common-factor model provides a reasonable approximation to the volatility process of the data.

Remark. In Example 10.8, we used a two-step estimation procedure. In the first step, a volatility model is built for the common factor. The estimated volatility is treated as given in the second step to estimate the multivariate volatility model. Such an estimation procedure is simple but may not be efficient. A more efficient estimation procedure is to perform a joint estimation. This can be done relatively easily provided that the common factors are known. For example, for the monthly log returns of Example 10.8, a joint estimation of Eqs. (10.38)–(10.40) can be performed if the common factor $x_t = 0.769r_{1t} + 0.605r_{2t}$ is treated as given. □

10.7 APPLICATION

We illustrate the application of multivariate volatility models by considering the value at risk (VaR) of a financial position with multiple assets. Suppose that an investor holds a long position in the stocks of Cisco Systems and Intel Corporation each worth \$1 million. We use the daily log returns for the two stocks from January 2, 1991, to December 31, 1999, to build volatility models. The VaR is computed using the 1-step-ahead forecasts at the end of data span and 5% critical values.

Let VaR_1 be the value at risk for holding the position on Cisco Systems stock and VaR_2 for holding Intel stock. Results of Chapter 7 show that the overall daily VaR for the investor is

$$\text{VaR} = \sqrt{\text{VaR}_1^2 + \text{VaR}_2^2 + 2\rho\text{VaR}_1\text{VaR}_2}.$$

In this illustration, we consider three approaches to volatility modeling for calculating VaR. For simplicity, we do not report standard errors for the parameters involved or model checking statistics. Yet all of the estimates are statistically significant at the 5% level, and the models are adequate based on the Ljung–Box statistics of the standardized residual series and their squared series. The log returns are in percentages so that the quantiles are divided by 100 in VaR calculations. Let r_{1t} be the return of Cisco stock and r_{2t} the return of Intel stock.

Univariate Models

This approach uses a univariate volatility model for each stock return and uses the sample correlation coefficient of the stock returns to estimate ρ. The univariate volatility models for the two stock returns are

$$\begin{aligned} r_{1t} &= 0.380 + 0.034 r_{1,t-1} - 0.061 r_{1,t-2} - 0.055 r_{1,t-3} + a_{1t}, \\ \sigma_{1t}^2 &= 0.599 + 0.117 a_{1,t-1}^2 + 0.814 \sigma_{1,t-1}^2 \end{aligned}$$

and

$$\begin{aligned} r_{2t} &= 0.187 + a_{2t}, \\ \sigma_{2t}^2 &= 0.310 + 0.032 a_{2,t-1}^2 + 0.918 \sigma_{2,t-1}^2. \end{aligned}$$

The sample correlation coefficient is 0.473. The 1-step-ahead forecasts needed in VaR calculation at the forecast origin $t = 2275$ are

$$\hat{r}_1 = 0.626, \qquad \hat{\sigma}_1^2 = 4.152, \qquad \hat{r}_2 = 0.187, \qquad \hat{\sigma}_2^2 = 6.087, \qquad \hat{\rho} = 0.473.$$

The 5% quantiles for both daily returns are

$$q_1 = 0.626 - 1.65\sqrt{4.152} = -2.736, \qquad q_2 = 0.187 - 1.65\sqrt{6.087} = -3.884,$$

where the negative sign denotes loss. For the individual stocks, $\mathrm{VaR}_1 = \$1000000 q_1/100 = \$27{,}360 and \mathrm{VaR}_2 = \$1000000 q_2/100 = \$38{,}840$. Consequently, the overall VaR for the investor is $\mathrm{VaR} = \$57{,}117$.

Constant-Correlation Bivariate Model

This approach employs a bivariate GARCH(1,1) model for the stock returns. The correlation coefficient is assumed to be constant over time, but it is estimated jointly with other parameters. The model is

$$\begin{aligned} r_{1t} &= 0.385 + 0.038 r_{1,t-1} - 0.060 r_{1,t-2} - 0.047 r_{1,t-3} + a_{1t}, \\ r_{2t} &= 0.222 + a_{2t}, \\ \sigma_{11,t} &= 0.624 + 0.110 a_{1,t-1}^2 + 0.816 \sigma_{11,t-1}, \\ \sigma_{22,t} &= 0.664 + 0.038 a_{2,t-1}^2 + 0.853 \sigma_{22,t-1}, \end{aligned}$$

and $\hat{\rho} = 0.475$. This is a diagonal bivariate GARCH(1,1) model. The 1-step-ahead forecasts for VaR calculation at the forecast origin $t = 2275$ are

$$\hat{r}_1 = 0.373, \qquad \hat{\sigma}_1^2 = 4.287, \qquad \hat{r}_2 = 0.222, \qquad \hat{\sigma}_2^2 = 5.706, \qquad \hat{\rho} = 0.475.$$

Consequently, we have $\mathrm{VaR}_1 = \$30{,}432$ and $\mathrm{VaR}_2 = \$37{,}195$. The overall 5% VaR for the investor is $\mathrm{VaR} = \$58{,}180$.

Time-Varying Correlation Model

Finally, we allow the correlation coefficient to evolve over time by using the Cholesky decomposition. The fitted model is

$$
\begin{aligned}
r_{1t} &= 0.355 + 0.039 r_{1,t-1} - 0.057 r_{1,t-2} - 0.038 r_{1,t-3} + a_{1t}, \\
r_{2t} &= 0.206 + a_{2t}, \\
g_{11,t} &= 0.420 + 0.091 b_{1,t-1}^2 + 0.858 g_{11,t-1}, \\
q_{21,t} &= 0.123 + 0.689 q_{21,t-1} - 0.014 a_{2,t-1}, \\
g_{22,t} &= 0.080 + 0.013 b_{2,t-1}^2 + 0.971 g_{22,t-1},
\end{aligned}
$$

where $b_{1t} = a_{1t}$ and $b_{2t} = a_{2t} - q_{21,t} a_{1t}$. The 1-step-ahead forecasts for VaR calculation at the forecast origin $t = 2275$ are

$$
\hat{r}_1 = 0.352, \quad \hat{r}_2 = 0.206, \quad \hat{g}_{11} = 4.252, \quad \hat{q}_{21} = 0.421, \quad \hat{g}_{22} = 5.594.
$$

Therefore, we have $\hat{\sigma}_1^2 = 4.252$, $\hat{\sigma}_{21} = 1.791$, and $\hat{\sigma}_2^2 = 6.348$. The correlation coefficient is $\hat{\rho} = 0.345$. Using these forecasts, we have $\text{VaR}_1 = \$30{,}504$, $\text{VaR}_2 = \$39{,}512$, and the overall VaR $= \$57{,}648$.

The estimated VaR values of the three approaches are similar. The univariate models give the lowest VaR, whereas the constant-correlation model produces the highest VaR. The range of the difference is about \$1100. The time-varying volatility model seems to produce a compromise between the two extreme models.

10.8 MULTIVARIATE *t* DISTRIBUTION

Empirical analysis indicates that the multivariate Gaussian innovations used in the previous sections may fail to capture the kurtosis of asset returns. In this situation, a multivariate Student-*t* distribution might be useful. There are many versions of the multivariate Student-*t* distribution. We give a simple version here for volatility modeling.

A k-dimensional random vector $\boldsymbol{x} = (x_1, \ldots, x_k)'$ has a multivariate Student-*t* distribution with v degrees of freedom and parameters $\boldsymbol{\mu} = \boldsymbol{0}$ and $\boldsymbol{\Sigma} = \boldsymbol{I}$ (the identity matrix) if its probability density function (pdf) is

$$
f(\boldsymbol{x}|v) = \frac{\Gamma[(v+k)/2]}{(\pi v)^{k/2}\Gamma(v/2)} (1 + v^{-1}\boldsymbol{x}'\boldsymbol{x})^{-(v+k)/2}, \tag{10.41}
$$

where $\Gamma(y)$ is the gamma function; see Mardia, Kent, and Bibby (1979, p. 57). The variance of each component x_i in Eq. (10.41) is $v/(v-2)$, and hence we define $\boldsymbol{\epsilon}_t = \sqrt{(v-2)/v}\,\boldsymbol{x}$ as the standardized multivariate Student-*t* distribution with v degrees of freedom. By transformation, the pdf of $\boldsymbol{\epsilon}_t$ is

$$
f(\boldsymbol{\epsilon}_t|v) = \frac{\Gamma[(v+k)/2]}{[\pi(v-2)]^{k/2}\Gamma(v/2)} [1 + (v-2)^{-1}\boldsymbol{\epsilon}_t'\boldsymbol{\epsilon}_t]^{-(v+k)/2}. \tag{10.42}
$$

For volatility modeling, we write $\boldsymbol{a}_t = \boldsymbol{\Sigma}_t^{1/2}\boldsymbol{\epsilon}_t$ and assume that $\boldsymbol{\epsilon}_t$ follows the multivariate Student-t distribution in Eq. (10.42). By transformation, the pdf of $\boldsymbol{a}_t$ is

$$f(\boldsymbol{a}_t|v, \boldsymbol{\Sigma}_t) = \frac{\Gamma[(v+k)/2]}{[\pi(v-2)]^{k/2}\Gamma(v/2)|\boldsymbol{\Sigma}_t|^{1/2}}[1+(v-2)^{-1}\boldsymbol{a}_t'\boldsymbol{\Sigma}_t^{-1}\boldsymbol{a}_t]^{-(v+k)/2}.$$

Furthermore, if we use the Cholesky decomposition of $\boldsymbol{\Sigma}_t$, then the pdf of the transformed shock $\boldsymbol{b}_t$ becomes

$$f(\boldsymbol{b}_t|v, \boldsymbol{L}_t, \boldsymbol{G}_t) = \frac{\Gamma[(v+k)/2]}{[\pi(v-2)]^{k/2}\Gamma(v/2)\prod_{j=1}^{k} g_{jj,t}^{1/2}} \times \left[1+(v-2)^{-1}\sum_{j=1}^{k}\frac{b_{jt}^2}{g_{jj,t}}\right]^{(v+k)/2},$$

where $\boldsymbol{a}_t = \boldsymbol{L}_t\boldsymbol{b}_t$ and $g_{jj,t}$ is the conditional variance of b_{jt}. Because this pdf does not involve any matrix inversion, the conditional-likelihood function of the data is easy to evaluate.

APPENDIX: SOME REMARKS ON ESTIMATION

The estimation of multivariate ARMA models in this chapter is done by using the time series program SCA of Scientific Computing Associates. The estimation of multivariate volatility models is done by using either the S-Plus package with FinMetrics or the Regression Analysis for Time Series (RATS) program or Matlab. Below are some run streams for estimating multivariate volatility models using the RATS program. A line starting with * means “comment” only.

Estimation of the Diagonal Constant-Correlation AR(2)–GARCH(1,1) Model for Example 10.5

The program includes some Ljung–Box statistics for each component and some fitted values for the last few observations. The data file is `m-ibmspln.txt`, which has two columns, and there are 888 observations.

```
all 0 888:1
open data m-ibmspln.txt
data(org=obs) / r1 r2
set h1 = 0.0
set h2 = 0.0
nonlin a0 a1 b1 a00 a11 b11 rho c1 c2 p1
frml a1t = r1(t)-c1-p1*r2(t-1)
frml a2t = r2(t)-c2
frml gvar1 = a0+a1*a1t(t-1)**2+b1*h1(t-1)
```

```
frml gvar2 = a00+a11*a2t(t-1)**2+b11*h2(t-1)
frml gdet = -0.5*(log(h1(t)=gvar1(t))+log(h2(t)=gvar2(t)) $
           +log(1.0-rho**2))
frml gln = gdet(t)-0.5/(1.0-rho**2)*((a1t(t)**2/h1(t)) $
    +(a2t(t)**2/h2(t))-2*rho*a1t(t)*a2t(t)/sqrt(h1(t)*h2(t)))
smpl 3 888
compute c1 = 1.22, c2 = 0.57, p1 = 0.1, rho = 0.1
compute a0 = 3.27, a1 = 0.1, b1 = 0.6
compute  a00 = 1.17, a11 = 0.13, b11 = 0.8
maximize(method=bhhh,recursive,iterations=150) gln
set fv1 = gvar1(t)
set resi1 = a1t(t)/sqrt(fv1(t))
set residsq = resi1(t)*resi1(t)
* Checking standardized residuals *
cor(qstats,number=12,span=4) resi1
* Checking squared standardized residuals *
cor(qstats,number=12,span=4) residsq
set fv2 = gvar2(t)
set resi2 = a2t(t)/sqrt(fv2(t))
set residsq = resi2(t)*resi2(t)
* Checking standardized residuals *
cor(qstats,number=12,span=4) resi2
* Checking squared standardized residuals *
cor(qstats,number=12,span=4) residsq
* Last few observations needed for computing forecasts *
set shock1 = a1t(t)
set shock2 = a2t(t)
print 885  888 shock1 shock2 fv1 fv2
```

Estimation of the Time-Varying Coefficient Model in Example 10.5

```
all 0 888:1
open data m-ibmspln.txt
data(org=obs) / r1 r2
set h1 = 45.0
set h2 = 31.0
set rho = 0.8
nonlin  a0 a1 b1 f1 a00 a11 b11 d11 f11 c1 c2 p1 p3 q0 q1 q2
frml a1t = r1(t)-c1-p1*r1(t-1)-p3*r2(t-2)
frml a2t = r2(t)-c2
frml gvar1 = a0+a1*a1t(t-1)**2+b1*h1(t-1)+f1*h2(t-1)
frml gvar2 = a00+a11*a2t(t-1)**2+b11*h2(t-1)+f11*h1(t-1) $
             +d11*a1t(t-1)**2
frml rh1 = q0 + q1*rho(t-1) $
 + q2*a1t(t-1)*a2t(t-1)/sqrt(h1(t-1)*h2(t-1))
frml rh = exp(rh1(t))/(1+exp(rh1(t)))
frml gdet = -0.5*(log(h1(t)=gvar1(t))+log(h2(t)=gvar2(t)) $
     +log(1.0-(rho(t)=rh(t))**2))
frml gln = gdet(t)-0.5/(1.0-rho(t)**2)*((a1t(t)**2/h1(t)) $
```

```
  +(a2t(t)**2/h2(t))-2*rho(t)*a1t(t)*a2t(t)/sqrt(h1(t)*h2(t)))
smpl 4 888
compute c1 = 1.4, c2 = 0.7, p1 = 0.1,  p3 = -0.1
compute a0 = 2.95, a1 = 0.08, b1 = 0.87, f1 = -.03
compute a00 = 2.05, a11 = 0.05
compute  b11 = 0.92, f11=-.06, d11=.04, q0 = -2.0
compute  q1 = 3.0, q2 = 0.1
nlpar(criterion=value,cvcrit=0.00001)
maximize(method=bhhh,recursive,iterations=150) gln
set fv1 = gvar1(t)
set resi1 = a1t(t)/sqrt(fv1(t))
set residsq = resi1(t)*resi1(t)
* Checking standardized residuals *
cor(qstats,number=16,span=4) resi1
* Checking squared standardized residuals *
cor(qstats,number=16,span=4) residsq
set fv2 = gvar2(t)
set resi2 = a2t(t)/sqrt(fv2(t))
set residsq = resi2(t)*resi2(t)
* Checking standardized residuals *
cor(qstats,number=16,span=4) resi2
* Checking squared standardized residuals *
cor(qstats,number=16,span=4) residsq
* Last few observations needed for computing forecasts *
set rhohat = rho(t)
set shock1 = a1t(t)
set shock2 = a2t(t)
print 885  888 shock1 shock2 fv1 fv2 rhohat
```

Estimation of the Time-Varying Coefficient Model in Example 10.5 Using Cholesky Decomposition

```
all 0 888:1
open data m-ibmspln.txt
data(org=obs) / r1 r2
set h1 = 45.0
set h2 = 20.0
set q = 0.8
nonlin  a0 a1 b1 a00 a11 b11 d11 f11 c1 c2 p1 p3 t0 t1 t2
frml a1t = r1(t)-c1-p1*r1(t-1)-p3*r2(t-2)
frml a2t = r2(t)-c2
frml v1 = a0+a1*a1t(t-1)**2+b1*h1(t-1)
frml qt = t0 + t1*q(t-1) + t2*a2t(t-1)
frml bt = a2t(t) - (q(t)=qt(t))*a1t(t)
frml v2 = a00+a11*bt(t-1)**2+b11*h2(t-1)+f11*h1(t-1) $
         +d11*a1t(t-1)**2
frml gdet = -0.5*(log(h1(t) = v1(t))+ log(h2(t)=v2(t)))
frml garchln = gdet-0.5*(a1t(t)**2/h1(t)+bt(t)**2/h2(t))
smpl 5 888
```

```
compute c1 = 1.4, c2 = 0.7, p1 = 0.1, p3 = -0.1
compute a0 = 1.0, a1 = 0.08, b1 = 0.87
compute a00 = 2.0, a11 = 0.05, b11 = 0.8
compute  d11=.04, f11=-.06, t0 =0.2, t1 = 0.1, t2 = 0.1
nlpar(criterion=value,cvcrit=0.00001)
maximize(method=bhhh,recursive,iterations=150) garchln
set fv1 = v1(t)
set resi1 = a1t(t)/sqrt(fv1(t))
set residsq = resi1(t)*resi1(t)
* Checking standardized residuals *
cor(qstats,number=16,span=4) resi1
* Checking squared standardized residuals *
cor(qstats,number=16,span=4) residsq
set fv2 = v2(t)+qt(t)**2*v1(t)
set resi2 = a2t(t)/sqrt(fv2(t))
set residsq = resi2(t)*resi2(t)
* Checking standardized residuals *
cor(qstats,number=16,span=4) resi2
* Checking squared standardized residuals *
cor(qstats,number=16,span=4) residsq
* Last few observations needed for forecasts *
set rhohat = qt(t)*sqrt(v1(t)/fv2(t))
set shock1 = a1t(t)
set shock2 = a2t(t)
set g22 = v2(t)
set q21 = qt(t)
set b2t = bt(t)
print 885  888 shock1 shock2 fv1 fv2 rhohat g22 q21 b2t
```

Estimation of Three-Dimensional Time-Varying Correlation Volatility Model in Example 10.7 Using Cholesky Decomposition

Initial estimates are obtained by a sequential modeling procedure.

```
all 0 2275:1
open data d-cscointc.txt
data(org=obs) / r1 r2 r3
set h1 = 1.0
set h2 = 4.0
set h3 = 3.0
set q21 = 0.8
set q31 = 0.3
set q32 = 0.3
nonlin  c1 c2 c3 p3 p21 p22 p31 p33 a0 a1 a2 t0 t1 t2 b0 b1 $
        b2 u0 u1 u2 w0 w1 w2 d0 d1 d2 d5
frml a1t = r1(t)-c1-p3*r1(t-3)
frml a2t = r2(t)-c2-p21*r1(t-2)-p22*r2(t-2)
frml a3t = r3(t)-c3-p31*r1(t-1)-p33*r3(t-1)
frml v1 = a0+a1*a1t(t-1)**2+a2*h1(t-1)
```

```
frml q1t = t0 + t1*q21(t-1) + t2*a2t(t-1)
frml bt = a2t(t) - (q21(t)=q1t(t))*a1t(t)
frml v2 = b0+b1*bt(t-1)**2+b2*h2(t-1)
frml q2t = u0 + u1*q31(t-1) + u2*a3t(t-1)
frml q3t = w0 + w1*q32(t-1) + w2*a2t(t-1)
frml b1t = a3t(t)-(q31(t)=q2t(t))*a1t(t)-(q32(t)=q3t(t))*bt(t)
frml v3 = d0+d1*b1t(t-1)**2+d2*h3(t-1)+d5*h2(t-1)
frml gdet = -0.5*(log(h1(t) = v1(t))+ log(h2(t)=v2(t)) $
            +log(h3(t)=v3(t)))
frml garchln = gdet-0.5*(a1t(t)**2/h1(t)+bt(t)**2/h2(t) $
            +b1t(t)**2/h3(t))
smpl 8 2275
compute c1 = 0.07, c2 = 0.33, c3 = 0.19, p1 = 0.1, p3 = -0.04
compute p21 =0.2, p22 = -0.1, p31 = -0.26, p33 = 0.06
compute a0 = .01, a1 = 0.05, a2 = 0.94
compute t0 = 0.28, t1 =0.82, t2 = -0.035
compute b0 = .17, b1 = 0.08, b2 = 0.89
compute u0= 0.04, u1 = 0.97, u2 = 0.01
compute  w0 =0.006, w1=0.98, w2=0.004
compute d0 =1.38, d1 = 0.06, d2 = 0.64, d5 = -0.027
nlpar(criterion=value,cvcrit=0.00001)
maximize(method=bhhh,recursive,iterations=250) garchln
set fv1 = v1(t)
set resi1 = a1t(t)/sqrt(fv1(t))
set residsq = resi1(t)*resi1(t)
* Checking standardized residuals *
cor(qstats,number=12,span=4) resi1
* Checking squared standardized residuals *
cor(qstats,number=12,span=4) residsq
set fv2 = v2(t)+q1t(t)**2*v1(t)
set resi2 = a2t(t)/sqrt(fv2(t))
set residsq = resi2(t)*resi2(t)
* Checking standardized residuals *
cor(qstats,number=12,span=4) resi2
* Checking squared standardized residuals *
cor(qstats,number=12,span=4) residsq
set fv3 = v3(t)+q2t(t)**2*v1(t)+q3t(t)**2*v2(t)
set resi3 = a3t(t)/sqrt(fv3(t))
set residsq = resi3(t)*resi3(t)
* Checking standardized residuals *
cor(qstats,number=12,span=4) resi3
* Checking squared standardized residuals *
cor(qstats,number=12,span=4) residsq
* print standardized residuals and correlation-coefficients
set rho21 = q1t(t)*sqrt(v1(t)/fv2(t))
set rho31 = q2t(t)*sqrt(v1(t)/fv3(t))
set rho32 = (q2t(t)*q1t(t)*v1(t) $
            +q3t(t)*v2(t))/sqrt(fv2(t)*fv3(t))
print 10  2275 resi1 resi2 resi3
```

```
print 10 2275 rho21 rho31 rho32
print 10 2275 fv1 fv2 fv3
```

EXERCISES

10.1. Consider the monthly simple returns, including dividends, of IBM stock, Hewlett-Packard (HPQ) stock, and the S&P composite index from January 1962 to December 2008 for 564 observations. The returns are in the file `m-ibmhpqsp6208.txt`. Transform into log returns in percentages. Use the exponentially weighted moving-average method to obtain a multivariate volatility series for the three return series. What is the estimated λ? Plot the three volatility series.

10.2. Focus on the monthly log returns of IBM and HPQ stocks from January 1962 to December 2008. Fit a DVEC(1,1) model to the bivariate return series. Is the model adequate? Plot the fitted volatility series and the time-varying correlations.

10.3. Focus on the monthly log returns of the S&P composite index and HPQ stock. Build a BEKK model for the bivariate series. What is the fitted model? Plot the fitted volatility series and the time-varying correlations.

10.4. Build a constant-correlation volatility model for the three monthly log returns of IBM stock, HPQ stock, and S&P composite index. Write down the fitted model. Is the model adequate? Why?

10.5. The file `m-geibmsp2608.txt` contains the monthly simple returns of General Electric stock, IBM stock, and the S&P composite index from January 1926 to December 2008. The returns include dividends. Transform into log returns in percentages. Focus on the monthly log returns in percentages of GE stock and the S&P 500 index. Build a constant-correlation GARCH model for the bivariate series. Check the adequacy of the fitted model, and obtain the 1-step-ahead forecast of the covariance matrix at the forecast origin December 2008.

10.6. Again, consider the monthly log returns of GE, IBM, and S&P composite index from January 1926 to December 2008. Build a dynamic correlation model for the three-dimensional series. For simplicity, use the sample correlation matrix for $\boldsymbol{\rho}$ in Eq. (10.32).

10.7. The file `m-spibmge.txt` contains the monthly log returns in percentages of the S&P composite index, IBM stock, and GE stock from January 1926 to December 1999. Focus on GE stock and the S&P 500 index. Build a time-varying correlation GARCH model for the bivariate series using a logistic function for the correlation coefficient. Check the adequacy of the fitted model, and obtain the 1-step-ahead forecast of the covariance matrix at the forecast origin December 1999.

10.8. Focus on the monthly log returns in percentages of GE stock and the S&P 500 index from January 1926 to December 1999. Build a time-varying correlation GARCH model for the bivariate series using the Cholesky decomposition. Check the adequacy of the fitted model, and obtain the 1-step-ahead forecast of the covariance matrix at the forecast origin December 1999. Compare the model with the other model built in the previous exercise.

10.9. Consider the three-dimensional return series of the previous exercise jointly. Build a multivariate time-varying volatility model for the data, using the Cholesky decomposition. Discuss the implications of the model and compute the 1-step-ahead volatility forecast at the forecast origin $t = 888$.

10.10. An investor is interested in daily value at risk of his position on holding long \$0.5 million of Dell stock and \$1 million of Cisco Systems stock. Use 5% critical values and the daily log returns from February 20, 1990, to December 31, 1999, to do the calculation. The data are in the file `d-dellcsco9099.txt`. Apply the three approaches to volatility modeling in Section 10.7 and compare the results.

REFERENCES

Bauwens, L., Laurent, S., and Rombouts, J. V. K. (2004). Multivariate GARCH models: A survey. *Journal of Applied Econometrics* **21**:79–109.

Bollerslev, T. (1990). Modeling the coherence in short-term nominal exchange rates: A multivariate generalized ARCH approach. *Review of Economics and Statistics* **72**:498–505.

Bollerslev, T., Engle, R. F., and Wooldridge, J. M. (1988). A capital-asset pricing model with time-varying covariances. *Journal of Political Economy* **96**:116–131.

Chib, S., Nardari, F. and Shephard, N. (1999). Analysis of high dimensional multivariate stochastic volatility models. Working paper, Washington University, St. Louis.

Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate GARCH models. *Journal of Business and Economic Statistics* **20**:339–350.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory* **11**:122–150.

Harvey, A, Ruiz, E., and Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies* **62**:247–264.

Li, W. K. (2004). *Diagnostic Checks in Time Series*. Chapman and Hall, London.

Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979). *Multivariate Analysis*. Academic, New York.

Pourahmadi, M. (1999). Joint mean-covariance models with applications to longitudinal data: Unconstrained parameterization. *Biometrika* **86**:677–690.

Tsay, R. S. (2006). Multivariate volatility models in H. C. Ho, C. K. Ing, and T. L. Lai (eds.). *Time Series and Related Topics* in memory of C.Z. Wei, Lecture Notes Monograph Series, Institute of Mathematical Statistics. Beachwood, Ohio.

Tse, Y. K. (2000). A test for constant correlations in a multivariate GARCH model. *Journal of Econometrics* **98**:107–127.

Tse, Y. K. and Tsui, A. K. C. (2002). A multivariate GARCH model with time-varying correlations. *Journal of Business & Economic Statistics* **20**:351–362.