Covariance and Correlation Models

FNCE 5321

Hang Bai

Overview

- This Chapter models dynamic covariance and correlation, which along with dynamic volatility models is used to construct covariance matrices.

Portfolio Variance and Covariance

- Consider a portfolio of n securities
- The return on the portfolio on date *t+1* is

$$r_{PF,t+1} = \sum_{i=1}^{n} w_{i,t} r_{i,t+1}$$

- The sum is taken over the *n* securities in portfolio
- $w_{i,t}$ denotes the relative weight of security *i* at the end of day *t*.

Portfolio Variance and Covariance

- The variance of the portfolio is

$$\sigma^2_{PF,t+1} = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,t} w_{j,t} \sigma_{ij,t+1}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,t} w_{j,t} \sigma_{i,t+1} \sigma_{j,t+1} \rho_{ij,t+1}$$

- Where $\sigma_{ij,t+1}$ and $\rho_{ij,t+1}$ are covariance and correlation respectively between security *i* and *j* on day *t+1*
- Note $\sigma_{ij,t+1} = \sigma_{ji,t+1}$ and $\rho_{ij,t+1} = \rho_{ji,t+1}$ for all i and j
- Also we have $\rho_{ii,t+1} = 1$ and $\sigma_{ii,t+1} = \sigma^2_{i,t+1}$ for all i

Portfolio Variance and Covariance

- Using vector notation we can write:

$$\sigma^2_{PF,t+1} = w_t' \Sigma_{t+1} w_t$$

- where w_t is the n by 1 vector of portfolio weights and Σ_{t+1} is the n by n covariance matrix of returns
- In the case where $n = 2$ we have

$$\begin{aligned} \sigma^2_{PF,t+1} &= \begin{bmatrix} w_{1,t} & w_{2,t} \end{bmatrix} \begin{bmatrix} \sigma^2_{1,t+1} & \sigma_{12,t+1} \\ \sigma_{12,t+1} & \sigma^2_{2,t+1} \end{bmatrix} \begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix} \\ &= w^2_{1,t}\sigma^2_{1,t+1} + w^2_{2,t}\sigma^2_{2,t+1} + 2w_{1,t}w_{2,t}\sigma_{12,t+1} \end{aligned}$$

Portfolio Variance and Covariance

- If we assume assets are multivariate normally distributed, then the portfolio is normally distributed and we can write,

$$VaR_{t+1}^{p} = -\sigma_{PF,t+1}\Phi_p^{-1}$$

$$ES_{t+1}^{p} = \sigma_{PF,t+1}\frac{\phi\left(\Phi_p^{-1}\right)}{p}$$

- Note: if we have *n* assets in the portfolio then we have to model *n*(n-1)/2 different correlations
- So if *n* is 100, then we'll have 4950 correlations to model, a daunting task that forces us to find methods that can easily handle large-dimensional portfolios

Exposure Mapping

- One way to reduce dimensionality of portfolio variance is to impose a factor structure using observed market returns as factors.
- In the extreme case we assume that portfolio return is the market return plus a portfolio specific risk term:

$$r_{PF,t+1} = r_{Mkt,t+1} + \varepsilon_{t+1}$$

- where we assume that the idiosyncratic risk term, ε_{t+1}, is independent of the market return and has constant variance.
- The portfolio variance in this case is

$$\sigma^2_{PF,t+1} = \sigma^2_{Mkt,t+1} + \sigma^2_{\varepsilon}$$

Exposure Mapping

- In a well diversified stock portfolio, for example, we can assume that the variance of the portfolio equals that of the S&P 500 market index.
- In this case, only one volatility needs to be modelled, and no correlation modelling is necessary.
- This is referred to as index mapping and written as:

$$\sigma^2_{PF,t+1} \approx \sigma^2_{Mkt,t+1}$$

- The 1-day *VaR* assuming normality is

$$VaR^p_{t+1} = -\sigma_{Mkt,t+1}\Phi_p^{-1}$$

Exposure Mapping

- In general, portfolios that contain systematic risk have

$$r_{PF,t+1} = \beta r_{Mkt,t+1} + \varepsilon_{t+1}$$

- So that $$\sigma^2_{PF,t+1} = \beta^2 \sigma^2_{Mkt,t+1} + \sigma^2_{\varepsilon}$$

- If the portfolio is well diversified then the portfolio-specific risk can be ignored, and we can pose a linear relationship between the portfolio and the market index and use the beta mapping as

$$\sigma^2_{PF,t+1} \approx \beta^2 \sigma^2_{Mkt,t+1}$$

- Here only an estimate of β is necessary and no further correlation modelling is needed.

Exposure Mapping

- The risk manager of a large-scale portfolio may consider risk coming from a reasonable number of factors n_F where $n_F << n$ so that we have many fewer risk factors than assets
- Let us assume that we need 10 factors.
- We can write the 10-factor return model as

$$r_{PF,t+1} = \beta_1 r_{F1,t+1} + \cdots + \beta_{10} r_{F10,t+1} + \varepsilon_{t+1}$$

- here ε_{t+1} is assumed to be independent of risk factors
- In this case, it makes sense to model the variances and correlations of these risk factors and assign exposures to each factor to get the portfolio variance.

Exposure Mapping

- The portfolio variance in this general factor structure can be written

$$\sigma^2_{PF,t+1} = \beta'_F \Sigma^F_{t+1} \beta_F + \sigma^2_\varepsilon$$

- where β_F is a vector of exposures to each risk factor and where Σ^F_{t+1} is the covariance matrix of the returns from the risk factors.
- Again, if the factor model explains a large part of the portfolio return variation, then we can assume that

$$\sigma^2_{PF,t+1} \approx \beta'_F \Sigma^F_{t+1} \beta_F$$

GARCH Conditional Covariance

- Suppose a portfolio contains *n* assets or factors. An *n*-dimensional covariance matrix must be estimated where *n* may be a large number.
- Now, we turn to various methods for constructing the covariance matrix directly, without first modeling the correlations.

  The simplest way to model time varying covariances is to rely on plain rolling averages.
- For the covariance between assets *i and j*, we can simply estimate

$$\sigma_{ij,t+1} = \frac{1}{m} \sum_{\tau=1}^{m} R_{i,t+1-\tau} R_{j,t+1-\tau}$$

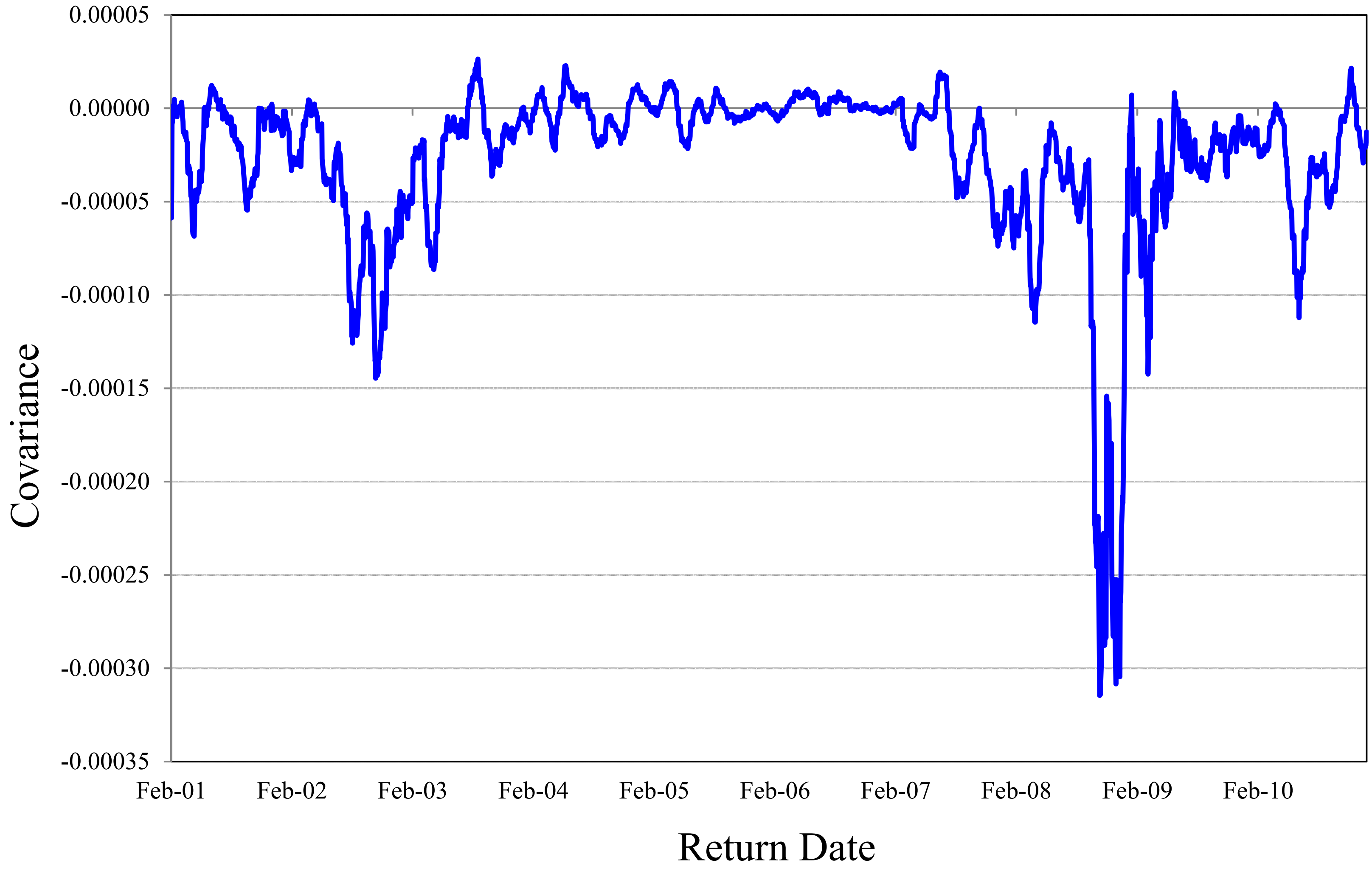
GARCH Conditional Covariance

- where m is the number of days used in the moving estimation window

- This estimate is very easy to construct but it is not satisfactory due to dependence on choice of m and equal weighting put on past cross products of returns

- We assume that the average expected return on each asset is zero

Figure 7.1: Rolling Covariance between S&P 500 and 10-Year Treasury Note Index



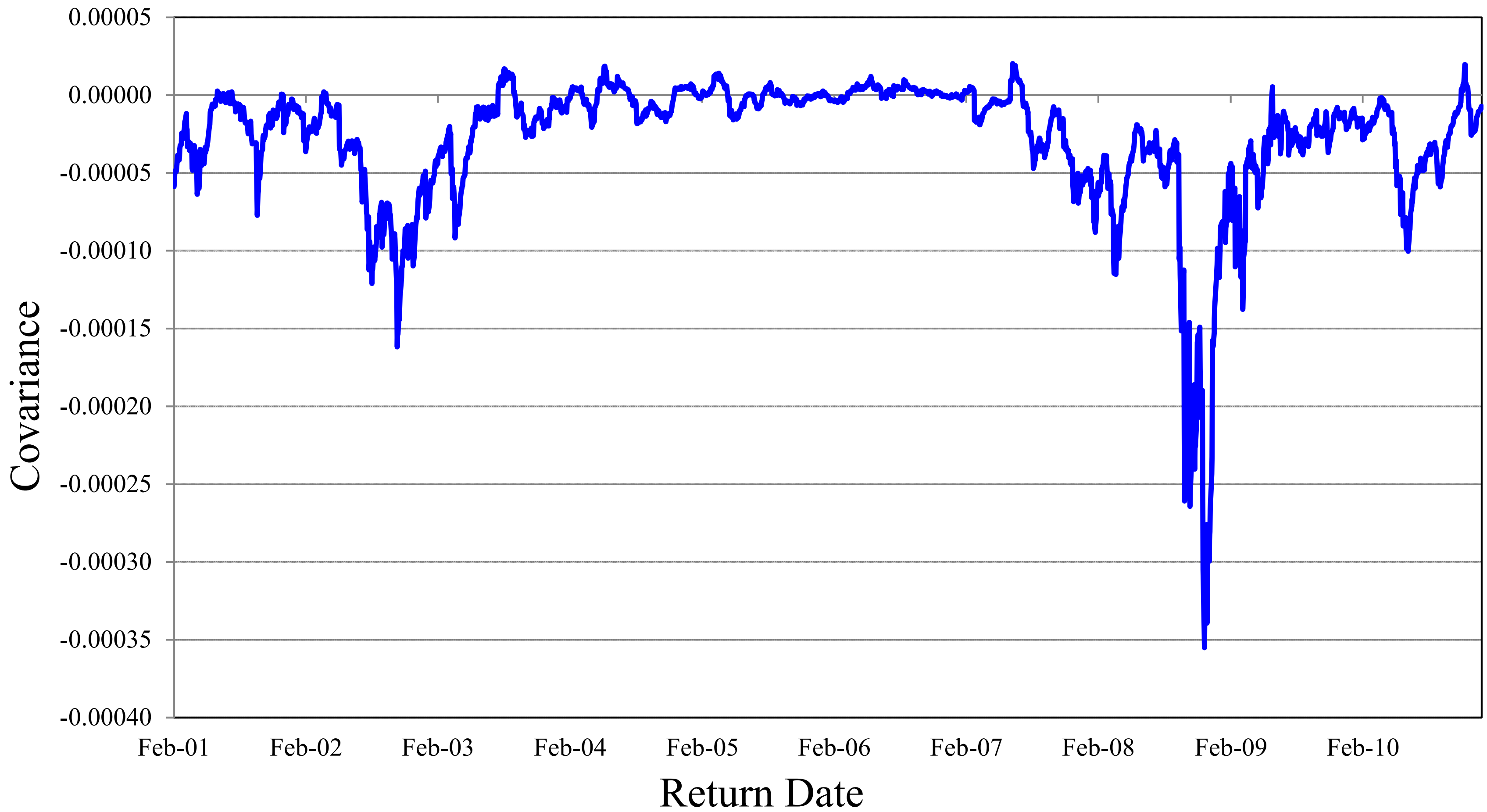
GARCH Conditional Covariance

- To avoid equal weighting we can use a simple exponential smoother model on the covariances, and let

$$\sigma_{ij,t+1} = (1-\lambda)\, R_{i,t} R_{j,t} + \lambda \sigma_{ij,t}$$

- where λ=0.94 as it was for the corresponding volatility model in the previous chapters

Figure 7.2: Exponentially Smoothed Covariance between S&P 500 and 10-year Treasury Note Index



GARCH Conditional Covariance

- The caveats which applied to the exponential smoother volatility model, apply to the exponential smoother covariance model as well
- The restriction that the coefficient (1-λ) on the cross product of returns $(R_{i,t}R_{j,t})$ and coefficient λ on the past covariance $(\sigma_{ij,t})$ sum to one is not desirable.
- The restriction implies that there is no mean-reversion in covariance
- If tomorrow's forecasted covariance is high then it will remain high for all future horizons, rather than revert back to its mean.

GARCH Conditional Covariance

- We can instead consider models with mean-reversion in covariance.
- For example, a GARCH-style specification for covariance would be

$$\sigma_{ij,t+1} = \omega_{ij} + \alpha R_{i,t} R_{j,t} + \beta \sigma_{ij,t}$$

- which will tend to revert to its long-run average covariance, which equals

$$\sigma_{ij} = \omega_{ij}/(1 - \alpha - \beta)$$

GARCH Conditional Covariance

- Note that we have not allowed for the persistence parameters λ, α and β to vary across pairs of securities in the covariance models
- This is done to guarantee that the portfolio variance will be positive regardless of the portfolio weights, w_t.
- A covariance matrix Σ_{t+1} is internally consistent if for all possible vectors w_t of portfolio weights we have

$$w_t' \Sigma_{t+1} w_t \geq 0$$

- Thus covariance matrix is positive semidefinite. It is ensured by estimating volatilities and covariances in an internally consistent fashion.

GARCH Conditional Covariance

- For example, relying on exponential smoothing using the same λ for every volatility and every covariance will work.
- Similarly, using a GARCH(1,1) model with α and β identical across variances and covariances will work as well.
- Unfortunately, it is not clear that the persistence parameters λ, α and β should be the same for all variances and covariance
- We therefore next consider methods that are not subject to this restriction

Dynamic Conditional Correlation

- Now we will model correlation rather than covariance
- Variances and covariance are restricted by the same persistence parameters
- Covariance is a confluence of correlation and variance. Could be time varying just from variances.
- Correlations increase during financial turmoil and thereby increase risk even further
- Therefore, modeling correlation dynamics is crucial to a risk manager
- Correlation is defined from covariance and volatility by

$$\rho_{ij,t+1} = \sigma_{ij,t+1}/(\sigma_{i,t+1}\sigma_{j,t+1})$$

Dynamic Conditional Correlation

- If we have the RiskMetrics model, then

$$\sigma_{ij,t+1} = (1-\lambda)\, R_{i,t} R_{j,t} + \lambda \sigma_{ij,t}, \text{ for all } i, j$$

- and then we get the implied dynamic correlations

$$\rho_{ij,t+1} = \frac{(1-\lambda)\, R_{i,t} R_{j,t} + \lambda \sigma_{ij,t}}{\sqrt{((1-\lambda)\, R_{i,t}^2 + \lambda \sigma_{i,t}^2)((1-\lambda)\, R_{j,t}^2 + \lambda \sigma_{j,t}^2)}}$$

- which isn't particularly intuitive, we therefore consider models where dynamic correlation is modeled directly
- The definition of correlation can be rearranged to provide the decomposition of covariance into volatility and correlation

$$\sigma_{ij,t+1} = \sigma_{i,t+1} \sigma_{j,t+1} \rho_{ij,t+1}$$

Dynamic Conditional Correlation

- In matrix notation, we can write

$$\Sigma_{t+1} = D_{t+1} \Upsilon_{t+1} D_{t+1}$$

- where D_{t+1} is a matrix of standard deviations, $\sigma_{i,t+1}$, on the ith diagonal and zero everywhere else, and where Υ_{t+1} is a matrix of correlations, $\rho_{ij,t+1}$, with ones on the diagonal.
- In the two-asset case, we have

$$\begin{aligned}\Sigma_{t+1} &= \begin{bmatrix} \sigma^2_{1,t+1} & \sigma_{12,t+1} \\ \sigma_{12,t+1} & \sigma^2_{2,t+1} \end{bmatrix} \\ &= \begin{bmatrix} \sigma_{1,t+1} & 0 \\ 0 & \sigma_{2,t+1} \end{bmatrix} \begin{bmatrix} 1 & \rho_{12,t+1} \\ \rho_{12,t+1} & 1 \end{bmatrix} \begin{bmatrix} \sigma_{1,t+1} & 0 \\ 0 & \sigma_{2,t+1} \end{bmatrix}\end{aligned}$$

Dynamic Conditional Correlation

- We will consider the volatilities of each asset to already have been estimated through GARCH or one of the other methods considered in Chapter 4 or 5
- We can then standardize each return by its dynamic standard deviation to get the standardized returns,

$$z_{i,t+1} = R_{i,t+1}/\sigma_{i,t+1} \text{ for all } i$$

- By dividing the returns by their conditional standard deviation, we create variables, $z_{i,t+1}$, $i = 1,2,\ldots,n$, which all have a conditional standard deviation of one

Dynamic Conditional Correlation

- The conditional covariance of the $z_{i,t+1}$ variables equals the conditional correlation of the raw returns as can be seen from

$$\begin{aligned} E_t\left(z_{i,t+1}z_{j,t+1}\right) &= E_t\left(\left(R_{i,t+1}/\sigma_{i,t+1}\right)\left(R_{j,t+1}/\sigma_{j,t+1}\right)\right) \\ &= E_t\left(R_{i,t+1}R_{j,t+1}\right)/\left(\sigma_{i,t+1}\sigma_{j,t+1}\right) \\ &= \sigma_{ij,t+1}/\left(\sigma_{i,t+1}\sigma_{j,t+1}\right) \\ &= \rho_{ij,t+1}, \text{ for all } i,j \end{aligned}$$

- Thus, modeling the conditional correlation of the raw returns is equivalent to modeling the conditional covariance of the standardized returns

Exponential Smoother correlations

- First we consider simple exponential smoothing correlation models.
- Let the correlation dynamics be driven by $q_{ij,t+1}$, which gets updated by the cross product of the standardized returns $z_{i,t}$ and $z_{j,t}$ as in

$$q_{ij,t+1} = (1-\lambda)\left(z_{i,t}z_{j,t}\right) + \lambda q_{ij,t}, \text{ for all } i, j$$

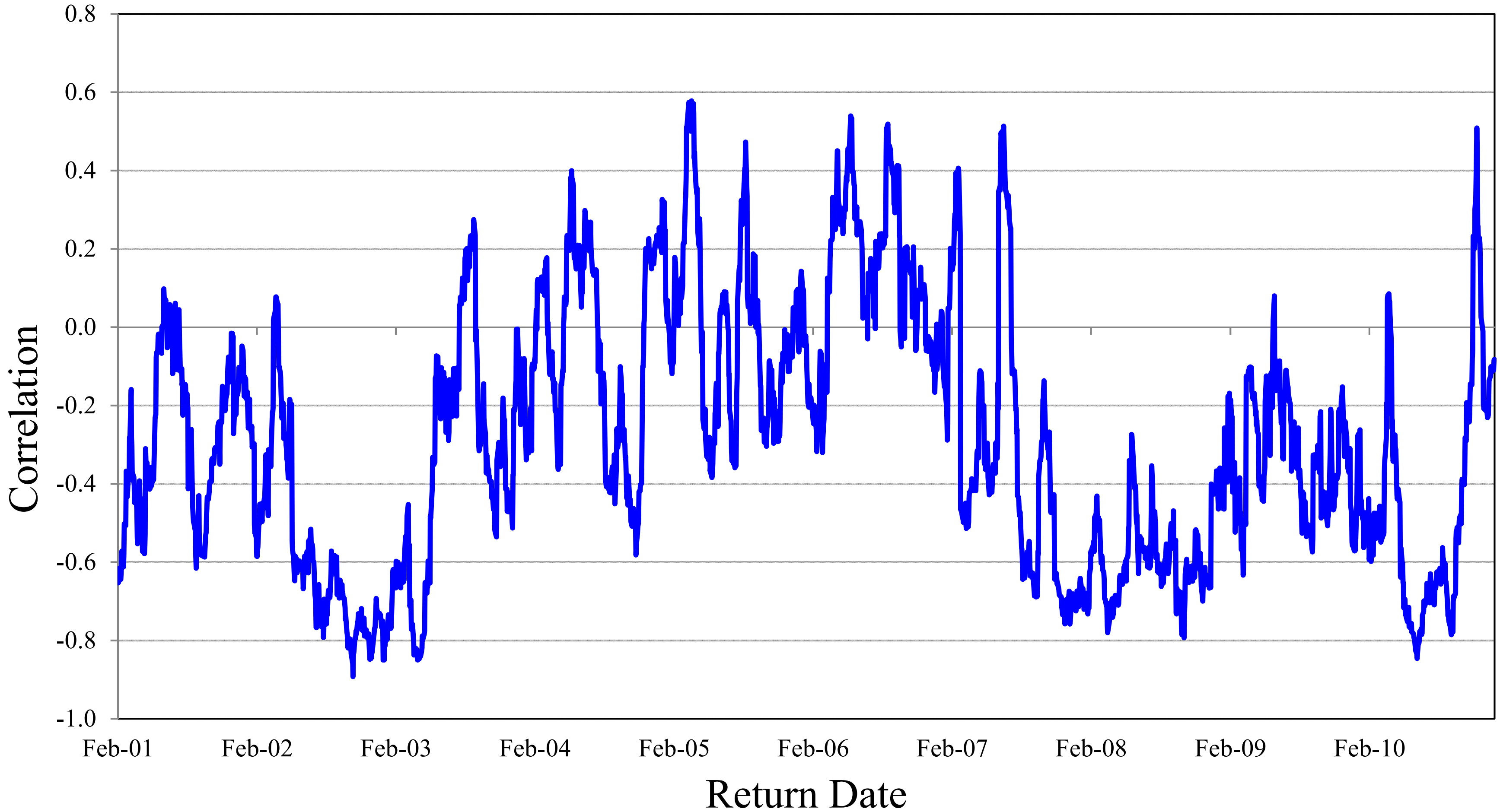
- The conditional correlation can now be obtained by normalizing the $q_{ij,t+1}$ variable as in

$$\rho_{ij,t+1} = \frac{q_{ij,t+1}}{\sqrt{q_{ii,t+1}q_{jj,t+1}}}$$

- Now we need to do the normalization to ensure that

$-1 < \rho_{ij,t+1} < +1$ on each day.

Figure 7.3: Exponentially Smoothed Correlation between S&P 500 and 10-year Treasury Note Index



Mean-Reverting Correlation

- Consider a generalization of exponential smoothing correlation model, which allows for correlations to revert to a long-run average correlation $\rho_{ij} = E\,[z_{i,t}z_{j,t}]$
- GARCH(1,1) type specification:

$$q_{ij,t+1} = \rho_{ij} + \alpha\left(z_{i,t}z_{j,t} - \rho_{ij}\right) + \beta\left(q_{ij,t} - \rho_{ij}\right)$$

- If we rely on correlation targeting, and set

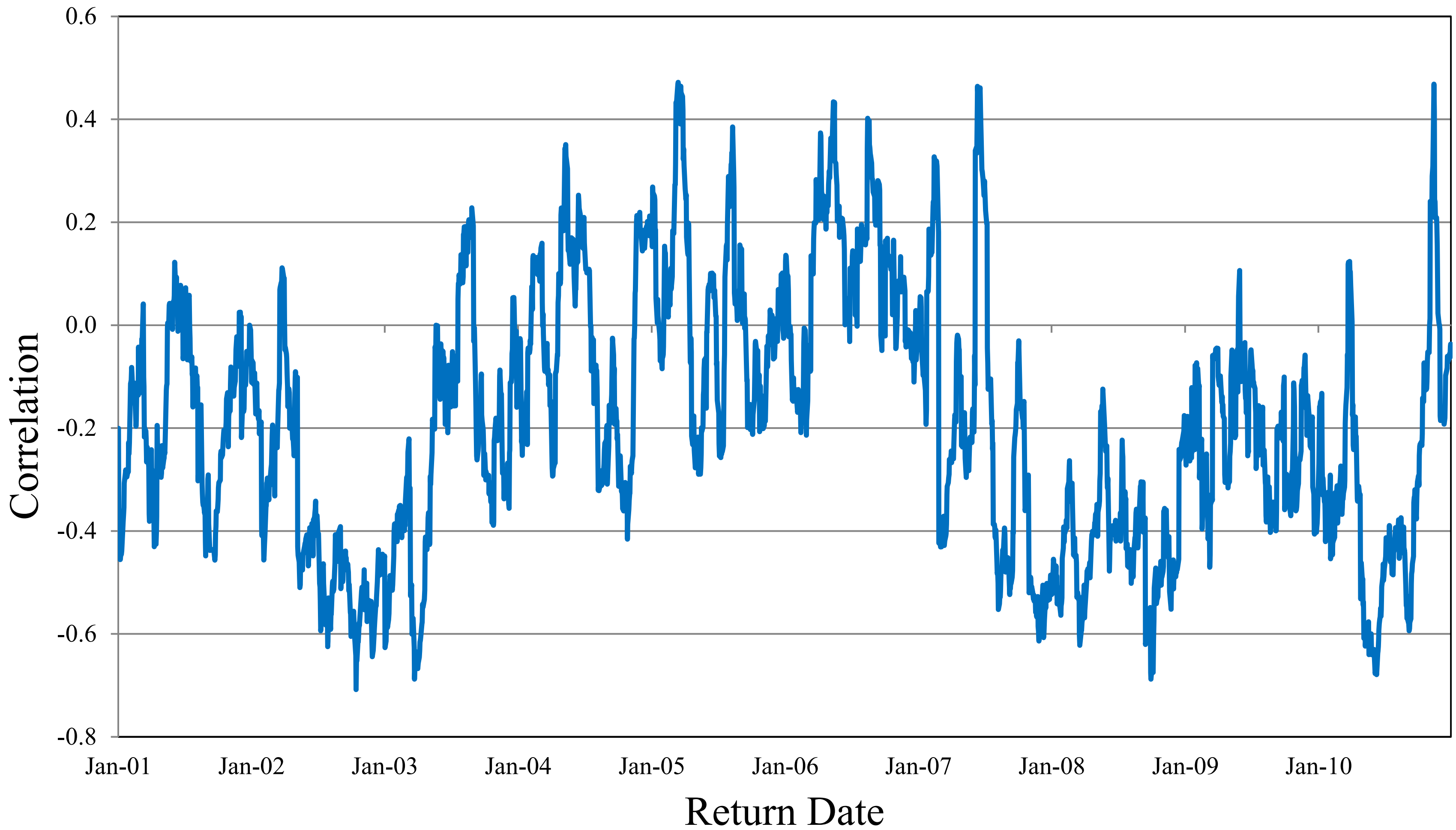
$\bar{\rho}_{ij} = \frac{1}{T}\sum_{t=1}^{T} z_{i,t}z_{j,t}$, then we have

$$q_{ij,t+1} = \bar{\rho}_{ij} + \alpha\left(z_{i,t}z_{j,t} - \bar{\rho}_{ij}\right) + \beta\left(q_{ij,t} - \bar{\rho}_{ij}\right)$$

- Again we have to normalize to get the conditional correlations

$$\rho_{ij,t+1} = \frac{q_{ij,t+1}}{\sqrt{q_{ii,t+1}q_{jj,t+1}}}$$

Figure 7.4: Mean-Reverting Correlation between S&P 500 and 10-Year Treasury Note



Mean-Reverting Correlation

- Note that correlation persistence parameters α and β are common across *i* and *j*.
- It does not imply that the level of correlations at any time are the same across pairs of assets.
- It does not imply that the persistence in correlation is the same persistence in volatility.
- The model does imply that the persistence in correlation is constant across assets
- Fig 7.4 shows the GARCH(1,1) correlations for the S&P500 and 10-year treasury note example.

Mean-Reverting Correlation

- We write the models in matrix notation as

$$Q_{t+1} = (1-\lambda)\left(z_t z_t'\right) + \lambda Q_t$$

- For the exponential smoother, and for the mean-reverting DCC, we can write

$$Q_{t+1} = E\left[z_t z_t'\right](1-\alpha-\beta) + \alpha\left(z_t z_t'\right) + \beta Q_t$$

- In two-asset case for mean-reverting model, we have

$$\begin{aligned} Q_{t+1} &= \begin{bmatrix} q_{11,t+1} & q_{12,t+1} \\ q_{12,t+1} & q_{22,t+1} \end{bmatrix} \\ &= \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix}(1-\alpha-\beta) + \alpha \begin{bmatrix} z_{1,t}^2 & z_{1,t}z_{2,t} \\ z_{1,t}z_{2,t} & z_{2,t}^2 \end{bmatrix} + \beta \begin{bmatrix} q_{11,t} & q_{12,t} \\ q_{12,t} & q_{22,t} \end{bmatrix} \end{aligned}$$

Mean-Reverting Correlation

- where ρ_{12} is the unconditional correlation between the two assets, which can be estimated as

$$\bar{\rho}_{12} = \frac{1}{T}\sum_{t=1}^{T} z_{1,t} z_{2,t}$$

- An important feature of these models is that the matrix Q_{t+1} is positive semi-definite as it is a weighted average of positive semi-definite and positive definite matrices.
- This will ensure that the correlation matrix Υ_{t+1} and the covariance matrix Σ_{t+1} will be positive semi-definite as required.

Mean-Reverting Correlation

- A key advantage of this model is that we can estimate the parameters in a sequential fashion.
- First all the individual variances are estimated one by one using one of the methods from Chapter 4 or 5
- Second, the returns are standardized and the unconditional correlation matrix is estimated.
- Third, the correlation persistence parameters α and β are estimated.
- The key issue is that only very few parameters are estimated simultaneously using numerical optimization. This makes the dynamic correlation models considered extremely tractable for risk management of large portfolios.

Bivariate Quasi-Maximum Likelihood Estimation

- We begin by analyzing a portfolio consisting of only two assets
- In this case, we can use the bivariate normal distribution function for $z_{1,t}$ and $z_{2,t}$ to write the log likelihood as

$$\ln\left(L_{c,12}\right) = -\frac{1}{2}\sum_{t=1}^{T}\left(\ln\left(1-\rho_{12,t}^{2}\right) + \frac{\left(z_{1,t}^{2}+z_{2,t}^{2}-2\rho_{12,t}z_{1,t}z_{2,t}\right)}{\left(1-\rho_{12,t}^{2}\right)}\right)$$

Bivariate Quasi-Maximum Likelihood Estimation

- Where $\rho_{12,t}$ is given from the particular correlation model being estimated, and the normalization rule
- In the simple exponential smoother example

$$\rho_{12,t} = \frac{q_{12,t}}{\sqrt{q_{11,t} q_{22,t}}}$$

where

$$\begin{aligned} q_{11,t} &= (1-\lambda)\left(z_{1,t-1}^2\right) + \lambda q_{11,t-1} \\ q_{12,t} &= (1-\lambda)\left(z_{1,t-1} z_{2,t-1}\right) + \lambda q_{12,t-1} \\ q_{22,t} &= (1-\lambda)\left(z_{2,t-1}^2\right) + \lambda q_{22,t-1} \end{aligned}$$

Bivariate Quasi-Maximum Likelihood Estimation

- We find the optimal correlation parameter(s), in this case λ, by maximizing the correlation log-likelihood function, $\ln(L_{c,12})$
- To initialize the dynamics, we set

$$q_{11,0} = 1, q_{22,0} = 1, \text{ and } q_{12,0} = \frac{1}{T}\sum_{t=1}^{T} z_{1,t} z_{2,t}.$$

- Notice that the variables that enter the likelihood are the standardized returns, z_t, and not the original raw returns, R_t themselves.
- We are essentially treating the standardized returns as actual observations here.

Bivariate Quasi-Maximum Likelihood Estimation

- To get efficient estimates, we are forced to rely on a stepwise QMLE method where we first estimate the volatility model for each of the assets and second estimate the correlation models

- This approach gives decent parameter estimates while avoiding numerical optimization in high dimensions.

Bivariate Quasi-Maximum Likelihood Estimation

- In the case of the mean-reverting GARCH correlations we have the same likelihood function and correlation definition but now

$$
\begin{aligned}
q_{11,t+1} &= 1 + \alpha\left(z_{1,t}^2 - 1\right) + \beta\left(q_{11,t} - 1\right) \\
q_{12,t+1} &= \bar{\rho}_{12} + \alpha\left(z_{1,t}z_{2,t} - \bar{\rho}_{12}\right) + \beta\left(q_{12,t} - \bar{\rho}_{12}\right) \\
q_{22,t+1} &= 1 + \alpha\left(z_{2,t}^2 - 1\right) + \beta\left(q_{22,t} - 1\right)
\end{aligned}
$$

- Where $\bar{\rho}_{12}$ can be estimated using

$$\bar{\rho}_{12} = \frac{1}{T}\sum_{t=1}^{T} z_{1,t}z_{2,t}$$

Bivariate Quasi-Maximum Likelihood Estimation

- Therefore we only have to find α and β using numerical optimization

- Again, in order to initialize the dynamics, we set

$$q_{11,0} = 1, q_{22,0} = 1$$ and

$$q_{12,0} = \bar{\rho}_{12} = \frac{1}{T}\sum_{t=1}^{T} z_{1,t} z_{2,t}.$$

Composite Likelihood Estimation in Large Systems

- In a portfolio *with n* assets, we rely on *n*-dimensional normal distribution function to write log likelihood as

$$\ln L_c = -\frac{1}{2}\sum_t \left(\log|\Upsilon_t| + z_t'\Upsilon_t^{-1}z_t\right)$$

- where $|\Upsilon_t|$ denotes the determinant of the correlation matrix, Υ_t
- Maximizing this likelihood can be very cumbersome if *n* is large
- The correlation matrix Υ_t must be inverted on each day and for many possible values of the parameters in the model when doing the numerical search for optimal parameter values

Composite Likelihood Estimation in Large Systems

- When n is large the inversion of Υ_t will be slow and inaccurate causing biases in parameter values
- To solve dimensionality problem, we can maximize the sum of the bivariate likelihoods rather than maximizing the n-dimensional log likelihood

$$\ln\left(CL_c\right) = -\frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j>i}\left(\ln\left(1-\rho_{ij,t}^2\right)+\frac{\left(z_{i,t}^2+z_{j,t}^2-2\rho_{ij,t}z_{i,t}z_{j,t}\right)}{\left(1-\rho_{ij,t}^2\right)}\right)$$

- Computationally this composite likelihood function is much easier to maximize than the likelihood function where the n-dimensional correlation matrix must be inverted numerically

An Asymmetric Correlation Model

- Now we model for a down-market effect in correlation
- This can be achieved using the asymmetric DCC model where

$$Q_{t+1} = (1 - \alpha - \beta) E[z_t z_t'] + \alpha (z_t z_t') + \beta Q_t + \gamma (\eta_t \eta_t' - E[\eta_t \eta_t'])$$

- where the $n_{i,t}$ for asset *i* is defined as the negative part of $z_{i,t}$ as follows

$$\eta_{i,t} = \begin{cases} z_{i,t}, \text{ if } z_{i,t} < 0 \\ 0, \text{ if } z_{i,t} > 0 \end{cases} \quad \text{for all } i$$

An Asymmetric Correlation Model

- Note that γ corresponds to a leverage effect in correlation: When γ is positive then the correlation for asset i and j will increase more when $z_{i,t}$ and $z_{j,t}$ are negative than in any other case.
- If we envision a scatterplot of $z_{i,t}$ and $z_{j,t}$, then $\gamma > 0$ will provide an extra increase in correlation when we observe an observation in the lower left quadrant of the scatterplot.
- This captures a phenomenon often observed in markets for risky assets: Their correlation increases more in down markets ($z_{i,t}$ and $z_{j,t}$ both negative) than in up markets ($z_{i,t}$ and $z_{j,t}$ both positive).

Summary

- For normally distributed returns, the covariance matrix is all that is needed to calculate the *VaR*
- First, we presented simple rolling estimates of covariance, followed by simple exponential smoothing and GARCH models of covariance
- We then discussed the important issue of estimating variances and covariances
- We then presented a simple framework for dynamic correlation modelling