22

The Variance-Covariance Method

The *variance-covariance method* makes use of covariances (volatilities and correlations) of the risk factors and the sensitivities of the portfolio values with respect to these risk factors with the goal of approximating the value at risk. This method leads directly to the final result, i.e., the portfolio's value at risk, based on the properties of the assumed portfolio value's probability distribution; no simulation of market data scenarios is involved. The variance-covariance method utilizes linear approximations of the risk factors themselves throughout the entire calculation, often neglecting the drift as well. In view of Eq. 21.25, we have

$$\delta S_i(t) \approx S_i(t)\,[\mu_i \delta t + \delta Z_i] \approx S_i(t)\delta Z_i \ . \tag{22.1}$$

The main idea characterizing this method, however, is that the portfolio value V is expanded in its Taylor series as a function of its risk factors $S_i, i = 1, \ldots n$, and approximated by breaking off after the first or second order term. Let

$$\mathbf{S}(t) = \begin{pmatrix} S_1(t) \\ \vdots \\ S_n(t) \end{pmatrix}$$

H.-P. Deutsch, M. W. Beinker, *Derivatives and Internal Models*, Finance and Capital Markets Series, https://doi.org/10.1007/978-3-030-22899-6_22

denote the vector of risk factors. The Taylor expansion for the change in portfolio value $\delta V(\mathbf{S})$ up to second order is

$$\begin{aligned}
\delta V(\mathbf{S}(t)) &= V(\mathbf{S}(t) + \delta\mathbf{S}(t)) - V(\mathbf{S}(t)) \\
&\approx \sum_i^n \frac{\partial V}{\partial S_i}\delta S_i(t) + \frac{1}{2}\sum_{i,j}^n \delta S_i(t)\frac{\partial^2 V}{\partial S_i \partial S_j}\delta S_j(t) \qquad (22.2) \\
&= \sum_i^n \Delta_i \delta S_i(t) + \frac{1}{2}\sum_{i,j}^n \delta S_i(t)\Gamma_{ij}\delta S_j(t) \\
&\approx \sum_i^n \widetilde{\Delta}_i\ [\mu_i\delta t + \delta Z_i] + \frac{1}{2}\sum_{i,j}^n [\mu_i\delta t + \delta Z_i]\,\widetilde{\Gamma}_{ij}\ \left[\mu_j\delta t + \delta Z_j\right] \qquad (22.3) \\
&\approx \sum_i^n \widetilde{\Delta}_i\,\delta Z_i + \frac{1}{2}\sum_{i,j}^n \delta Z_i\widetilde{\Gamma}_{ij}\,\delta Z_j\ . \qquad (22.4)
\end{aligned}$$

In Eq. 22.2, the portfolio value has been approximated by its Taylor expansion. In Eq. 22.3, the risk factor changes were linear approximate according to Eq. 22.1, and finally, in Eq. 22.4, the average returns have been neglected. The last line in 22.4 is referred to as the *delta-gamma approximation*. Taking the Taylor expansion up to linear order only is called *delta approximation* correspondingly, resulting in an approximation solely consisting of the first of the two sums appearing in the last equation in 22.4.

The abbreviations Δ_i and Γ_{ij}, as usual, denote the *sensitivities* (at time t) of V with respect to the risk factors

$$\Delta_i := \frac{\partial V}{\partial S_i},\ \Gamma_{ij} := \frac{\partial^2 V}{\partial S_i \partial S_j},\ i, j = 1, \ldots n\ .$$

Note here that the mixed partial derivatives arise in the expression for Γ_{ij}. In the literature, the matrix Γ_{ij} is sometimes called the *Hessian matrix*.

We will see below that the sensitivities usually appear in connection with the current levels $S_i(t)$ and $S_j(t)$. The notation $\widetilde{\Delta}_i$ and $\widetilde{\Gamma}_{ij}$ will be used to denote these sensitivities multiplied by the current levels:

$$\widetilde{\Delta}_i := S_i(t)\frac{\partial V}{\partial S_i},\ \widetilde{\Gamma}_{ij} := S_i(t)S_j(t)\frac{\partial^2 V}{\partial S_i \partial S_j}\ . \tag{22.5}$$

Using $\widetilde{\Delta}_i$ and $\widetilde{\Gamma}_{ij}$ will prove to substantially simplify the notation. Interpreting the $\widetilde{\Delta}_i$ as components of a vector $\widetilde{\mathbf{\Delta}}$, and $\widetilde{\Gamma}_{ij}$ as the elements of a matrix $\widetilde{\mathbf{\Gamma}}$, Eq. 22.4 can be written in vector form as

$$\begin{aligned}\delta V(\mathbf{S}(t)) &= \left(\widetilde{\Delta}_1 \cdots \widetilde{\Delta}_n\right)\begin{pmatrix}\delta Z_1\\ \vdots\\ \delta Z_n\end{pmatrix}\\ &\quad + \frac{1}{2}\left(\delta Z_1 \cdots \delta Z_n\right)\begin{pmatrix}\widetilde{\Gamma}_{1,1} \cdots \widetilde{\Gamma}_{1,n}\\ \vdots \quad \ddots \quad \vdots\\ \widetilde{\Gamma}_{n,1} \cdots \widetilde{\Gamma}_{n,n}\end{pmatrix}\begin{pmatrix}\delta Z_1\\ \vdots\\ \delta Z_n\end{pmatrix}\\ &= \widetilde{\mathbf{\Delta}}^T\delta\mathbf{Z} + \frac{1}{2}\delta\mathbf{Z}^T\widetilde{\mathbf{\Gamma}}\delta\mathbf{Z}\ .\end{aligned}$$

The approximation of $V(\mathbf{S})$ through its Taylor series expansion up to second order is presented in Fig. 22.1 for a *straddle* (a portfolio made up of a call and a put option) on a risk factor S. The figure has been extracted from the Excel workbook Straddle.xlsm available in the download section [50]. We can recognize that the delta-gamma approximation for a simple payoff profile is quite a good approximation. For somewhat more complicated portfolios, however, the delta-gamma approximation fails to be a reasonable representation of the payoff profile. In such cases, we recommend using one of the simulation methods presented in Chap. 23 instead of the Delta-Gamma method. However, the drawback of the simulation approach is of course the significant increase of numerical computations required.

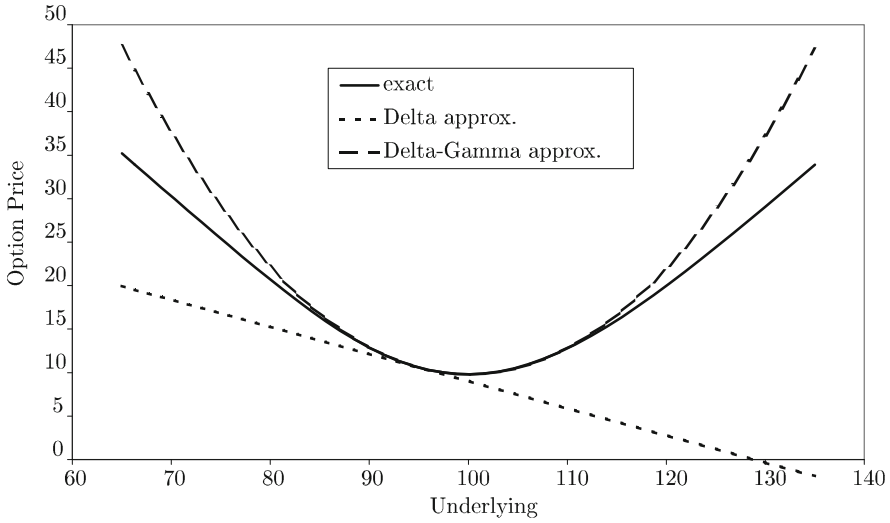

Fig. 22.1 Black-Scholes price of a straddle (strike = 100, time to maturity = 1 year) on an underlying S (volatility 25%, dividend yield 6%, repo rate 3%). The dashed line is the delta-gamma proxy, the dotted line is the simple delta proxy. The Taylor expansion was done about $S = 95$

22.1 Portfolios vs. Financial Instruments

Although we continually refer to *portfolio* sensitivities, the same results hold for individual *financial instruments* as well. In fact, sensitivity of a portfolio composed of financial instruments on the same underlying as described in Sect. 12.4.1 can be obtained by simply adding together the sensitivities of the individual instruments. This is the approach most commonly taken when calculating portfolio sensitivities. For the sake of clarity, we once again present this method explicitly here.

Consider a portfolio with a value V consisting of M different financial instruments with values V_k ,$k = 1, \ldots M$. N_k denotes the number of each instrument with value V_k held in the portfolio. The total value of the portfolio is naturally the sum of the values of each individual position

$$V(t) = \sum_{k=1}^{M} N_k V_k(\mathbf{S}(t)) \ . \tag{22.6}$$

The change in value $\delta V(t)$ of this portfolio is (approximated up to second order)

$$\begin{aligned}\delta V(\mathbf{S}(t)) &= \sum_{k=1}^{M} N_k \delta V_k(\mathbf{S}(t)) \\ &\approx \sum_{k=1}^{M} N_k \left[\sum_{i}^{n} \frac{\partial V_k}{\partial S_i} \delta S_i(t) + \frac{1}{2} \sum_{i,j}^{n} \delta S_i(t) \frac{\partial^2 V_k}{\partial S_i \partial S_j} \delta S_j(t) \right] \\ &= \sum_{k=1}^{M} N_k \sum_{i}^{n} \Delta_i^k \delta S_i(t) + \frac{1}{2} \sum_{k=1}^{M} N_k \sum_{i,j}^{n} \delta S_i(t) \Gamma_{ij}^k \delta S_j(t) \ ,\end{aligned} \tag{22.7}$$

where the sensitivities of the *financial instruments* have been introduced in the last step. For example, Δ_i^k is the linear sensitivity of the kth financial instrument in the portfolio with respect to the ith risk factor, etc.:

$$\Delta_i^k := \frac{\partial V_k}{\partial S_i}, \ \Gamma_{ij}^k := \frac{\partial^2 V_k}{\partial S_i \partial S_j}, \ i, j = 1, \ldots n; \ k = 1, \ldots M \ .$$

Simply rearranging the terms makes it clear that the summing over the index k (which denotes the different financial instruments) yields the portfolio sensitivities:

$$\begin{aligned}\delta V(\mathbf{S}(t)) &= \sum_{i}^{n} \delta S_i(t) \sum_{k=1}^{M} N_k \Delta_i^k + \frac{1}{2} \sum_{i,j}^{n} \delta S_i(t) \delta S_j(t) \sum_{k=1}^{M} N_k \Gamma_{ij}^k \\ &= \sum_{i}^{n} \delta S_i(t) \Delta_i + \frac{1}{2} \sum_{i,j}^{n} \delta S_i(t) \delta S_j(t) \Gamma_{ij} \ ,\end{aligned}$$

Thus, a portfolio sensitivity like Δ_i, for example, contains the sensitivities of all instruments in the portfolio (including all the position sizes N_k) with respect to the considered risk factor. This then yields (as is intuitively clear) the sensitivities of the entire portfolio as sums over the sensitivities of all instruments in the portfolio:

$$\Delta_i = \sum_{k=1}^{M} N_k \Delta_i^k, \ \Gamma_{ij} = \sum_{k=1}^{M} N_k \Gamma_{ij}^k \ .$$

In practice, this procedure is usually referred to as *position mapping*. Using the approximation in Eq. 22.1 for the change in risk factor, we finally obtain an expression for the portfolio's change in value as

$$\begin{aligned}\delta V(\mathbf{S}(t)) &\approx \sum_i^n \widetilde{\Delta}_i \, [\mu_i \delta t + \delta Z_i] + \frac{1}{2} \sum_{i,j}^n [\mu_i \delta t + \delta Z_i] \widetilde{\Gamma}_{ij} \, \left[\mu_j \delta t + \delta Z_j\right] \\ &\approx \sum_i^n \widetilde{\Delta}_i \, \delta Z_i + \frac{1}{2} \sum_{i,j}^n \delta Z_i \widetilde{\Gamma}_{ij} \, \delta Z_j \ ,\end{aligned}$$

using the modified portfolio sensitivities defined as in Eq. 22.5:

$$\widetilde{\Delta}_i := S_i(t) \Delta_i = S_i(t) \sum_{k=1}^M N_k \Delta_i^k \tag{22.8}$$

$$\widetilde{\Gamma}_{ij} := S_i(t) S_j(t) \Gamma_{ij} = S_i(t) S_j(t) \sum_{k=1}^M N_k \Gamma_{ij}^k \ .$$

Adding the sensitivities of the financial instruments and multiplying by the current levels of the risk factors to obtain these modified portfolio sensitivities is sometimes referred to as *VaR mapping*.

22.2 The Delta-Normal Method

In the *delta-normal method,* the Taylor series 22.4 of the portfolio value is broken off after the linear term.

$$\begin{aligned}\delta V(\mathbf{S}(t)) &\approx \sum_i^n \frac{\partial V}{\partial S_i} \delta S_i(t) \qquad (22.9) \\ &\approx \sum_i^n \widetilde{\Delta}_i \, [\mu_i \delta t + \delta Z_i] \\ &\approx \sum_i^n \widetilde{\Delta}_i \, \delta Z_i = \widetilde{\boldsymbol{\Delta}}^T \delta \mathbf{Z} \ .\end{aligned}$$

22.2.1 The Value at Risk with Respect to a Single Risk Factor

For a single risk factor this means

$$\delta V(S(t)) \approx \Delta \delta S(t) \tag{22.10}$$

with the sensitivity $\Delta := \partial V / \partial S$. The change in the portfolio's value is thus approximated to be a linear function of the change in the underlying risk factor. This corresponds to the situation described in Sect. 21.3. There, the constant of proportionality (the sensitivity) was not Δ but N or $-N$ for a long or short position, respectively. The linear approximation implies intuitively that a portfolio with a linear sensitivity Δ with respect to a risk factor can be interpreted as a portfolio consisting of Δ risk factors. The only subtlety in this argumentation is that in Sect. 21.3, we distinguished between a long and a short position, treating the two cases differently on the basis of whether the proportionality constant N was positive or negative (which leads to the two different VaRs in Eq. 21.16). However, we cannot know a priori whether Δ is greater or less than 0. We do know, however, that V is linear and in consequence, a monotone function of S. Therefore, in the sense of Eq. 21.6, the results following from Eq. 21.16 hold with the correspondence $\Delta \widehat{=} N$ for $\Delta > 0$, and $\Delta \widehat{=} -N$ for $\Delta < 0$. Using this fact allows us to write

$$\begin{aligned} \mathrm{VaR}_V(c) \approx \max \Big\{ & \widetilde{\Delta} \left[1 - \exp\left(\mu \delta t + Q_{1-c} \sigma \sqrt{\delta t} \right) \right], \\ & \widetilde{\Delta} \left[1 - \exp\left(\mu \delta t - Q_{1-c} \sigma \sqrt{\delta t} \right) \right] \Big\} \end{aligned} \tag{22.11}$$

using the notation $\widetilde{\Delta} := S(t)\Delta = S(t)\partial V / \partial S$. As usual, Q_{1-c} is the $(1-c)$ percentile of the standard normal distribution.[1] The maximum function in Eq. 21.6 effects the correct choice for the VaR. If $\Delta > 0$, the *lower bound* of the confidence interval of the risk factor[2] is relevant and consequently the VaR function as defined above takes on the value corresponding to a *long* position in Eq. 21.16. Likewise for $\Delta < 0$, the *upper bound* of the confidence interval of the risk factor[3] is relevant and the above defined maximum function takes on the value corresponding to the VaR of the short position in Eq. 21.16.

[1] For all relevant confidence levels, this percentile is a negative number, see Eq. 21.13.

[2] This corresponds to the percentile Q_{1-c} of the standard normal distribution.

[3] This corresponds to the percentile $-Q_{1-c} = Q_c$ of the standard normal distribution.

In all our deliberations up to this point, only the portfolio value has been approximated with Eq. 22.10. The change in the *risk factor* in Eq. 22.11 is still exact. Approximating this risk factor change with Eq. 22.1, the VaR becomes

$$\begin{aligned}\mathrm{VaR}_V(c) \approx \max\{&\widetilde{\Delta}\left[-\mu\delta t - Q_{1-c}\sigma\sqrt{\delta t}\right],\\ &\widetilde{\Delta}\left[-\mu\delta t + Q_{1-c}\sigma\sqrt{\delta t}\right]\}\end{aligned}$$

This corresponds exactly to Eq. 21.19, since $\widetilde{\Delta} \widehat{=} N$ for $\widetilde{\Delta} > 0$ and $\widetilde{\Delta} \widehat{=} -N$ for $\widetilde{\Delta} < 0$.

The common summand $-\widetilde{\Delta}\mu\delta t$ can now be taken out of the maximum function

$$\begin{aligned}\mathrm{VaR}_V(c) &\approx \max\left\{-\widetilde{\Delta}Q_{1-c}\sigma\sqrt{\delta t}, +\widetilde{\Delta}Q_{1-c}\sigma\sqrt{\delta t}\right\} - \widetilde{\Delta}\mu\delta t\\ &= \left|\widetilde{\Delta}Q_{1-c}\sigma\sqrt{\delta t}\right| - \widetilde{\Delta}\mu\delta t \qquad (22.12)\end{aligned}$$

In this approximation, the maximum function produces precisely the absolute value of the risk which is caused by the volatility of the risk factor. A positive drift μ of the risk factor reduces the portfolio risk when $\widetilde{\Delta} > 0$ (intuitively, the portfolio then represents a long position). If, on the other hand, the portfolio sensitivity is negative, i.e., $\widetilde{\Delta} < 0$, a positive drift *increases* the portfolio risk (intuitively, the portfolio represents a short position). The drift's influence is of course lost if the drift is neglected in the approximation. The value at risk then reduces to

$$\mathrm{VaR}_V(c) \approx \left|\widetilde{\Delta}Q_{1-c}\sigma\sqrt{\delta t}\right| \qquad (22.13)$$

where the absolute value makes it immediately clear that the sign of the portfolio sensitivity no longer plays a role.

22.2.2 The Value at Risk with Respect to Several Risk Factors

In the previous section, linear approximations enabled us to reduce the VaR with respect to a single risk factor to that of a position consisting of Δ instruments representing the risk factor. We were then able, as was done in Sect. 21.3, to deduce information about the unknown distribution of the

portfolio's value V from the known distribution of the *risk factor* S. The extension of these results to the case of several risk factors is not trivial even for the delta-normal approximation. Only by using the roughest approximation in Eq. 22.1 for the change in the risk factors, namely $\delta S_i(t) \approx S_i(t)\delta Z_i$, can we manage to avoid involving the distribution of V in the discussion.

The approximation Eq. 21.26, i.e. $\delta S(t) \approx S(t)\delta Z$, led to the value at risk Eq. 22.13 with respect to a single risk factor. Squaring both sides of this equation yields

$$\begin{aligned}
\mathrm{VaR}_V^2(c) &\approx \widetilde{\Delta}^2 \left(Q_{1-c'}\right)^2 \sigma^2 \delta t \\
&= \Delta^2 \left(Q_{1-c'}\right)^2 S(t)^2 \sigma^2 \delta t \\
&= \Delta^2 \left(Q_{1-c'}\right)^2 S(t)^2 \,\mathrm{var}\left[\delta Z\right] \\
&= \Delta^2 \left(Q_{1-c'}\right)^2 \mathrm{var}\left[\delta S(t)\right] \ ,
\end{aligned}$$

since the approximation in Eq. 21.26 allows the approximation of the variance of $\delta S(t)$ with $S(t)^2 \,\mathrm{var}\left[\delta Z\right]$. On the other hand, the variance of V can be calculated from Eq. 22.10 simply as

$$\mathrm{var}\left[\delta V\right] \approx \mathrm{var}\left[\Delta \delta S(t)\right] = \Delta^2 \,\mathrm{var}\left[\delta S(t)\right] \ .$$

This means that this approximation can be used to express the square of the value at risk in terms of a multiple of the *variance of the portfolio*:

$$\begin{aligned}
\mathrm{VaR}_V^2(c) &\approx \left(Q_{1-c'}\right)^2 \Delta^2 \,\mathrm{var}\left[\delta S(t)\right] \\
&= \left(Q_{1-c'}\right)^2 \mathrm{var}\left[\delta V\right] \ .
\end{aligned}$$

Now, only the variance of the portfolio's value needs to be determined for the computation of the VaR and not its *distribution* or its percentiles.

If several risk factors are involved, Eq. 22.9 can be used to write the portfolio's change in value, δV, in the approximation given by Eq. 21.26, as a linear combination of normally distributed random variables δZ_i (with deterministic coefficients $\widetilde{\Delta}_i$)

$$\delta V \approx \sum_{i=1}^{n} \widetilde{\Delta}_i \delta Z_i \ .$$

The variance of a sum of random variables is equal to the sum of the covariances of these random variable as can be seen from Eq. A.17. The

variance of the portfolio value is thus

$$\begin{aligned}\operatorname{var}[\delta V] &\approx \sum_{i,j=1}^{n} \widetilde{\Delta}_i \widetilde{\Delta}_j \operatorname{cov}\left[\delta Z_i, \delta Z_j\right] \\ &= \sum_{i,j=1}^{n} \widetilde{\Delta}_i \delta \Sigma_{ij} \widetilde{\Delta}_j \\ &= \delta t \sum_{i,j=1}^{n} \widetilde{\Delta}_i \sigma_i \rho_{ij} \sigma_j \widetilde{\Delta}_j \ ,\end{aligned} \tag{22.14}$$

where the definition of the covariance matrix in Eq. 21.22 was used in the last step. This means that the value at risk can be approximated as

$$\begin{aligned}\mathrm{VaR}_V(c) &\approx |Q_{1-c}| \sqrt{\operatorname{var}[\delta V]} \\ &= |Q_{1-c}| \sqrt{\widetilde{\boldsymbol{\Delta}} \delta \boldsymbol{\Sigma} \widetilde{\boldsymbol{\Delta}}} \\ &= |Q_{1-c}| \sqrt{\delta t} \sqrt{\sum_{i,j=1}^{n} \widetilde{\Delta}_i \sigma_i \rho_{ij} \sigma_j \widetilde{\Delta}_j} \ .\end{aligned} \tag{22.15}$$

This is the central equation for the delta-normal method. It summarizes all assumptions, approximations, and computation methods of the delta-normal method.[4]

[4]If all portfolio sensitivities are non-negative (which is often the case for instance for a private investor's portfolio containing only long positions), then this equation can be rewritten in an alternative and quite intuitive form. Let $\mathrm{VaR}_i(c)$ denote the value at risk of the portfolio with respect to a particular risk factor S_i. Using Eq. 22.13, we can approximate this by

$$\mathrm{VaR}_i(c) \approx \left| \widetilde{\Delta}_i Q_{1-c} \sigma_i \sqrt{\delta t} \right| \ .$$

Thus, for the special case that *none* of the portfolio deltas is negative, the VaR with respect to all risk factors can be obtained by computing the square root of the weighted sum of the products of all the VaRs with respect to the individual risk factors. The weights under consideration are the respective correlations between the risk factors:

$$\mathrm{VaR}_V(c) \approx \sqrt{\sum_{i,j=1}^{n} \mathrm{VaR}_i(c) \rho_{ij} \mathrm{VaR}_j(c)} \quad \text{falls} \widetilde{\Delta}_i \geq 0 \forall i \in \{1, \ldots, n\} \ . \tag{22.16}$$

In the linear approximation, the effect of the drifts can be subsequently introduced into the approximation. The *expected* change in portfolio value is calculated using the deltas and the drifts of the risk factors. Analogously to Eq. 22.12 for a single risk factor, this expected change is then subtracted from the value at risk of the portfolio given in Eq. 22.15:

$$\mathrm{VaR}_V(c) \approx |Q_{1-c}| \sqrt{\delta t} \sqrt{\sum_{i,j=1}^{n} \widetilde{\Delta}_i \sigma_i \rho_{ij} \sigma_j \widetilde{\Delta}_j} - \delta t \sum_i \widetilde{\Delta}_i \, \mu_i \ . \tag{22.17}$$

The delta-normal approach to the calculation of the value at risk can be summarized as follows:

- Calculate the sensitivities of the portfolio with respect to all risk factors.
- Multiply the covariance matrix with the sensitivities of the portfolio and the current values of the risk factors as in Eq. 22.14 to obtain the variance of the portfolio. The covariance matrix's elements consist of the product of volatilities and correlations of the risk factors as defined in Eq. 21.22.
- Multiply the portfolio variance as in Eq. 22.15 by the liquidation period and the square of the percentile corresponding to the desired confidence interval (for example, -2.326 for 99% confidence).
- The square root of the thus obtained number is the value at risk of the entire portfolio, neglecting the effect of the drifts of the risk factors.
- The effect of the drifts can be taken into account using Eq. 22.17.

For future reference we re-write the final Value at Risk in Eq. 22.17 in terms of the covariances for the logarithmic changes and in terms of the covariances of the returns. According to Eqs. 21.27 and 21.28 the difference is an overall factor δt:

$$\begin{aligned} \mathrm{VaR}_V(c) &\approx |Q_{1-c}| \sqrt{\sum_{i,j=1}^{n} \widetilde{\Delta}_i \widetilde{\Delta}_j \operatorname{cov}\left[\delta \ln S_i, \delta \ln S_j\right] - \sum_i \widetilde{\Delta}_i \mathrm{E}\left[\delta \ln S_i\right]} \\ &= \delta t \ |Q_{1-c}| \sqrt{\sum_{i,j=1}^{n} \widetilde{\Delta}_i \widetilde{\Delta}_j \operatorname{cov}\left[r_i, r_j\right] - \delta t \sum_i \widetilde{\Delta}_i \mathrm{E}\left[r_i\right]} \ . \end{aligned} \tag{22.18}$$

Here, the r_i are the historic portfolio returns over holding periods of length δt, each return annualized.

These two forms of the VaR are very important in practice, when one is given historical time series of risk factor prices S_i or *annualized* risk factor returns r_i rather then the volatilities and correlations needed in Eq. 22.17, which otherwise would have to be calculated or bought from some market data vendor.

22.3 The Delta-Gamma Method

The delta-gamma method for calculating the portfolio's VaR makes use of the Taylor series expansion of the value of the portfolio up to and including the second order terms along with the approximation in Eq. 22.1 for the risk factors. The starting point for the delta-gamma method is thus the last line in Eq. 22.4, which when written in vector notation is given by

$$\delta V(\mathbf{S}(t)) = \widetilde{\mathbf{\Delta}}^T \delta\mathbf{Z} + \frac{1}{2}\delta\mathbf{Z}^T \widetilde{\mathbf{\Gamma}} \delta\mathbf{Z} \tag{22.19}$$

where

$$\delta\mathbf{Z} \sim \mathrm{N}(\mathbf{0}, \delta\mathbf{\Sigma}) \implies \mathrm{cov}[\delta Z_i, \delta Z_j] = \sigma_i \rho_{ij} \sigma_j \delta t, \ \mathrm{E}[\delta Z_i] = 0 \, .$$

The right-hand side of Eq. 22.19 can *not* be written as the sum of the contributions of each risk factor as was the case for the delta-normal method:

$$\sum_i^n \widetilde{\Delta}_i \, \delta Z_i + \frac{1}{2} \sum_{i,j}^n \delta Z_i \widetilde{\Gamma}_{ij} \, \delta Z_j \neq \sum_i^n (\text{contribution of the } i\text{-th risk factor}) \ .$$

The contributions of the individual risk factors can not be considered separately since they are *coupled* in the above equation by the matrix $\widetilde{\mathbf{\Gamma}}$. Furthermore the random variables δZ_j are not independent. They are *correlated* through the covariance matrix $\delta\mathbf{\Sigma}$. Two essential elements of the method presented here are[5]:

- Using the Cholesky decomposition of the covariance matrix $\delta\mathbf{\Sigma}$ to transform the δZ_j into independent random variables.

[5]This goes back to a paper by Rouvinez, see [166].

- Diagonalizing the gamma matrix $\widetilde{\mathbf{\Gamma}}$ thereby decoupling the contributions of the individual risk factors.

Also after the Cholesky decompositions, and in contrast to the delta-normal case, it is still not possible in the situation of Eq. 22.19 to reduce the VaR with respect to a risk factor to the VaR of a position consisting of Δ of these risk factors. Thus the (unknown) distribution of the *portfolio value* can no longer be substituted by the (known) distribution of the individual *risk factors*, as could still be done in Sect. 21.3. Instead, the distribution of δV must be determined *directly* in order to calculate the value at risk defined in Eqs. 21.2 or 21.3. A third essential step of the delta-gamma method presented here involves the determination of the distribution of δV.

22.3.1 Decoupling of the Risk Factors

Motivated by Eq. 21.40, we first introduce a matrix $\mathbf{A}$ satisfying the property 21.35. This matrix can be constructed through the Cholesky decomposition of the covariance matrix as described in detail in Sect. 21.5.3. This matrix transforms the correlated δZ_i into uncorrelated random variables. With this goal in mind, we rewrite Eq. 22.19, first introducing identity matrices into the equation and then replacing them with $\mathbf{A}\mathbf{A}^{-1}$ or $(\mathbf{A}^T)^{-1}\mathbf{A}^T$ as follows:

$$\begin{aligned}
\delta V(\mathbf{S}(t)) &= \widetilde{\mathbf{\Delta}}^T \mathbf{1}\,\delta\mathbf{Z} + \frac{1}{2}\delta\mathbf{Z}^T \mathbf{1}\,\widetilde{\mathbf{\Gamma}}\,\mathbf{1}\,\delta\mathbf{Z} \\
&= \widetilde{\mathbf{\Delta}}^T \mathbf{A}\mathbf{A}^{-1}\,\delta\mathbf{Z} + \frac{1}{2}\delta\mathbf{Z}^T (\mathbf{A}^T)^{-1}\mathbf{A}^T\,\widetilde{\mathbf{\Gamma}}\,\mathbf{A}\mathbf{A}^{-1}\,\delta\mathbf{Z} \\
&= \widetilde{\mathbf{\Delta}}^T \mathbf{A}\mathbf{A}^{-1}\,\delta\mathbf{Z} + \frac{1}{2}\delta\mathbf{Z}^T (\mathbf{A}^{-1})^T\mathbf{A}^T\,\widetilde{\mathbf{\Gamma}}\,\mathbf{A}\mathbf{A}^{-1}\,\delta\mathbf{Z} \\
&= \widetilde{\mathbf{\Delta}}^T \mathbf{A}\,(\mathbf{A}^{-1}\,\delta\mathbf{Z}) + \frac{1}{2}(\mathbf{A}^{-1}\delta\mathbf{Z})^T\mathbf{A}^T\widetilde{\mathbf{\Gamma}}\mathbf{A}\,(\mathbf{A}^{-1}\delta\mathbf{Z})\,.
\end{aligned}$$

In the penultimate step, the property 21.39 has been used; recall that this holds for every invertible matrix. In the last step, the parentheses are intended to emphasize the fact that $\delta\mathbf{Z}$ only appears in combination with $\mathbf{A}^{-1}$. We have shown before that the components of the vector $\mathbf{A}^{-1}\delta\mathbf{Z}$ are *iid* random

variables, see Eq. 21.40. We can therefore write

$$\begin{aligned} \delta V(\mathbf{S}(t)) &= \widetilde{\boldsymbol{\Delta}}^T \mathbf{A}\, \delta\mathbf{Y} + \frac{1}{2}\delta\mathbf{Y}^T \mathbf{M}\ \delta\mathbf{Y} \\ &\text{mit}\ \ \delta\mathbf{Y} := \mathbf{A}^{-1}\delta\mathbf{Z} \sim \mathrm{N}(\mathbf{0},\mathbf{1}),\ \ \textit{iid} \\ &\text{und}\ \ \mathbf{M} := \mathbf{A}^T \widetilde{\boldsymbol{\Gamma}} \mathbf{A}\ . \end{aligned} \tag{22.20}$$

Thus, the first goal has been accomplished. The δZ_i have been transformed into *iid* random variables δY_i.

Because $\widetilde{\boldsymbol{\Gamma}}$ is by definition a symmetric matrix, i.e., $\widetilde{\Gamma}_{ij} = \widetilde{\Gamma}_{ji}$, we can show that the newly defined matrix $\mathbf{M}$ is symmetric as well:

$$\begin{aligned} M_{ij} &= \left(\mathbf{A}^T \widetilde{\boldsymbol{\Gamma}} \mathbf{A}\right)_{ij} = \sum_k \sum_m (A^T)_{ik} \widetilde{\Gamma}_{km} A_{mj} \\ &= \sum_k \sum_m (A^T)_{ik} \widetilde{\Gamma}_{mk} A_{mj} = \sum_k \sum_m A_{mj} \widetilde{\Gamma}_{mk} (A^T)_{ik} \\ &= \sum_k \sum_m (A^T)_{jm} \widetilde{\Gamma}_{mk} A_{ki} = \left(\mathbf{A}^T \widetilde{\boldsymbol{\Gamma}} \mathbf{A}\right)_{ji} = M_{ji}\ . \end{aligned}$$

22.3.2 Diagonalization of the Gamma Matrix

The next step is to decouple the contributions to δV of the individual random variables in Eq. 22.20. This is accomplished by diagonalizing the gamma matrix, or more precisely, the transformed gamma matrix $\mathbf{M}$ introduced in Eq. 22.20. Diagonalizing a matrix is a standard procedure in *linear algebra*. We refer the reader to the relevant literature.[6] We nevertheless take the opportunity to demonstrate the fundamental operations for *diagonalizing a matrix* here since they contain essential elements of the practical value at risk computations to be performed in the delta-gamma method.

The *eigenvectors* $\mathbf{e}^i$ of a matrix $\mathbf{M}$ are the non-zero vectors which are mapped by $\mathbf{M}$ to the same vector multiplied by a number (called a *scalar* in algebra):

$$\begin{aligned} \mathbf{M}\mathbf{e}^i &= \lambda_i \mathbf{e}^i \ \Leftrightarrow \\ (\mathbf{M} - \lambda_i \mathbf{1})\, \mathbf{e}^i &= \mathbf{0}\ . \end{aligned} \tag{22.21}$$

[6] The most important results required for the analysis here receive a clear and concise treatment in [78], for example.

These scalars λ_i are called *eigenvalues* of the matrix. As known from linear algebra, an equation of this kind has a non-trivial solution $\mathbf{e}^i \neq \mathbf{0}$ if and only if the matrix $(\mathbf{M} - \lambda_i \mathbf{1})$ is *singular*. For this to be the case, the *determinant* of this matrix must be zero:

$$\det(\mathbf{M} - \lambda_i \mathbf{1}) = 0 \,. \tag{22.22}$$

The solutions of Eq. 22.22 are the eigen*values* λ_i. Having determined these values, they can be substituted into Eq. 22.21 to determine the eigen*vectors* $\mathbf{e}^i$ of the matrix. The eigenvectors have as yet only been defined up to a multiplicative scalar since if $\mathbf{e}^i$ solves Eq. 22.21 then $c\mathbf{e}^i$ does as well, for any arbitrary scalar c. The uniqueness of the eigenvalues can be guaranteed by demanding that the eigenvectors have norm 1:

$$(\mathbf{e}^i)^T \mathbf{e}^i = 1 \tag{22.23}$$

As is known from linear algebra, a symmetric, non-singular $n \times n$ matrix has n linearly independent eigenvectors which are *orthogonal*. This means that the *inner product* of each pair of different eigenvectors equals zero (graphically: the angle formed by the two vectors is 90 degrees). Together with the normalization the eigenvectors thus have the following property

$$(\mathbf{e}^i)^T \mathbf{e}^j = \sum_k e_k^i e_k^j = \delta_{ij} \tag{22.24}$$

where δ_{ij} denotes the well-known *Kronecker delta*. A collection of vectors satisfying this property is called *orthonormal*. Since we have shown that the $n \times n$ matrix $\mathbf{M}$ in Eq. 22.20 is symmetric, we can be sure that it indeed has n orthonormal eigenvectors satisfying Eq. 22.24.

To clarify the notation for these eigenvectors: The subscript k and the superscript i of e_k^i identify this value as the k-th component of the i-th eigenvector $\mathbf{e}^i$.

$$\mathbf{e}^j = \begin{pmatrix} e_1^j \\ e_2^j \\ \vdots \\ e_n^j \end{pmatrix}, \quad (\mathbf{e}^j)^T = \begin{pmatrix} e_1^j & e_2^j & \cdots & e_n^j \end{pmatrix}, \quad j = 1, \ldots, n \,.$$

A matrix $\mathbf{O}$ can now be constructed whose column vectors are composed of the eigenvectors of $\mathbf{M}$:

$$\mathbf{O} = \begin{pmatrix} \mathbf{e}^1 & \mathbf{e}^2 & \dots & \mathbf{e}^n \end{pmatrix} = \begin{pmatrix} e_1^1 & e_1^2 & \cdots & e_1^n \\ e_2^1 & e_2^2 & & \vdots \\ \vdots & & \ddots & \vdots \\ e_n^1 & \cdots & \cdots & e_n^n \end{pmatrix} \Rightarrow O_{ij} = e_i^j \ . \tag{22.25}$$

The jth eigenvector $\mathbf{e}^j$ is in the jth column of the matrix. In the ith row, we find the ith components of all eigenvectors. As we will soon see, this matrix is an indispensable tool for the purpose of the diagonalization. As can be immediately verified[7]

$$\mathbf{O}^T\mathbf{O} = \mathbf{1} \tag{22.26}$$

and therefore also

$$\mathbf{O}^T = \mathbf{O}^{-1} \Rightarrow \mathbf{O}\mathbf{O}^T = \mathbf{1} \ . \tag{22.27}$$

Equation 22.26 characterizes a group of matrices known as *orthonormal transformations.* Applying such a matrix to a vector effects a rotation of the vector.

The eigen*values* of the matrix $\mathbf{M}$ can be used to construct a matrix as well, namely the diagonal matrix

$$\boldsymbol{\lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{pmatrix} \ . \tag{22.28}$$

From Eq. 22.21, it follows immediately that the relation

$$\mathbf{M}\mathbf{O} = \mathbf{O}\boldsymbol{\lambda} \tag{22.29}$$

[7] $(\mathbf{O}^T\mathbf{O})_{ij} = \sum_k (O^T)_{ik} O_{kj} = \sum_k e_k^i e_k^j = \delta_{ij}$.

holds for the matrices $\mathbf{M}$, $\mathbf{O}$ and $\boldsymbol{\lambda}$**.** Such matrix equations can often be verified quite easily by comparing the matrix elements individually:

$$\begin{aligned}(\mathbf{MO})_{ij} &= \sum_k M_{ik} O_{kj} = \sum_k M_{ik} e_k^j = \left(\mathbf{M}\mathbf{e}^j\right)_i \\ &= \left(\lambda_j \mathbf{e}^j\right)_i = \lambda_j e_i^j = O_{ij}\lambda_j = (\mathbf{O}\boldsymbol{\lambda})_{ij} \ .\end{aligned}$$

The decisive step in the above proof is the first equality in the second line where the eigenvector equation 22.21 was used. Multiplying both sides of this equation *from the left* by the matrix $\mathbf{O}^T$ and using Eq. 22.26 directly yields the desired diagonalization of $\mathbf{M}$, since $\boldsymbol{\lambda}$ is a diagonal matrix:

$$\mathbf{O}^T\mathbf{M}\mathbf{O} = \mathbf{O}^T\mathbf{O}\boldsymbol{\lambda} = \boldsymbol{\lambda} \ . \tag{22.30}$$

Multiplying both sides of Eq. 22.29 *from the right* by the matrix $\mathbf{O}^T$ and using Eq. 22.27, also yields a very useful representation of $\mathbf{M}$, namely the *spectral representation* (also referred to as *eigenvector decomposition*)

$$\mathbf{M} = \mathbf{O}\boldsymbol{\lambda}\mathbf{O}^T = \sum_k \lambda_k \left(\mathbf{e}^k (\mathbf{e}^k)^T\right) \ .$$

We are now in a position to introduce the diagonalized matrix $\mathbf{O}^T\mathbf{M}\mathbf{O}$ into Eq. 22.20 by inserting identity matrices in Eq. 22.20 and subsequently replacing them by $\mathbf{O}\mathbf{O}^T$. Equation 22.27 ensures that equality is maintained**.**

$$\begin{aligned}\delta V(\mathbf{S}(t)) &= \widetilde{\boldsymbol{\Delta}}^T \mathbf{A}\,\mathbf{1}\,\delta\mathbf{Y} + \frac{1}{2}\delta\mathbf{Y}^T\,\mathbf{1}\,\mathbf{M}\,\mathbf{1}\,\delta\mathbf{Y} \\ &= \widetilde{\boldsymbol{\Delta}}^T \mathbf{A}\mathbf{O}\mathbf{O}^T\delta\mathbf{Y} + \frac{1}{2}\delta\mathbf{Y}^T\mathbf{O}\underbrace{\mathbf{O}^T\mathbf{M}\mathbf{O}}_{\boldsymbol{\lambda}}\mathbf{O}^T\delta\mathbf{Y} \\ &= \widetilde{\boldsymbol{\Delta}}^T \mathbf{A}\mathbf{O}\left(\mathbf{O}^T\delta\mathbf{Y}\right) + \frac{1}{2}\left(\mathbf{O}^T\delta\mathbf{Y}\right)^T \boldsymbol{\lambda}\left(\mathbf{O}^T\delta\mathbf{Y}\right) \ .\end{aligned}$$

In the last equality, the parentheses are meant to emphasize that $\delta\mathbf{Y}$ appears only in combination with $\mathbf{O}^T$. In consequence, we can write

$$\begin{aligned}\delta V(\mathbf{S}(t)) &= \widetilde{\boldsymbol{\Delta}}^T \mathbf{A}\mathbf{O}\,\delta\mathbf{X} + \frac{1}{2}\delta\mathbf{X}^T\boldsymbol{\lambda}\delta\mathbf{X} \\ \text{mit } \ \delta\mathbf{X} &:= \mathbf{O}^T\delta\mathbf{Y} = \mathbf{O}^T\mathbf{A}^{-1}\delta\mathbf{Z} \\ \text{und } \ \boldsymbol{\lambda} &:= \mathbf{O}^T\mathbf{M}\mathbf{O} = \mathbf{O}^T\mathbf{A}^T\widetilde{\boldsymbol{\Gamma}}\mathbf{A}\mathbf{O} \ .\end{aligned} \tag{22.31}$$

The δY_i were *iid*, standard normally distributed random variables. This was accomplished in the previous section by the mapping $\mathbf{A}^{-1}$. Now to achieve the diagonalization of the gamma matrix, the random variables must undergo a further transformation under the mapping $\mathbf{O}^T$. The question remains as to whether the accomplishments of the previous section was undone by this new transformation, in other words, whether the transformed random variables have remained independent. We therefore consider the covariance of the new random variables

$$\begin{aligned}
\mathrm{cov}[\delta X_i, \delta X_j] &= \mathrm{cov}\left[\sum_k O^T_{ik}\delta Y_k, \sum_m O^T_{jm} Y_m\right] \\
&= \sum_k \sum_m O^T_{ik} O^T_{jm} \underbrace{\mathrm{cov}[\delta Y_k, \delta Y_m]}_{\delta_{km}} \\
&= \sum_k O^T_{ik} O^T_{jk} \\
&= \sum_k O^T_{ik} O_{kj} = \left(\mathbf{O}^T\mathbf{O}\right)_{ij} = \mathbf{1}_{ij} = \delta_{ij} \ .
\end{aligned}$$

The covariances have remained invariant under the transformation $\mathbf{O}^T$. Thus, the new random variables are also uncorrelated and all have variance 1. Also, the zero expectation does not change under the transformation $\mathbf{O}^T$:

$$\mathrm{E}[\delta X_i] = \mathrm{E}\left[\sum_k O^T_{ik}\delta Y_k\right] = \sum_k O^T_{ik} \underbrace{\mathrm{E}[\delta Y_k]}_{0} \ .$$

Since matrix multiplication is a linear transformation (we are operating in the realm of *linear* algebra), the form of the distribution remains the same as well. In summary, the new random variables are uncorrelated and all have the same standard-normal distribution. We have argued after Eq. 21.40 that such variables are indeed *iid* random variables. Summarizing the above deliberations, we can write for the δX_i:

$$\delta\mathbf{X} \sim \mathrm{N}(\mathbf{0}, \mathbf{1}), \ \textit{iid} \ .$$

We have only used property 22.26, i.e., the *iid* property of random variables remains invariant under *every* orthonormal transformation.

If we define a "transformed sensitivity vector" as

$$\mathbf{L} := \mathbf{O}^T \mathbf{A}^T \widetilde{\boldsymbol{\Delta}}$$

(which implies for its transposed $\mathbf{L}^T = \widetilde{\boldsymbol{\Delta}}^T \mathbf{AO}$), the portfolio-change Eq. 22.31 can be brought into the following simple form

$$\delta V(\mathbf{S}(t)) = \mathbf{L}^T \delta\mathbf{X} + \frac{1}{2}\delta\mathbf{X}^T \boldsymbol{\lambda} \delta\mathbf{X} \quad \text{with} \quad \delta\mathbf{X} \sim \mathrm{N}(\mathbf{0}, \mathbf{1}) \tag{22.32}$$

or, expressed component-wise

$$\delta V(\mathbf{S}(t)) = \sum_i \left[L_i \, \delta X_i + \frac{1}{2}\lambda_i \delta X_i^2 \right] = \sum_i \delta V_i$$

$$\text{mit } \delta V_i = L_i \, \delta X_i + \frac{1}{2}\lambda_i \delta X_i^2, \; i = 1, \ldots, n \; . \tag{22.33}$$

The change in the portfolio's value is now the decoupled sum of the individual contributions of *iid* random variables as was our original intention.

At this stage, we collect all transformations involved in mapping the original random variables in Eq. 22.19 into the *iid* random variables in the above expression Eq. 22.32:

$$\delta\mathbf{X} := \mathbf{O}^T \mathbf{A}^{-1} \delta\mathbf{Z}, \; \boldsymbol{\lambda} := \mathbf{O}^T \mathbf{A}^T \widetilde{\boldsymbol{\Gamma}} \mathbf{AO}, \; \mathbf{L} := \mathbf{O}^T \mathbf{A}^T \widetilde{\boldsymbol{\Delta}} \; .$$

From Eq. 22.27 we know that $\mathbf{O}^T \mathbf{A}^{-1} = \mathbf{O}^{-1} \mathbf{A}^{-1} = (\mathbf{AO})^{-1}$. We now recognize that all of these transformations can be represented with a single matrix defined as

$$\mathbf{D} := \mathbf{AO} \tag{22.34}$$

With this matrix, the transformations become simply

$$\delta\mathbf{X} := \mathbf{D}^{-1} \delta\mathbf{Z}, \; \boldsymbol{\lambda} := \mathbf{D}^T \widetilde{\boldsymbol{\Gamma}} \mathbf{D}, \; \mathbf{L} := \mathbf{D}^T \widetilde{\boldsymbol{\Delta}} \; . \tag{22.35}$$

The matrix $\mathbf{D}$ *directly* diagonalizes the gamma matrix $\widetilde{\boldsymbol{\Gamma}}$ (as opposed to $\mathbf{O}$, which diagonalizes the matrix $\mathbf{M}$). *In addition*, $\mathbf{D}$ is by definition, an orthonormal transformation (a "rotation") of $\mathbf{A}$, the Cholesky decomposition of the covariance matrix. $\mathbf{D}$ is likewise a "square root" of the covariance

matrix, since the square of a matrix remains invariant under orthonormal transformations of the matrix. Explicitly:

$$\mathbf{D}\mathbf{D}^T = \mathbf{A}\mathbf{O}(\mathbf{A}\mathbf{O})^T = \mathbf{A}\mathbf{O}\mathbf{O}^T\mathbf{A}^T = \mathbf{A}\mathbf{1}\mathbf{A}^T = \mathbf{A}\mathbf{A}^T = \delta\boldsymbol{\Sigma} \ . \tag{22.36}$$

Therefore, the matrix $\mathbf{D}$ satisfies *both* tasks, namely the decoupling of the gamma matrix *and* the transformation of the correlated random variables into uncorrelated ones.

As a little consistence check, using the matrix $\mathbf{D}$ we immediately recognize the equivalence of Eqs. 22.32 and 22.19:

$$\begin{aligned}\delta V(\mathbf{S}(t)) &= \left(\mathbf{D}^T\widetilde{\boldsymbol{\Delta}}\right)^T \mathbf{D}^{-1}\delta\mathbf{Z} + \frac{1}{2}\left(\mathbf{D}^{-1}\delta\mathbf{Z}\right)^T \mathbf{D}^T\,\widetilde{\boldsymbol{\Gamma}}\,\mathbf{D}\mathbf{D}^{-1}\delta\mathbf{Z} \\ &= \widetilde{\boldsymbol{\Delta}}^T \underbrace{\mathbf{D}\mathbf{D}^{-1}}_{\mathbf{1}}\delta\mathbf{Z} + \frac{1}{2}\delta\mathbf{Z}^T \underbrace{\left(\mathbf{D}^T\right)^{-1}\mathbf{D}^T}_{\mathbf{1}}\,\widetilde{\boldsymbol{\Gamma}}\,\underbrace{\mathbf{D}\mathbf{D}^{-1}}_{\mathbf{1}}\delta\mathbf{Z} \ ,\end{aligned}$$

where Eq. 21.39 was again used in verifying this equivalence.

22.3.3 The Distribution of the Portfolio Value Changes

Having decoupled the individual contributions to δV in Eq. 22.33 into standard normally distributed *iid* random variables, we can now determine the distribution of the sum. δV is nevertheless not simply the sum of normally distributed random variables alone, since the expression also includes the *square* of normally distributed random variables. These additional random variables represent the difference in complexity compared to the delta-normal method. According to Sect. A.4.6, the square of a standard normally distributed random variable is χ^2-distributed with one degree of freedom. We can thus write Eq. 22.33 as

$$\delta V(\mathbf{S}(t)) = \sum_{i=1}^{n} L_i\,\delta X_i + \frac{1}{2}\sum_{i=1}^{n}\lambda_i\widetilde{X}_i \quad \text{with} \quad \delta X_i \sim \mathrm{N}(0,1),\ \widetilde{X}_i \sim \chi^2(1) \ .$$

However, the $\widetilde{X}_i$ are not independent random variables since we obviously have

$$\widetilde{X}_i = (\delta X_i)^2 \ \ \forall i \ .$$

We need to re-write δV in such a way that every term appearing is statistically independent of every other term. First note that a δX_i is independent of every other term in δV if and only if the corresponding eigenvalue λ_i is zero, since in this case the corresponding $\widetilde{X}_i$ does not appear in the sum. We will emphasize this by introducing the index set J which contains only the indices of non-zero eigenvalues:

$$J = \left\{1, \ldots, n \,\middle|\, \lambda_j \neq 0\right\} \,. \tag{22.37}$$

With this index set we can write[8]

$$\begin{aligned} \delta V(\mathbf{S}(t)) &= \sum_{i \notin J} L_i \, \delta X_i + \sum_{j \in J} L_j \, \delta X_j + \frac{1}{2} \sum_{j \in J} \lambda_j \delta X_j^2 \\ &= \sum_{i \notin J} L_i \, \delta X_i + \sum_{j \in J} \left[L_j \, \delta X_j + \frac{1}{2} \lambda_j \delta X_j^2 \right] \,. \end{aligned} \tag{22.38}$$

The first sum in Eq. 22.38 is actually a sum of normally distributed random variables and as such is again a normally distributed random variable which we denote by u_0. The expectation of this random variable can be calculated as

$$\mathrm{E}\left[u_0\right] = \mathrm{E}\left[\sum_{i \notin J} L_i \, \delta X_i\right] = \sum_{i=1}^{n} L_i \underbrace{\mathrm{E}\left[\delta X_i\right]}_{0} = 0 \,.$$

and the variance is

$$\begin{aligned} \mathrm{var}\left[u_0\right] &= \mathrm{var}\left[\sum_{i \notin J} L_i \, \delta X_i\right] = \sum_{i,j \notin J} \mathrm{cov}\left[L_i \, \delta X_i, \, L_j \, \delta X_j\right] \\ &= \sum_{i,j \notin J} L_i L_j \underbrace{\mathrm{cov}\left[\delta X_i, \, \delta X_j\right]}_{\delta_{ij}} = \sum_{i \notin J} L_i^2 \,. \end{aligned}$$

[8] The notation $i \notin J$ denotes all indices i with eigenvalue $\lambda_i = 0$, i.e., the set $\{1, \ldots, n \,|\lambda_i = 0\}$.

with the components L_i of the transformed sensitivity vector $\mathbf{L}$ defined in Eq. 22.35. Thus

$$u_0 := \sum_{i \notin J} L_i \, \delta X \sim \mathrm{N}\left(0, \sum_{i \notin J} L_i^2\right) .$$

Consider now the sums over $j \in J$ in Eq. 22.38. To combine the *dependent* random numbers δX_j and δX_j^2 in the square brackets of Eq. 22.38 into one single random number, we complete the square for each $j \in J$:

$$\frac{1}{2}\lambda_j \delta X_j^2 + L_j \, \delta X_j = \frac{1}{2}\lambda_j \left(\delta X_j^2 + 2\frac{L_j}{\lambda_j}\, \delta X_j\right) = \frac{1}{2}\lambda_j \left(\delta X_j + \frac{L_j}{\lambda_j}\right)^2 - \frac{L_j^2}{2\lambda_j} .$$

Since δX_j is a standard normal random variable, we have

$$\delta X_j \sim \mathrm{N}(0, 1) \implies \delta X_j + \frac{L_j}{\lambda_j} \sim \mathrm{N}\left(\frac{L_j}{\lambda_j}, 1\right) .$$

Therefore, according to Eq. A.94 in Sect. A.4.6 $u_j := \left(\delta X_j + L_j/\lambda_j\right)^2$ has a *non-central* χ^2-distribution with one degree of freedom and non-central parameter L_j^2/λ_j^2:

$$\left(\delta X_j + \frac{L_j}{\lambda_j}\right)^2 =: u_j \sim \chi^2\left(1, \frac{L_j^2}{\lambda_j^2}\right) \quad \forall j \in J .$$

In summary, δV has now become a sum of non-central χ^2-distributed random variables u_j plus a normally distributed random variable u_0 (plus a constant), where all the random variables appearing are independent of each other:

$$\delta V(\mathbf{S}(t)) = u_0 + \frac{1}{2}\sum_{j \in J} \lambda_j u_j \underbrace{- \frac{1}{2}\sum_{j \in J} L_j^2/\lambda_j}_{\text{constant}} \tag{22.39}$$

$$\text{mit } \; u_0 \sim \mathrm{N}(0, \sum_{i \notin J} L_i^2) , \;\; u_j \sim \chi^2\left(1, (L_j/\lambda_j)^2\right) , \;\; j \in J .$$

The problem now consists in determining the distribution of the sum of independent but *differently* distributed random variables (or at least its percentiles). According to Eq. 21.4, the value at risk at a specified confidence level is then computed with precisely these percentiles or, equivalently, by inverting the cumulative distribution function of δV.

22.3.4 Moments of the Portfolio Value Distribution

We begin by calculating the moments of the random variable δV. The first moments, the expectation, the variance, etc. (see Eqs. A.22 and A.23) have intuitive interpretations and their explicit forms provide an intuitive conception of the distribution. In addition, approximations for the distribution (the Johnson approximation) and for the percentiles (the Cornish-Fisher approximation) will later be presented which can be computed with the moments.

The *Moment Generating Function* is a very useful tool for calculating the moments. The moment generating function (abbreviated as *MGF*) G_x of a random variable x with density function $\mathrm{pdf}(x)$ is defined in Sect. A.3.1 by

$$G_x(s) \equiv \mathrm{E}[e^{sx}] = \int_{-\infty}^{\infty} e^{sx}\,\mathrm{pdf}(x)dx = \sum_{n=0}^{\infty} \frac{s^n}{n!}\,\mathrm{E}[x^n]\,, \tag{22.40}$$

where in the last step the exponential function e^{sx} has been expanded in its Taylor series. The name *moment generating function* stems from the fact, stated mathematically in Eq. A.27, that the derivatives of the function $G_x(s)$ with respect to s evaluated at $s = 0$ generate the moments of the random variable x

$$\mathrm{E}[x^n] = \left.\frac{\partial^n G_x(s)}{\partial s^n}\right|_{s=0}\,. \tag{22.41}$$

As demonstrated in Sect. A.3.1, the MGF can be explicitly computed for many distributions from their integral representation, Eq. A.25. In particular, according to Eq. A.57, the MGF of u_0, i.e., the MGF of a $\mathrm{N}(0, \sum_{i\notin J} L_i^2)$ distributed random variable is given by

$$G_{u_0}(s) = \exp\left(\frac{1}{2}\,s^2 \sum_{i\notin J} L_i^2\right) \tag{22.42}$$

while Eq. A.95 gives the MGF of u_j, i.e., of a non-central χ^2-distribution with one degree of freedom $\chi^2(1, (L_j/\lambda_j)^2)$

$$G_{u_j}(s) = \frac{1}{\sqrt{1-2s}} \exp\left\{\frac{s}{1-2s}\frac{L_j^2}{\lambda_j^2}\right\}, \quad j \in J \ . \tag{22.43}$$

This function is well-defined for $s < 1/2$, which is sufficient for our needs since as is clear from Eq. 22.41 that we are particularly interested in values of s in a neighborhood of zero.

The usefulness of the MGF in the calculation of the distribution of δV in Eq. 22.39 stems from the fact that according to Eq. A.30 the MGF of the sum of *independent* random variables x, y is simply the product of the MGFs of each of the random variables:

$$G_{x+y}(s) = G_x(s)\, G_y(s) \tag{22.44}$$

and that furthermore, from Eq. A.31

$$G_{ax+b}(s) = e^{bs} G_x(as) \ . \tag{22.45}$$

for all non-stochastic values a, b and random variables x. The MGF of δV can thus be written as the product of the each of the MGFs appearing in the sum:

$$G_{\delta V}(s) = \exp\left\{-s\sum_{j\in J}\frac{L_j^2}{2\lambda_j}\right\} G_{u_0}(s) \prod_{j\in J} G_{u_j}\left(\frac{1}{2}\lambda_j s\right) \ .$$

The MGFs of each of the individual random variables are given explicitly in Eqs. 22.42 and 22.43. Substituting s by $\lambda_j s/2$ in Eq. 22.43 yields the required MGF for the argument $\lambda_j s/2$:

$$G_{u_j}\left(\frac{1}{2}\lambda_j s\right) = \frac{1}{\sqrt{1-\lambda_j s}} \exp\left\{\frac{L_j^2}{2\lambda_j}\frac{s}{1-\lambda_j s}\right\} \ .$$

We thus obtain an *explicit* expression for the moment generating function of the distribution of the portfolio's value changes:

$$\begin{aligned} G_{\delta V}(s) &= \exp\left\{-s\sum_{j\in J}\frac{L_j^2}{2\lambda_j}\right\}\exp\left(\frac{1}{2}s^2\sum_{i\notin J}L_i^2\right)\prod_{j\in J}\frac{1}{\sqrt{1-\lambda_j s}}\exp\left\{\frac{L_j^2}{2\lambda_j}\frac{s}{1-\lambda_j s}\right\} \\ &= \exp\left(\frac{1}{2}s^2\sum_{i\notin J}L_i^2\right)\prod_{j\in J}\left[\frac{1}{\sqrt{1-\lambda_j s}}\exp\left\{\frac{L_j^2}{2\lambda_j}\frac{s}{1-\lambda_j s}\right\}\exp\left\{-s\frac{L_j^2}{2\lambda_j}\right\}\right] \\ &= \exp\left(\frac{1}{2}s^2\sum_{i\notin J}L_i^2\right)\prod_{j\in J}\left[\frac{1}{\sqrt{1-\lambda_j s}}\exp\left\{\frac{L_j^2}{2\lambda_j}s\left(\frac{1}{1-\lambda_j s}-1\right)\right\}\right]. \end{aligned}$$

Using the same denominator in the second exp-function finally yields

$$G_{\delta V}(s) = \exp\left(\frac{1}{2}s^2\sum_{i\notin J}L_i^2\right)\prod_{j\in J}\frac{1}{\sqrt{1-\lambda_j s}}\exp\left\{\frac{1}{2}L_j^2\frac{s^2}{1-\lambda_j s}\right\}. \tag{22.46}$$

This can be simplified even further by the following trick: Since $\lambda_i = 0$ for all $i \notin J$, we can re-write the first exp-function in the following way:

$$\exp\left(\frac{1}{2}s^2\sum_{i\notin J}L_i^2\right) = \prod_{i\notin J}\exp\left(\frac{1}{2}s^2L_i^2\right) = \prod_{i\notin J}\frac{1}{\sqrt{1-\lambda_i s}}\exp\left\{\frac{1}{2}L_i^2\frac{s^2}{1-\lambda_i s}\right\}.$$

Using this form in Eq. 22.46 allows us to write δV very compactly as a product over *all* indexes $j = 1, \ldots, n$

$$G_{\delta V}(s) = \prod_{j=1}^{n}\frac{1}{\sqrt{1-\lambda_j s}}\exp\left\{\frac{1}{2}L_j^2\frac{s^2}{1-\lambda_j s}\right\}. \tag{22.47}$$

This function is well-defined for all $s < \min_{j\in J}\left(\frac{1}{2|\lambda_i|}\right)$, which is sufficient for our needs since because of Eq. 22.41, we are only interested in values of s which are close to zero.

Now, using Eq. 22.41, arbitrary moments of δV can be computed. We start the calculation of the first moment by introducing the abbreviation

$$a_j := \frac{1}{2} L_j^2 \frac{s^2}{1 - \lambda_j s} .$$

Application of the well-known *product rule* yields

$$\begin{aligned} \mathrm{E}[\delta V] &= \left. \frac{\partial G_{\delta V}(s)}{\partial s} \right|_{s=0} \\ &= \left. \frac{\partial}{\partial s} \prod_{j=1}^{n} \frac{e^{a_j}}{\sqrt{1 - \lambda_j s}} \right|_{s=0} \\ &= \left. \sum_{j=1}^{n} \left(\frac{\partial}{\partial s} \frac{e^{a_j}}{\sqrt{1 - \lambda_j s}} \right) \prod_{k=1, k \neq j}^{n} \frac{e^{a_k}}{\sqrt{1 - \lambda_k s}} \right|_{s=0} . \end{aligned}$$

The derivative we need to calculate is

$$\begin{aligned} \frac{\partial}{\partial s} \frac{e^{a_j}}{\sqrt{1 - \lambda_j s}} &= e^{a_j} \frac{\partial}{\partial s} \frac{1}{\sqrt{1 - \lambda_j s}} + \frac{1}{\sqrt{1 - \lambda_j s}} \frac{\partial}{\partial s} e^{a_j} \\ &= \frac{\frac{1}{2} \lambda_j e^{a_j}}{\left(1 - \lambda_j s\right)^{3/2}} + \frac{\frac{1}{2} L_j^2 e^{a_j}}{\sqrt{1 - \lambda_j s}} \left(\frac{2s}{1 - \lambda_j s} + \frac{\lambda_j \, s^2}{\left(1 - \lambda_j s\right)^2} \right) \\ &= \frac{1}{2} \frac{e^{a_j}}{\left(1 - \lambda_j s\right)^{3/2}} \left(\lambda_j + 2 L_j^2 s + \lambda_j L_j^2 \frac{s^2}{1 - \lambda_j s} \right) . \end{aligned}$$

For $s = 0$ almost all terms vanish and we are left with $\lambda_j / 2$. Thus $\mathrm{E}[\delta V]$ becomes simply

$$\mathrm{E}[\delta V] = \left. \sum_{j=1}^{n} \frac{1}{2} \lambda_j \prod_{k \in J, k \neq j} \frac{e^{a_k}}{\sqrt{1 - \lambda_k s}} \right|_{s=0} = \frac{1}{2} \sum_{j=1}^{n} \lambda_j .$$

The expectation of the portfolio's value changes in the delta-gamma approximation Eq. 22.4 is thus just half the sum of the eigenvalues of the transformed gamma matrix $\mathbf{M}$. This is by definition half the *trace* of the eigenvalue matrix $\boldsymbol{\lambda}$.

With Eqs. 22.35 and 22.36 we arrive at the conclusion that the expectation of δV equals half the trace of the product of the gamma matrix and the covariance matrix[9]:

$$\mathrm{E}[\delta V] = \frac{1}{2}\,\mathrm{tr}\,(\boldsymbol{\lambda}) = \frac{1}{2}\,\mathrm{tr}\left(\mathbf{D}^T\,\widetilde{\boldsymbol{\Gamma}}\,\mathbf{D}\right) = \frac{1}{2}\,\mathrm{tr}\left(\widetilde{\boldsymbol{\Gamma}}\,\mathbf{D}\mathbf{D}^T\right) = \frac{1}{2}\,\mathrm{tr}\left(\widetilde{\boldsymbol{\Gamma}}\,\delta\boldsymbol{\Sigma}\right)\,. \tag{22.48}$$

Note that the drifts of all risk factors have been neglected (see Eqs. 22.19 and 22.1). The *risk factors* are thus all approximated to be *drift-free*. But then, for a portfolio depending only *linearly* on the risk factors (or in the linear approximation of the delta-normal method) the expectation (the drift) of the portfolio value changes also equals zero. In Eq. 22.48, the expectation (the drift) of the portfolio changes is *not* zero because *non-linear* effects were taken into consideration. It is readily seen that the gamma matrix gives rise to the drift of δV in contrast to the linear sensitivities $\widetilde{\boldsymbol{\Delta}}$ which do *not* appear in 22.48.

To find out more about the distribution of δV, we proceed by computing its variance. According to Eq. A.5 the variance is the second central moment which can be calculated via Eq. A.29:

$$\begin{aligned}
\mathrm{var}[\delta V] &= \mathrm{E}[(\delta V - \mathrm{E}[\delta V])^2] \\
&= \frac{\partial^2}{\partial s^2}\exp\left(-s\mathrm{E}[\delta V]\right) G_{\delta V}(s)\Big|_{s=0} \\
&= \frac{\partial^2}{\partial s^2}\exp\left(-s\frac{1}{2}\sum_{i=1}^{n}\lambda_i\right)\prod_{j=1}^{n}\frac{\exp\left(\frac{1}{2}L_j^2\frac{s^2}{1-\lambda_j s}\right)}{\sqrt{1-\lambda_j s}}\Bigg|_{s=0} \\
&= \frac{\partial^2}{\partial s^2}\prod_{j=1}^{n}\frac{1}{\sqrt{1-\lambda_j s}}\exp\left(\frac{1}{2}L_j^2\frac{s^2}{1-\lambda_j s} - \frac{1}{2}\lambda_j s\right)\Bigg|_{s=0} \\
&= \frac{\partial^2}{\partial s^2}\prod_{j=1}^{n} a_j\Bigg|_{s=0}
\end{aligned} \tag{22.49}$$

[9]Here the well-known cyclic property of the trace has been used: $\mathrm{tr}\,(\mathbf{ABC}) = \mathrm{tr}\,(\mathbf{BCA})$ for arbitrary matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$.

with the abbreviation

$$a_j := \frac{1}{\sqrt{1-\lambda_j s}} \exp\left(\frac{1}{2} L_j^2 \frac{s^2}{1-\lambda_j s} - \frac{1}{2}\lambda_j s\right) .$$

The second derivative of this product is quite involved. We nonetheless present it explicitly here to demonstrate how moments of δV can be determined in practice. Such moments are needed quite often, for instance for the Cornish-Fisher expansion. We start by repeatedly applying the product rule to arrive at

$$\begin{aligned}
\frac{\partial^2}{\partial s^2} \prod_{j=1}^{n} a_j &= \frac{\partial}{\partial s} \left(\sum_{j=1}^{n} \frac{\partial a_j}{\partial s} \prod_{k=1, k\neq j}^{n} a_k \right) \\
&= \sum_{j=1}^{n} \left(\frac{\partial}{\partial s} \frac{\partial a_j}{\partial s} \right) \prod_{k=1, k\neq j}^{n} a_k + \sum_{j=1}^{n} \frac{\partial a_j}{\partial s} \frac{\partial}{\partial s} \prod_{k=1, k\neq j}^{n} a_k \\
&= \sum_{j=1}^{n} \frac{\partial^2 a_j}{\partial s^2} \prod_{k=1, k\neq j}^{n} a_k + \sum_{j=1}^{n} \frac{\partial a_j}{\partial s} \sum_{k=1, k\neq j}^{n} \frac{\partial a_k}{\partial s} \prod_{\substack{m=1, \\ m\neq k,\, m\neq j}}^{n} a_m .
\end{aligned} \tag{22.50}$$

Thus, we mainly have to differentiate a_j. For ease of notation, we introduce yet another abbreviation, namely

$$b_j := \frac{1}{2} L_j^2 \frac{s^2}{1-\lambda_j s} - \frac{1}{2}\lambda_j s \implies a_j = \frac{e^{b_j}}{\sqrt{1-\lambda_j s}} .$$

The first derivative with respect to a_j is now calculated as

$$\begin{aligned}
\frac{\partial a_j}{\partial s} &= \frac{1}{\sqrt{1-\lambda_j s}} \frac{\partial e^{b_j}}{\partial s} + e^{b_j} \frac{\partial}{\partial s} \frac{1}{\sqrt{1-\lambda_j s}} \\
&= \frac{e^{b_j}}{\sqrt{1-\lambda_j s}} \left(\frac{s\, L_j^2}{1-\lambda_j s} + \frac{\frac{1}{2} s^2 L_j^2 \lambda_j}{\left(1-\lambda_j s\right)^2} - \frac{\lambda_j}{2} \right) + \frac{\frac{1}{2} e^{b_j} \lambda_j}{\left(1-\lambda_j s\right)^{3/2}} \\
&= \frac{s L_j^2 e^{b_j}}{\left(1-\lambda_j s\right)^{3/2}} + \frac{\frac{1}{2} s^2 L_j^2 e^{b_j} \lambda_j}{\left(1-\lambda_j s\right)^{5/2}} - \frac{\frac{1}{2} e^{b_j} \lambda_j}{\left(1-\lambda_j s\right)^{1/2}} + \frac{\frac{1}{2} e^{b_j} \lambda_j}{\left(1-\lambda_j s\right)^{3/2}} .
\end{aligned} \tag{22.51}$$

For $s = 0$ the first two terms in the last line vanish and the last two terms just compensate each other so that $\partial a_j/\partial s$ vanishes completely at $s = 0$:

$$\left.\frac{\partial a_j}{\partial s}\right|_{s=0} = -\frac{1}{2}e^{b_j}\lambda_j + \frac{1}{2}e^{b_j}\lambda_j = 0\,. \tag{22.52}$$

Therefore only the term involving the second derivative in Eq. 22.50 contributes to Eq. 22.49. Using the result 22.51, this second derivative is explicitly:

$$\begin{aligned}
\frac{\partial^2 a_j}{\partial s^2} &= \frac{\partial}{\partial s}\left[\frac{sL_j^2 e^{b_j}}{(1-\lambda_j s)^{3/2}} + \frac{\frac{1}{2}s^2 L_j^2 e^{b_j}\lambda_j}{(1-\lambda_j s)^{5/2}} - \frac{\frac{1}{2}e^{b_j}\lambda_j}{(1-\lambda_j s)^{1/2}} + \frac{\frac{1}{2}e^{b_j}\lambda_j}{(1-\lambda_j s)^{3/2}}\right] \\
&= L_j^2\frac{\partial e^{b_j}}{\partial s}\frac{s}{(1-\lambda_j s)^{3/2}} + L_j^2 e^{b_j}\frac{3}{2}\frac{\lambda_j s}{(1-\lambda_j s)^{5/2}} + L_j^2 e^{b_j}\frac{1}{(1-\lambda_j s)^{3/2}} \\
&\quad + \frac{1}{2}L_j^2\frac{\partial e^{b_j}}{\partial s}\frac{\lambda_j s^2}{(1-\lambda_j s)^{5/2}} + \frac{5}{4}L_j^2 e^{b_j}\frac{\lambda_j^2 s^2}{(1-\lambda_j s)^{7/2}} + L_j^2 e^{b_j}\frac{\lambda_j s}{(1-\lambda_j s)^{5/2}} \\
&\quad - \frac{1}{2}\frac{\partial e^{b_j}}{\partial s}\frac{\lambda_j}{(1-\lambda_j s)^{1/2}} - \frac{1}{4}e^{b_j}\frac{\lambda_j^2}{(1-\lambda_j s)^{3/2}} \\
&\quad + \frac{1}{2}\frac{\partial e^{b_j}}{\partial s}\frac{\lambda_j}{(1-\lambda_j s)^{3/2}} + \frac{3}{4}e^{b_j}\frac{\lambda_j^2}{(1-\lambda_j s)^{3/2}}\,.
\end{aligned}$$

Most terms above have s as a factor. They all vanish for $s = 0$. The only terms remaining are:

$$\begin{aligned}
\left.\frac{\partial^2 a_j}{\partial s^2}\right|_{s=0} &= L_j^2 e^{b_j} - \frac{1}{2}\frac{\partial e^{b_j}}{\partial s}\lambda_j - \frac{1}{4}e^{b_j}\lambda_j^2 + \frac{1}{2}\frac{\partial e^{b_j}}{\partial s}\lambda_j + \frac{3}{4}e^{b_j}\lambda_j^2 \\
&= \underbrace{e^{b_j}}_{1\text{ for }s=0}\left(L_j^2 + \frac{1}{2}\lambda_j^2\right)\,.
\end{aligned} \tag{22.53}$$

Inserting all these results into Eq. 22.49 finally yields

$$\mathrm{var}[\delta V] = \sum_{j=1}^{n}\frac{\partial^2 a_j}{\partial s^2} + \sum_{j=1}^{n}\frac{\partial a_j}{\partial s}\sum_{k=1,k\neq j}^{n}\left.\frac{\partial a_k}{\partial s}\right|_{s=0}$$

$$= \sum_{j=1}^{n} \frac{\partial^2 a_j}{\partial s^2}\bigg|_{s=0}$$
$$= \sum_{j=1}^{n} \left(L_j^2 + \frac{1}{2}\lambda_j^2 \right) ,$$

In the first step we used Eq. 22.50 and $a_k|_{s=0} = 1$. In the second step the result 22.52 was inserted and in the third step the result 22.53. The sum $\sum L_j^2$ is just the square of the transformed sensitivity *vector* and $\sum \lambda_j^2$ is the trace of the square of the *matrix* of eigenvalues, i.e.,

$$\text{var}[\delta V] = \mathbf{L}^T\mathbf{L} + \frac{1}{2}\,\text{tr}\left(\boldsymbol{\lambda}^2\right) .$$

Finally, making use of the transformations in Eq. 22.35 and applying Eq. 22.36, the variance of the portfolio's value change in the framework of the delta-gamma method becomes

$$\begin{aligned}\text{var}[\delta V] &= \widetilde{\boldsymbol{\Delta}}^T \mathbf{D}\mathbf{D}^T \widetilde{\boldsymbol{\Delta}} + \frac{1}{2}\,\text{tr}\left(\mathbf{D}^T \widetilde{\boldsymbol{\Gamma}}\, \mathbf{D}\mathbf{D}^T \widetilde{\boldsymbol{\Gamma}}\, \mathbf{D}\right) \\ &= \widetilde{\boldsymbol{\Delta}}^T \delta\boldsymbol{\Sigma} \widetilde{\boldsymbol{\Delta}} + \frac{1}{2}\,\text{tr}\left(\widetilde{\boldsymbol{\Gamma}}\, \delta\boldsymbol{\Sigma} \widetilde{\boldsymbol{\Gamma}}\, \delta\boldsymbol{\Sigma}\right) . \end{aligned} \tag{22.54}$$

Note that the first term resulting from the linear portfolio sensitivities $\widetilde{\boldsymbol{\Delta}}$ is identical to the portfolio variance in the delta-normal method (see Eq. 22.14). The non-linear sensitivities $\widetilde{\boldsymbol{\Gamma}}$ effect a correction of the linear portfolio variance which has a form similar to the drift correction from the non-linear term in Eq. 22.48. While in Eq. 22.48, the trace of the product of the gamma matrix and the covariance matrix was relevant, now the trace of the *square* of this product is required for computing the variance.

The variance is the second *central* moment of the random variable. The *central moments* μ_i of a random variable are defined in general terms in Eq. A.23 as the "expectation of powers of the deviation from the expectation":

$$\mu_i := \text{E}[(\delta V - \text{E}[\delta V])^i], \;\; i > 1 .$$

Analogously to the approach for the first two moments demonstrated above, we can continue to calculate the further moments of δV. The first (central)

moments are compiled here:

$$\begin{aligned}
\mu &:= \mathrm{E}[\delta V] = \frac{1}{2}\,\mathrm{tr}\,\left(\widetilde{\mathbf{\Gamma}}\,\delta\mathbf{\Sigma}\right)\\
\mu_2 &= \mathrm{E}[(\delta V - \mathrm{E}[\delta V])^2] = \widetilde{\mathbf{\Delta}}^T\delta\mathbf{\Sigma}\widetilde{\mathbf{\Delta}} + \frac{1}{2}\,\mathrm{tr}\,\left((\widetilde{\mathbf{\Gamma}}\,\delta\mathbf{\Sigma})^2\right)\\
\mu_3 &= \mathrm{E}[(\delta V - \mathrm{E}[\delta V])^3] = 3\widetilde{\mathbf{\Delta}}^T\delta\mathbf{\Sigma}\,\widetilde{\mathbf{\Gamma}}\,\boldsymbol{\delta}\mathbf{\Sigma}\,\widetilde{\Delta} + \mathrm{tr}\,\left((\widetilde{\mathbf{\Gamma}}\,\delta\mathbf{\Sigma})^3\right) \qquad (22.55)\\
\mu_4 &= \mathrm{E}[(\delta V - \mathrm{E}[\delta V])^4] = 12\widetilde{\mathbf{\Delta}}^T\delta\mathbf{\Sigma}\,(\widetilde{\mathbf{\Gamma}}\,\boldsymbol{\delta}\mathbf{\Sigma})^2\,\widetilde{\mathbf{\Delta}} + 3\,\mathrm{tr}\,\left((\widetilde{\mathbf{\Gamma}}\,\delta\mathbf{\Sigma})^4\right) + 3\mu_2^2\,.
\end{aligned}$$

In this way, a great deal of additional information about the distribution of δV can be generated. For instance *skewness* and *kurtosis* of the distribution of δV are[10]

$$\begin{aligned}
\text{Schiefe} &\equiv \frac{\mu_3}{\mu_2^{3/2}} = \frac{3\widetilde{\mathbf{\Delta}}^T\delta\mathbf{\Sigma}\,\widetilde{\mathbf{\Gamma}}\,\boldsymbol{\delta}\mathbf{\Sigma}\,\widetilde{\mathbf{\Delta}} + \mathrm{tr}\,(\widetilde{\mathbf{\Gamma}}\,\delta\mathbf{\Sigma})^3}{\left(\widetilde{\mathbf{\Delta}}^T\delta\mathbf{\Sigma}\widetilde{\mathbf{\Delta}} + \frac{1}{2}\,\mathrm{tr}\,(\widetilde{\mathbf{\Gamma}}\,\delta\mathbf{\Sigma})^2\right)^{3/2}}\\
\text{Kurtosis} &\equiv \frac{\mu_4}{\mu_2^2} = \frac{12\widetilde{\mathbf{\Delta}}^T\delta\mathbf{\Sigma}\,(\widetilde{\mathbf{\Gamma}}\,\boldsymbol{\delta}\mathbf{\Sigma})^2\,\widetilde{\mathbf{\Delta}} + 3\,\mathrm{tr}\,\left((\widetilde{\mathbf{\Gamma}}\,\delta\mathbf{\Sigma})^4\right) + 3\mu_2^2}{\left(\widetilde{\mathbf{\Delta}}^T\delta\mathbf{\Sigma}\widetilde{\mathbf{\Delta}} + \frac{1}{2}\,\mathrm{tr}\,(\widetilde{\mathbf{\Gamma}}\,\delta\mathbf{\Sigma})^2\right)^2}\,.
\end{aligned}$$

A *percentile*, however, is needed for the computation of the value at risk as given in Eq. 21.4.

Johnson Transformation

Computation of a percentile necessitates knowledge of the distribution of the random variable *directly* and not of its moments. In order to be able to proceed, we could assume a particular functional form of the distribution and then establish a relation between the parameters of this functional form and the moments of the random variable via *moment matching*. Since the moments, as shown above, can be explicitly computed, the parameters of the assumed distribution can thus be determined. For example, if we assume that a random variable is normally or lognormally distributed, we would take Eqs. 22.48 and 22.54 as parameter values. Additional functional forms for

[10] Recall that a normal distribution has skewness 0 and kurtosis 3, see Eq. A.59.

approximating the distribution of δV were suggested by Johnson [114]. These *Johnson transformations* have four parameters which can be determined from the first four moments in Eq. 22.55. They represent a substantially better approximation than, for example a lognormal distribution.

Cornish-Fisher Expansion

One possibility of approximating the *percentiles* of a distribution from its *central moments* and the percentiles $Q^{\mathrm{N}(0,1)}$ of the standard normal distribution is the *Cornish-Fisher expansion*. Since this expansion makes use of the *standard* normal distribution, we must first transform δV into a centered and normalized random variable $\widetilde{\delta V}$ with expectation 0 and variance 1. This is accomplished by defining

$$\widetilde{\delta V} := \frac{\delta V - \mathrm{E}[\delta V]}{\sqrt{\mathrm{var}[\delta V]}} = \frac{\delta V - \mu}{\sqrt{\mu_2}} \; .$$

The percentile of the distribution of $\widetilde{\delta V}$ can now be approximated with the Cornish-Fisher expansion [41, 196] as follows

$$\begin{aligned} Q^{\mathrm{cpf}_{\widetilde{\delta V}}} \approx\; & Q^{\mathrm{N}(0,1)} + \frac{1}{6}\left[(Q^{\mathrm{N}(0,1)})^2 - 1\right]\frac{\mu_3}{\mu_2^{3/2}} \\ & + \frac{1}{24}\left[(Q^{\mathrm{N}(0,1)})^3 - 3Q^{\mathrm{N}(0,1)}\right]\left(\frac{\mu_4}{\mu_2^2} - 3\right) \\ & - \frac{1}{36}\left[2\,(Q^{\mathrm{N}(0,1)})^3 - 5Q^{\mathrm{N}(0,1)}\right]\left(\frac{\mu_3}{\mu_2^{3/2}}\right)^2 \; , \end{aligned} \tag{22.56}$$

where the expansion is taken up to the order, which uses only the first four moments from Eq. 22.55. The probability that $\widetilde{\delta V}$ is less than a number a is, naturally, the same as the probability that δV is less than $\mu + \sqrt{\mu_2}\,a$. Thus

$$Q^{\mathrm{cpf}_{\delta V}} = \mu + \sqrt{\mu_2}\, Q^{\mathrm{cpf}_{\widetilde{\delta V}}} \; .$$

holds for the percentiles. From Eq. 21.4, the value at risk is thus

$$\mathrm{VaR}(c) = -Q_{1-c}^{\mathrm{cpf}_{\delta V}} = -\mu - \sqrt{\mu_2}\, Q_{1-c}^{\mathrm{cpf}_{\widetilde{\delta V}}} \; ,$$

where we now can use the approximation 22.56 for $Q_{1-c}^{\widetilde{\delta V}}$ since the percentiles of the standard normal distribution for the confidence level $(1 - c)$ are well known (see for instance Eq. 21.13).

22.3.5 Fourier-Transformation of the Portfolio Value Distribution

While a great deal of information about the distribution function can be gleaned from the computation of the moments with the moment generating function, we are still not able to calculate the distribution itself directly. *Characteristic functions* (CFs), however, generate the distribution directly (this can at least be accomplished numerically). As defined in Sect. A.3.2, the *characteristic function* Φ_x of a random variable x with density function $\mathrm{pdf}(x)$ is[11]

$$\Phi_x(s) \equiv \mathrm{E}[e^{isx}] = \int_{-\infty}^{\infty} e^{isx}\,\mathrm{pdf}(x)dx\ . \tag{22.57}$$

This is precisely the definition of the *Fourier transformation* of the density function. As demonstrated in Sect. A.3.2, the CFs of many random variables with density functions can be computed explicitly. In particular, the CF of u_0, i.e., the CF of a normally distributed random variable $\mathrm{N}(0, \sum_{i\notin J} L_i^2)$ is given by

$$\Phi_{u_0}(s) = \exp\left(-\frac{1}{2}s^2 \sum_{i\notin J} L_i^2\right) \tag{22.58}$$

And according to Eq. A.97, the CF of the u_j , i.e., of a non-central χ^2-distributed random variable with one degree of freedom $\chi^2(1, (L_j/\lambda_j)^2)$ is given by

$$\Phi_{u_j}(s) = \frac{1}{\sqrt{1-2is}} \exp\left\{\frac{is}{1-2is}\frac{L_j^2}{\lambda_j^2}\right\},\ j \in J,\ i \equiv \sqrt{-1}\ . \tag{22.59}$$

[11]Here, i denotes the *imaginary number* satisfying the property $i^2 = -1$, thus intuitively $i = \sqrt{-1}$.

Similarly to the moment generating function, the usefulness of the CF in computing the distribution of δV in Eq. 22.39 stems from the property A.34. The CF of a sum of independent random variables x, y is simply the product of the CFs of each of these variables:

$$\Phi_{x+y}(s) = \Phi_x(s)\,\Phi_y(s) \tag{22.60}$$

Furthermore, according to Eq. A.35,

$$\Phi_{ax+b}(s) = e^{ibs}\,\Phi_x(as)\ . \tag{22.61}$$

holds for all non-stochastic values a, b and random variables x. Thus the CF of δV can be expressed as the product of the CFs of the random variables appearing in the definition of δV:

$$\Phi_{\delta V}(s) = \exp\left\{-is\sum_{j\in J}\frac{L_j^2}{2\lambda_j}\right\}\Phi_{u_0}(s)\prod_{j\in J}\Phi_{u_j}\left(\frac{1}{2}\lambda_j s\right)\ . \tag{22.62}$$

The CFs of the individual random variables are given explicitly in Eqs. 22.58 and 22.59. We now have all information at our disposal to calculate an *explicit* expression for the characteristic function of the distribution of δV. The result is of course the same as Eq. 22.46 for the MGF with the obvious substitution $s \to is$:

$$\Phi_{\delta V}(s) = \exp\left(-\frac{1}{2}s^2\sum_{i\notin J}L_i^2\right)\prod_{j\in J}\frac{1}{\sqrt{1-i\lambda_j s}}\exp\left\{-\frac{1}{2}L_j^2\frac{s^2}{1-i\lambda_j s}\right\}\ . \tag{22.63}$$

Since $\lambda_i = 0$ for $i \notin J$,we can—as we did with the MGF—write the first exp-function as

$$\exp\left(-\frac{1}{2}s^2\sum_{i\notin J}L_i^2\right) = \prod_{i\notin J}\frac{1}{\sqrt{1-i\lambda_i s}}\exp\left\{-\frac{1}{2}L_i^2\frac{s^2}{1-i\lambda_i s}\right\}\ .$$

Thus, the characteristic function can be written as a product over *all* indexes j:

$$\Phi_{\delta V}(s) = \prod_{j=1}^{n} \frac{1}{\sqrt{1 - i\lambda_j s}} \exp\left\{-\frac{1}{2} L_j^2 \frac{s^2}{1 - i\lambda_j s}\right\} \text{ mit } i \equiv \sqrt{-1} \,. \tag{22.64}$$

In contrast to the moment generating function, there exists an inverse transformation for the characteristic function, namely the *inverse Fourier Transformation* (see Sect. A.3.2). From Eq. 22.64 the density function $\text{pdf}(\delta V)$ can thus be computed (at least numerically).

$$\begin{aligned} \text{pdf}_{\delta V}(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-isx} \Phi_{\delta V}(s) ds \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-isx} \prod_{j=1}^{n} \frac{\exp\left\{-\frac{1}{2} L_j^2 \frac{s^2}{1 - i\lambda_j s}\right\}}{\sqrt{1 - i\lambda_j s}} ds \,. \end{aligned} \tag{22.65}$$

The *cumulative* probability function of δV can now be obtained through the (numerical) integration of this probability density

$$\begin{aligned} \text{cpf}_{\delta V}(c) &\equiv \int_{-\infty}^{c} \text{pdf}_{\delta V}(x) dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{c} \int_{-\infty}^{\infty} e^{-isx} \prod_{j=1}^{n} \frac{\exp\left\{-\frac{1}{2} L_j^2 \frac{s^2}{1 - i\lambda_j s}\right\}}{\sqrt{1 - i\lambda_j s}} ds\, dx \,. \end{aligned}$$

A method which can likewise be applied in practice does not use the Fourier transformation of the density, but the Fourier transformation of the *cumulative* distribution directly:

$$F_x(s) \equiv \int_{-\infty}^{\infty} e^{isx} \,\text{cpf}(x) dx \,.$$

This Fourier transformation has the analogous properties to those indicated in Eqs. 22.61 and 22.60. Hence, the Fourier transformation of the cumulative distribution function of the portfolio's change is (up to a constant)

$$F_{\delta V}(s) \sim F_{u_0}(s) \prod_{j \in J} F_{u_j}\left(\frac{1}{2}\lambda_j s\right) \,. \tag{22.66}$$

analogous to Eq. 22.62. Here, the individual Fourier transformations of the *cumulative* distribution function of a normally distributed random variable (for u_0) and the *cumulative* distribution of the non-central χ^2 random variables with one degree of freedom (for the u_j) appear. These can only be computed numerically. Then, the results as shown in Eq. 22.66 are multiplied in Fourier space. Finally, the function $F_{\delta V}(s)$ thus obtained must be transformed back with the inverse Fourier transformation to obtain the cumulative probability function of δV as

$$\mathrm{cpf}_{\delta V}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-isx} F_{\delta V}(s) ds \ . \tag{22.67}$$

The Wiener-Chintchine theorem states that the approach just described is equivalent to taking the *convolution* of the cumulative distribution functions. The twice computed Fourier transformation is numerically preferable to computing the convolution. The recommended method for numerically performing Fourier transformations (or inverse Fourier transformations like Eq. 22.67) is the *fast Fourier transformation* (FFT). This method requires significantly fewer computations as compared to other common procedures for numerical integration.[12]

22.3.6 Monte Carlo Simulation of the Portfolio Value Distribution

Calculating the cumulative distribution of δV with characteristic functions involves complicated numerical procedures. Using moment-generating functions to calculate the moments we need additional assumptions and approximations to establish a relation between those moments and the distribution or the percentiles.[13] All methods introduced here therefore offer sufficient scope for error to creep into the calculations. Additionally, significant difficulties are often involved in calculating the gamma and covariance matrices. Hence, it is by all means legitimate to apply a simple Monte Carlo simulation, instead of the often complicated methods described above, to generate the distribution of δV. The statistical error in doing so is often no larger than the errors of the

[12] In contrast to other common numerical procedures, the FFT reduces the number of necessary multiplications from order $O(N^2)$ to $O(N \ \ln(N))$. See, for example [20] or [156].

[13] It should not be forgotten that the Delta-Gamma method itself is only an approximation of the portfolio's value obtained from the second-order Taylor series approximation.

above methods assuming of course that a sufficient number of simulation runs have been carried out.

For a Monte Carlo Simulation we proceed directly from Eq. 22.33. n normally distributed random numbers are generated with which the simulated change in the portfolio's value can be immediately computed with the expression

$$\delta V = \sum_{i=1}^{n} \left[L_i \, \delta X_i + \frac{1}{2} \lambda_i \delta X_i^2 \right] .$$

This procedure is repeated N times (as a rule, several thousand times) thus obtaining N simulated changes in the portfolio's value from which the distribution of δV can be approximated. The percentiles of this distribution can be approximated by simply sorting the simulated values of δV in increasing order as described in Sect. 23.1 (a detailed discussion of value at risk computations using the Monte Carlo Method can be found in this section).

Here, in contrast to the method described in Sect. 23.1, we do not simulate each single risk factor separately. Instead, the portfolio change is directly calculated based on Eq. 22.33. Therefore, no time consuming revaluation of the portfolio for the each generated scenarios is required. However, before the simulation can be performed, the eigenvalues of the transformed gamma matrix must be calculated by solving Eq. 22.22, and the transformed sensitivities L_i need to be determined as well. Because of Eqs. 22.35 and 22.34, both the Cholesky decomposition of the covariance matrix as well as the eigenvectors of the gamma matrix must be computed. The eigenvectors are determined by solving Eq. 22.21.