Quadratic VAR model

Michael Carter

■ Preliminaries

Introduction

Suppose we wish to quantify the value-at-risk of a Japanese metals trading firm that has exposure to forward and option positions in platinum. Some of the positions are denominated in USDs. We identify three key risk factors

- spot price of platinum in yen (X_1)
- a representative implied volatility of platinum (X_2)
- spot JPY/USD exchange rate (X_3)

and assume that $X = (X_1, X_2, X_3)$ is normally distributed $N(\mu, \Sigma)$ with μ =(53.150, 0.2670 ,107.80) and

$$\Sigma = \begin{pmatrix} 799600 & 1.074 & -48.91 \\ 1.074 & 7.056\ 10^{-5} & -3.875\ 10^{-5} \\ -48.91 & -3.875\ 10^{-5} & 0.4343 \end{pmatrix}$$

We value the portfolio using applicable forward and option pricing formulas, and quadratically approximate this as

$$Y = a + \Delta^T X + X^T \Gamma X$$

Our objective is to estimate certain quantiles of Y.

In general, we are presented with a quadratic function of a random vector X

$$Y = a + \Delta^T X + X^T \Gamma X$$

where $X \sim N_m(\mu, \Sigma)$. This might arise as above from a delta-gamma approximation to a VAR measure of a portfolio containing derivatives, or some alternative approximation. Then, as we will show, Y can be expressed as a linear combination of *independent* chi-squared and normal random variables. Based on this representation, we can calculate the moments and the distribution function of Y.

For exposition, let us start with a numerically simpler specific example. Suppose that X is multivariate normal with mean and variance

$$\mu = \{1, -1, 0\}; \ \Sigma = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 5 \end{pmatrix};$$

and that Y is given by

$$Y = a + \Delta^T X + X^T \Gamma X$$

with

```
a = 12; Δ = {18, 32, -12}; Γ = ( 3   6  -3
                                 6  16  -6
                                -3  -6   3 );
```

In other words, Y is determined by the following quadratic function of correlated normal random variables.

$$Y = 12 + 18 X_1 + 3 X_1^2 + 32 X_2 + 12 X_1 X_2 + 16 X_2^2 - 12 X_3 - 6 X_1 X_3 - 12 X_2 X_3 + 3 X_3^2$$

Define new random variables as follows

$$Z_1 = X_2 - \mu_2, \quad Z_2 = X_1 - \mu_1 + 2(X_2 - \mu_2) - (X_3 - \mu_3), \; Z_3 = -2(X_2 - \mu_2) + (X_3 - \mu_3)$$

The Z_i are *independent* standard normal random variables.

Exercise: Verify that the Z_1, Z_2, Z_3 are independent standard normal random variables.

The random vector $Z = (Z_1, Z_2, Z_3)$ is defined by the transformation

$$Z = A(X - \mu)$$

where

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -2 & 1 \end{pmatrix}$$

Z has mean 0 and variance-convariance matrix $A \Sigma A^T = I$.

Solving for X gives

$$X = A^{-1} Z + \mu$$

Substituting in the equation defining Y and simplifying, we have

$$\begin{aligned} Y &= 12 + 18 X_1 + 3 X_1^2 + 32 X_2 + 12 X_1 X_2 + 16 X_2^2 - 12 X_3 - 6 X_1 X_3 - 12 X_2 X_3 + 3 X_3^2 \\ &= 5 + 4 Z_1^2 + 12 Z_2 + 3 Z_2^2 + 6 Z_3 \end{aligned}$$

By "completing the square" for Z_2, we can express this as

$$\begin{aligned} Y &= 5 + 4 Z_1^2 + 3(Z_2^2 + 4 Z_2) + 6 Z_3 \\ &= 5 + 4 Z_1^2 + 3(Z_2^2 + 2)^2 - 12 + 6 Z_3 \\ &= -7 + 4 Z_1^2 + 3(Z_2^2 + 2)^2 + 6 Z_3 \end{aligned}$$

We have expressed Y as a linear combination of 3 *independent* random variables: $Z_1^2 \sim \chi^2(1, 0)$, $Z_2^2 \sim \chi^2(1, 4)$, $Z_3 \sim N(0, 1)$.

In mathematical terms, we have achieved two things.

- We have expressed Y in terms of independent standard normal random variables Z_i (*Cholesky decomposition*).
- We have diagonalized the quadratic form with respect to these variables so that there are no cross-terms $Z_i Z_j$ (*Principal axis theorem*).

The transformation A is the composition of these two steps. We now examine the conditions required for this in general.

Suppose that Y is a quadratic function of random vector X

$$Y = a + \Delta^T X + X^T \Gamma X$$

where $X \sim N_m(\mu, \Sigma)$.

Without loss of generality, we can assume that $\mu = 0$, since any mean effect can be incorporated into a new constant term α. Further, we can assume that Σ is positive definite. There exists a lower triangular matrix H such that $\Sigma = H H^T$ and $X = H \tilde{Z}$ where $\tilde{Z}$ is standard normal. This is known as the *Cholesky decomposition*. Substituting

$$Y = \alpha + \Delta^T H \tilde{Z} + (H \tilde{Z})^T \Gamma H \tilde{Z} = \alpha + D^T \tilde{Z} + \tilde{Z}^T G \tilde{Z}$$

where

$$D = H^T \Delta \qquad \text{and} \qquad G = H^T \Gamma H$$

Since Γ is symmetric, so is G and there exists an orthogonal matrix P such that

$$P^T G P = \Lambda \quad \text{or} \quad G = P \Lambda P^T$$

where Λ is a diagonal matrix containing the eigenvalues of G (*Spectral theorem*). Therefore

$$Y = \alpha + D^T \tilde{Z} + \tilde{Z}^T G \tilde{Z} = \alpha + D^T P P^T \tilde{Z} + \tilde{Z}^T P \Lambda P^T \tilde{Z} = \alpha + \mathrm{B}^T Z + Z^T \Lambda Z$$

where

$$\mathrm{B} = P^T D = P^T H^T \Delta \qquad \text{and} \qquad Z = P^T \tilde{Z}$$

Since P is orthogonal, $Z = (Z_1, Z_2, \ldots, Z_m)$ is also a vector of independent standard normal random variables. Therefore, Y can be expressed as a linear combination of normal and chi-squared random variables

$$Y = \alpha + \mathrm{B}^T Z + Z^T \Lambda Z = \alpha + \sum_{i=1}^{m} (\beta_i Z_i + \lambda_i Z_i^2)$$

We summarize in the following theorem.

Theorem *Suppose that Y is a quadratic function of random vector X*

$$Y = a + \Delta^T X + X^T \Gamma X$$

where $X \sim N_m(\mu, \Sigma)$*, and* Σ *is positive definite. Then there exists a linear transformation A such that*

$$X = A^{-1} Z + \mu \qquad \text{and} \qquad Y = \alpha + \mathrm{B}^T Z + Z^T \Lambda Z$$

where Λ *is a diagonal matrix and* $Z_1, Z_2, \ldots, Z_m$ *are independent standard normal random variables. Consequently Y can be written as*

$$Y = \alpha + \sum_{i=1}^{m} (\beta_i Z_i + \lambda_i Z_i^2)$$

This theorem shows that Y is a linear combination of independent normal and chi-square random variables *or* non-central chi-square random variables (see complementary lecture note *Quadratic functions of normal random variables*). Consequently, the characteristic function of Y can be readily determined, from which the exact distribution function can be calculated. Alternatively, a number of computationally easier approximations for the quantiles are available, including

- Cornish-Fisher expansion
- saddlepoint approximation

Both of these approximations are based upon the cumulant generating function.

The distribution of *Y*

Our first goal is to determine the moment generating function, from which we can derive both the cumulant generating function and the characteristic function. From the latter, we can obtain the distribution function by Fourier inversion.

■ Moment generating function

By direct integration, we can readily determine that the moment generating function of a single term of the form $\beta_i Z + \lambda_i Z^2$ is

$$M_i(t) = E\left[e^{\beta_i Z + \lambda_i Z^2}\right] = \int_{-\infty}^{\infty} e^{(\beta_i z + \lambda_i z^2) t} \phi(z)\, dz = \frac{1}{\sqrt{1 - 2 t \lambda_i}} e^{\frac{1}{2} \frac{\beta_i^2 t^2}{1 - 2 \lambda_i t}}$$

where

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

is the standard normal density function. Therefore, the moment generating function of $Y = \alpha + \sum_{i=1}^{m} (\beta_i Z_i + \lambda_i Z_i^2)$ is

$$M_Y(t) = E[e^{tY}] = e^{\alpha t} M_1(t) M_2(t) \ldots M_n(t) =$$
$$e^{\alpha t} \prod_{i=1}^{n} \left(\frac{1}{\sqrt{1 - 2 t \lambda_i}} e^{\frac{1}{2} \frac{\beta_i^2 t^2}{1 - 2 \lambda_i t}} \right) = e^{\alpha t} \operatorname{Exp}\left(\frac{1}{2} \sum_{i=1}^{n} \frac{\beta_i^2 t^2}{1 - 2 \lambda_i t} \right) \prod_{i=1}^{n} \left(\frac{1}{\sqrt{1 - 2 t \lambda_i}} \right)$$

from which the moments can be immediately derived by differentiation. (Holton 2003: 149 gives a complicated recursive formula for the moments of Y.)

In the previous example, $\alpha = 5$, $\mathrm{B} = (0, 12, 6)$ and $\Lambda = (4, 3, 0)$. The moments of Y are

EV[Y]	12
EV[Y^2]	374
EV[Y^3]	13328
EV[Y^4]	615900
EV[Y^5]	33217840

The variance of Y is therefore $E[Y^2] - E[Y]^2 = 374 - 144 = 230$.

■ The cumulant generating function

Cumulants are analogous to moments. The first cumulant is the same as the first moment (the expected value); the second and third cumulants are respectively the second (variance) and third central moments; but the higher cumulants are neither moments nor central moments, but rather more complicated polynomial functions of the moments.

The cumulant generating function is given by the log of the moment generating function, that is

$$K(t) = \log M(t) = \alpha t - \frac{1}{2} \sum_{i=1}^{n} \log(1 - 2 \lambda_i t) + \frac{1}{2} \sum_{i=1}^{n} \frac{\beta_i^2 t^2}{1 - 2 \lambda_i t}$$

from which the cumulants can be derived by differentiation.

In the previous example, with $\alpha = 5$, $\mathrm{B} = (0, 12, 6)$ and $\Lambda = (4, 3, 0)$, the cumulants of Y are

κ_1	12
κ_2	230
κ_3	3320
κ_4	78384
κ_5	2352768

■ Characteristic function

The characteristic function of any random variable Y is

$$\Psi(w) = E[e^{i\,w\,Y}]$$

Consequently, it can be obtained by substituting $t = i\,w$ into the moment generating function. The characteristic function of Y is

$$\Psi(w) = M(i\,w) = \operatorname{Exp}\left(i\,a\,w - \frac{1}{2}\sum_{i=1}^{n}\frac{w^2\,b_i^2}{1 - 2\,i\,w\,\lambda_i}\right)\prod_{i=1}^{n}\frac{1}{\sqrt{1 - 2\,i\,w\,\lambda_i}}$$

■ The distribution function

The distribution function of Y can be obtained by inverting the characteristic function (Holton 2003: 159, *Quadratic functions of normal random variables*)

$$F(y) = \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty}\frac{\operatorname{Im}(\Psi(w)\,e^{-i\,w\,y})}{w}\,dw$$

Substituting the characteristic function and simplifying gives

$$F(y) = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty}\frac{e^{A}\,\sin(B + C)}{D}\,dw$$

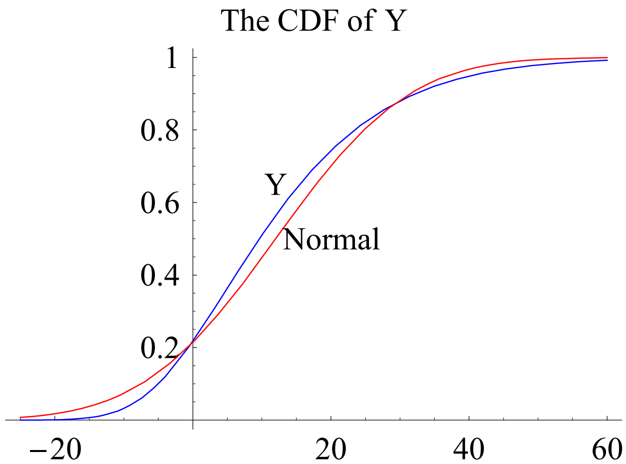
where

$$A = -\frac{w^2}{2}\sum_{i=1}^{m}\frac{\beta_i^2}{1 + 4\,\lambda_i^{\,2}\,w^2}\,, \qquad B = w\left(y - \alpha + w^2\sum_{i=1}^{m}\frac{\beta_i^2\,\lambda_i}{1 + 4\,\lambda_i^{\,2}\,w^2}\right),$$

$$C = \frac{1}{2}\sum_{i=1}^{m}\tan^{-1}(-2\,\lambda_i\,w), \qquad D = w\prod_{i=1}^{m}(1 + 4\,\lambda_i^{\,2}\,w^2\,)^{1/4}$$

For given values of β_i, λ_i ($i = 1, 2, \ldots, m$) *and* y, this expression can be numerically integrated to give a very accurate approximation of distribution function.

The cumulative distribution function of Y for the example with $\alpha = 5$, $\mathrm{B} = (0, 12, 6)$ and $\Lambda = (4, 3, 0)$ is depicted below, together with the CDF of a normal distribution with the same mean and variance (red).



From the distribution function, the quantiles can be computed directly. The *Fast Fourier Transform* provides a potential alternative method of computing the distribution function of Y.

Approximating the quantiles of *Y*

A number of methods have been suggested for approximating the quantiles of Y. Two of the most promising are the Cornish-Fisher expansion and saddlepoint approximations. Comparative studies suggest that the Cornish-Fisher expansion is computationally faster, while the saddlepoint approximation is more accurate, especially in the tails.

■ Cornish-Fisher expansion

The Cornish-Fisher expansion approximates the quantiles of a distribution as a polynomial of its cumulants. That is, the α quantile of Y is

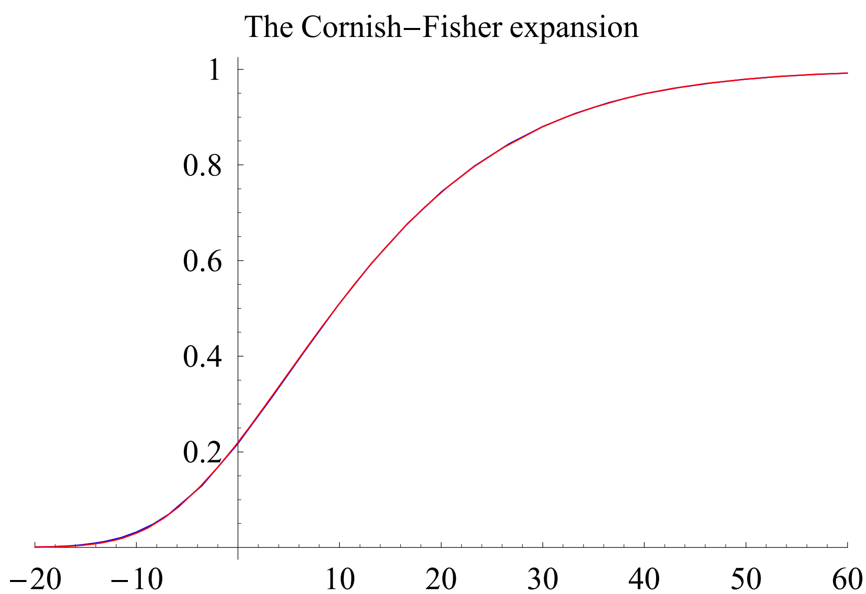
$$F_Y^{-1}(\alpha) \approx a_0 + a_1 \kappa_3 + a_2 \kappa_4 + a_3 \kappa_3^2 + a_4 \kappa_5 + a_5 \kappa_3 \kappa_4 + a_6 \kappa_3^3$$

where the coefficients $a_0,\ a_1,\ \ldots,\ a_6$ are polynomials of the quantiles of the standard normal distribution.

$$a_0 = \Phi^{-1}(\alpha),\ a_1 = \frac{1}{6}\left(\Phi^{-1}(\alpha)^2 - 1\right),\ a_2 = \frac{1}{24}\left(\Phi^{-1}(\alpha)^3 - 3\,\Phi^{-1}(\alpha)\right),\ a_3 = -\frac{1}{36}\left(2\,\Phi^{-1}(\alpha)^3 - 5\,\Phi^{-1}(\alpha)\right)$$
$$a_4 = \frac{1}{120}\left(\Phi^{-1}(\alpha)^4 - 6\,\Phi^{-1}(\alpha)^2 + 3\right),\ a_5 = -\frac{1}{24}\left(\Phi^{-1}(\alpha)^4 - 5\,\Phi^{-1}(\alpha) + 2\right),$$
$$a_6 = \frac{1}{324}\left(12\,\Phi^{-1}(\alpha)^4 - 53\,\Phi^{-1}(\alpha)^2 + 17\right)$$

The cumulants are readily obtained from the cumulant generating function.

The following graph shows the Cornish-Fisher expansion superimposed on the exact distribution function of Y.



■ Saddlepoint approximation

The saddlepoint approximation to the cumulative distribution function of Y (due to Lugannani and Rice) is given by

$$F_Y(y) \approx \Phi(r) - n(r)\left(\frac{1}{u} - \frac{1}{r}\right) \tag{1}$$

where Φ and n are respectively the distribution function and density function of the standard normal distribution, and

$$r = \sqrt{2}\,\sqrt{\phi\,y - K(\phi)} \quad \text{and} \quad u = \phi\,\sqrt{K''(\phi)}$$

K is the cumulant generating function

$$K(t) = \log M(t) = \alpha\,t - \frac{1}{2}\sum_{i=1}^{m}\log(1 - 2\,\lambda_i\,t) + \frac{1}{2}\sum_{i=1}^{m}\frac{\beta_i^2\,t^2}{1 - 2\,\lambda_i\,t}$$

and the saddlepoint ϕ solves

$$K'(\phi) = y \tag{2}$$

An alternative approximation, due to Barndorff-Nielsen, is given by

$$F_Y(y) \approx \Phi\left(r - \frac{1}{r}\log\frac{1}{u}\right)$$

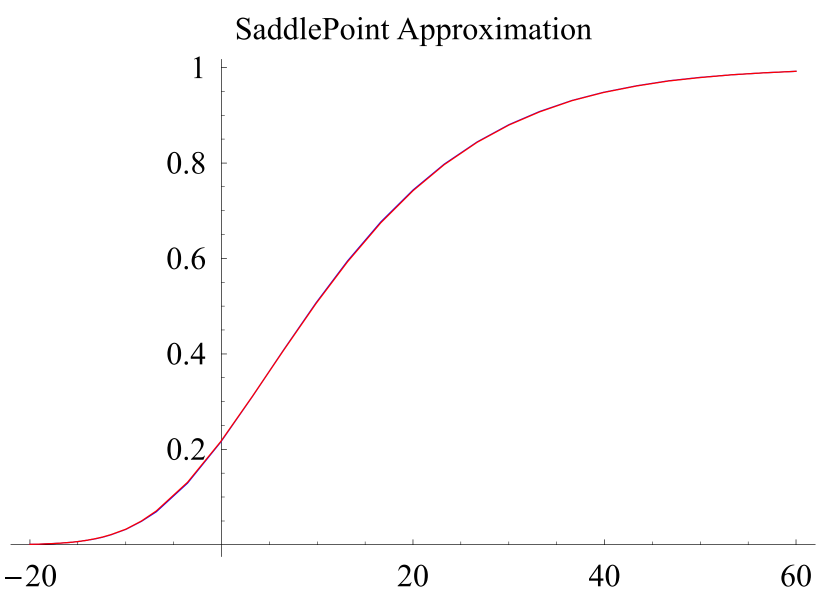
Without loss of generality, assume that $\lambda_1 = \min_i \lambda_i$ and $\lambda_m = \max_i \lambda_i$. If $\lambda_1 < 0$, then we must have $t > 1/2\lambda_1$. Similarly, if $\lambda_m > 0$, then we must have $t < 1/2\lambda_m$. In any case, K is defined on an interval around the origin.

Computing the saddlepoint approximation for a specific y involves solving equation (2) for the saddlepoint ϕ, and then calculating $F_Y(y)$ using equation (1).

The first and second derivatives of K are

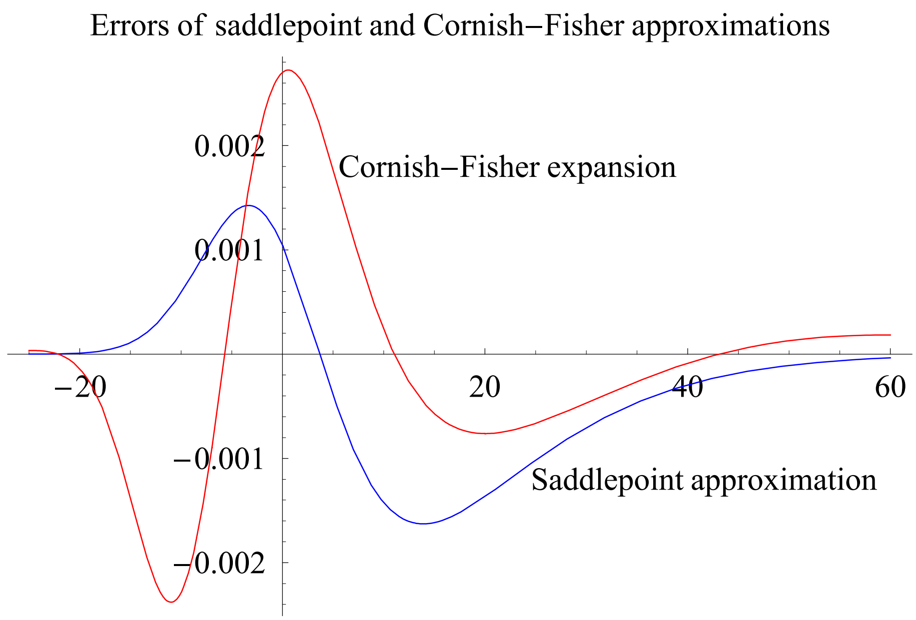
$$K'(t) = \alpha + \sum_{i=1}^{m} \frac{\lambda_i}{1 - 2\lambda_i t} + \sum_{i=1}^{m} \frac{\beta_i^2 t(1 - \lambda_i t)}{(1 - 2\lambda_i t)^2}$$

$$K''(t) = \sum_{i=1}^{m} \frac{2\lambda_i^2}{(1 - 2\lambda_i t)^2} + \sum_{i=1}^{m} \frac{\beta_i^2}{(1 - 2\lambda_i t)^3}$$



■ Comparison

The following graph compares the errors of the two approximation methods, relative to the cumulative distribution function of Y.



Example (adapted from Holton 2003)

A Japanese metals trading firm has exposure to forward and option positions in platinum. Some of the positions are denominated in USDs. We identify three key risk factors X

- spot price of platinum (JPY)
- a representative implied volatility of platinum
- spot JPY/USD exchange rate

and assume that X is normally distributed $N(\mu, \Sigma)$ with

$$\mu = \{53.150, 0.2670, 107.80\}; \Sigma = \begin{pmatrix} 799600 & 1.074 & -48.91 \\ 1.074 & 7.056\,10^{-5} & -3.875\,10^{-5} \\ -48.91 & -3.875\,10^{-5} & 0.4343 \end{pmatrix};$$

We value the portfolio using applicable forward and option pricing formulas, and quadratically approximate this as

$$P = a + \Delta^T X + X^T \Gamma X$$

with

$$a = 1.2110 * 10^{10}; \Delta = \{-459700, -4.819 * 10^8, -2.605 * 10^7\};$$

$$\Gamma = \begin{pmatrix} 4.305 & 3921 & 257.1 \\ 3921 & 8.407 * 10^7 & 3.647 * 10^6 \\ 257.1 & 3.647 * 10^6 & -5673 \end{pmatrix};$$

```
(z = CholeskyDecomposition[Σ] // Transpose) // MatrixForm
```

$$\begin{pmatrix} 894.204 & 0. & 0. \\ 0.00120107 & 0.00831369 & 0. \\ -0.0546967 & 0.00324098 & 0.656733 \end{pmatrix}$$

```
({Λ, u} = Eigensystem[Transpose[z].Γ.z] ) ;
{Λ, MatrixForm[u]}
```

$$\left\{\{3.43235\times 10^{6}, -21858.1, 18245.9\}, \begin{pmatrix} -0.998955 & -0.00874194 & -0.0448574 \\ -0.032341 & -0.558283 & 0.82902 \\ 0.0322904 & -0.829604 & -0.557417 \end{pmatrix}\right\}$$

```
Transpose[u].u // Chop // TableForm
```

1.	0	0
0	1.	0
0	0	1.

```
(A = u.Inverse[z]) // MatrixForm
```

$$\begin{pmatrix} -0.00111995 & -1.02488 & -0.068304 \\ 0.000131906 & -67.6444 & 1.26234 \\ 0.000117781 & -99.4569 & -0.848773 \end{pmatrix}$$

Using the substitution

$$Z = A(X - \mu)$$

we transform P into a quadratic function of independent chi-squared random variables, namely

```
Chop[
 a + Δ. X + X . Γ . X /. X → Inverse[A]. {Z₁, Z₂, Z₃} + μ // Expand, 10^-8]
```

$$9.30179\times 10^{9} + 3.57761\times 10^{8}\, Z_1 + 3.43235\times 10^{6}\, Z_1^2 - 3.79171\times 10^{6}\, Z_2 - 21858.1\, Z_2^2 - 4.6239\times 10^{6}\, Z_3 + 18245.9\, Z_3^2$$

That is

$$P = \alpha + \sum_{i=1}^{m} (\beta_i Z_i + \lambda_i Z_i^2)$$

with $\alpha = 9.30179\times 10^{9}$, $\mathrm{B} = (3.57761\times 10^{8}, -3.79171\times 10^{6}, -4.6239\times 10^{6})$ and $\Lambda = (3.43235\times 10^{6}, -21858.1, 18245.9)$.

```
A.Σ.Transpose[A] // Chop // MatrixForm
```

$$\begin{pmatrix} 1. & 0 & 0 \\ 0 & 1. & 0 \\ 0 & 0 & 1. \end{pmatrix}$$

```
parameters = {m = 3, α = 9.30179*10^9,
   β1 = 3.57761*10^8, β2 = -3.79171*10^6, β3 = -4.6239*10^6,
   λ1 = 3.43235*10^6, λ2 = -21858.1, λ3 = 18245.9};
```

The cumulants of *P* are

```
Table[{κk, D[K[t], {t, k}] /. t → 0}, {k, 5}] // TableForm
```

κ_1	9.30522×10^{9}
κ_2	1.28052×10^{17}
κ_3	2.63622×10^{24}
κ_4	7.23853×10^{31}
κ_5	2.48447×10^{39}

The standard deviation is

$\sqrt{K''[0]}$

3.57844×10^{8}

The portfolio's 1-day standard deviation is JPY 358 million. Applying the Cornish-Fisher expansion, the 5% quantile of *P* is approximately JPY 8.72 billion.

```
CFQuantile[0.05]
```

8.72252×10^{9}

Based on the expected mean of JPY 9.3 billion, the portfolio has a 1-day 95% VAR of approximately JPY 583 million, which is close to 1.65 standard deviations.

```
K'[0] - CFQuantile[0.05]
```

5.82694×10^{8}

$1.65 * \sqrt{K''[0]}$

5.90442×10^{8}

Suppose that we simply ignored the non-normality, and calculated the standard deviation linearly, in effect ignoring the quadratic term.

$$P = a + \Delta^T X + X^T \Gamma X$$

In this case, we over-estimate the standard deviation.

$$\sqrt{\Delta . \Sigma . \Delta}$$

$$4.10596 \times 10^8$$

Our 95% VAR estimate is then

$$1.65 * \sqrt{\Delta . \Sigma . \Delta}$$

$$6.77484 \times 10^8$$

which overestimates the risk by 16%.

$$100 \frac{6.7748386 - 5.82694312}{5.82694312}$$

$$16.2675$$