A Short Guide to Managing Risk in Worst Case Scenarios

THOMAS BREUER

PPE Research Centre, Hochschulstrasse 1, A-6850 Dornbirn
thomas.breuer@fhv.at

Summary. Stress tests have emerged as an important complement to Value at Risk. The usefulness of analysing portfolio behaviour in stress scenarios is only ensured (1) if the scenarios considered are not too implausible and (2) if no plausible scenarios are missed in which the portfolio is more dangerous than in the scenarios considered. These two requirements are satisfied when among all scenarios above some plausibility threshold those scenarios are identified in which the portfolio suffers the most severe loss. Plausibility of scenarios is measured in a dimension-independent way by the Mahalanobis distance. The location of the worst case scenario indicates the key risk factors of the portfolio. This information can be used to construct risk mitigating positions which reduce worst case losses deemed unacceptable.

Key words: risk measures, maximum loss, worst case scenarios, risk mitigation, dimensional dependence

1 Introduction

In the 1990ies Value at Risk (VaR) set out to conquer the risk management community. Originally intended as a reporting tool for senior management, it soon entered other core areas of banking such as capital allocation, portfolio optimization or risk limitation. With its increasing importance, also regulators acknowledged VaR as a risk measure when they allowed the calculation of capital requirements to be based on VaR. In this case, however, they required that a rigorous and comprehensive stress testing program is in place in order to complement the statistical model [1, 2]. In this context, two questions arise quite naturally: (1) Why is there a need to complement VaR models? and (2) What can be regarded as a rigorous and comprehensive stress testing program?

Other capital calculation rules are based on scenario analysis. One example is SPAN developed by the Chicago Mercantile Exchange [3], described in [4]. First, fourteen standard scenarios are considered. Next, two additional scenarios relate to extreme up or down moves of the futures price. The measure of risk is the maximum loss incurred, using the full loss for the first fourteen scenarios and only 35% of the loss for the two extreme scenarios.

Both the VaR based capital calculation and the scenario based capital calculation have some shortcomings. For VaR one main shortcoming, apart
from possible deficiencies in statistical modeling of some implementations, is that it does not provide information about potential losses exceeding the specified quantile of the profit-loss distribution. Furthermore, Artzner et al. [4, 5] showed that VaR lacks the desirable property of sub-additivity in general, a fact that should be alarming for both risk managers and regulators. The following example is from [6].

**Example:** Assume that a bank sets itself a VaR-limit of EUR 70 mill. Management give Trading Desk A a VaR-limit of EUR 50 mill and Trading Desk B a limit of EUR 20 mill. Some might consider this limit system to be conservative, since diversification effects are expected to bring down the joint VaR of the two trading desks well below the sum of the separate VaR of the two trading desks. But assume now that both trading desks hold option positions in the same underlying. The current value of the underlying is 10.000, Trading Desk A holds a million short European puts with a strike of 9.200 and Trading Desk B holds a million short European calls with a strike of 11.300. Assume both options mature in three months, the volatility of logarithmic returns is 5% and the risk free interest rate is also 5%. If neither desk has any other positions in its portfolio, the VaR at the 95%-level will be EUR 42.92 mill for Trading Desk A and EUR 18.47 mill for Trading Desk B. So both desks are comfortably below their VaR-limits. And yet, the joint VaR of both portfolios is EUR 80.91 mill—well above the VaR-limit the bank set itself.

What is so awkward about the lack of sub-additivity is the fact that this can give rise to regulatory arbitrage. If regulation allows the capital requirement of a firm to be calculated as the sum of the requirements of its subsidiaries and if the requirements are based on VaR, the firm could create artificial subsidiaries in order to save regulatory capital. By the same token, this amounts to a break-down of global risk management within one single firm.

Criticism of SPAN and other scenario-based capital calculation schemes points to the plausibility of scenarios and the possibility on missing important scenarios. This kind of criticism is valid for any kind of stress testing with predefined standard scenarios. Standardised stress tests do not take into account the composition of portfolios. As a consequence, it is quite likely that there exist scenarios which are much more harmful to a portfolio than the ones used in standardized stress tests.

Studer [7, 8] proposed Maximum Loss (MaxLoss) as an alternative to both VaR and stress testing with predefined standard scenarios. Instead of using only predefined scenarios, the set of all scenarios with a given minimal amount of plausibility should be considered when performing stress tests [9]. The loss in the worst scenario satisfying the plausibility constraint is MaxLoss. The size of the worst loss can be the basis of capital requirements. Additionally,
the location of the worst case scenario gives important information about the vulnerable spots of a portfolio.

The paper is structured as follows. First in Section 2 we introduce MaxLoss as a systematic way to perform stress tests. In Section 3 we discuss how to choose the set of scenarios over which to take MaxLoss. Section 4 gives a practical guide to measuring and managing risk with MaxLoss. In that Section we also show how the worst case scenario indicates possible action to mitigate risk.

2 Stress Tests and Maximum Loss

In the sequel we will use the following notation. The value of a portfolio depends on risk factors \(r_1, \ldots, r_n\) which can be combined into one single vector \(\mathbf{r} := (r_1, \ldots, r_n)\). The relevant risk factors could be interest rates, stock indices, exchange rates, and the like. The function determining the value of the portfolio when the values of the risk factors are given is denoted by \(P\). The values of the risk factors characterize the market situation as far as it is of relevance to the portfolio. \(\mathbf{r}_{CM}\) shall denote the vector representing the current values of the risk factors, i.e. the current market situation. \(P(\mathbf{r}_{CM})\) therefore represents the current value of the portfolio.

In our setup stress testing consists of selecting scenarios \(\mathbf{r}\) according to specific criteria and calculating the values of the current portfolio under these scenarios. These values are given by \(P(\mathbf{r})\). Each \(\mathbf{r}\) can be interpreted as one possible state of the market at the end of the holding period. By comparing \(P(\mathbf{r})\) with the current value of the portfolio \(P(\mathbf{r}_{CM})\) one can assess the losses that would be incurred if the market moved from \(\mathbf{r}_{CM}\) to \(\mathbf{r}\) without allowing for re-balancing the portfolio:

\[
\text{MaxLoss}_A(P) := P(\mathbf{r}_{CM}) - \min_{\mathbf{r} \in A} P(\mathbf{r}),
\]

where \(A\) is the set of selected scenarios. \(P(\mathbf{r}_{CM}) - P(\mathbf{r})\) gives only an upper bound for the loss potential due to market moves from \(\mathbf{r}_{CM}\) to \(\mathbf{r}\), because counter-actions taken during the holding period could reduce the loss.

In contrast to VaR, Maximum Loss is essentially a coherent risk measure for any choice of \(A\). Especially the sub-additivity property holds in general [7, p. 23]. Therefore if capital requirements defined via MaxLoss exclude regulatory arbitrage and make firmwide risk management possible.

In some sense, MaxLoss is typical of coherent risk measures. Every coherent risk measure can be represented as Maximum Expected Loss over some set of generalised scenarios [10, 11]. (Generalised scenarios are probability distributions rather than points in the sample space.) This basically is the dual representation of the risk measure. Taking as set of generalised scenarios all probability distributions with support in the trust region \(A\) we recover MaxLoss.
3 Choice of the trust region

The next question is which trust region $A$ should be chosen. Since we require of stress scenarios not only that they should lead to serious losses, but also that they be plausible to a certain extent, it is natural to take as trust region all scenarios above a certain plausibility threshold. But how should we measure plausibility? Certainly the plausibility of a scenario should be a probabilistic concept: The higher the probability of a move from the present market state $r_{CM}$ to a scenario $r$, the higher the plausibility of $r$. This implies that scenarios which are more distant from expected value of the market state will be less plausible. And scenarios which involve a move against the prevailing correlations will be less plausible than scenarios involving a move along with the correlations. If risk factor changes follow an elliptical distribution, trust regions will be the volumes of ellipsoids of constant probability density.

Denote by $\Sigma$ the covariance matrix of deviations of $r$ from its expectation $\mu$. The Mahalanobis distance of $r$ from $\mu$ is defined as

$$Maha(r) := \sqrt{(r - \mu)^T \cdot \Sigma^{-1} \cdot (r - \mu)}.$$  \hfill (1)

For elliptical distributions the density is a function of the Mahalanobis distance. The level surfaces of the density function are ellipsoids with constant Mahalanobis distance from the expected market state.

As trust regions we propose to take the ellipsoid of some given Mahalanobis radius $k$:

$$\text{Ell}_k := \{ r : \text{Maha}(r) \leq k \}.$$

$k$ is proportional to the lengths of the main axes of the ellipsoid. For an intuitive interpretation of $k$, note that $\text{Ell}_k$ contains only (but not all) moves in which all risk factor move $k$ standard deviations or less.

This choice of trust regions has the advantage of avoiding the problem of dimensional dependence which Studer [7, p. 44] pointed out. At the root of the problem is the fact that the probability mass of an ellipsoid depends on both the Mahalanobis radius and the number of dimensions. With increasing number of dimensions the probability mass of ellipsoids with constant Mahalanobis radius is decreasing. If we wanted ellipsoids of constant probability mass, then with increasing number of dimensions we also would have to increase the Mahalanobis radius. Therefore, in a model with more risk factors individual risk factors would be allowed larger moves and MaxLoss would be higher. This is an argument against the use of MaxLoss with trust regions of specified probability mass. MaxLoss with trust regions of specified Mahalanobis radius is preferable.

Finding a scenario in the ellipsoid $\text{Ell}_k$ which has minimal portfolio value is a deterministic optimisation problem with non-linear boundary conditions and often non-convex objective function. Some numerical algorithms for this problem are compared in Pistovčák et al. [12].
4 A Practical Guide to Managing Risk with Stress Tests

Let us turn to the practical aspects of managing risk with stress tests. First we present a way of reporting MaxLoss and the results of a systematic search for worst case scenarios. Then we show how knowledge of the worst case scenario suggests specific action to reduce risk if desired.

How can the results of a systematic search of a worst-case scenario be presented in a concise and readily understandable manner? It is certainly not enough to simply report the values of the risk factors in the worst-case scenario. Listing the values of say 1000 risk factors in the worst-case scenario would hopelessly overtax the capacity of any recipient of the report. Consequently, reports should include only the most important risk factors of the worst-case scenario.

What are the “most important” risk factors of a worst-case scenario? Sensitivities are certainly not an appropriate indicator of the importance of a risk factor: sensitivities in the present market state are completely unrelated to the worst-case scenario to be characterized; and all sensitivities will be zero in the worst-case scenario if it is a local minimum.

Table 1: Maximum Loss and Worst Case Scenario of the sample portfolio at various plausibility levels.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Current value in worst case (10^7 EUR)</th>
<th>Max. Maha loss (10^7 EUR)</th>
<th>(% of curr. val.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBP.SE</td>
<td>5.126</td>
<td>4.464</td>
<td>9.36%</td>
</tr>
<tr>
<td>USD.SE</td>
<td>5.126</td>
<td>4.44</td>
<td>13.37%</td>
</tr>
<tr>
<td>USD.XS</td>
<td>5.126</td>
<td>4.166</td>
<td>18.72%</td>
</tr>
<tr>
<td>GBP.R030</td>
<td>5.126</td>
<td>3.928</td>
<td>23.35%</td>
</tr>
<tr>
<td>GBP.XS</td>
<td>5.126</td>
<td>3.709</td>
<td>27.64%</td>
</tr>
</tbody>
</table>

Table 1: Worst Case Scenario at Maha ≤ 6

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GBP.SE</td>
<td>5402</td>
<td>5834</td>
<td>8.00%</td>
<td>4.84</td>
<td>62.12%</td>
</tr>
<tr>
<td>USD.SE</td>
<td>1139</td>
<td>1113</td>
<td>-2.35%</td>
<td>-1.37</td>
<td>24.54%</td>
</tr>
<tr>
<td>USD.XS</td>
<td>1.120</td>
<td>1.088</td>
<td>-2.79%</td>
<td>-3.41</td>
<td>15.00%</td>
</tr>
<tr>
<td>GBP.R030</td>
<td>0.995</td>
<td>0.995</td>
<td>0.0036%</td>
<td>-1.43</td>
<td>0.05%</td>
</tr>
<tr>
<td>GBP.XS</td>
<td>1.600</td>
<td>1.578</td>
<td>-1.38%</td>
<td>-2.60</td>
<td>0.05%</td>
</tr>
<tr>
<td>USD.R030</td>
<td>0.996</td>
<td>0.995</td>
<td>-0.02%</td>
<td>-3.19</td>
<td>0.00%</td>
</tr>
</tbody>
</table>
The following approach appears more useful: The search for the key risk factors is a search for a subset of risk factors which have a certain explanatory power, i.e. which explain the loss under the worst-case scenario up to a previously defined degree. For example, an explanatory power of 80% means that we are looking for a subset of the risk factors which will be able to explain at least 80% of the loss under the worst-case scenario. This means: Let us assume that, instead of the complete worst-case scenario $r_{WC} = (r_{WC,1}, \ldots, r_{WC,n})$, only the values of a subset of $w$ risk factors are reported. This corresponds to a simplified report scenario

$$r_{report} = (r_{1}^{i_{1}}, \ldots, r_{WC,i_{1}}, \ldots, r_{WC,i_{2}}, \ldots, r_{WC,i_{w}}, \ldots, r_{n}^{i_{w}}),$$

where the risk factors $i_{1}, i_{2}, \ldots, i_{w}$ have their worst-case values $r_{WC,i_{1}}, \ldots, r_{WC,i_{w}}$, and all other risk factors have their conditional expectation values $r_{j}^{*}$, given the worst case values of $i_{1}, i_{2}, \ldots, i_{w}$.

**Proposition 1.** The report scenario $r_{report}$ has the same plausibility as the full worst case scenario $r_{WC}$:

$$\text{Maha}(r_{report}) = \text{Maha}(r_{WC}).$$

A proof can be found in the Appendix.

| Table 2: Representing the key risk factors of the worst case scenario in condensed management reports. |
|---|---|---|---|
| Risk Factors | Relative Changes | Loss | Explanatory Power |
| Report 1 | GBP.SE | 4.85 | 9.09% | 50.06% |
| Report 2 | GBP.SE | 4.85 | 18.29% | 97.66% |
| | USD.SE | -1.32 | | |
| Report 3 | GBP.SE | 4.85 | 18.29% | 99.98% |
| | USD.SE | -1.39 | | |
| | USD.R030 | -3.21% | | |

The task is to find a fairly small set of $w$ risk factors which has a high explanatory power. The **explanatory power** of the risk factors $i_{1}, i_{2}, \ldots, i_{w}$ is

$$P(r_{CM}) - P(r_{report}) \over P(r_{CM}) - P(r_{WC}).$$

This can be solved by optimization algorithms in a discrete $w$-dimensional space. The **loss contribution** of a risk factor is the loss in the scenario where
the specified risk factor has its worst case value and all other risk factors have their current values, expressed as a fraction of the full worst case loss. For most portfolios the a first very good guess of the $w$ key risk factors is to take the $w$ risk factors with the highest MaxLoss contribution. (The MaxLoss contribution of risk factor $i$ is defined to be

$$\frac{P(r_{CM}) - P(r_{CM,1}, \ldots, r_{WC,i}, \ldots, r_{CM,n})}{P(r_{CM}) - P(r_{WC})},$$

which is the fraction of MaxLoss achieved in a scenario where only risk factor $i$ takes its worst case value and all the other risk factors remain at their present value.)

As an example of how this reporting method works, Table 2 shows results of identifying the one, two, and three most important risk factors in the worst case scenario of Portfolio 1. This kind of information can be used to report stress test results in an understandable way. For example Report 1 would be:

“A move of $+4.85\sigma$ in the GBP.SE would lead to a relative loss of 9.09% of the total portfolio value, assuming the other risk factors take their conditional expectation values. This explains 50.06% of the loss in the worst case scenario.”

The left plot in Figure 1 shows a plot of the portfolio value as a function of relative changes in the risk factor which is responsible for 74% of losses in the worst case scenario:

From Table 2 one could formulate a slightly more detailed Report 2:

“A move of $+4.85\sigma$ in the GBP.SE and of $-1.32\sigma$ in the USD.SE would lead to a loss of 18.293% of the total portfolio value, assuming
the other risk factors take their conditional expectation values. This explains 97.66% of the loss in the worst case scenario.”

One could formulate a more detailed Report 3:

“A move of +4.85σ in the GBP.SE, of -1.32σ in the USD.SE, and of -3.21σ in the USD.R030 would lead to a relative loss of 18,295% of the total portfolio value, assuming the other risk factors take their conditional expectation values. This explains 97.66% of the loss in the worst case scenario.”

But since two risk factors already explain 97.66% of the loss in the worst case scenario, there is no point in reporting more risk factors. Figure 1 shows a plot of the portfolio value as a function of relative changes in the two risk factors which are responsible for the bulk of losses in the worst case scenario: It is evident from this plot (and also from the first two rows of Table 2) that the portfolio value depends more sensitively on the GBP.SE than on the USD.SE. We also see that a move of roughly the value 6000 in the GBP.SE would be very harmful to the portfolio.

If a bank decides that it does not want to take the risk of this portfolio it can buy insurance in the form of risk reducing positions. Knowing that the key risk factor is GBP.SE and that its worst case value (at a plausibility level of Maha ≤ 6 is 5834, the bank can construct an insurance position which is exactly targeted at this value of the key risk factor. As insurance position we take a so-called condor in the GBP.SE with a peak around the worst case value of 5800. The composition of the insurance position is shown in Table 3. Figure 2 shows the payoff profile of the insurance position and of the insured portfolio as a function of GBP.SE. The insurance position produces profits if the GBP.SE rised to around 5800.

<table>
<thead>
<tr>
<th>amount</th>
<th>kind</th>
<th>underlying</th>
<th>strike</th>
<th>expiry</th>
</tr>
</thead>
<tbody>
<tr>
<td>200000</td>
<td>Eur. call</td>
<td>GBP.SE</td>
<td>5700</td>
<td>17d</td>
</tr>
<tr>
<td>-200000</td>
<td>Eur. call</td>
<td>GBP.SE</td>
<td>5710</td>
<td>17d</td>
</tr>
<tr>
<td>-200000</td>
<td>Eur. call</td>
<td>GBP.SE</td>
<td>5890</td>
<td>17d</td>
</tr>
<tr>
<td>200000</td>
<td>Eur. call</td>
<td>GBP.SE</td>
<td>5900</td>
<td>17d</td>
</tr>
</tbody>
</table>

The cost of such an insurance is close to zero: it is 3.94 times 10^{−39}EUR. The effect of the insurance is nicely displayed when we compare the payoff profile of the insured portfolio (bottom in Fig. 2) to the payoff profile of the insured portfolio (right in Fig. 1). Note that the worst case scenario of the insured portfolio differs from the one of the original portfolio, see Table 4.
Fig. 2: Top: Payoff profile of the insurance position as a function of GBP.SE. The insurance position produces profits if the GBP.SE100 drops by 10%. Bottom: Payoff profile of the insurance position as a function of GBP.SE.

The effect of the insurance position on the worst case loss (at a plausibility level of Maha ≤ 6) is enormous. It reduces MaxLoss from 9.59 mEUR to 6.32 mEUR. This is the effect of an insurance position which has a value of practically zero in the current market state. The more precisely we know the vulnerable market states for a portfolio, the easier it is to take up risk reducing positions which at the same time are affordable and do not affect the gain potential in other market states.

Can more insurance positions achieve a further reduction of MaxLoss? The payoff profile of the insured portfolio in dependence of GBP.SE suggests that at a plausibility level of Maha ≤ 6, at further smoothening of the profile of GBP.SE has no great effect on the total portfolio. A MaxLoss calculation of the insured portfolio confirms this. The results are shown in Table 4. We observe that GBP.SE is not any more the key risk factor. For the insured portfolio the risk factor with the highest contribution MaxLoss is USD.SE.
Table 4: Maximum Loss and Worst Case Scenario of the insured portfolio at a plausibility level of Maha \leq 6.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Current value in worst case ($10^7$ EUR)</th>
<th>Portfolio</th>
<th>max. Maha</th>
<th>loss (% of curr. val.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.126</td>
<td>4.493</td>
<td>6</td>
<td>12.34%</td>
<td></td>
</tr>
</tbody>
</table>

Worst Case Scenario at Maha \leq 6

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>USD.SE</td>
<td>1,139.830</td>
<td>-4.16%</td>
<td>-2.47</td>
<td>60.22%</td>
</tr>
<tr>
<td>USD.XS</td>
<td>1.120</td>
<td>-3.77%</td>
<td>-4.60</td>
<td>30.66%</td>
</tr>
<tr>
<td>GBP.SE</td>
<td>5,402.300</td>
<td>4.35%</td>
<td>2.63</td>
<td>11.73%</td>
</tr>
<tr>
<td>GBP.R030</td>
<td>0.995</td>
<td>0.00%</td>
<td>-1.22</td>
<td>0.08%</td>
</tr>
<tr>
<td>GBP.XS</td>
<td>1.600</td>
<td>-1.80%</td>
<td>-3.40</td>
<td>0.08%</td>
</tr>
<tr>
<td>USD.R030</td>
<td>0.996</td>
<td>-0.01%</td>
<td>-2.30</td>
<td>0.03%</td>
</tr>
</tbody>
</table>

As a function of the key risk factor USD.SE the payoff of the insured portfolio is more or less linear, apart from a bump around USD.SE = 1100 (see Fig. 3). For risk factors in which the portfolio behaves more or less linearly there are no cheap insurance positions available. The only way to reduce risk is to reduce exposure to the risk factor. Of course this also reduces the potential of gains.

5 Conclusion

There are two main benefits of considering MaxLoss instead of VaR. First MaxLoss is a coherent risk measure whereas VaR is not in general coherent. Especially the violation of the sub-additivity property can cause problems from the point of view of global risk management and regulation.

Secondly, MaxLoss provides worst-case scenario which leads to this loss figure. The knowledge of the worst-case scenario can be the basis of informed risk decisions and suggest possible risk reducing transactions.

References

1. Basel Committee on Banking Supervision: Amendment to the capital accord to

Fig. 3: Payoff profile of the insured portfolio as a function of the key risk factor USD.SE. The payoff is more or less linear, apart from a bump around USD.SE = 1100.
A Proof of Proposition 1

Let us assume that we have \( n \) risk factors, whose change is governed by a multivariate elliptically symmetric distribution with covariance matrix \((E(R^2)/n)\Sigma\) and mean \( \mu \). Denote the \( n \)-dimensional density function by \( f^n \). Let us assume that the risk factors are indexed in such a way that \( \{i_1, i_2, \ldots, i_w\} = \{1, 2, \ldots, w\} \). Step by step the remaining \( n-w \) risk factors will be set to their conditional expectation values without changing the Mahalanobis distance.

When we fix the value of some remaining risk factor, say risk factor \( n \), the distribution of the remaining \( n-1 \) risk factors is described by the marginal distribution \( f^{n-1}(r') \), resulting from integration over \( r_n \). Here \( r' \) is the \( n-1 \)-dimensional vector resulting from deleting the last component from \( r \). The conditional distribution of \( r_n \) given some fixed \( r' \) is

\[
h(r_n|r') = \frac{f^n(r', r_n)}{f^{n-1}(r')},
\]

Call \( r^*_n \) the expectation value of the conditional distribution \( h(r_n|r') \).

**Lemma 1.** Assume \( \Sigma \) is a positive definite \( n \times n \)-matrix. Then we have

\[
\text{Maha}(r)^2 - \text{Maha}(r')^2 = \left( \sum_{i=1}^{n} C(\Sigma)^{-1}_{in} r_i \right)^2, \tag{2}
\]

where \( C(\Sigma)^{-1}_{in} \) is the element in row \( i \) and column \( n \) of the inverse matrix of the Cholesky decomposition of \( \Sigma \).

**Proof.** For an arbitrary symmetric positive definite matrix \( M \) denote by \( C(M) \) its Cholesky decomposition. \( C(M) \) is the upper triangular matrix satisfying

\[
M = C(M)^T C(M). \tag{3}
\]

Here are some properties of the Cholesky decomposition.

\[
C(M)^{-1}C(M)^{-1}T = M^{-1} \tag{4}
\]

In other words, the transpose of the inverse of the Cholesky decomposition of \( M \) is the Cholesky decomposition of \( M^{-1} \).

Furthermore we have \( C(M)'T C(M)' = M' \). So we may write
Deleting the $n$-th row and column of $M$ and then making the Cholesky decomposition amounts to the same as making the Cholesky decomposition of $M$ and then deleting the $n$-th row and column.

For an arbitrary triangular matrix $C$ we have
\[ (C^{-1})' = (C')^{-1}. \] (6)

Now let us calculate the squares of the Mahalanobis distances.

\[ r' \cdot \Sigma^{-1} \cdot r' = r'^T \cdot (C(\Sigma')^{-1} \cdot C(\Sigma')^{-1})' \cdot r' \]
\[ = r'^T \cdot C(\Sigma')^{-1}C(\Sigma')^{-1}T \cdot r' \]
\[ = r'^T \cdot C(\Sigma)^{-1}C(\Sigma)^{-1}T \cdot r'. \] (7)

Similarly we have

\[ r \cdot \Sigma^{-1} \cdot r = r^T \cdot C(\Sigma)^{-1}C(\Sigma)^{-1}T \cdot r \]
\[ = (r^T \cdot C(\Sigma)^{-1}) (r^T \cdot C(\Sigma)^{-1})^T. \] (8)

The inverse of the triangular matrix $C(\Sigma)$ is again triangular, so we can write

\[ C(\Sigma)^{-1} = \begin{pmatrix} C(\Sigma)^{-1}_{1n} \\ \vdots \\ \vdots \\ 0 \cdots 0 C(\Sigma)^{-1}_{nn} \end{pmatrix}. \]

Writing $r^T = (r'^T, r_n)$ equation (8) reads

\[ r^T \cdot \Sigma^{-1} \cdot r = \left( r'^T C(\Sigma)^{-1} , \sum_{i=1}^{n} C(\Sigma)^{-1}_{in} r_i \right) \left( r'^T C(\Sigma)^{-1} , \sum_{i=1}^{n} C(\Sigma)^{-1}_{in} r_i \right)^T 
\]
\[ = r'^T \cdot C(\Sigma)^{-1}C(\Sigma)^{-1}T \cdot r' + \left( \sum_{i=1}^{n} C(\Sigma)^{-1}_{in} r_i \right)^2. \] (9)

Subtracting (7) from (9) yields the Lemma. \[ \square \]

**Lemma 2.** The expectation value of the conditional distribution $h(r_n|r')$ is given by

\[ r_n^* = -\frac{\sum_{i=1}^{n-1} C(\Sigma)^{-1}_{in} r_i}{C(\Sigma)^{-1}_{nn}}. \]
Proof: As a function of \( r_n \), the conditional density \( h(r_n | r') = f^n(r', r_n)/f^{n-1}(r') \) is a constant times the \( n \)-dimensional density \( f^n(r', r_n) \). \( f^n(r) = (\det \Sigma)^{-1/2} \)
\( g(\text{Maha}(r)) \) was assumed to be a strictly decreasing function of \( \text{Maha}(r) \). By eq. (2) \( \text{Maha}(r) \) as a function of \( r_n \) is minimal, namely equal to \( \text{Maha}(r') \), at
\[
 r^*_n = -\frac{\left(\sum_{i=1}^{n-1} C(\Sigma)^{-1} \, \text{in} \, r_i \right)}{C(\Sigma)^{-1} \, \text{nn}}.
\]
So the conditional density \( h(r_n | r') \) is maximal at \( r^*_n \), where \( \text{Maha}(r) \) is minimal.

We also have
\[
f^n(r', r_n) = cg(\text{Maha}(r)^2) = cg \left( \text{Maha}(r')^2 + \left( \sum_{i=1}^{n} C(\Sigma)^{-1} \, \text{in} \, r_i \right)^2 \right)
\]
\[
= cg \left( \text{Maha}(r')^2 + \left( \sum_{i=1}^{n-1} C(\Sigma)^{-1} \, \text{in} \, r_i + C(\Sigma)^{-1} \, \text{nn} \, r_n \right)^2 \right)
\]
\[
= cg \left( \text{Maha}(r')^2 + (-C(\Sigma)^{-1} r^*_n + C(\Sigma)^{-1} \, \text{nn} \, r_n)^2 \right)
\]
\[
= cg \left( \text{Maha}(r')^2 + C(\Sigma)^{-1} \, \text{nn} \, (r_n - r^*_n)^2 \right)
\]
Here \( c \) is \( (\det \Sigma)^{-1/2} \). This implies that \( f^n(r', r_n) \), and consequently \( h(r_n | r') \) is symmetric around its maximum \( r^*_n \). Thus the expectation value of \( h(r_n | r') \) is \( r^*_n \).

Lemma 2 together with eq. (2) implies that the scenarios \( r = (r', r^*_n)^T \) and \( r' \) have the same Mahalanobis distance from the current state, \( \text{Maha}(r', r^*_n) = \text{Maha}(r') \). This finishes the proof of Proposition 1.