Journal of Risk **20**(3), 83–93
DOI: 10.21314/JOR.2018.369

Research Paper

The quickest way to lose the money you cannot afford to lose: reverse stress testing with maximum entropy

Riccardo Rebonato

EDHEC-Risk Institute/EDHEC Business School, EDHEC, 10 Fleet Place, London EC4M 7RB, UK; email: riccardo.rebonato@edhec.edu

(Received August 9, 2016; revised November 3, 2016; accepted November 4, 2016)

ABSTRACT

We extend a technique devised by Saroka and Rebonato to "optimally" deform a yield curve in order to deal with a common and practically relevant class of optimization problems subject to linear constraints. In particular, we show how the idea can be applied to the case of reverse stress testing, and we present a case study to illustrate how it works. Finally, we point out a maximum-entropy interpretation of (or justification for) the procedure and present some obvious generalizations.

Keywords: reverse stress testing; maximum entropy; risk management; stress testing; yield curve deformation.

1 MOTIVATION

Saroka and Rebonato (2015) recently introduced a simple method to produce the likeliest (in some sense to be precisely defined) deformation of a yield curve, conditional on some exogenous views about some of its key rates. The central insight was that, even when the views expressed were very simple (say, an upward move of the ten-year rate by 20 basis points), the most likely deformation of the yield curve compatible with this view was, in general, not given by a parallel move of the yield curve, or a move in the first principal component only. A close analogy with the least energetically demanding deformation of a rod was presented, where the eigenvalues were related to the stiffness of the associated springs (large eigenvalue, weak spring).

This paper extends and generalizes the approach in several directions.

First of all, the procedure can be applied to produce the most likely deformation of a set of risk factors compatible with a variety of economically interesting constraints, of which the exogenous views dealt with in Saroka and Rebonato (2015) are a very special case. (This can be done as long as the constraints can be exactly or approximately expressed in linear form.) We study as an example of these wider applications the case of reverse stress testing, which we again apply to a fixed-income portfolio.

Given the way the more general method is presented, it becomes clear that the underlying set of risk factors need not be limited to yield curve variables but can encompass any financial (or indeed macroeconomic) variables: again, as long as the constraints of interest can be expressed in linear form.

Finally, it is straightforward to show that, if one decides to make use in one's analysis only of the information in the principal components obtained from the covariance matrix (as one often does), then the solution we propose is a maximum-entropy ("least-committal") solution, conditional on the mean-and-covariance information that we have.

2 THE SETUP

Let the value of a portfolio be affected linearly by a number of risk factors. Let Σ be the covariance matrix for changes in the risk factors ("yields") $\overrightarrow{\Delta y}$. We stress again that – to stick with the notation of the previous paper – we denote the changes in factors by the symbol $\overrightarrow{\Delta y}$, but that these need not be interpreted as yields. The orthogonalization of the covariance matrix gives

$$\Sigma = U \Lambda U^{\mathrm{T}}, \tag{2.1}$$

with U an orthogonal matrix: $UU^{\mathrm{T}} = I$.

Principal components, $\overrightarrow{\Delta x}$, are defined by

$$\overrightarrow{\Delta x} = U^{\mathrm{T}} \overrightarrow{\Delta y}. \tag{2.2}$$

Because of orthogonality we have

$$\overrightarrow{\Delta y} = U \overrightarrow{\Delta x}. \tag{2.3}$$

As in Saroka and Rebonato (2015), we now proceed to define normalized principal components, $\overrightarrow{\Delta p}$, ie, orthogonal quantities proportional to the principal components that have unit variance. By analogy with (2.2), we want to write

$$\overrightarrow{\Delta p} = V^{\mathrm{T}} \overrightarrow{\Delta y}. \tag{2.4}$$

To this effect, set

$$V = U \Lambda^{-1/2}. \tag{2.5}$$

Then,

$$V^{\mathrm{T}} = \Lambda^{-1/2} U^{\mathrm{T}}. \tag{2.6}$$

From (2.4), we then have

$$\overrightarrow{\Delta p} = V^{\mathrm{T}} \overrightarrow{\Delta y} = \underbrace{\Lambda^{-1/2} U^{\mathrm{T}}}_{V^{\mathrm{T}}} \overrightarrow{\Delta y} = \underbrace{\Lambda^{-1/2} U^{\mathrm{T}}}_{V^{\mathrm{T}}} \underbrace{U \overrightarrow{\Delta x}}_{\overrightarrow{\Delta y}} = \Lambda^{-1/2} \overrightarrow{\Delta x}, \tag{2.7}$$

which confirms that the $\overrightarrow{\Delta p}$ are indeed the variance-normalized principal components. The reason for introducing these normalized quantities is to show that minimizing the sum of the squares of the normalized principal components (under a given constraint) produces a maximum-entropy (and, under some conditions, maximum-probability) result. For the moment, let us just assume that minimizing S, defined as

$$S \equiv \sum_i \Delta p_i^2 \tag{2.8}$$

subject to some constraint, is a reasonable thing to do.

3 APPLICATION TO REVERSE STRESS TESTING

We want to show an application of the procedure to the now-topical case of reverse stress testing. With reverse stress testing – as advocated in, for example, Basel Committee on Banking Supervision (2009) – one attempts to find out what has to happen to a portfolio for a certain loss (usually deemed to be "unacceptable") to be incurred. The problem is that, for all but the simplest portfolios, completely different changes in the risk factors can give rise to the same loss. Take as an example the case of an interest rate delta-hedged option book. A very large loss could theoretically be incurred by changing the rates used for discounting by a fantastic amount. It is clear, however, that it would be much "easier" to incur the same large loss if, say, the implied volatilities were to change by a significant, but more reasonable, amount. Alternatively, if the delta-hedged portfolio had a negative gamma, a large move in the underlying could give rise to the prespecified loss. So, in this very simple example, the very real question arises of which combination of moves in the implied volatilities and in the underlying (and yes, perhaps in the discounting rate as well) may give rise to the unacceptable loss with "greatest ease" (ie, with the highest probability). This is why the title of this paper makes reference to the easiest way to incur an unacceptable loss.

In a linear setting the problem can be formalized as follows. Consider a profit-or-loss function, z, given by

$$z = \overrightarrow{\beta}^{\mathrm{T}} \overrightarrow{\Delta y}, \tag{3.1}$$

where $\overrightarrow{\beta}^{\mathrm{T}}$ is a vector of linear sensitivities to the yield changes. (The generalization to situations when the vector $\overrightarrow{\Delta y}$ contains not yield changes but changes in generic risk factors is obvious.)

We want to express the vector $\overrightarrow{\Delta y}$, and ultimately the loss function z, as a function of $\overrightarrow{\Delta p}$ rather than $\overrightarrow{\Delta x}$. From

$$\overrightarrow{\Delta p} = V^{\mathrm{T}} \overrightarrow{\Delta y} \tag{3.2}$$

it follows from (2.6) that

$$\overrightarrow{\Delta p} = \Lambda^{-1/2} U^{\mathrm{T}} \overrightarrow{\Delta y}, \tag{3.3}$$

and since

$$\overrightarrow{\Delta y} = U \overrightarrow{\Delta x}, \tag{3.4}$$

one has

$$\overrightarrow{\Delta p} = \Lambda^{-1/2} U^{\mathrm{T}} U \overrightarrow{\Delta x} = \Lambda^{-1/2} \overrightarrow{\Delta x} \implies \overrightarrow{\Delta x} = \Lambda^{1/2} \overrightarrow{\Delta p} \tag{3.5}$$

and, therefore,

$$z = \overrightarrow{\beta}^{\mathrm{T}} \overrightarrow{\Delta y} = \overrightarrow{\beta}^{\mathrm{T}} U \Lambda^{1/2} \overrightarrow{\Delta p}. \tag{3.6}$$

For compactness of notation, define

$$\overrightarrow{\beta}^{\mathrm{T}} U \Lambda^{1/2} \equiv \overrightarrow{A}^{\mathrm{T}}. \tag{3.7}$$

With this definition, one has

$$z = \overrightarrow{A}^{\mathrm{T}} \overrightarrow{\Delta p} = \sum a_k \Delta p_k, \tag{3.8}$$

where the a_i are the elements of the vector A. For a reverse stress testing application, we require the profit-or-loss function, z, to attain the (negative) value K. We then construct the Lagrangian, $\mathcal{L}$, given by

$$\mathcal{L} = -S + \lambda\Big(\sum a_k \Delta p_k - K\Big) = -\sum_i \Delta p_i^2 + \lambda\bigg(\sum_j a_j \Delta p_j - K\bigg). \tag{3.9}$$

Setting the derivatives to zero, one gets

$$\frac{\partial \mathcal{L}}{\partial p_k} \implies -2\Delta p_k + \lambda a_k = 0, \tag{3.10}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} \implies \sum_k a_k \Delta p_k = K. \tag{3.11}$$

This can be written in matrix form as

$$\begin{bmatrix} -2\mathbb{I}_{n\times n} & \overrightarrow{A} \\ \overrightarrow{A}^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} \overrightarrow{\Delta p} \\ \lambda \end{bmatrix} = \begin{bmatrix} \overrightarrow{0} \\ K \end{bmatrix}. \tag{3.12}$$

To simplify the notation, denote

$$\begin{bmatrix} -2\mathbb{I}_{n\times n} & \overrightarrow{A} \\ \overrightarrow{A}^{\mathrm{T}} & 0 \end{bmatrix}$$

by Γ and

$$\begin{bmatrix} \overrightarrow{0} \\ K \end{bmatrix}$$

by δ.

The solution for the change in the value of the normalized principal components, $\overrightarrow{\Delta p}^*$, that gives the specified loss (and for the Lagrange multiplier, λ) is then given by

$$\begin{bmatrix} \overrightarrow{\Delta p}^* \\ \lambda \end{bmatrix} = \Gamma^{-1}\delta. \tag{3.13}$$

Once the vector $\overrightarrow{\Delta p}^*$ has been obtained, one can get the corresponding changes in yields from

$$\overrightarrow{\Delta y}^* = U\overrightarrow{\Delta x}^* \tag{3.14}$$

with

$$\overrightarrow{\Delta x}^* = \Lambda^{1/2}\overrightarrow{\Delta p}^*. \tag{3.15}$$

This concludes the derivation.

4 WHY MINIMIZE S?

We have not explained why it should be desirable to perform a constrained minimization of the quantity S. In Saroka and Rebonato (2015), a justification was given in terms of maximizing the loglikelihood of the joint distribution of the normalized principal components, once the assumption is made that the distribution of the risk factors is jointly Gaussian. Since the normalized principal components have unit variances, it is immediate to show that the loglikelihood is proportional to $\sum_i \Delta p_i^2$, which is just the definition of S.

One can give a slightly more general interpretation, however. As is well known, the derivation of principal components is not predicated on any distributional assumption about the original variables or, indeed, the principal components themselves. When the covariance matrix is orthogonalized, we are simply making use of information about the vector of means and the covariance matrix of the original variables.

Now, when all we know about a set of variables is their mean and their covariance matrix, it is also well known that the associated maximum-entropy solution is the multivariate Gaussian. So, if we choose to solve a problem using principal components (and we therefore implicitly neglect additional information that we may have about higher moments or dependencies not captured by the covariance matrix), the maximum-entropy distribution compatible with the information we do use is the joint Gaussian one.

The solution we have found can therefore be interpreted as the maximum-entropy solution compatible with the constraints imposed and with our knowledge about the original variables being limited to their first two moments. Or, in other words, if we decide to look for a solution to our problem using principal components, then the solution we offer is the maximum-entropy solution.

5 A STYLIZED CASE STUDY

In this section, we present a stylized example in order to get an intuitive feel for how the approach works. For the sake of simplicity, we consider the case of a yield curve described by three yields of maturity: two, five and ten years. We assume that today's yield curve is given by $y(2) = 0.015$, $y(5) = 0.0200$ and $y(10) = 0.0225$. The volatilities of the three yields are assumed to be 100, 90 and 84 basis points (bps), respectively. Their correlation matrix is given by

$$\rho = \begin{array}{c c} & \begin{array}{ccc} y(2) & y(5) & y(10) \end{array} \\ \begin{array}{c} y(2) \\ y(5) \\ y(10) \end{array} & \begin{bmatrix} 1 & 0.84 & 0.75 \\ 0.84 & 1 & 0.90 \\ 0.75 & 0.90 & 1 \end{bmatrix} \end{array}. \tag{5.1}$$

The investor sets a reverse stress testing limit of, say, US\$10 million and is assumed to have a linear portfolio of exposures described by a β vector of the form

$$\beta = \begin{bmatrix} \beta_{2\mathrm{y}} \\ \beta_{5\mathrm{y}} \\ \beta_{10\mathrm{y}} \end{bmatrix}. \tag{5.2}$$

Figure 1 shows the three eigenvectors from the orthogonalization of the covariance matrix. The shape of the three eigenvectors lends itself to the usual interpretation in terms of "level", "slope" and "curvature".

FIGURE 1 The three eigenvectors from the orthogonalization of the covariance matrix obtained with the volatilities and correlations given in the text.

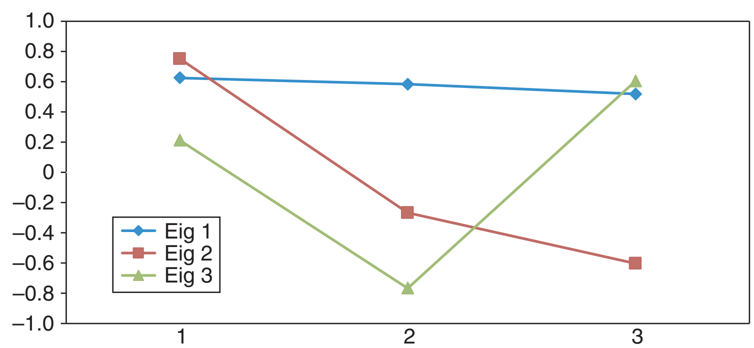


TABLE 1 The reverse stress testing solution expressed in terms of the changes in the normalized principal components ($\mathrm{d}p^*$), in the principal components ($\mathrm{d}x^*$) and in the yields ($\mathrm{d}y^*$) corresponding to the vector β as in (5.3) (in all cases, the $*$ superscript indicates the extremum solution).

	$\mathrm{d}p^*$	$\mathrm{d}x^*$	$\mathrm{d}y^*$
1	−0.0225	−0.0003	−0.0009
2	−0.2344	−0.0011	+0.0013
3	+0.4314	+0.0011	+0.0002

We then look at the changes in the principal components, in the normalized principal components and in the risk factors (the yields) corresponding to several portfolios of exposure (ie, to several vectors β).

5.1 "Butterfly" exposure

First, we assume that the portfolio of exposures is a "barbell" ("butterfly") of yields:

$$\beta_{\text{butterfly}} = \begin{bmatrix} \beta_{2\text{y}} = -3000 \\ \beta_{5\text{y}} = +6000 \\ \beta_{10\text{y}} = -3000 \end{bmatrix}. \tag{5.3}$$

The resulting solution is shown in Table 1 and Figure 2.

Given the nature of the exposure vector β, note that the extremum reverse stress testing solution strongly engages not only the third ("curvature") principal components

FIGURE 2 The original and deformed yield curves corresponding to the butterfly position in (5.3).

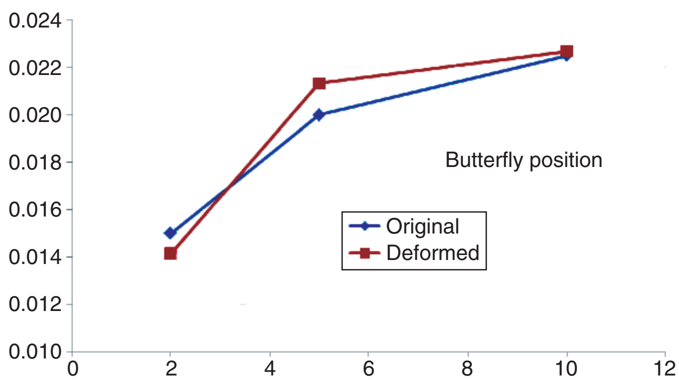


TABLE 2 Same as Table 1 but for the vector β as in (5.4).

	$\mathrm{d}p^*$	$\mathrm{d}x^*$	$\mathrm{d}y^*$
1	−0.0710	−0.0011	+0.0017
2	+0.2876	+0.0014	+0.0003
3	+0.0446	+0.0001	−0.0003

(as one would have expected) but also the second to a similar extent, and, to a smaller degree, the first.

5.2 "Steepening/flattening" exposure

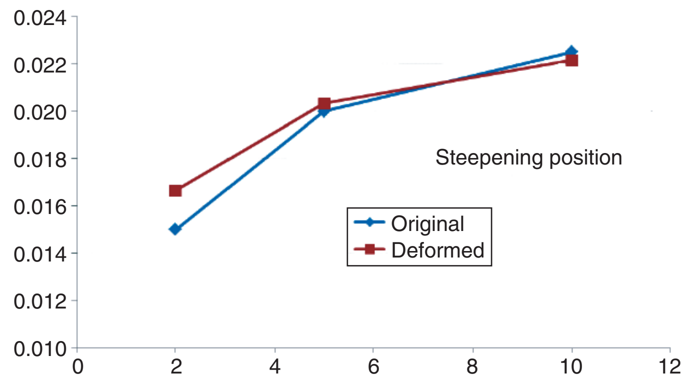
The next vector β that we examine describes a steepening or flattening bias, such as

$$\beta_{\text{slope}} = \begin{bmatrix} \beta_{2\text{y}} = +5000 \\ \beta_{5\text{y}} = 0 \\ \beta_{10\text{y}} = -5000 \end{bmatrix}. \tag{5.4}$$

The solution and the attending deformation of the yield curve are shown in Table 2 and Figure (3).

Also in this case, note that the mode of deformation one would have expected (the "slope") is indeed the most actively engaged one. However, the first and third modes of deformation also provide a nonnegligible contribution to the solution. In particular,

FIGURE 3 Same as Figure 1 but for the steepening vector of exposures in (5.4).



the vector $\mathrm{d}y^*$ clearly shows that the most likely flattening involves a substantial upward move of the two-year yield and a very modest downward move of the ten-year yield, despite the fact that the exposures to the two- and five-year yields were identical. Intuitively, this is in part due to the higher volatility of the two-year yield than the ten-year yield (100bps versus 84bps). However, it would have been difficult to guess this solution simply from the ratio of the volatilities (which is 0.84, while the ratio of the yield moves is 0.20): it is the full covariance structure that determines the most likely solution.

5.3 "Long-duration" exposure

The last case we discuss is the most straightforward, in that it represents an identical long-duration exposure in each maturity bucket:

$$\beta_{\text{long duration}} = \begin{bmatrix} \beta_{2\text{y}} = +5000 \\ \beta_{5\text{y}} = +5000 \\ \beta_{10\text{y}} = +5000 \end{bmatrix}. \tag{5.5}$$

For the sake of brevity, we only show the solution in tabular form: see Table 3.

Given the almost exactly flat structure of the first eigenvector, the solution in this case presents no surprises. We stress, however, that this would not be the case if one were dealing with a more complex and realistic case, with the risk factors relating to different asset classes and currencies.

TABLE 3 Same as Table 1 but for the vector β as in (5.5).

	$\mathrm{d}p^*$	$\mathrm{d}x^*$	$\mathrm{d}y^*$
1	+0.0775	−0.0012	+0.0007
2	−0.0017	+0.0000	+0.0007
3	−0.0004	+0.0000	+0.0007

6 OTHER APPLICATIONS

We have shown above how the idea presented in Saroka and Rebonato (2015) can be adapted to deal with reverse stress testing. In this section, we suggest another application.

In the case of stress testing or scenario analysis, the user is often interested in the impact of a macrofinancial event on the risk factors of a large and complex portfolio. Rebonato and Denev (2014) present one such approach based on the Bayesian net technology. Given the nature of a stress testing program, the shocks to the relevant risk factors are rarely obtained from frequentist information, and instead rely on the output of structural models or, more generally, on expert knowledge. At the same time, complex portfolios are affected by a very large number of risk factors, and dimensionality-reduction techniques such as principal component analysis are often used to make the problem tractable and reduce the risk of overparameterization.

The problem with doing this, however, is that experts are then required to assign stresses not to quantities such as yields or equity indexes – about which they can have well-formed ideas – but to principal components, which may well combine different asset classes in a complex way.

The technique we have presented can be of great help in this situation. Indeed, the shocks assigned by the expert relating to a handful of visible market risk factors (such as equity indexes) now play the role of “exogenous views” and the resulting shocks to all the risk factors of the associated maximum-entropy deformation of the complex portfolio.

7 A SIMPLE GENERALIZATION

We have assumed so far that the joint distribution for the normalized principal components was jointly Gaussian. This strong assumption can be relaxed to a considerable extent by allowing each principal component to have an arbitrary marginal distribution. Each of these arbitrary marginal distributions can be transformed via inversion of the empirical cumulative distribution to the equivalent normal distributions. The associated uniform variates can in turn be transformed to equivalent Gaussian variables.

Once these transformations are carried out, the covariance matrix among the equivalent Gaussian variables can be calculated, and its eigenvectors, eigenvalues and (normalized) principal components can be extracted.

All of the treatment above then applies to the transformed and then normalized principal components associated with the equivalent Gaussian variables. The important remaining assumption is that a Gaussian copula should describe the transformed variables well.

8 CONCLUSIONS

We have presented a generalization of the approach introduced in Saroka and Rebonato (2015) to deal with a class of problems for which linear constraints are appropriate, and we have given a maximum-entropy justification for the results.

We have shown how the technique can be used to tackle the problem of reverse stress testing for (exactly or approximately) linear portfolios. We have presented a case study to illustrate the workings of the approach and shown that not-easily-guessed deformation modes are engaged in the reverse stress solutions associated with three simple portfolios.

We have generalized the results to the case where the yields (the risk factors) have arbitrary marginal distributions, as long as their codependence can be described by a Gaussian copula.

Finally, we have suggested how the technique can be of use in handling the mapping problem in stress testing programs for complex portfolios that make use of subjective (expert) knowledge for their input.

DECLARATION OF INTEREST

The author reports no conflicts of interest. The author alone is responsible for the content and writing of the paper.

REFERENCES

Basel Committee on Banking Supervision (2009). Principles for sound stress testing practices and supervision. Report, May, Bank for International Settlements.

Rebonato, R., and Denev, A. (2014). *Portfolio Management Under Stress*. Cambridge University Press.

Saroka, I., and Rebonato, R. (2015). How to deform a yield curve: the most likely deformation of a yield curve consistent with subjective views. *Journal of Fixed Income* **24**(3), 30–38 (https://doi.org/10.3905/jfi.2014.24.3.030).