Mixed-Integer Credit Portfolio Optimization: an application to Italian segregated funds

L. Passalacqua

Università di Roma *La Sapienza*

e-mail: Luca.Passalacqua@Uniroma1.It

phone: + [39] (06) 4991 9559
fax: + [39] (06) 8535 2601
Keywords: Credit Risk, Optimization, Simulated Annealing

Abstract

We have solved the mixed-integer optimization problem of finding the efficient frontier in the $\{risk\ capital, current\ yield\}$ plane for a typical Italian life insurance segregated fund portfolio of medium size (90 assets) with a minimum number of assets and a fixed duration constraint. The solution is obtained using a metaheuristic technique known as adaptive simulated annealing. A comparison of the simplified quadratic programming problem with a standard hill-climbing algorithm has also been performed.

1 Introduction

The impressive reduction of interest rates witnessed in continental Europe in the last few years has driven the attention of the management of life insurance segregated funds towards corporate bonds. Therefore portfolios traditionally designed to meet asset-allocation requirements have now to be optimised both with respect to variations of interest rates and with respect to credit risk, as measured by some kind of capital requirements.

In this work we have investigated the asset-allocation problem for a typical medium size Italian life insurance segregated fund portfolio imposing realistic semicontinuous constraints, such as a minimal number of assets composing the optimal portfolio. The optimization is performed in the $\{risk\ capital, current\ yield\}$ plane where Risk Capital is defined as the difference between the loss at the $\alpha = 99\%$ confidence level and the expected (average) loss and current yield is defined as the ratio of the coupon to the current price. Since we consider the reduction of credit risk a second order requirement with respect to interest rate risk control, an additional constraint of constant portfolio duration is also imposed.

2 Credit Risk modelling

Among different models of credit risk we have decided to implement the CreditRisk$^+$ [1] model of Credit Suisse Financial Products. CreditRisk$^+$ is a pure default model: at the end of the valuation period each of the N obligors can be found either in the default or non-default state. In case of default, for each contract j issued by the defaulted obligor, the lender suffers a fixed loss, the exposure, that is determined as a $1 - rr_j$ fraction of the value of the contract, rr_j being the recovery rate assigned on the basis of contractual guarantees (seniority and security).

The unconditional default probability p_i of the i-th obligor is assumed to be known. Correlated defaults are described following a typical actuarial approach: it is assumed that, conditional on the values of K auxiliary variables $(x_1, \ldots, x_K)$, the defaults are independent and thus binomially distributed. Each risk factor x_k is distributed according to a Gamma distribution $\Gamma(\alpha_k, \beta_k)$ with mean μ_k and standard deviation σ_k. These parameters are calibrated in order to recover, on average, the unconditional probability p_i with standard deviation σ_i.

The distribution of defaults is found by convolving the Gamma distributions with the Poisson distribution approximating the true Binomial distribution. This "Poisson approximation" introduces some distortions, noticeably the fact that each obligor can default many times, which become more important as p_i increases.

The loss distribution is derived from the distribution of defaults after rounding the exposures in units of an arbitrary unit scale $\mathcal{U}$. This rounding allows to lower the combinatorial complexity of the problem since the loss suffered in case of default of n out of m contracts having the same discrete exposure ν is simply $n\,\nu$. The effects of this second approximation are generally easier to keep under control than those of the Poisson approximation.

The convolutions needed to obtain the portfolio loss distribution are performed by standard $\mathcal{Z}$ (*i.e.* discrete Laplace) transforms, obtaining the portfolio loss probability generating function $\mathcal{G}(z)$ (z being the auxiliary variable of the transform):

$$\mathcal{G}(z) = Exp\left[\sum_{i=1}^{N} \omega_{0,i}\, p_i\, z^{\nu_i} - \mu_0\right] \times \prod_{k=1}^{K} \left(\frac{1-\delta_k}{1 - \frac{\delta_k}{\mu_k} \sum_{i=1}^{N} \omega_{k,i}\, p_i\, z^{\nu_i}} \right)^{\alpha_k} \quad (1)$$

where $\delta_k = \beta_k/(1+\beta_k)$, ν_i is the exposure of the i-th issuer, ω_{ki} is the weight of the k-th risk driver x_k on the default of the i-th issuer and ω_{0i} is the weight of the idiosyncratic component of risk, which is clearly indentifiable as the source of the first term in eq. (1). The loss distribution is obtained from $\mathcal{G}(z)$ performing the inverse transform by mean of a numerical algorithm originally due to Panjer [2]. As this algorithm is potentially exposed to large numerical errors, after the original formulation of the model, other algorithms have been proposed [3]. However, as discussed by Gordy in [4], it is possible to compute the moments of the loss distribution directly from $\mathcal{G}(z)$ without the Panjer inversion and use them as benchmarks. Moreover these errors are less important when a single sector is considered.

From the loss distribution different measure of risk can be derived, the most popular being the α-quantile Value-at-Risk VaR_α, the expected shortfall E_α and the risk capital RC_α, where the usual *caveats* on the lack of sub-additivity of VaR_α apply [5].

The auxiliary variables x_k can be interpreted as market factors driving credit risk in different independent industry sectors. The possibility of allocating each title to several sectors by mean of the weigths ω allows to implement a complex structure of correlations. The limiting cases of the model are those in which there is a single market factor, and thus correlations are maximal, and that in which there is only idiosyncratic risk, *i.e.* the defaults are independent.

In this analysis we have implemented the model with maximal correlations (*i.e.* with a single industry sector and no idiosyncratic risk), and chosen $\mathcal{U}$ as 1/1000 of the largest exposure. The default probabilities p_i and the corresponding standard deviations σ_i have been taken from the historical determinations provided by Moody's on the basis of the rating of the obligor [6]. Similarly the same source has been used to determine the values of the recovery rates. Risk Capital is defined as the difference between the loss at the $\alpha = 99\%$ confidence level and the expected (average) loss.

3 Portfolio Optimisation

The asset allocation problem we shall address is the following: given a portfolio initially composed of N assets and having defined an additional basket of M assets, we look for the configuration which minimize the Risk Capital, subject to constraints on the total portfolio value, duration and current yield. In addition, for the sake of diversification, we shall also impose that the fraction of assets whose issuers belong to a given industry sector is lower than a fixed threshold and that the optimal portfolio is composed by -at least- a minimum number of assets N_{min}. Thus the problem can be formally written as:

Pb1 *minimize* Risk Capital

$$
\begin{aligned}
\textit{subject to} \quad & \textstyle\sum_i^{N+M} \omega_i = 1 \\
& \textstyle\sum_i^{N+M} \omega_i d_i = D \\
& \textstyle\sum_i^{N+M} \omega_i r_i = R \\
& 0 \leq s_j \leq s_{max} \qquad\qquad s_j = \textstyle\sum_{i \in S_j}^{N+M} \omega_i \\
& \textstyle\sum_i^{M+N} \mathbb{I}_i \geq N_{min} \qquad\quad \mathbb{I}_i = \begin{cases} 0 & \text{if} \quad \omega_i = 0 \\ 1 & \text{if} \quad \omega_i > 0 \end{cases} \\
& \omega_i \in \{0\} \cup [\omega_{min}, \omega_{max}]
\end{aligned}
$$

where D and R are the (Macaulay) duration and the current yield of the portfolio, ω_i, d_i and r_i are respectively the weight, duration and current yield of the i-*th* asset, and finally S_j is the set of assets whose issuers belong to the j-*th* industry sector.

Notice that the last requirement cannot be enforced in a linear or quadratic programming approach and since ω is a *semicontinuous* variable, the problem can be classified as a mixed-integer optimisation problem. This kind of problem cannot be solved by most commercial packages (*e.g.* like NAG, IMSL, MatLab, etc.) but requires either *ad hoc* solutions or the use of metaheuristic techniques.

Other difficulties derive from the fact that the objective function, the Risk Capital, and its derivatives can only assume integer values which, in addition, have to be computed numerically, and from the large difference in the value of assets between the defaulted and non-defaulted states, which is likely to create deep local minima from which it could be difficult to escape.

Metaheuristic techniques as genetic algorithms, simulated annealing and tabu search are generally indicated [8] as the most suited to solve mixed-integer optimisation problems. In fact in [7] a genetic algorithm has been implemented for a risk-return analysis, although with the use of derivatives and without the numerosity constraint. Differently, we have addressed our attention to numerical methods that do not require derivatives, and in particular to simulated annealing.

4 Simulated Annealing

Simulated annealing is a random-search optimisation technique developed in 1983 [9] which exploits the analogy between the way in which a molten metal cools and freezes into a minimum energy crystalline structure (the annealing process) and the search for a minimum in a more general system: The two pillars of the method are the idea that transition from favourable to unfavourable states (that is from lower to higher values of energy, which is the objective function) can take place when the temperature of the system is sufficiently high (thus escaping from local minima in search of the global minimum) and that - lowering the temperature in the appropriate way - all possible minima are attained. More precisely a simulated annealing algorithm can be described as composed by the following steps:

1. starting from the D-dimensional initial point x_k a new point y_{k+1} is sampled from the *next candidate* distribution $D(x_k)$;

2. the new point is randomly accepted or rejected: the acceptance criterion is defined by the *acceptance* function $A(x_k, y_{k+1}, t_k)$:

$$x_{k+1} = \begin{cases} y_{k+1} & \text{if} \quad u \leq A(x_k, y_{k+1}, t_k) \\ x_k & \text{otherwise} \end{cases}$$

   where u is a random number uniformly distributed in $[0, 1]$ and t_k is an additional parameter which, due to the analogy with physics, is called the temperature;

3. update the temperature using the *cooling schedule* $C(\mathcal{F}_k)$:

$$t_{k+1} = C(\mathcal{F}_k)$$

   where $\mathcal{F}_k$ is the information collected up to iteration k;

4. check the *stopping criterion* and if it fails set $k = k+1$ and go back to the first step.

Since the evolution from x_k to x_{k+1} is a random process this kind of algorithm can be viewed as the implementation of a discrete-parameter finite Markov chain. In fact in 1984 [10] it was proven that simulated annealing could statistically find the best minimum, although in an infinite time. As expected the choice of $D(x_k)$, $A(x_k, y_{k+1}, t_k)$, $C(\mathcal{F}_k)$ and of the stopping criterion play a decisive role in determining the speed of convergence. Particularly important is the relationship between $D(x_k)$ and $C(\mathcal{F}_k)$ since the convergence is assured only when a sufficiently large domain is sampled before lowering the temperature. The analogy with physics had originally suggested the choices of the so-called Metropolis-Boltzmann annealing:

$$D(x, y, T) = (2\pi T)^{-D/2} e^{-\Delta x^2/(2T)} \quad \Delta x = x - y$$
$$A(x, y, T) = min\{1, e^{-[E_{k+1} - E_k]/T}\} = min\{1, e^{-\Delta E/T}\}$$
$$C(\mathcal{F}_k) \quad \text{slower than} \quad T_k = \frac{T_0}{\ln k}$$

where -traditionally- $E_k = E(x_k)$ indicates the value in x_k of the *cost function* $f(x)$. Notice that in this case also $D(x)$ is a function of the temperature.

Since sampling is fundamental it is not surprising that a next candidate function with heavier tails than those of the Boltzmann distribution is able to assure a faster convergence. In fact in 1984 a method of fast annealing [11] was developed, which permitted to lower the temperature exponentially faster and guaranteed the convergence in a finite time. Shortly thereafter an even faster method [12], originally called very fast re-annealing and now known as adaptive simulated annealing (ASA) [13] was developed. In this case one has:

$$D(x, y, T) = \prod_i^D \frac{1}{2(|\Delta x|_i + T_i) \ln(1 + 1/T_i)}$$
$$A(x, y, T) = \frac{e^{-E_{k+1}T}}{e^{-E_{k+1}T} + e^{-E_k T}} = \frac{1}{1 + e^{\Delta E T}}$$
$$C(\mathcal{F}_k) \quad \text{slower than} \quad T_k = T_0\, e^{k^{1/D}}$$

where now there are $D+1$ temperatures (one for each parameter and one for the acceptance function). The ASA code is probably one of the most tested codes in the literature and applications in a large variety of disciplines have been reported [14].

Unfortunately, whilst simulated annealing has no need for derivatives and can implement very easily combinatorial requirements, one of its major drawbacks is that it could be painful to implement equality constraints in high-dimensional problems. This is because when the search is limited to a finite subspace the hit-or-miss ratio can expand the computing time beyond reasonable limits. In this work we have adapted the ASA code to cope with the linear constraints of our problem as explained in the next paragraph.

5 The optimisation procedure

Instead of changing the next candidate generation in order to satisfy the equality constraint, we have developed a simple yet effective procedure to "project" the ASA candidate point onto the surface representing the equality constraints. Thus there will be two (in fact three) parameters arrays evolving in parallel, one according to the principles of simulated annealing, and another obtained from the first with simple algebra. In practice the optimisation algorithm is composed of the following steps:

1. given the initial parameter array ω provided by the ASA generation mechanism two new arrays are computed: Ω, which is ω normalized to 1 in such a way that every component is either null or belonging to the interval $[\omega_{min}, \omega_{max}]$ and η which is obtained by changing any three components of Ω in such a way to satisfy the three linear constraints of the problem:

$$\begin{aligned} ||\eta|| &= \textstyle\sum_i^{N+M} \eta_i = 1 \\ d(\eta) &= \textstyle\sum_i^{N+M} \eta_i \left(\frac{d_i}{D}\right) = 1 \\ r(\eta) &= \textstyle\sum_i^{N+M} \eta_i \left(\frac{r_i}{R}\right) = 1 \end{aligned}$$

2. if η exists and satisfy all the other requirements, the risk capital is computed using a portfolio defined by η and this value is returned to ASA; otherwise the procedure returns to ASA a penalty function of the type

$$P = c\left\{Abs(1 - ||\Omega_k||) + Abs(1 - d(\Omega_k)) + Abs(1 - r(\Omega_k)) + min(s(\Omega_k)/s_{max} - 1, 0)\right\}$$

   where c is an arbitrary large constant (*e.g.* 10000) and $s(\Omega_k)$ is the largest of the contributions to a given industry sector; notice that to impose that at least N_{min} components of η are larger than ω_{min} is trivial;

3. independently of η, for all the k components of Ω belonging to $[\omega_{min}, \omega_{max}]$, the corresponding components of ω are changed on exit to those of Ω, that is:

$$\omega_k = \begin{cases} \Omega_k & \text{if } \Omega_k \in [\omega_{min}, \omega_{max}] \\ \omega_k & \text{otherwise} \end{cases}$$

this is done to speed up the convergence of ω to a normalized array, assuring at the same time the continuity of the components that can fall below the ω_{min} threshold;

4. the algorithm is stopped after 500 iterations if no improvements have been found in the last 100, otherwise other 500 iterations are allowed.

The code is written in ANSI C and Fortan90 and run on a Pentium IV IBM laptop with the windows XP/cygwin operating system, without optimisation. The determination (with companion diagnostics) of a single point takes about 15 minutes (real time).

6 Results

For a numerical investigation we have considered an ideal portfolio mimicking the typical composition of Italian life insurance segregated fund portfolios. The portfolio is composed by 80 bonds identified by their issuer, the industry sector of the issuer, the Moody's rating of the issuer, the Macaulay duration and the current yield. The additional basket of opportunities is composed by other 11 bonds. The figures of merit of the initial composition of the portfolio are described in table 1, together with the analysis constraints. Notice that the initial configuration is diversified in a roughly uniform way over four industry sectors (finance, industrial, etc.), with a small addition of bonds in a fifth one; finally, to further increase diversification, the basket of new opportunities contains bonds from a sixth sector. Since the total value of the portfolio is irrelevant for the analysis and for the sake of comparison, the risk capital has been scaled to 1. The loss distribution is shown in figure 1. The loss at the 99% confidence level correspond to about 5.56 times the standard deviation of the loss distribution, while the probability of no losses is about 75%.

	Initial Value	Constraint
Current Yield	4.00%	-
Risk Capital	1 a.u.	-
Duration	5 years	5 years
$max(\omega_i)$	8.25%	$\leq$ 8.68%
$min(\omega_i)$	0.33%	$\geq$ 0.32%
Sector 1	38.1%	$\leq$ 50%
Sector 2	3.8%	$\leq$ 50%
Sector 3	20.6%	$\leq$ 50%
Sector 4	18.0%	$\leq$ 50%
Sector 5	19.5%	$\leq$ 50%
Sector 6	0.0%	$\leq$ 50%
N_{min}	-	64

Table 1: Figures of the initial portfolio configuration and analysis constraints; notice that the risk capital has been scaled to 1 and thus it is expressed in arbitrary units.

Before solving the mixed-integer optimisation problem we have performed a check on the ability of procedure to solve the simplified problem (Pb 2):

$$\begin{array}{lll} \textbf{Pb2} & \textit{minimize} & \text{Risk Capital} \\ & \textit{subject to} & \sum_i^{N+M} \omega_i = 1 \\ & & \sum_i^{N+M} \omega_i d_i = D \\ & & \sum_i^{N+M} \omega_i r_i = R \\ & & 0 \le s_j \le s_{max} \qquad s_j = \sum_{i \in S_j}^{N+M} \omega_i \\ & & \omega_i \in [\omega_{min}, \omega_{max}] \end{array}$$

which is Pb 1 for the case $N_{min} = N + M$. As a benchmark test we will compare the results obtained with our technique with those obtained with a standard gradient-driven approach, namely the DCONF algorithm of the IMSL library [15], which implements a modified version of the well-known TOLMIN algorithm [16].

Current Yield	3.65%	3.80%	4.00%	4.20%	4.40%	4.50%	4.60%	4.70%
$\#\omega_i = 0$	27	27	26	27	27	27	27	24
$max(\omega_i)$	8.54%	6.39%	6.79%	8.32%	7.43%	7.86%	8.49%	8.30%
Sector 1	46.7%	48.4%	47.7%	44.5%	45.0%	37.6%	38.2%	36.0%
Sector 2	4.3%	4.1%	3.7%	4.0%	3.5%	3.7%	4.4%	4.8%
Sector 3	11.6%	11.2%	10.1%	10.8%	16.6%	21.4%	20.4%	22.0%
Sector 4	20.9%	20.2%	18.2%	19.4%	17.2%	18.3%	19.7%	18.5%
Sector 5	15.4%	14.8%	19.1%	20.3%	16.7%	18.0%	16.6%	17.9%
Sector 6	1.2%	1.1%	1.0%	1.1%	1.0%	1.1%	0.7%	0.8%

Table 2: Figures of the efficient frontier points of Pb1: the table reports the number of assets discarded, the largest asset weight and the contributions of each of the six industry sectors.

Figure 2 resumes the results of the three analyses. Table 2 reports the composition of the portfolios on the efficient frontier of Pb 1. As one can see there is a substancial similarity in the frontiers obtained with the two methods for Pb2. On the contrary the frontier of Pb 1 is significantly different from that of Pb 2.

7 Conclusions

We have successfully investigated the ability of adaptive simulated annealing in addressing the portfolio optimisation problem in presence of "standard" linear constraints and additional numerosity constraints for a typical medium size life insurance segregated fund portfolios. The presence of a constraint on portfolio duration ensures that the minimal risk portfolios also meet the requirements of a classical asset-liability management. Extensions of the method to include other constraints should be straightforward. Relevant differences are found when the number of assets in the portfolio is allowed to vary.

References

[1] CreditRisk$^+$ technical document,
Credit Suisse Financial Products, available at *http:://www.csfb.com/creditrisk*

[2] Insurance Risk Models,
H. H. Panjer, G. E. Wilmott, Society of Actuaries, Schaumberg (1992)

[3] Numerically stable computation of CreditRisk$^+$,
H. Haaf *et al.*, in Credit Risk$^+$ in the Banking Industry, Springer-Verlag (2004)

[4] Saddlepoint approximation of CreditRisk$^+$,
M. Gordy, Journal of Banking & Finance 26, 1335-1353 (2002)

[5] Coherent measures of risk,
P. Artzner *et al.*, Mathematical Finance 9, 3, 203-228 (1999)

[6] Default and Recovery Rates of Corporate Bond Issuers, 1920-2004,
D. T. Hamilton *et al.*, Moody's Investor Service (2005)
available at *http:://www.moodyskmv.com*

[7] Risk-return analysis of credit portfolios,
F. Schlottmann *et al.*, in Credit Risk$^+$ in the Banking Industry, Springer-Verlag (2004)

[8] Heuristics for cardinality constrained portfolio optimization,
T. J. Chang *et al.*, Computers and Operations Research, 27, 1271-1302 (2000)

[9] Optimization by simulated annealing,
S. Kirkpatrick *et al.*, Science 220 (4598), 671-680 (1983)

[10] Stochastic relaxation, Gibbs distribution and the Bayesian restoration in images,
S. Geman *et al.*, IEEE Trans. on Pattern Analysis and Machine Intell. 6, 721-741 (1984)

[11] Fast simulated annealing,
H. Szu *et al.*, Physics Letters A 122, 157-162 (1984)

[12] Very fast simulated re-annealing,
L. Ingber, Mathematical Computer Modelling 12, 967-973 (1989)

[13] The ASA code is available at *http:://www.ingber.com*

[14] Simulated Annealing: practice versus theory,
L. Ingber, Mathematical Computer Modelling 18, 29-57 (1993)

[15] http://www.absoft.com/Products/Libraries/imsl.html

[16] TOLMIN: A Fortran Package for Linearly Constrained Optimization Calculation,
DAMTP, 1989/NA2, (1989).

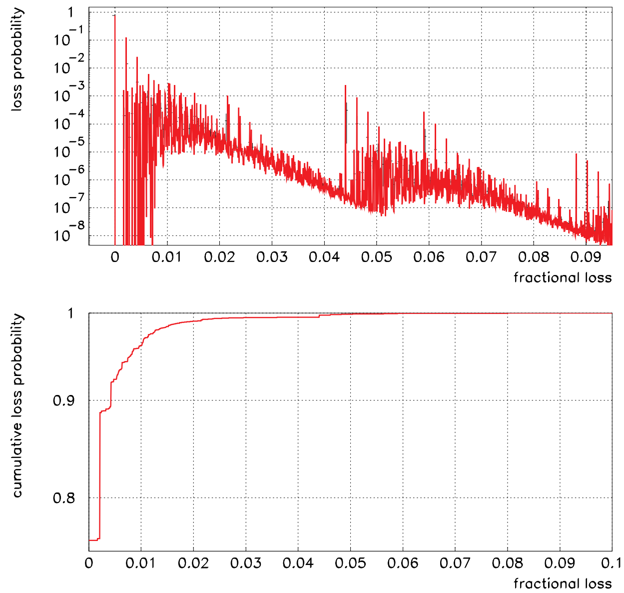

Figure 1: Fractional loss (loss divided total portfolio value) distribution and its cumulative for the initial portfolio configuration.

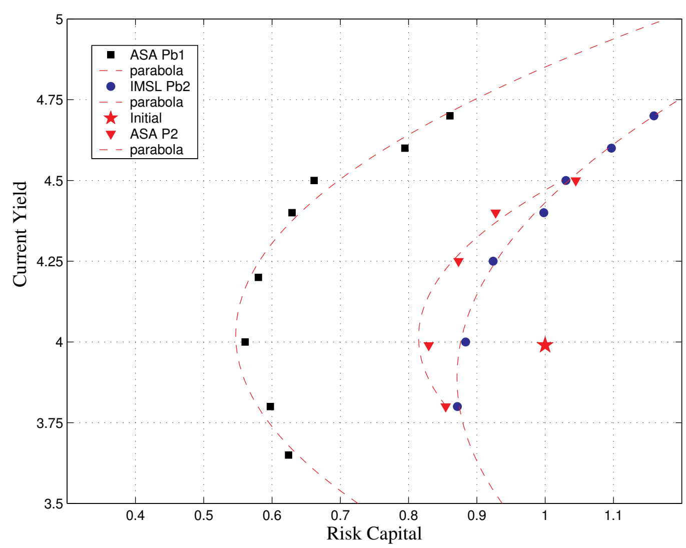

Figure 2: Efficient frontier in $\{risk\ capital, current\ yield\}$ plane for Pb1 (squares) and Pb2 (triangles and dots for simulated annealing and IMSL respectively); lines are parabolas drawn to guide the eye.