AFIR 2005
Zurich, Sep. 9^{th} 2005

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

L. Passalacqua

Dipartimento di Scienze Attuariali e Finanziarie
Università di Roma *La Sapienza*
luca.passalacqua@uniroma1.it

Outline

- Credit Portfolio Optimization $\rightsquigarrow$ the goal
- Credit Risk Modelling $\rightsquigarrow$ the cost function
- Mixed-Integer Constraints $\rightsquigarrow$ "the" problem
- Simulated Annealing (SA) $\rightsquigarrow$ the method
- Modified Adaptive SA $\rightsquigarrow$ a taylored solution
- Results
- Comments & Outlook

Credit Portfolio Optimization

In this work we address the problem of capital allocation when a constraint of minimal numerosity of the optimized portfolio is imposed. This condition transforms the classical Markowitz problem in a combinatorial problem, much harder to solve.

In other words we like to impose that the contribution ω_i of the i-th asset ($i = 1, \ldots, N$) to the optimized portfolio is either null or belonging to a fixed interval:

$$\omega_i \in \{0\} \cup [\omega_{min}, \omega_{max}]_i$$

which is a non-convex set, and impose that

$$\sum_{i=1}^{N} \mathbb{I}_i \geq N_{min} \quad \mathbb{I}_i = \begin{cases} 0 & \text{if } \omega_i = 0 \\ 1 & \text{if } \omega_i > 0 \end{cases}$$

In particular we have addressed the problem of optimizing the composition of a typical Italian life insurance segregated fund portfolio in the plane **(risk capital, current yield)**, imposing an additional constraint on portfolio duration.

From the methodological point of view, the actual choice of the variables does not affect the combinatorial nature of the problem. One can easily substitute risk capital with *e.g.* value-at-risk, expected shortfall or other risk measures, as well as use a different measure of return or a risk-adjusted measure of performance.

For our particular choice the problem is then:

$$\textbf{Pb1} \quad \textit{minimize} \quad \text{Risk Capital} = RC(\{\omega\})$$

$$\textit{subject to} \quad \sum_{i=1}^{N+M} \omega_i = 1$$

$$\sum_{i=1}^{N+M} \omega_i d_i = D$$

$$\sum_{i=1}^{N+M} \omega_i r_i = R$$

$$s_j \in [0, s_{max}]_j \quad \forall j \qquad\qquad s_j = \sum_{i \in S_j}^{N+M} \omega_i$$

$$\sum_{i=1}^{N+M} \mathbb{I}_i \geq N_{min} \qquad\qquad \mathbb{I}_i = \begin{cases} 0 & \text{if } \omega_i = 0 \\ 1 & \text{if } \omega_i > 0 \end{cases}$$

$$\omega_i \in \{0\} \cup [\omega_{min}, \omega_{max}]_i$$

where we have considered a portfolio initially composed of N assets together with a basket of M additional assets; d_i and r_i are respectively the Macauley duration and current yield of the i-th asset, s_j is the contribution of the portfolio to the S_j ($j = 1, \ldots, s$) industry sectors and the variables on the right side of the equations are the problem constraints.

For a numerical example we have chosen CreditRisk$^+$ to compute risk capital. The extension to CreditMetrics-*like* models and stochastic duration does not add further complexity to the problem.

CreditRisk$^+$

CreditRisk$^+$ [Cre97] is a very well known model developed by CSFB. Mostly for the sake of notation we recall the following:

1. the probability-generating function $\mathcal{G}(z)$ of the loss distribution is given by:

$$\mathcal{G}(z) = exp\left[\sum_{i=1}^{N} \omega_{i0}\, p_i\, z^{\nu_i} - \mu_0\right] \prod_{k=1}^{K} \left(\frac{1-\delta_k}{1 - \frac{\delta_k}{\mu_k} \sum_{i=1}^{N} \omega_{ik}\, p_i\, z^{\nu_i}}\right)^{\alpha_k}$$

   where for each obligor $i\,(i = 1, \ldots, N)$ p_i and $\nu_i \in \mathbb{N}$ are respectively the unconditional default probability and the exposure expressed in units of an arbitrary base unit $\mathcal{U}$; the first term is the contribution of the idiosyncratic risk and the product extends to the number K of common market latent variables affecting the default, which are $\Gamma(\alpha_k, \beta_k)$ distributed random variables with mean μ_k and variance σ_k^2; finally ω_{ik} are the factor loadings for obligor i and latent variable k while δ_k stands for $\beta_k/(1+\beta_k)$.

2. The n-th order moments of the loss distributions μ_n are obtained directly from $\mathcal{G}(z)$:

$$\mu_n = \left(z\frac{\partial}{\partial z}\right)^n G(z)\Big|_{z=1}$$

   and (at least in our case) are used to control the numerical precision of the inverse transform algorithm.

3. Since moments of the loss distributions are analytically obtained, also the sensitivities to portfolio composition $\partial\mu_n/\partial\nu_i$ are known. Unfortunately no exact expression for the sensitivity of the Risk Capital $\partial RC/\partial\nu_i$ exists. An approximated expression is suggested in the original formulation by setting:

$$RC = \mu_L + \xi\sigma_L$$

where μ_L and σ_L are respectively the mean and standard deviation of the loss distribution, and ξ is a constant implicitly defined by the above equation.

In some works this expression is used to avoid the numerical computation of the derivative, in order to save computing time. However we prefer to avoid using it.

In this work we have used the model in its original formulation, that is operating the inverse $\mathcal{Z}$ transform by mean of the Panjer algorithm. Other inversion algorithms are described by various authors, *e.g.* in [GL04].

In this work the Risk Capital is computed in a conservative way, that is with maximal correlation in default, corresponding to the choice of a single market variable and no idiosyncratic risk. This configuration has the additional advantage that the Panjer algorithm is more robust.

Mixed-Integer Constraints

Metaheuristic techniques as genetic algorithms, simulated annealing and tabu search are usually indicated [CMBY00] as the most suited methods to solve problems with so-called mixed-integer constraints:

$$\textbf{Pb1} \quad \textit{minimize} \quad \text{Risk Capital} = RC(\{\omega\})$$

$$\textit{subject to} \quad \sum_{i=1}^{N+M} \omega_i = 1$$

$$\sum_{i=1}^{N+M} \omega_i d_i = D$$

$$\sum_{i=1}^{N+M} \omega_i r_i = R$$

$$s_j \in [0, s_{max}]_j \quad \forall j \qquad\qquad s_j = \sum_{i \in S_j}^{N+M} \omega_i$$

$$\sum_{i=1}^{N+M} \mathbb{I}_i \geq N_{min} \qquad\qquad \mathbb{I}_i = \begin{cases} 0 & \text{if } \omega_i = 0 \\ 1 & \text{if } \omega_i > 0 \end{cases}$$

$$\omega_i \in \{0\} \cup [\omega_{min}, \omega_{max}]_i$$

Recently, in [GL04] a **genetic algorithm** has been implemented for a risk-return analysis, although without the numerosity constraint.

In this work we have explored the use of **simulated annealing**.

Simulated Annealing (SA)

Annealing: treatment of a metal or alloy by heating to a predetermined temperature, holding for a certain time, and then cooling to room temperature to improve ductility and reduce brittleness. (from *Encyclopaedia Britannica*)

Simulated annealing is an heuristic optimization technique pioneered by Kirkpatrick *et al.* [KJV83] and rooted in the work of Metropolis *et al.* [MRR+53] originally aimed to simulate the evolution of a solid to thermal equilibrium. Born to solve physical problems it has spread out in many other fields (circuit design, imaging, biology, etc.) still keeping the original "physicists "jargon.

SA has the following advantages:

⇒ is rigorously [LA88] described in terms of **Markov Chains**
⇒ **finite time convergence** has been proven
⇒ it does not make use of derivatives
⇒ the basic algorithm is relatively **simple to code** [PVTF92]
⇒ complex inequality constraints are very easily implemented
⇒ has been **successfully employed** to solve a large variety of combinatorial problems (*e.g.* the travelling salesman)
⇒ there is a **large community** using and still developing it

SA has the following disadvantages:

⇒ convergence speed is a matter of tuning
⇒ strict equality constraints are less "natural "and their implementation is poorly documented in the literature

Simulated Annealing (SA)

The basic SA algorithm proceeds as follows:

$\Rightarrow$ Start from iteration k at point $\{x_k\} \in \Omega \subseteq \mathbb{R}^n$
$\Rightarrow$ Choose a **candidate** next point $\{y_{k+1}\}$ from distribution $D(x_k)$
$\Rightarrow$ Decide to accept/reject $\{y_{k+1}\}$ on the basis of $A(x_k, y_{k+1}, t_k)$

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

$u \in [0,1] \sim uniform$
t_k is the a control parameter called **temperature**

$\Rightarrow$ Change t_k according to the **cooling schedule** $C(\mathcal{F}_k)$
where $\mathcal{F}_k$ is the information up to the iteration k
$\Rightarrow$ Check the **stopping criterion**

Historical choices for $D(x)$, $A(x, y, t)$, $C(\mathcal{F}_k)$ are those of so-called Metropolis-Boltzmann annealing:

$$D(x, y, t) = \frac{e^{-\Delta x^2/2t}}{(2\pi t)^{n/2}} \qquad \Delta x^2 = |x - y|^2$$
$$A(x, y, t) = min\{1, e^{-\Delta E/t}\} \qquad \Delta E = |E(x) - E(y)|$$
$$C(\mathcal{F}_k) \qquad \text{slower than } t_k = t_0/\ln k$$

where $E_k = E(x_k)$ is the value in x_k of the **cost function**. Notice that $D(x)$ is a function of $\{x\}$ and of the temperature.

Adaptive SA (ASA)

The $\{x_k\} \rightarrow \{x_{k+1}\}$ move allows to escape from local minima as much as the temperature t is large.

The algorithm is started at high temperature, spends a **proper** number of trials at each temperature, while the temperature is slowly decreased until the system is **frozen** in the global minimum.

$\Rightarrow$ the speed of convergence depends on $D(x)$, $A(x, y, t)$ and $C(\mathcal{F}_k)$

Since sampling Ω is fundamental, similarly to **importance sampling**, it is convenient to use next candidate functions with heavy tails.

Improvements were obtained in 1987 [SH87], which permitted to lower the temperature exponentially faster and in 1989 [Ing89], with a method presently known as adaptive simulated annealing (ASA). In this case one has $n + 1$ temperatures:

$$D(x, y, t) = \prod_{i=1}^{n} \left[2(|\Delta x|_i + t_i) \, \ln\left(1 + 1/t_i\right) \right]^{-1} \quad \text{“Cauchy-like ”}$$

$$A(x, y, T) = \frac{1}{1 + e^{\Delta E T}}$$

$C(\mathcal{F}_k)$ slower than $t_k = t_0 \, e^{k^{1/n}}$

Modified ASA

Standard SA algorithms do not foresee equality constraints. These can be implemented in three ways:

- ▷ with Lagrange multipliers $\leftrightsquigarrow$ works only for **small** n
- ▷ modifying $D(x, y.t)$ $\leftrightsquigarrow$ under investigation
- ▷ projecting onto Ω_C $\leftrightsquigarrow$ **this work**

where $\Omega_C \subset \Omega$ is the hyper-surface defined by the constraints.

Technically the projection is made in two steps:

$$\{\omega_k\} \rightarrow \{\Omega_k\} \rightarrow \{\eta_k\} \rightarrow \text{cost function}$$

$$\{\omega_k\} \rightarrow \{\tilde{\omega}_k\} \rightarrow \text{next SA move}$$

where:

$\{\omega_k\}$	$\in$	Ω	$\lvert\omega_k\rvert \neq 1$
$\{\Omega_k\} \simeq \{\omega_k\}$	$\in$	$\{0\} \cup \left[\{\omega_{min}\}, \{\omega_{max}\}\right]$	$\lvert\Omega_k\rvert = 1$
$\{\eta_k\}$(the "shadow"of $\{\Omega_k\}$)	$\in$	Ω_C	$\lvert\eta_k\rvert = 1$

$$\{\tilde{\omega}_k\}_i = \begin{cases} \{\Omega_k\}_i & \text{if } \{\Omega_k\}_i \in [\{\omega_{min}\}_i, \{\omega_{max}\}_i] \\ \{\omega_k\}_i & \text{otherwise} \end{cases} \quad \in \Omega \quad |\tilde{\omega}_k| \neq 1$$

η_k is obtained by changing n_C components of Ω_k in order to satisfy the n_C linear constraints. Notice that:

- $\Rightarrow$ $\{\Omega_k\}, \{\tilde{\omega}_k\}$ are always well defined
- $\Rightarrow$ the move $\{\omega_k\} \rightarrow \{\tilde{\omega}_k\}$ brings "as near as possible" to Ω_C it is effective mainly at the beginning as ω_0 can be very far from Ω_C
- $\Rightarrow$ $\{\eta_k\}$ does not always exist in that case a penalty function is returned

Numerical Application

For a numerical application we have mimicked a typical Italian life insurance segregated fund portfolio, initially composed of $N = 80$ assets together with a basket of $M = 10$ additional assets. The portfolio is chosen so to have a duration of 5 years and a current yield of 4%.

For simplicity sake the constraints imposed on the asset contributions and on the industy sector contributions are - respectively- all equal:

$$[\omega_{min}, \omega_{max}] = [0.3, 9]\% \qquad s_{max} = 50\%$$

The principal figures of the portfolio are resumed in the following table.

	Initial Value	Constraint
Current Yield	4.00%	-
Risk Capital	1 a.u.	-
Duration	5 years	5 years
ω_{max}	8%	$\leq 9\%$
ω_{min}	0.3%	$\geq 0.3\%$
Sector 1	38%	$\leq 50\%$
Sector 2	4%	$\leq 50\%$
Sector 3	20%	$\leq 50\%$
Sector 4	18%	$\leq 50\%$
Sector 5	20%	$\leq 50\%$
Sector 6	0.0%	$\leq 50\%$
N_{min}	-	64

Validation of the algorithm

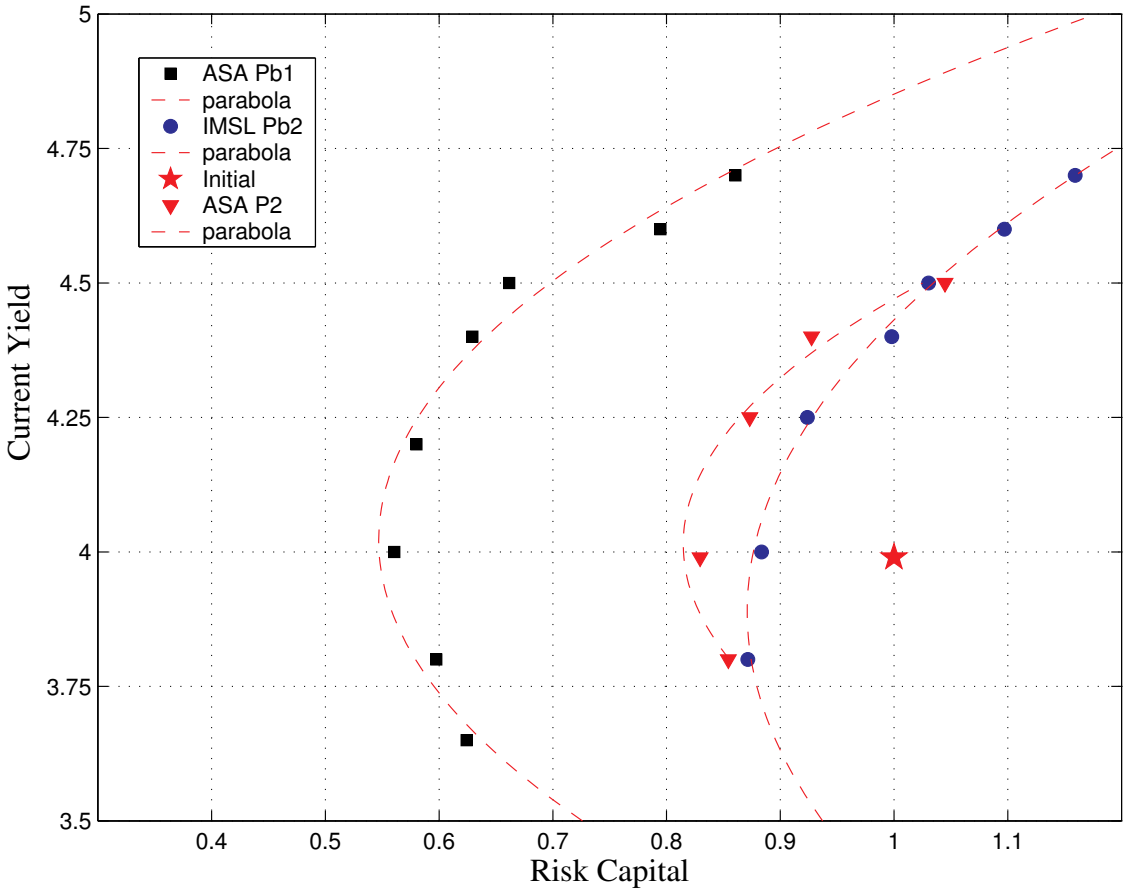
The algorithm is written in C and Fortan90 and run on a Pentium IV IBM laptop with the MSFT windows XP/cygwin operating system. The determination (with companion diagnostics) of a single point takes about 15 minutes of real time.

To validate the optimization procedure we first considered the following simplified problem:

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

which is nothing but **Pb 1** when $N_{min} = N + M$. As a performance test we compared the results obtained with our technique with those obtained with a standard gradient-driven algorithm, namely DCONF of the IMSL library [IMS].

Results



Efficient frontier in (risk capital, current yield) plane for Pb1 (squares) and Pb2 (triangles and dots for simulated annealing and IMSL respectively). The risk capital is expressed as fraction of the initial risk capital, the current yield in percent. Lines are parabolas drawn to guide the eye.

▷▷ the **Pb2** frontier obtained with SA and a standard technique are comparable $\rightsquigarrow$ SA validated.

▷▷ In most points of the frontier of **Pb 1** the solution is such that $N + M = N_{min}$.

Comments & Outlook

▶ As expected simulated annealing is a viable solution for solving capital allocation problems with numerosity constraints.

▶ Similarly to the classical Markowitz problem the procedure seems willing to discard as much as "bad "assets as possible; a numerosity constraint could therefore be used as an emprical handle to ensure diversification.

▶ It is very likely that modifications of the variables or the constraints, like the use of risk-adjusted performance variables or the introduction of transaction costs, will not alter sigificantly the difficulty of the problem.

▶ evolution of the technique are under study in order to modify the next-candidate distribution so to satisfy the constraints automatically.

▶ a second generation improvement could possibly come by the use of the technique known as *parralell tempering* which could increase the speed of convergence by making the problem less sensitive to the starting point.

References

[CMBY00] T.J. Chang, N. Meade, J.E. Beasley, and Y.M.Sharaiha. Heuristics for cardinality constrained portfolio optimization. *Computers and Operations Research*, 13(27):1271–1302, 2000.

[Cre97] Credit Suisse First Boston. *CreditRisk$^+$ technical document*, 1997.

[GL04] M. Gundlach and F. Lehrbass. *CreditRisk$^+$ in the Banking Industry*. Springer-Verlag, 2004.

[IMS] http://www.absoft.com/products/libraries/imsl.html.

[Ing89] L. Ingber. Very fast simulated re-annealing. *Mathematical Computer Modelling*, 3, 1989.

[KJV83] S. Kirkpatrick, C.D. Gelatt Jr, and M.P. Vecchi. Optimization by simulated annealing. *Science*, (220):671–680, 1983.

[LA88] P. J. M. Laarhoven and E. H. L. Aarts. *Simulated Annealing: Theory and Applications*. Kluwer Academic Publisher, 1988.

[MRR$^+$53] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. Equation of state calculation by fast computing machine. *Journal of Chemical Physics*, (21):1087–1092, 1953.

[PVTF92] W.H. Press, W.T. Vetterling, S.A. Teukolsky, and B.P. Flannery. *Numerical Recipes in C*. Cambridge University Press, second edition, 1992.

[SH87] S.Szu and R. Hartley. Fast simulated annealing. *Physics Letters A*, (122):157–162, 1987.