Stochastic Quasi-Gradient Techniques in VaR-Based ALM Models

Alexander V. Mertens

Kiev National Taras Shevchenko University, Department of Economic Cybernetics;

International Center of Privatization, Investment and Management (International Institute of Business)

Abstract

The paper presents VaR-based stochastic optimization model of asset-liability management with no particular assumptions about distribution of returns and other random parameters. VaR is widely accepted measure of portfolio risk and a number of research on implementation of the VaR measure in portfolio optimization models appeared recently. Proposed approach allows to include VaR constraints into the optimization model using combination of Monte-Carlo simulation and stochastic quasi-gradient techniques.

1. Value-at-Risk measure in the optimization framework

In the last decade Value-at-Risk (*VaR*) became industry standard as measure of risk of investment portfolios and widely used tool for risk evaluation and control. The important problem which is considered in this context is the construction of the portfolio with pre-determined constraints on *VaR* or with minimum possible *VaR*. As a consequence the task of including *VaR* measure into optimization problem appears.

Traditionally *VaR* is determined as lowest amount L such that with probability α the loss in portfolio value will not exceed L within some time interval t. Approaches for calculation of *VaR* can be divided into two groups. The first approach is based on assumption of probability distribution of risk factors (normal for returns or log-normal for values). The second approach use Monte-Carlo simulations generating scenarios and does not depend on particular distributional assumptions. The latter is very often the only possible technique for large portfolios especially when it include instruments with options properties. But including *VaR* into optimization problem is difficult due to bad mathematical properties (e.g. non-convexity).

One approach for optimizing *VaR* which is calculated from scenarios was proposed in Rocafellar and Uryasev (1999). It use alternative risk measure, Conditional Value-at-Risk (*CVaR*). *CVaR* is defined as conditional expectation of losses above amount L (where L is defined as in above definition of *VaR*). *CVaR* is better than *VaR* in terms of its properties, but at the same time, optimizing *CVaR* is very close to optimizing *VaR*.

According to Rocafellar and Uryasev (2000) if we denote $f(\mathbf{x}, \mathbf{r}(\theta))$ - the function of losses in portfolio value which depends on decision vector $\mathbf{x}$ and random parameters vector $\mathbf{r}(\theta)$, and $\Phi(\mathbf{x}, L)$ - the probability that losses will not exceed L:

$$\Phi(\mathbf{x}, L) = \int_{\mathbf{r}:\, f(\mathbf{x},\mathbf{r}) \leq L} p(\mathbf{r}) d\mathbf{r}, \qquad (1)$$

VaR and *CVaR* could be defined respectively as:

$$VaR(\mathbf{x}, \alpha) = \min\{L : \Phi(\mathbf{x}, L) \geq \alpha\}, \qquad (2)$$

$$CVaR(\mathbf{x}, \alpha) = \frac{1}{1-\alpha} \int_{\mathbf{r}:\, f(\mathbf{x},\mathbf{r}) \geq VaR(\mathbf{x},\alpha)} f(\mathbf{x}, \mathbf{r}) p(\mathbf{r}) d\mathbf{r}. \qquad (3)$$

With function $F_\alpha(\mathbf{x}, L)$ defined as:

$$F_\alpha(\mathbf{x}, L) = L + \frac{1}{1-\alpha} \int_{\mathbf{r} \in R^m} \max\{f(\mathbf{x}, \mathbf{r}) - L, 0\} p(\mathbf{r}) d\mathbf{r} \qquad (4)$$

(which is convex and continuously differentiable as function of α), *CVaR* and *VaR* could be expressed as following:

$$CVaR(\mathbf{x}, \alpha) = \min_{L \in R} F_\alpha(\mathbf{x}, L), \qquad (5)$$

$$VaR = \min\left\{\arg\min_{L \in R} F_\alpha(\mathbf{x}, L)\right\}. \qquad (6)$$

Theorem 2 in Rocafellar and Uriasev (2000) postulates that minimizing of *CVaR* over $\mathbf{x} \in X$ is equivalent to minimizing $F_\alpha(\mathbf{x}, L)$ over $(x, L) \in X \times R$ so that:

$$\min_{\mathbf{x} \in X} CVaR(\mathbf{x}, \alpha) = \min_{(\mathbf{x}, L) \in X \times R} F_\alpha(\mathbf{x}, L), \qquad (7)$$

and while $F_\alpha(\mathbf{x}, L)$ is convex on $(\mathbf{x}, L)$, and if X is convex set, minimization of $F_\alpha(\mathbf{x}, L)$ is convex problem.

In practice, having the set of observations over random vector $\mathbf{r}(\theta)$:

$$\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N,$$

we can approximate the function $F_\alpha(\mathbf{x}, L)$:

$$\overline{F}_\alpha(\mathbf{x},L)=L+\frac{1}{(1-\alpha)N}\sum_{i=1}^{N}\max\{f(\mathbf{x},\mathbf{r}_i)-L,0\}. \tag{8}$$

Function $\overline{F}_\alpha(\mathbf{x},L)$ could be included into optimization problem either as objective:

$$\min_{\mathbf{x},L}\overline{F}_\alpha(\mathbf{x},L), \tag{9}$$

or as constraint:

$$\overline{F}_\alpha(\mathbf{x},L)\le\overline{L}\ . \tag{10}$$

In order to linearize the problem, one can use auxiliary variables, so that:

$$\overline{F}_\alpha(\mathbf{x}.L)=L+\frac{1}{(1-\alpha)N}\sum_{i=1}^{N}u_i\ , \tag{11}$$

$$f(\mathbf{x},\mathbf{r}_i)-L\le u_i,\quad u_i\ge 0,\quad i=1,\ldots,N\ . \tag{12}$$

2. VaR-based optimization in the two-stage model

In many practical problems of asset-liability management some corrections of decision parameters could be made after obtaining information about realizations of random parameters. It means two-stage (or multi-stage in general case) setting of decision problem. Some decision variables should be chosen *before* observation of the state of nature, but some corrections are possible *after* observation. In this case, for example, the loss function could be presented as $f(\mathbf{x},\mathbf{y}(\mathbf{x},\theta),\mathbf{r}(\theta))$, where $\mathbf{x}$ is *ex ante* (strategic) decision and $\mathbf{y}(\mathbf{x},\theta)$ is correction (or *adaptation*) which is dependent on the decision $\mathbf{x}$ and the state of nature.

The problem for minimizing *CVaR* could be written as follows:

$$\min_{\mathbf{x},\mathbf{y},L} L+\frac{1}{1-\alpha}E[u(\theta)] \tag{13}$$

s.t.:

$$f(\mathbf{x},\mathbf{y}(\theta),\mathbf{r}(\theta))-L\le u(\theta)\quad a.s., \tag{14}$$

$$u(\theta)\ge 0\quad a.s.,\quad \mathbf{x}\in X,\quad \mathbf{y}(\mathbf{x},\theta)\in Y(\theta) \tag{15}$$

If loss function $f(\mathbf{x},\mathbf{y}(\mathbf{x},\theta),\mathbf{r}(\theta))$ is linear in $\mathbf{x}$ and $\mathbf{y}$ and sets X and Y include only linear in $\mathbf{x}$ and $\mathbf{y}$ constraints, problem (13) - (14) become well known *two-stage linear stochastic programming problem*. In its general form this problem can be written as:

$$
\begin{aligned}
&\max_{\mathbf{x}} \quad \mathbf{x}^T \mathbf{a} + E\left[\mathbf{y}(\mathbf{x},\theta)^T \mathbf{b}(\theta)\right],\\
&s.t.:\\
&\mathbf{D}\mathbf{x} \le \mathbf{d}, \quad \mathbf{x} \ge 0,\\
&\mathbf{y}(\mathbf{x},\theta) = \arg\max_{\mathbf{y}} \left\{\mathbf{y}^T \mathbf{b}(\theta) : \mathbf{B}(\theta)\mathbf{y} \le \mathbf{q}(\theta) - \mathbf{A}(\theta)\mathbf{x}, \quad \mathbf{y} \ge 0\right\}.
\end{aligned}
\tag{16}
$$

where $\mathbf{x}$ and $\mathbf{y}(\mathbf{x},\theta)$ are first and second stage decisions respectively, $\mathbf{a}$, $\mathbf{d}$ and $\mathbf{D}$ - deterministic parameters, $\mathbf{b}(\theta)$, $\mathbf{q}(\theta)$, $\mathbf{A}(\theta)$ and $\mathbf{B}(\theta)$ - random parameters, θ - state of nature.

3. Approaches for solving two-stage model: combining of Monte-Carlo simulation with optimization

The approach which is most often used in practice to solve problem like (16) is based on Benders (1962) decomposition (see e.g. Infager (1994)). Modern techniques utilize Monte-Carlo simulation for generating scenarios and efficient algorithms which allow solution of very large scale problems.

In this paper we propose to use alternative technique which is well-known stochastic quasi-gradient algorithm (see e.g. Ermoliev (1976), Ermoliev and Yastremsky (1979), Ermoliev and Wets (1988)). Despite widely accepted thought about slow convergence rate of quasi-gradient methods, our experience suggests that this approach is fully operational for practical problems while considerably more flexible (in terms of class of problems which could be solved) comparing to Benders decomposition and related algorithms.

For the problem (16) so-called “linearization” quasi-gradient method could be used. On each iteration we calculate new approximation of the optimum, which is based on random direction calculated using stochastic quasi-gradient:

$$
\begin{aligned}
&\mathbf{x}^{s+1} = \mathbf{x}^s + \rho_s\left(\overline{\mathbf{x}}^s - \mathbf{x}^s\right)\\
&\mathbf{z}^{s+1} = \mathbf{z}^s + \delta_s\left(\xi^s - \mathbf{z}^s\right)\\
&\overline{\mathbf{x}}^s = \arg\max_{\mathbf{x}} \left\{\mathbf{x}^T \mathbf{z}^s : \mathbf{D}\mathbf{x} \le \mathbf{d}, \mathbf{x} \ge 0, \mathbf{x} \in I^s\right\}
\end{aligned}
\tag{17}
$$

where $\mathbf{x}^s$ is approximation of the solution on the iteration s, ξ^s is stochastic quasi-gradient - random vector which satisfy following conditions:

$$
\begin{aligned}
&E\left[\xi_s / \mathbf{x}^0, \mathbf{x}^1, \ldots, \mathbf{x}^s\right] = \nabla F\left(\mathbf{x}^s\right) + \mathbf{b}^s\\
&\mathbf{b}^s \to 0 \quad a.s. \quad (s \to \infty)
\end{aligned}
,
\tag{18}
$$

$\mathbf{z}^s$ is average of stochastic quasi-gradient over all iterations (it is necessary to guarantee convergence), I^s is the set of so-called induced constraints (we need induced constraints to guarantee non-empty set for the second-stage problem).

Stochastic quasi-gradient for the problem (16) is calculated from the following conditions (actually from the solution of the dual to the second-stage problem):

$$\begin{aligned} &\boldsymbol{\xi}^s = \mathbf{a} - \mathbf{A}(\theta^s)^T \mathbf{v}(\mathbf{x}^s, \theta^s) \\ &\mathbf{v}(\mathbf{x}^s, \theta^s) = \arg\min_{\mathbf{v}} \left\{ (\mathbf{v}^s)^T (\mathbf{q} - \mathbf{A}(\theta^s)\mathbf{x}^s) : \mathbf{B}(\theta^s)\mathbf{v}, \mathbf{v} \geq 0 \right\}. \end{aligned} \tag{19}$$

The set of induced constraint could be written as follows:

$$I^s = \left\{ \mathbf{x} : \mathbf{A}(\theta^s)\mathbf{x} \leq \mathbf{q}(\theta^s) \right\} \tag{20}$$

The algorithm consists in generating random scenarios θ^s on each iteration, calculating quasi-gradient using (19) (it requires solution of linear programming problem) and calculating next approximation of the solution using (17) (one more linear problem).

It was proved in Ermoliev (1976) that process (17) converges to the optimal solution if step coefficients ρ_s and δ_s satisfy following conditions:

$$\begin{aligned} &\rho_s \geq 0, \quad \delta_s \geq 0, \quad \rho_s / \delta_s \to 0 \quad a.s. \quad (s \to \infty), \\ &\sum_{s=0}^{\infty} \rho_s = \infty, \quad \sum_{s=0}^{\infty} \delta_s = \infty \quad a.s., \\ &\sum_{s=0}^{\infty} \rho_s \left\| \mathbf{b}^s \right\| < \infty \quad a.s., \quad \sum_{s=0}^{\infty} E\left[\rho_s^2 + \delta_s^2 \right] < \infty \end{aligned} \tag{21}$$

An example of such coefficients is as follows:

$$\begin{aligned} &\rho_s = K_1 / (1+s)^{\alpha}, \quad \delta_s = K_2 / (1+s)^{\beta}, \\ &\alpha > \beta, \quad 1/2 < \alpha \leq 1, \quad 1/2 < \beta \leq 1, \\ &0 < \underline{K} \leq K_1 \leq K_2 \leq \overline{K} < \infty \end{aligned} \tag{22}$$

The software for implementing algorithm (17) for the problem (16) was developed at Economic Cybernetics department of the Kiev State University in 1991 (see Mertens (1991)). This software (called ROS, i.e. Risk Optimization System), which is developed using platform-independent C-code, solves general form of the two-stage stochastic linear problem (16) and utilizes number of techniques which improve efficiency of the algorithm (for example, speed up of the solution of linear problems on each iteration and adaptive algorithms for the step size - see e.g. Uryasev (1991)). The stopping criteria, which is one of the main difficulties in quasi-gradient algorithms, is based on observations on the statistics of objective function in (16):

$$
\begin{aligned}
f_s &= \frac{1}{(1+s)} \sum_{i=1}^{s} f\left(x^i, \theta^i\right), \\
F_s &= \frac{1}{(1+L)} \sum_{i=s-L}^{s} f\left(x^i, \theta^i\right)
\end{aligned}
\tag{23}
$$

where $f\left(x^i, \theta^i\right)$ is the random value of objective on s-th iteration. The number of stopping criteria are used in the ROS software including comparison with estimates of value of objective functions in dual problem, e.g.:

$$
\text{i)} \quad
\begin{aligned}
&\left| F_s - G_s \right| < \varepsilon \\
&G_s = \frac{1}{(1+L)} \sum_{i=s-L}^{s} \left(\left(\mathbf{u}^s\right)^T \mathbf{d} + \mathbf{v}\left(\mathbf{x}^s, \theta^s\right) \right) \\
&\mathbf{u}^s = \arg\min_{\mathbf{u}} \left\{ \mathbf{u}^T \mathbf{d} : \mathbf{D}^T \mathbf{u} \geq \mathbf{a} - \mathbf{A}\left(\theta^s\right)^T \mathbf{v}\left(\mathbf{x}^s, \theta^s\right) \right\} \\
&\mathbf{v}\left(\mathbf{x}^s, \theta^s\right) = \arg\min_{\mathbf{v}} \left\{ \left(\mathbf{v}^s\right)^T \left(\mathbf{q} - \mathbf{A}\left(\theta^s\right) \mathbf{x}^s\right) : \mathbf{B}\left(\theta^s\right) \mathbf{v}, \mathbf{v} \geq 0 \right\}
\end{aligned}
\tag{24}
$$

$$
\text{ii)} \quad
\begin{aligned}
&\left| \left(\mathbf{x}^s\right)^T \left(\mathbf{a} - \mathbf{V}_s\right) \right| < \varepsilon \\
&V_s = \frac{1}{s} \sum_{i=1}^{s} \left(\mathbf{A}\left(\theta^i\right)^T \mathbf{v}\left(\mathbf{x}^i, \theta^i\right) \right)
\end{aligned}
\tag{25}
$$

(for additional discussion on stopping criteria see e.g. Pflug (1996)).

5. Model of optimization of inter-bank loans portfolio

Proposed approach was used for the real-life asset-liability management problem, namely management of the inter-bank loans portfolio of commercial bank. The problem consists in establishing of tomorrow limits of operation for the instruments with different maturities so that the Value-at-Risk of inter-bank loans portfolio was at pre-determined level and the expected return on entire portfolio was at maximum.

Let us define:

$W(\theta) = \overline{W} + w(\theta)$ is exogenous to the model general limit on tomorrow inter-bank operations which is defined form financing needs or excess liquidity of the bank;

t - time to maturity of particular instrument;

$r_t(\theta)$ - tomorrow spot rate at inter-bank market (continuously compounded);

$p_t(\theta) = \exp(-r_t(\theta) \cdot t)$ - tomorrow price of one currency unit of inter-bank loan with maturity t;

s_t - existing position in particular instrument;

x_t - tomorrow limit (decision);

$y_t(\mathbf{x}, \theta)$ - tomorrow correction of the limit which is bounded in some way.

The loss function which is necessary to build *VaR* constraint is defined as dollar duration of total inter-bank loans position:

$$f(x, y(x,\theta), \theta) = -\sum_t t \cdot p_t(\theta) \cdot (s_t + x_t + y_t(\theta)).$$

The possible objectives are minimizing of the Value-at-Risk of inter-bank loans portfolio or maximizing of the total return:

$$\max_{\mathbf{x}} E\left[\sum_t r_t(\theta) \cdot (x_t + y_t(\theta))\right].$$

The problem was solved for more than one hundred instrument (inter-bank loans with 120 different maturities) for the real-life situation at the Ukrainian inter-bank loans market. The number of utilized scenarios (iterations in quasi-gradient algorithm) was up to 100,000, but for the relatively good approximation of the optimal solution it was enough about 10,000 iterations. The time of calculations at ordinary Pentium III 1 GHz processor was about 30 min for 10,000 iterations.

6. Conclusion

The paper demonstrates possibilities of using quasi-gradient techniques in *VaR*-based asset-liability management optimization models. The approach was used for real-life problem of inter-bank loans portfolio management. The obtained result demonstrate that this approach is fully operational and efficient while allowing to solve more general (comparing to traditional approach) form of two-stage stochastic programming problem.

The main directions of future research are (1) development and practical implementation of wider range of asset-liability management problems (including credit risk management, etc.) using proposed approach, and (2) precise comparison in terms of efficiency of numerical algorithms between traditional (based on Benders decomposition) techniques and quasi-gradient methods.

References

1. Benders, J.R. (1962). Partitioning Procedure for Solving Mixed-Variable Programming Problems, *Numerische Mathematik* 4, 238 - 252.

2. Birge, J.R. and F. Louveaux (1997). *Introduction to Stochastic Programming*, Springer Verlag.

3. Dantzig, G.B. and M. Madansky (1961) On the Solution of Two-Staged Linear Programs under Uncertainty. *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability* I. 165 - 176.

4. Ermoliev, Yu. and R. J.-B. Wets (eds.) (1988). *Numerical Techniques for Stochastic Optimization*, Springer Verlag.

5. Ermoliev, Yu. and A. Yastremsky (1979). *Stochastic Models and Methods in Economic Planning*. Moscow: Nauka.

6. Ermoliev, Yu. (1976) *Methods of Stochastic Programming*. Moscow: Nauka.

7. Infanger, G. (1994) *Planning Under Uncertainty: Solving Large-Scale Stochastic Linear Programs*. Boyd & Fraser Pub. Co.

8. Jorion, Ph. (1996). *Value at Risk: A New Benchmark for Measuring Derivatives Risk*. Irwin Professional Pub.

9. Kall, P. and S. W. Walace (1995). *Stochastic Programming*. John Wiley & Sons.

10. Mertens, A.V. (1992) The Risk Optimization Software for Economic Models. *Operations Research and Automatic Control Systems*, #32, 102 - 107.

11. Pflug, G.Ch. (1996) *Optimization of Stochastic Models: The Interface Between Simulation and Optimization*. Kluwer Academic Publishers.

12. RiskMetrics. *Technical Document* (1996). J.P.Morgan.

13. Rocafellar T. and S. Uryasev. (2000) Optimization of Conditional Value-at-Risk. *The Journal of Risk*.

14. Uryasev S. (1990) *Adaptive algorithms of Stochastic Optimization and Game Theory*. Nauka.

15. Zenios S.A. (Ed.) (1996). *Financial Optimization*. Cambridge University Press.