RUTCOR
RESEARCH
REPORT

Conditional Mean-Conditional Variance Portfolio Selection Model

András Prékopa[a]

RRR 34-2007, December 2007

RUTCOR
Rutgers Center for
Operations Research
Rutgers University
640 Bartholomew Road
Piscataway, New Jersey
08854-8003
Telephone: 732-445-3804
Telefax: 732-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/∼rrr

[a]RUTCOR, Rutgers Center for Operations Research 640 Bartholomew Road, Piscataway, NJ 08854–8003

Rutcor Research Report
RRR 34-2007, December 2007

Conditional Mean-Conditional Variance Portfolio Selection Model

András Prékopa

Abstract. Markowitz's mean-variance model (1952, 1959) is one of the most widely used model for portfolio selection, even today. Kataoka (1963) gave another formulation for the same problem. The introduction of the risk measure: Value at Risk or VaR and the formulation of a corresponding portfolio selection model essentially rephrases Kataoka's ideas. Conditional Value at Risk or CVaR is another popular risk measure, typically used jointly with VaR, providing us with a probabilistic constrained-conditional expectation model. The purpose of the present paper is to introduce the conditional mean-conditional variance or CVaR-CVAR model that takes us back to Markowitz's ideas in a modified form. The optimization problems are presented and their mathematical properties are explored.

1 Introduction

Consider n assets with random returns on unit investments $R_1, \ldots, R_n$, investment amounts $x_1, \ldots, x_n$ and introduce the notations:

$$R = (R_1, \ldots, R_n)^T, \quad x = (x_1, \ldots, x_n)^T$$
$$E(R_i) = m_i, \quad i = 1, \ldots, n, \quad m = (m_1, \ldots, m_n)^T$$
$$X = R^T x, \quad C = E[(R - m)(R - m)^T].$$

With these notations we have the equations:

$$E(X) = m^T x, \quad \mathrm{Var}\,(X) = x^T C x.$$

Markowitz's mean-variance model (1952, 1959) is usually formulated in three different ways:

Model I.
$$\begin{aligned} &\text{maximize } m^T x \\ &\text{subject to} \\ &x^T C x \le V \\ &\sum_{i=1}^{n} x_i = 1 \\ &\quad x \ge 0, \end{aligned} \tag{1.1}$$

Model II.
$$\begin{aligned} &\text{minimize } x^T C x \\ &\text{subject to} \\ &m^T x \le M \\ &\sum_{i=1}^{n} x_i = 1 \\ &\quad x \ge 0, \end{aligned} \tag{1.2}$$

Model III.
$$\begin{aligned} &\text{maximize } \{m^T x - \beta x^T C x\} \\ &\text{subject to} \\ &\sum_{i=1}^{n} x_i = 1 \\ &\quad x \ge 0, \end{aligned} \tag{1.3}$$

where M, V are some constant upper bounds, for the expectation and variance, respectively, of the total return and $\beta > 0$ is a constant. Sometimes lower and upper bounds are imposed individually on the variables x_i, $i = 1, \ldots, n$.

Kataoka's (1963) model is similar to Model III. We can state it as follows:

$$\begin{aligned} &\text{maximize } v \\ &\text{subject to} \\ &P(R^T x \ge v) \ge p \\ &\sum_{i=1}^{n} x_i = 1 \\ &x \ge 0, \end{aligned} \tag{1.4}$$

where $p \in (0,1)$ is a fixed probability chosen by ourselves.

If $R_1, \ldots, R_n$ have a normal joint distribution, then, no matter if it is degenerate or nondegenerate, problem (1.4) can be rewritten in the following form:

$$\begin{aligned} &\text{maximize } \left\{ m^T x + \Phi^{-1}(1-p)\sqrt{x^T C x} \right\} \\ &\text{subject to} \\ &\sum_{i=1}^{n} x_i = 1 \\ &x \ge 0. \end{aligned} \tag{1.5}$$

In what follows φ, Φ designate the standard normal density and distribution functions, respectively.

If $p \ge 1/2$, then $\Phi^{-1}(1-p) \le 0$ and thus problem (1.5) is a linearly constrained convex optimization problem.

The advantage of problem (1.5) over problem (1.3) is that we combine the standard deviation, rather than the variance of the total return with its expectation and both terms in the objective function have the same dimension. Problem (1.3), on the other hand, is computationally more tractable because it is a linearly constrained convex quadratic programming problem widely studied in the literature. The objective function in problem (1.5) is the p-quantile of the probability distribution of $R^T x$. The quantile recently acquired other name in the financial literature: Value at Risk.

The term Value at Risk was proposed by Till Guldimann at J.P. Morgan in the late 1980's. The Group of Thirty which had a representative from J.P. Morgan had a discussion on best risk management practices. The term VaR found its way through the G–30 report published in 1993.

The definition of Value at Risk is not unique in the literature. In this paper we use the following

Definition 1 Value at risk of confidence level 100 p% is defined as the optimum value of the optimization problem:

$$\begin{aligned} &\text{maximize } v \\ &\text{subject to} \\ &P(X \ge v) \ge p, \end{aligned} \tag{1.6}$$

where X is a random variable and $p \in (0,1)$ is a constant. Let $\mathrm{VaR}_p(X)$ designate this value. Sometimes we simply write VaR.

Another definition (see Pflug, 2000) takes it as the optimum value of the problem ($0 < \alpha < 1$):

$$\begin{aligned} &\inf v \\ &\text{subject to} \\ &P(X \leq v) > \alpha. \end{aligned} \tag{1.7}$$

The two definitions provide us with the same optimum value if $\alpha = 1 - p$. However, the practical use of the two definitions are different. We use (1.6) if X means revenue or return and use (1.7) if X means loss. Both p and α are chosen large, in practice, because we want to impose lower bound on revenue, or return, and upper bound on loss, by large probabilities.

Note that the c.d.f. of a random variable X, i.e., the function $F(v) = P(X \leq v)$, $-\infty < v < \infty$, is right continuous. If we define its inverse by the equation:

$$F^{-1}(v) = \inf\{z \mid F(z) > v\}, \quad -\infty < v < \infty,$$

then the optimum value of problem (1.7) is $F^{-1}(\alpha)$ and we have the equality $\mathrm{VaR}_p(X) = F^{-1}(1-p)$.

If $X = R^T X$, where R has normal distribution with $E(R) = m$ and covariance matrix C, then we easily derive the equation:

$$\mathrm{VaR}_p(X) = m^T x + \Phi^{-1}(1-p)\sqrt{x^T C x}.$$

The definition of the Conditional Value at Risk depends on the definition of the Value at Risk. Since VaR is defined by (1.6), we use the following

DEFINITION 2 (ROCKAFELLAR, URYASEV, 2002) Conditional Value at Risk is the value

$$E(X \mid X \leq v), \qquad v = \mathrm{VaR}_p(X). \tag{1.8}$$

We designate it by $\mathrm{CVaR}_p(X)$, or simply by CVaR, if it is clear what X and p we are dealing with.

If VaR were defined by (1.7), then we would define CVaR as

$$E(X \mid X \geq v), \tag{1.9}$$

where v is the optimum value of problem (1.7).

Under the condition of a normally distributed R we have the formula:

$$\mathrm{CVaR}_p(X) = m^T x - \frac{\varphi(\Phi^{-1}(1-p))}{1-p}\sqrt{x^T C x}. \tag{1.10}$$

Note that $\Phi^{-1}(1-p) = -\Phi^{-1}(p)$ which implies $\varphi(\Phi^{-1}(1-p)) = \varphi(\Phi^{-1}(p))$.

CVaR is closely connected with risk measures introduced earlier in stochastic optimization. If in the underlying problem we have a constraint

$$Tx \geq \xi,$$

where T is an $r \times n$ matrix and ξ is an r-component random vector, then in the probabilistic constrained formulation we prescribe that the constraint

$$P(Tx \geq \xi) \geq p \tag{1.11}$$

should be satisfied for every x that we consider feasible. As a relaxation of the constraint (1.11) Prékopa (1973b) introduced the conditional expectation constraints:

$$E(\xi_i - T_i x \mid \xi_i - T_i x > 0) \leq d_i, \qquad i = 1, \ldots, r, \tag{1.12}$$

where T_i is the ith row of T and ξ_i is the ith component of ξ. Prékopa (1973b) has also shown that if each ξ_i has continuous probability distribution with logarithmically concave p.d.f., then the constraints (1.12) are linear. In fact, it is shown in the above-mentioned paper that if (dropping the subscript i) any $\xi \in R^1$ has continuous distribution with logarithmically concave p.d.f., then

$$g(u) = E(\xi - u \mid \xi - u > 0), \quad -\infty < u < \infty$$

is a decreasing function of u. Thus, returning to $\xi = (\xi_1, \ldots, \xi_n)^T$, the constraints (1.12) can be written as

$$T_i x \geq g_i^{-1}(d_i), \qquad i = 1, \ldots, r.$$

Note that it is also reasonable to use both constraints (1.11) and (1.12) in one model (see Prékopa, 1973b). Constraint (1.11) ensures that $Tx \geq \xi$ occurs with a large probability while constraints (1.12) imposes upper bounds on the expectations of the unfavorable deviations, given that they occur.

A related constraint type, called integrated chance constraint, was introduced by Klein Haneveld (1986). As applied to our case, and as another relaxation of the single probabilistic constraint (1.11), we write

$$E([\xi_i - T_i x]_+) \leq d_i, \qquad i = 1, \ldots, r. \tag{1.13}$$

The constraining functions are linear in x, without any limitation regarding the probability distributions of the ξ_i, $i = 1, \ldots, r$.

Definition 3 The functions of the variable v $(-\infty < v < \infty)$:

$$E(X \mid X \geq v) \tag{1.14}$$
$$E(X \mid X \leq v) \tag{1.15}$$

will be called conditional expectation functions.

In the next definition new risk measures are introduced.

DEFINITION 4 The functions of the variable v $(-\infty < v < \infty)$:

$$E(X^2 \mid X \geq v) - E^2(X \mid X \geq v) \tag{1.16}$$

$$E(X^2 \mid X \leq v) - E^2(X \mid X \leq v) \tag{1.17}$$

will be called conditional variance or CVAR functions.

The definitions of VaR and CVaR are motivated by the intention to characterize the return of an investment by two risk measures. One is of probability type (VaR) that we want to make large, it expresses the probability that the return is above a lower bound. The other one (CVaR) is of expectation type in which an average is taken of the small (mostly negative, in practice) return values. We want to make the absolute value of the latter small. Thus, $\mathrm{Var}_p(X)$ and $\mathrm{CVaR}_p(X)$ are connected in this way.

In the above-mentioned model of Prékopa (1973b) as well as in the models of Pflug (2000) and Rockafellar and Uryasev (2002) the probability of the favorable outcome of the "experiment" is ensured to be large, and, given that it is violated, the conditional expectations of the opposite deviations are prescribed to be small. In the latter two papers this means, in our setting, that function (1.14) is used, either in the objective function or in the constraints, with $v = \mathrm{VaR}_p(X)$.

In the models presented in this paper, however, we combine the functions (1.14) and (1.16) and impose a constraint on $\mathrm{VaR}_p(X)$, in the same model, for the case of a $v = v_0$. We look at revenues or returns. Similar models can be formulated for the case of losses in which case the functions (1.15) and (1.17) have to be used, together with a separate constraint for the Value at Risk.

The objective of the paper is to formulate counterparts of Markowitz's mean-variance models, using conditional expectation and conditional variance functions. This is enabled by nice properties of those functions, especially the decreasing property of the function (1.16) under the condition of logconcavity of the probability distribution of the return random variable.

In Section 2 we derive some mathematical properties for the functions (1.16), (1.17). In Section 3 the case of a normally distributed random variable is considered and formulas are derived for all functions in (1.14), (1.15), (1.16) and (1.17). In Section 4 we formulate our portfolio selection models with the conditional expectation and conditional variance functions. Finally, in Section 5 the conclusions are summarized.

2 Mathematical Properties of the Conditional Variance Functions

The most important theorem in connection with CVAR is the following

Theorem 2.1 *If the random variable X has continuous distribution and logconcave p.d.f., then*

$$E(X^2 \mid X > v) - (E(X \mid X > v))^2, \quad -\infty < v < \infty$$

is a decreasing function of the variable v.

Proof. Let g, G designate the p.d.f. and c.d.f. of X, respectively. If for a v we have $G(v) = 1$, then, by definition, both terms in the CVAR are zero. Hence, we may restrict ourselves to the maximal open interval, where $G(v) < 1$.

We also remark that the existence of the first two moments of X implies that

$$\lim_{v\to-\infty} v^k G(v) = 0, \qquad \lim_{v\to\infty} v^k(1 - G(v)) = 0$$

for $k = 1, 2$ and

$$\int_v^\infty (t - v)g(t)\,\mathrm{d}t = \int_v^\infty [1 - G(t)]\,\mathrm{d}t,$$
$$\int_v^\infty (t - v)^2 g(t)\,\mathrm{d}t = 2\int_v^\infty \int_t^\infty [1 - G(x)]\,\mathrm{d}x\,\mathrm{d}t.$$

We have the relations

$$\begin{aligned}
&E(X^2 \mid X \geq v) - (E(X \mid X \geq v))^2\\
&\quad = E((X - v)^2 \mid X \geq v) - (E(X - v \mid X \geq v))^2\\
&\quad = \frac{\int_v^\infty (t - v)^2 g(t)\,\mathrm{d}t}{1 - G(v)} - \left(\frac{\int_v^\infty (t - v)g(t)\,\mathrm{d}t}{1 - G(v)}\right)^2\\
&\quad = \frac{2\int_v^\infty \int_t^\infty [1 - G(x)]\,\mathrm{d}x\,\mathrm{d}t}{1 - G(v)} - \left(\frac{\int_v^\infty [1 - G(t)]\,\mathrm{d}t}{1 - G(v)}\right)^2.
\end{aligned}$$

Hence it follows that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}v}(E(X^2 \mid X \geq v) - E^2(X \mid X \geq v))\\
&\quad = \frac{-2\int_v^\infty [1 - G(t)]\,\mathrm{d}t\,[1 - G(v)] + 2g(v)\int_v^\infty \int_t^\infty [1 - G(x)]\,\mathrm{d}x\,\mathrm{d}t}{[1 - G(v)]^2}\\
&\qquad -2\frac{\int_v^\infty [1 - G(t)]\,\mathrm{d}t}{1 - G(v)}\,\frac{-[1 - G(v)]^2 + g(v)\int_v^\infty [1 - G(t)]\,\mathrm{d}t}{[1 - G(v)]^2}\\
&\quad = \frac{2g(v)}{[1 - G(v)]^3}\left([1 - G(v)]\int_v^\infty \int_t^\infty [1 - G(x)]\,\mathrm{d}x\,\mathrm{d}t - \left(\int_v^\infty [1 - G(t)]\,\mathrm{d}t\right)^2\right).
\end{aligned}$$

By Theorem 2 in Prékopa (1973a), the logconcavity of g implies the logconcavity of

$$\int_v^\infty [1 - G(x)]\,\mathrm{d}x$$

and, in turn, the logconcavity of

$$h(v) = \int_v^\infty \int_t^\infty [1 - G(x)]\,\mathrm{d}x\,\mathrm{d}t.$$

This implies that

$$\begin{aligned}&[1 - G(v)] \int_v^\infty \int_t^\infty [1 - G(x)]\,\mathrm{d}x\,\mathrm{d}t - \left(\int_v^\infty [1 - G(t)]\,\mathrm{d}t\right)^2 \\ &\quad = h''(v)h(v) - (h'(v))^2 \le 0\end{aligned}$$

and the theorem is proved. □

Similar is the proof of the following

THEOREM 2.2 *If the random variable X has continuous distribution and logconcave p.d.f., then*

$$E(X^2 \mid X \le v) - (E(X \mid X \le v))^2, \qquad -\infty < v < \infty$$

is an increasing function of the variable v.

Note that in the above theorems the condition $X > v$ can be replaced by $X \ge v$ and the condition $X \le v$ can be replaced by $X < v$, the values of the functions are unchanged. Not as trivial is

THEOREM 2.3 *If the p.d.f. of X is strictly logconcave in the entire real line, then the function in Theorem 2.1 is strictly decreasing and the function in Theorem 2.2 is strictly increasing in the entire real line.*

Proof. If we use the same proofs what we have used in connection with Theorems 2.1 and 2.2 and refer to Theorem 5, rather than to Theorem 2, in Prékopa (1973a), then the assertion follows. □

Burridge (1982) stated Theorem 2.1 without proof with the remark that it can be derived from Prékopa's (1971, 1973a) logconcavity results. In this section we have proved and sharpened the theorem: the function is strictly decreasing if strict logconcavity holds for the p.d.f.

3 The Case of the Normal Distribution

In this section we look at the case, where the random variable has nondegenerate normal distribution. We derive special formulas for the CVAR functions. Let g, G designate the p.d.f. and the c.d.f. of X, respectively. Then

$$g(v) = \frac{1}{\sigma}\varphi\left(\frac{v-\mu}{\sigma}\right) \text{ and } G(v) = \Phi\left(\frac{v-\mu}{\sigma}\right).$$

THEOREM 3.1 *Let X be a normally distributed random variable with expectation μ and variance $\sigma^2 > 0$. Then for every real v we have the equations*

$$E(X \mid X \geq v) = \mu + \frac{\varphi\left(\frac{v-\mu}{\sigma}\right)}{1-\Phi\left(\frac{v-\mu}{\sigma}\right)}\sigma \tag{3.1}$$

and

$$E(X \mid X \leq v) = \mu - \frac{\varphi\left(\frac{v-\mu}{\sigma}\right)}{\Phi\left(\frac{v-\mu}{\sigma}\right)}\sigma. \tag{3.2}$$

Proof. We derive (3.1), the derivation of (3.2) is similar. We have the equations:

$$\begin{aligned} E(X \mid X \geq v) &= \frac{\int_v^\infty t g(t)\,\mathrm{d}t}{1-G(v)} = \frac{\int_v^\infty t\frac{1}{\sigma}\varphi\left(\frac{t-\mu}{\sigma}\right)\mathrm{d}t}{1-\Phi\left(\frac{v-\mu}{\sigma}\right)} = \frac{\int_{\frac{v-\mu}{\sigma}}^\infty (\mu+z\sigma)\varphi(z)\,\mathrm{d}z}{1-\Phi\left(\frac{v-\mu}{\sigma}\right)} \\ &= \frac{\left[1-\Phi\left(\frac{v-\mu}{\sigma}\right)\right]\mu + \varphi\left(\frac{v-\mu}{\sigma}\right)\sigma}{1-\Phi\left(\frac{v-\mu}{\sigma}\right)} = \mu + \frac{\varphi\left(\frac{v-\mu}{\sigma}\right)}{1-\Phi\left(\frac{v-\mu}{\sigma}\right)}\sigma. \end{aligned}$$

□

THEOREM 3.2 *Let X be a normally distributed random variable with expectation μ and variance $\sigma^2 > 0$. Then for every real v we have the equations:*

$$\begin{aligned} &E(X^2 \mid X \geq v) - E^2(X \mid X \geq v) \\ &= \sigma^2\left(1 + \frac{\varphi\left(\frac{v-\mu}{\sigma}\right)}{1-\Phi\left(\frac{v-\mu}{\sigma}\right)}\frac{v-\mu}{\sigma} - \left(\frac{\varphi\left(\frac{v-\mu}{\sigma}\right)}{1-\Phi\left(\frac{v-\mu}{\sigma}\right)}\right)^2\right) \end{aligned} \tag{3.3}$$

and

$$
\begin{aligned}
&E(X^2 \mid X \le v) - E^2(X \mid X \le v) \\
&= \sigma^2 \left(1 - \frac{\varphi\left(\frac{v-\mu}{\sigma}\right)}{\Phi\left(\frac{v-\mu}{\sigma}\right)} \frac{v-\mu}{\sigma} - \left(\frac{\varphi\left(\frac{v-\mu}{\sigma}\right)}{\Phi\left(\frac{v-\mu}{\sigma}\right)} \right)^2 \right).
\end{aligned} \tag{3.4}
$$

Proof. We present the proof of (3.3), the proof of (3.4) is similar. We have the equations:

$$
\begin{aligned}
&E(X^2 \mid X \ge v) - E^2(X \mid X \ge v) \\
&= \frac{\int_v^\infty t^2 g(t)\,\mathrm{d}t}{1 - G(v)} - \left(\mu + \frac{\varphi\left(\frac{v-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{v-\mu}{\sigma}\right)} \sigma \right)^2 \\
&= \frac{\int_{\frac{v-\mu}{\sigma}}^\infty (\mu + z\sigma)^2 \varphi(z)\,\mathrm{d}z}{1 - \Phi\left(\frac{v-\mu}{\sigma}\right)} - \left(\mu + \frac{\varphi\left(\frac{v-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{v-\mu}{\sigma}\right)} \sigma \right)^2 \\
&= \frac{1}{1 - \Phi\left(\frac{v-\mu}{\sigma}\right)} \left(\mu^2 \left[1 - \Phi\left(\frac{v-\mu}{\sigma}\right) \right] + 2\mu\sigma\varphi\left(\frac{v-\mu}{\sigma}\right) \right. \\
&\quad \left. + \frac{v-\mu}{\sigma}\varphi\left(\frac{v-\mu}{\sigma}\right)\sigma^2 + \left[1 - \Phi\left(\frac{v-\mu}{\sigma}\right) \right] \sigma^2 \right) \\
&\quad - \left(\mu + \frac{\varphi\left(\frac{v-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{v-\mu}{\sigma}\right)} \sigma \right)^2 \\
&= \sigma^2 \left(1 + \frac{\varphi\left(\frac{v-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{v-\mu}{\sigma}\right)} \frac{v-\mu}{\sigma} - \left(\frac{\varphi\left(\frac{v-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{v-\mu}{\sigma}\right)} \right)^2 \right),
\end{aligned}
$$

where we have used the relation: $\varphi'(z) = -z\varphi(z)$. □

If $\mu = 0$ and $\sigma = 1$, then equation (3.3) specializes to the equation:

$$
E(X^2 \mid X \ge v) - E^2(X \mid X \ge v) = 1 - h'(v), \tag{3.5}
$$

where $h(v)$ is the hazard rate function:

$$
h(v) = \frac{\varphi(v)}{1 - \Phi(v)}. \tag{3.6}
$$

Similarly, equation (3.4) specializes to the equation:

$$E(X^2 \mid X \leq v) - E^2(X \mid X \leq v) = 1 + k'(v), \tag{3.7}$$

where

$$k(v) = \frac{\varphi(v)}{\Phi(v)} = h(-v). \tag{3.8}$$

For functions h and k we have

THEOREM 3.3 *Both h' and k' are strictly increasing functions on the entire real line, further, for any real v*

$$0 < h'(v) < 1, \qquad -1 < k'(v) < 0,$$

finally,

$$\lim_{v\to-\infty} h'(v) = 0, \qquad \lim_{v\to\infty} h'(v) = 1$$
$$\lim_{v\to-\infty} k'(v) = -1, \qquad \lim_{v\to\infty} k'(v) = 0.$$

Proof. Since φ is strictly logconcave on the entire real line, Theorem 2.3 implies the strict monotonicity of $h'(v)$ and $k'(v)$ on R'. The other assertions follow from this fact and relations (3.5)–(3.8).

The inequality $h'(v) > 0$ implies that

$$\frac{\varphi(v)}{1-\Phi(v)} < v, \quad -\infty < v < \infty. \tag{3.9}$$

The latter inequality has an important consequence for VaR and CVaR of the total return $X = R^T x$ of a portfolio, where R has normal distribution with expectation $\mu = m^T x$ and variance $\sigma^2 = x^T C x$. Assuming $x^T C x > 0$, by (3.2) and the expression for $\mathrm{Var}_p(X)$ we have:

$$\begin{aligned}
\mathrm{CVaR}_p(X) &= m^T x - \frac{\varphi(\Phi^{-1}(1-p))}{1-p}\sqrt{x^T C x} \\
&= m^T x - \frac{\varphi(\Phi^{-1}(p))}{1-\Phi(\Phi^{-1}(p))}\sqrt{x^T C x} \\
&> m^T x - \Phi^{-1}(p)\sqrt{x^T C x} \\
&= m^T x + \Phi^{-1}(1-p)\sqrt{x^T C x} = \mathrm{VaR}_p(X).
\end{aligned} \tag{3.10}$$

If v is equal to VaR, i.e., $v = \mu + \Phi^{-1}(1-p)\sigma$, then (3.3) and (3.4) specialize in the following way:

$$\begin{aligned}
&E(X^2 \mid X \geq v) - E^2(X \mid X \geq v) \\
&\quad = \sigma^2\left(1 + \frac{1}{p}\varphi(\Phi^{-1}(1-p))\Phi^{-1}(1-p) - \frac{1}{p^2}\varphi^2(\Phi^{-1}(1-p))\right)
\end{aligned} \tag{3.11}$$

and

$$\begin{aligned}
&E(X^2 \mid X \leq v) - E^2(X \mid X \leq v) \\
&\quad = \sigma^2\left(1 - \frac{1}{1-p}\varphi(\Phi^{-1}(1-p))\Phi^{-1}(1-p) - \frac{1}{(1-p)^2}\varphi^2(\Phi^{-1}(1-p))\right).
\end{aligned} \tag{3.12}$$

4 Conditional Mean-Conditional Variance Portfolio Selection Models

If $X = R^T x$ and R has normal distribution with $E(R) = m$ and covariance matrix C, then $E(X) = m^T x$, $\mathrm{Var}\,(X) = x^T Cx$. If we replace $m^T x$ for μ and $x^T Cx$ for σ^2 in (3.1), (3.2), (3.3) and (3.4), then we obtain the formulas:

$$\begin{aligned}
&E(X \mid X \geq v) \\
&\quad = m^T x + \frac{\varphi\left(\dfrac{v - m^T x}{\sqrt{x^T Cx}}\right)}{1 - \Phi\left(\dfrac{v - m^T x}{\sqrt{x^T Cx}}\right)} \sqrt{x^T Cx} \qquad (4.1) \\
&E(X \mid X \leq v) \\
&\quad = m^T x - \frac{\varphi\left(\dfrac{v - m^T x}{\sqrt{x^T Cx}}\right)}{\Phi\left(\dfrac{v - m^T x}{\sqrt{x^T Cx}}\right)} \sqrt{x^T Cx} \qquad (4.2) \\
&E(X^2 \mid X \geq v) - E^2(X \mid X \geq v) \\
&\quad = x^T Cx \left(1 + \frac{\varphi\left(\dfrac{v - m^T x}{\sqrt{x^T Cx}}\right)}{1 - \Phi\left(\dfrac{v - m^T x}{\sqrt{x^T Cx}}\right)} \frac{v - m^T x}{\sqrt{x^T Cx}} - \left(\frac{\varphi\left(\dfrac{v - m^T x}{\sqrt{x^T Cx}}\right)}{1 - \Phi\left(\dfrac{v - m^T x}{\sqrt{x^T Cx}}\right)} \right)^2 \right) \qquad (4.3) \\
&E(X^2 \mid X \leq v) - E^2(X \mid X \leq v) \\
&\quad = x^T Cx \left(1 - \frac{\varphi\left(\dfrac{v - m^T x}{\sqrt{x^T Cx}}\right)}{\Phi\left(\dfrac{v - m^T x}{\sqrt{x^T Cx}}\right)} \frac{v - m^T x}{\sqrt{x^T Cx}} - \left(\frac{\varphi\left(\dfrac{v - m^T x}{\sqrt{x^T Cx}}\right)}{\Phi\left(\dfrac{v - m^T x}{\sqrt{x^T Cx}}\right)} \right)^2 \right). \qquad (4.4)
\end{aligned}$$

If $v = \mathrm{VaR}_p(X) = m^T x + \Phi^{-1}(1-p)\sqrt{x^T Cx}$, then the above formulas specialize in the following way:

$$\begin{aligned}
&E(X \mid X \geq v) \\
&\quad = m^T x + \frac{1}{p}\varphi\left(\Phi^{-1}(1-p)\right)\sqrt{x^T Cx}, \qquad (4.5) \\
&E(X \mid X \leq v) \\
&\quad = m^T x - \frac{1}{1-p}\varphi\left(\Phi^{-1}(1-p)\right)\sqrt{x^T Cx}, \qquad (4.6)
\end{aligned}$$

$$\begin{aligned} &E(X^2 \mid X \geq v) - E^2(X \mid X \geq v) \\ &\quad = x^T C x \left(1 + \frac{1}{p}\varphi(\Phi^{-1}(1-p))\Phi^{-1}(1-p) - \frac{1}{p^2}\varphi^2(\Phi^{-1}(1-p))\right), \end{aligned} \tag{4.7}$$

$$\begin{aligned} &E(X^2 \mid X \leq v) - E^2(X \mid X \leq v) \\ &\quad = x^T C x \left(1 - \frac{1}{1-p}\varphi(\Phi^{-1}(1-p))\Phi^{-1}(1-p) - \frac{1}{(1-p)^2}\varphi^2(\Phi^{-1}(1-p))\right). \end{aligned} \tag{4.8}$$

In this section we formulate models that provide us with more favorable solutions than Markowitz's models. We achieve it in such a way that we use conditional expectation and conditional variance, rather than absolute expectation and absolute variance in the models. We take X as the random return, or revenue and work with the formulas (4.1) (4.3).

Given that $X \geq v_0$, for some v_0, the conditional expectation of X is larger and the conditional variance of X is smaller than the corresponding unconditional expectation and variance, respectively. This is a favorable fact from the point of view of the portfolio selection problem and its optimal solution.

In practice, however, the above advantage is realized only if we ensure that the event $X \geq v_0$ occurs by a large probability. In order to achieve this goal we include the inequality $\mathrm{Var}_p(X) \geq v_0$ among the constraints of the portfolio selection problem(s). If the random return R has normal distribution, then this constraint takes the form: $m^T x + \Phi^{-1}(1-p)\sqrt{x^T C x} \geq v_0$, where p is a large probability.

Our models are the following:

Model I.

$$\begin{aligned} &\text{maximize } E(X \mid X \geq v_0) \\ &\text{subject to} \\ &E(X^2 \mid X \geq v_0) - E^2(X \mid X \geq v_0) \leq V \\ &\qquad \mathrm{VaR}_p(X) \geq v_0 \\ &\qquad \sum_{i=1}^n x_i = 1 \\ &\qquad x \geq 0, \end{aligned} \tag{4.9}$$

Model II.

$$\begin{aligned} &\text{minimize } \{E(X^2 \mid X \geq v_0) - E^2(X \mid X \geq v_0)\} \\ &\text{subject to} \\ &E(X \mid X \geq v_0) \geq M \\ &\qquad \mathrm{VaR}_p(X) \geq v_0 \\ &\qquad \sum_{i=1}^n x_i = 1 \\ &\qquad x \geq 0, \end{aligned} \tag{4.10}$$

$$\begin{aligned} \text{Model III.} \quad & \text{maximize } \{E(X \mid X \geq v_0) - \beta[E(X^2 \mid X \geq v_0) - E^2(X \mid X \geq v_0)]\} \\ & \text{subject to} \\ & \mathrm{VaR}_p(X) \geq v_0 \\ & \sum_{i=1}^n x_i = 1 \\ & x \geq 0, \end{aligned} \tag{4.11}$$

where $\beta > 0$ is some constant. In Model III. we may replace the conditional variance by its square root. In the general case the above models are nonconvex optimization problems but using suitable relaxation we can obtain problems the solutions of which are easier.

4.1 Discussion of Model I.

The detailed form of Model I. is obtained by replacing the function in (4.1) for the objective function of problem (4.9) and the function (4.3) for the constraining function in the first constraint of the same problem, writing $v = v_0$.

If $v_0 = \mathrm{VaR}_p(X)$, then the function (4.1) reduces to the function (4.5) and Model I. reduces to the following:

$$\begin{aligned} \text{Model I. (a)} \quad & \text{maximize } \left\{ m^T x + \frac{1}{p}\varphi(\Phi^{-1}(1-p))\sqrt{x^T C x} \right\} \\ & \text{subject to} \\ & x^T C x \left(1 + \frac{1}{p}\varphi(\phi^{-1}(1-p))\Phi^{-1}(1-p) - \frac{1}{p^2}\varphi^2(\Phi^{-1}(1-p))\right) \leq V \\ & m^T x + \Phi^{-1}(1-p)\sqrt{x^T C x} \geq v_0 \\ & \sum_{i=1}^n x_i = 1 \\ & x \geq 0. \end{aligned} \tag{4.12}$$

The constant that multiplies $x^T C x$ in the first constraint of problem (4.12) is positive, for any p. If $p \geq 1/2$, then the constant that multiplies $\sqrt{x^T C x}$ in the second constraint is nonpositive. Since $x^T C x$ and $\sqrt{x^T C x}$ are convex functions, the set of feasible solutions is convex. However, the objective function that is to be maximized is convex, hence (4.12) is not a convex problem.

We can create a more tractable problem from the general problem (4.9) if we choose some fixed $v = v_0$ and prescribe that

$$E(X^2 \mid X \geq v_0) - E^2(X \mid X \geq v_0) \leq d x^T C x, \tag{4.13}$$

where d is some constant, chosen by ourselves in such a way that $0 < d < 1$. Let $h_1(v) = 1 - h'(v)$. Then the constraint (4.13) can be written in the form

$$x^T C x h_1\left(\frac{v_0 - m^T x}{\sqrt{x^T C x}}\right) \leq d x^T C x$$

and since h_1 is a decreasing function, the constraint can be given another form:

$$m^T x + h_1^{-1}(d)\sqrt{x^T C x} \geq v_0. \tag{4.14}$$

Since h_1 is a decreasing function and $h_1(-\infty) = 1$, $h(\infty) = 0$, $h(0) = \frac{2}{\pi}$, it follows that $h_1^{-1}(d) < 0$, if $d > \frac{2}{\pi}$. Assuming this to be the case, the constraining function in (4.14) is concave and the inequality determines a convex set of x vectors. Our new problem is:

Model I. (b)

$$\begin{aligned}
&\text{maximize } \left\{ m^T x + \frac{\varphi\left(\frac{v_0 - m^T x}{\sqrt{x^T C x}}\right)}{1 - \Phi\left(\frac{v_0 - m^T x}{\sqrt{x^T C x}}\right)} \sqrt{x^T C x} \right\} \\
&\text{subject to} \\
&\quad m^T x + h_1^{-1}(d)\sqrt{x^T C x} \geq v_0 \\
&\quad m^T x + \Phi^{-1}(1-p)\sqrt{x^T C x} \geq v_0 \\
&\quad \sum_{i=1}^{n} x_i = 1 \\
&\quad x \geq 0.
\end{aligned}$$

The problem is still nonconvex because the objective function is not concave in x, in general.

4.2 Discussion of Model II.

If we interchange the objective function and the constraining function in the first constraint in Models I. (a), (b), then we obtain the new versions of Model II. We may disregard the constant factor of $x^T C x$ in problem (4.12). Thus, our new problems are the following:

Model II. (a)

$$\begin{aligned}
&\text{minimize } x^T C x \\
&\text{subject to} \\
&\quad m^T x + \frac{1}{p}\varphi(\Phi^{-1}(1-p))\sqrt{x^T C x} \geq M \\
&\quad m^T x + \Phi^{-1}(1-p)\sqrt{x^T C x} \geq v_0 \\
&\quad \sum_{i=1}^{n} x_i = 1 \\
&\quad x \geq 0
\end{aligned}$$

and

Model II. (b)

$$\begin{aligned}
&\text{maximize } \{m^T x + h_1^{-1}(d)\sqrt{x^T C x}\}\\
&\text{subject to}\\
&m^T x + \frac{\varphi\left(\dfrac{v_0 - m^T x}{\sqrt{x^T C x}}\right)}{1 - \Phi\left(\dfrac{v_0 - m^T x}{\sqrt{x^T C x}}\right)}\sqrt{x^T C x} \geq M\\
&m^T x + \Phi^{-1}(1-p)\sqrt{x^T C x} \geq v_0\\
&\sum_{i=1}^{n} x_i = 1\\
&x \geq 0.
\end{aligned}$$

None of the above problems is convex. If, however, we remove the first constraints from both problems, then they become convex.

4.3 Discussion of Model III.

If we take into account (4.1) and (4.3), then we can write Model III. as follows:

Model III.

$$\begin{aligned}
\text{maximize } &\left\{ m^T x + \frac{\varphi\left(\dfrac{v_0 - m^T x}{\sqrt{x^T C x}}\right)}{1 - \Phi\left(\dfrac{v_0 - m^T x}{\sqrt{x^T C x}}\right)}\sqrt{x^T C x} \right.\\
&-\beta x^T C x \left(1 + \frac{\varphi\left(\dfrac{v_0 - m^T x}{\sqrt{x^T C x}}\right)}{1 - \Phi\left(\dfrac{v_0 - m^T x}{\sqrt{x^T C x}}\right)} \frac{v_0 - m^T x}{\sqrt{x^T C x}} \right.\\
&\left.\left. - \left(\frac{\varphi\left(\dfrac{v_0 - m^T x}{\sqrt{x^T C x}}\right)}{1 - \Phi\left(\dfrac{v_0 - m^T x}{\sqrt{x^T C x}}\right)}\right)^2\right)\right\}
\end{aligned}$$

$$\begin{aligned}
&\text{subject to}\\
&m^T x + \Phi^{-1}(1-p)\sqrt{x^T C x} \geq v_0\\
&\sum_{i=1}^{n} x_i = 1\\
&x \geq 0,
\end{aligned}$$

where β is some positive constant. Another model is obtained if we take the square root of the factor that multiplies β in the objective function. The problem is nonconvex.

If we take $v_0 = \mathrm{VaR}_p(X)$, then the model specializes as follows:

Model III. (a)
$$\begin{aligned}
&\text{maximize } \Big\{ m^T x + \frac{1}{p}\varphi(\Phi^{-1}(1-p))\sqrt{x^T C x} \\
&\quad -\beta x^T C x \left(1 + \frac{1}{p}\varphi(\Phi^{-1}(1-p))\Phi^{-1}(1-p) - \frac{1}{p^2}\varphi^2(\Phi^{-1}(1-p))\right)\Big\} \\
&\text{subject to} \\
&\sum_{i=1}^{n} x_i = 1 \\
&\quad x \ge 0,
\end{aligned}$$

where β is some positive constant. Model III. (a) is not a convex optimization problem because $\sqrt{x^T C x}$ is a convex function of the variable x.

The value of p should be chosen large, e.g., 0.9, 0.95, 0.99. The constraints (4.13) and (4.14) when included into the models, as described above may produce infeasibility. In that case we decrease v_0 or p or both to achieve feasibility.

Efficient frontiers, based on Models I. and II. can be constructed in a similar way as suggested by Markowitz (1952, 1959).

5 Conclusions

We have formulated new portfolio selection models, where instead of the expectation and variance of the return conditional expectation and conditional variance are used, given that the return is greater than or equal to some fixed value. The new models are counterparts of Markowitz's models. A lower bound on the Value at Risk, that we include among the constraints, ensures the fulfilment of the condition. Since the conditional expectation is greater than or equal to the unconditional expectation and the conditional variance is smaller than or equal to the unconditional variance, the new models offer more favourable optimal solutions than Markowitz's original models. Most of the new problems are, however, nonconvex and therefore difficult to solve. New numerical solution techniques have to be developed. Research in this respect is underway.

References

[1] Burridge, J. (1982). Some unimodality properties of likelihoods derived from grouped data. *Biometrika* **69**, 145–151.

[2] Kataoka, S. (1963). A stochastic programming model. *Econometrica* **31**, 181–196.

[3] Klein Haneveld, W.K. (1986). *Duality in Stochastic Linear and Dynamic Programming.* Lecture Notes in Economics and Mathematical Systems 274. Springer, New York.

[4] Markowitz, H. (1952). Portfolio Selection. *Journal of Finance* **7**, 77–91.

[5] Markowitz, H. (1959). *Portfolio Selection.* Wiley, New York.

[6] Pflug, G. Ch. (2000). Some Remarks on the Value-at-Risk and Conditional Value-at-Risk. In: *Probabilistic Constrained Optimization* (S.P. Uryasev, ed.). Kluwer, Dordrecht, 272–281.

[7] Prékopa, A. (1971). Logarithmic Concave Measures with Application to Stochastic Programming. *Acta Sci. Math.* (Szeged) **32**, 301–316.

[8] Prékopa, A. (1973a). On Logarithmic Concave Measures and Functions. *Acta Sci. Math.* (Szeged), **34**, 335–343.

[9] Prékopa, A. (1973b). Contributions to the Theory of Stochastic Programming. *Math. Prog.* **4**, 202–221.

[10] Rockafellar, R.T., S. Uryasev (2002). Conditional Value-at-Risk for General Loss Distribution. *Journal of Banking and Finance* **26**, 1443–1471.