A REMARK ON NONLINEAR FUNCTIONALS AND EMPIRICAL ESTIMATES

VLASTA KAŇKOVÁ

Abstract. Optimization problems depending on a probability measure correspond very often to economic situations. Since the probability measure is there very often completely unknown, statistical estimates (based on date) have to replace mostly the unknown probability measure to obtain at least an approximate solution and an approximate optimal value. Properties of such statistical estimates have been investigated many times in the case of linear dependence of an objective function on the probability measure. However, this assumption is not fulfilled just in many economic models, see e.g. Markowitz model or some risk measures. We try to cover some of these complicated cases.

Keywords: Stochastic optimization problems, nonlinear functionals, empirical estimates, convergence rate, portfolio selection, Markowitz model.

1. INTRODUCTION

Let \((Ω, S, P)\) be a probability space; \(ξ := (ξ_1(ω), \ldots, ξ_s(ω))\) an \(s\)-dimensional random vector defined on \((Ω, S, P)\); \(F := F(z), z ∈ R^s\) the distribution function of \(ξ; F_i, i = 1, \ldots, s\) one-dimensional marginal distribution functions corresponding to \(F; P_F, Z := Z_F\) the probability measure and support corresponding to \(F\). Let, moreover, \(g_0 := g_0(x, z)\) be a real-valued (say continuous) function defined on \(R^n × R^s; X ⊂ R^n\) be a nonempty “deterministic” set. If the symbol \(E_F\) denotes the operator of mathematical expectation corresponding to \(F,\) then many economic applications correspond to a class of stochastic optimization problems that can be introduced in the form:

Find

\[ \varphi(F) = \inf \{ E_F g_0(x, ξ) | x ∈ X \}. \]  

(1)

In applications very often the “underlying” probability measure \(P_F\) has to be replaced by empirical one; evidently, then the solution is sought w.r.t. the “empirical problem”:

Find

\[ \varphi(F_N) = \inf \{ E_{F_N} g_0(x, ξ) | x ∈ X \}, \]  

(2)

where \(F_N\) denotes an empirical distribution function determined by (mostly) an independent random sample \(\{ξ_i\}_{i=1}^N\) corresponding to the distribution function \(F.\) If \(X(F), X(F_N)\) denote the optimal solution sets of the problems (1) and (2), then under rather general assumptions \(\varphi(F_N), X(F_N)\) are “good” stochastic estimates of \(\varphi(F), X(F)\) (see e.g. \([2, 5, 6, 8, 12, 15, 16, 17, 18, 19])\). There were introduced assumptions guaranteeing the consistency, asymptotic normality and convergence rate. Especially, it means in the last case that there were introduced assumptions under which

\[ P\{ω : N^β|\varphi(F) - \varphi(F_N)| > t\} → (N → ∞) 0 \quad \text{for} \quad t > 0, \ β ∈ (0, \frac{1}{2}). \]  

(3)
To obtain this relation the Hoeffding inequality (see e.g. [4], [6]), large deviation (see e.g. [5]), Talagrand approach (see e.g. [13]) and the stability results (see e.g. [10]) have been employed. Furthermore, let us consider a simple “underlying” classical portfolio problem:

Find

$$\max \sum_{k=1}^{n} \xi_k x_k \quad \text{s.t.} \quad \sum_{k=1}^{n} x_k \leq 1, \quad x_k \geq 0, \quad k = 1, \ldots, n, \quad s = n, \quad (4)$$

where $x_k$ is a fraction of the unit wealth invested in the asset $k$, $\xi_k$ denotes the return of the asset $k \in \{1, 2, \ldots n\}$. If $\xi_k$, $k = 1, \ldots, n$ are known, then (4) is a linear programming problem. However, $\xi_k$, $k = 1, \ldots, n$ are mostly random variables with unknown realizations in a time decision. If we denote

$$\mu_k = E_F \xi_k, \quad c_{k,j} = E_F (\xi_k - \mu_k) (\xi_j - \mu_j), \quad k, j = 1, \ldots n, \quad (5)$$

then it is reasonable to set to the portfolio selection two–objective optimization problem:

Find

$$\max \sum_{k=1}^{n} \mu_k x_k, \quad \min \sum_{k=1}^{n} \sum_{j=1}^{n} x_k c_{k,j} x_j \quad \text{s.t.} \quad \sum_{k=1}^{n} x_k \leq 1, \quad x_k \geq 0, \quad k = 1, \ldots, n. \quad (6)$$

Evidently, there exists only rarely a possibility to find an optimal solution simultaneously with respect to the both criteria. Markowitz suggested (see e.g. [3]) to replace the problem (6) by one–criterion optimization problem of the form:

Find

$$\varphi^M(F) = \max \left\{ \sum_{k=1}^{n} \mu_k x_k - K \sum_{k=1}^{n} \sum_{j=1}^{n} x_k c_{k,j} x_j \right\} \quad \text{s.t.} \quad \sum_{k=1}^{n} x_k \leq 1, \quad x_k \geq 0, \quad k = 1, \ldots, n, \quad (7)$$

where $K \geq 0$ is a constant. Obviously, for every $K \geq 0$ there exists $\lambda \in (0, 1)$ such that the problem (7) is equivalent to the following one:

Find

$$\varphi^\lambda(F) = \max \left\{ \lambda \sum_{k=1}^{n} \mu_k x_k - (1-\lambda) \sum_{k=1}^{n} \sum_{j=1}^{n} x_k c_{k,j} x_j \right\} \quad \text{s.t.} \quad \sum_{k=1}^{n} x_k \leq 1, \quad x_k \geq 0, \quad k = 1, \ldots, n. \quad (8)$$

Evidently,

$$\sigma^2(x) = \sum_{k=1}^{n} \sum_{j=1}^{n} x_k c_{k,j} x_j = E_F \left( \sum_{j=1}^{n} \xi_j x_j - E_F [\sum_{j=1}^{n} \xi_j x_j] \right)^2, \quad x = (x_1, \ldots, x_n)$$

can be considered as a risk measure, that can be replaced by

$$\sigma(x) = \sqrt{E_F \left[ \sum_{j=1}^{n} \xi_j x_j - E_F [\sum_{j=1}^{n} \xi_j x_j] \right]^2}, \quad x = (x_1, \ldots, x_n). \quad (9)$$

For more details, see e.g. [1], where an analysis of the corresponding relationship (according to multiobjective optimization theory) is introduced. Konno and Yamazaki introduced in [11] another risk measure $w(x)$ by

$$w(x) = E_F \left| \sum_{k=1}^{n} \xi_k x_k - E_F [\sum_{k=1}^{n} \xi_k x_k] \right|. \quad (10)$$
Moreover, they have proven that \( w(x) = \sqrt{\frac{2}{n}} \sigma(x) \) in the case of mutually normally distributed random vector \( \xi = (\xi_1, \ldots, \xi_n) \).

Evidently, \( \mathbb{E}_F \left| \sum_{k=1}^n \xi_k x_k - y \right| \) is a Lipschitz function of \( y \), as well as the objective function \( [\lambda \sum_{k=1}^n \mu_k x_k + (1 - \lambda)w(x)] \), \( \lambda \in (0, 1) \) is a Lipschitz function of \( y := \mathbb{E}_F \left[ \sum_{k=1}^n \xi_k x_k \right] \). Some others risk measures fulfilling the Lipschitz property can be found e.g. in [14].

To introduce more general problems covering (10), let \( h(x, z) = (h_1(x, z), \ldots, h_m(x, z)) \) be an \( m \times n \)-dimensional vector function defined on \( R^n \times R^s \), \( g_0^1(x, z, y) \) be a function defined on \( R^n \times R^s \times \mathbb{R}^{m_1} \). According to this new situation we replace the problem (1) by a stochastic programming problem in the form:

Find

\[
\varphi(F) := \varphi^1(F) = \inf \{ \mathbb{E}_F g_0^1(x, \xi, \mathbb{E}_F h(x, \xi)) | x \in X \}. \tag{11}
\]

### 2. PROBLEM ANALYSIS

To investigate the empirical estimates of the problem (11), evidently, we can employ the assertion introduced in [10].

**Lemma 1.** [10] Let \( G \) be an arbitrary \( s \) dimensional distribution function. If

1. \( g_0^1(x, z, y) \) is for every \( x \in X, z \in R^s \) a Lipschitz function of \( y \in Y \) with a Lipschitz constant \( L^y(x, z) \); where \( Y = \{ y \in R^{m_1} : y = h(x, z) \text{ for some } x \in X, z \in R^s \} \),

2. for every \( x \in X, y \in Y \) there exist finite mathematical expectations

\[
\mathbb{E}_F L^y(x, \xi), \quad \mathbb{E}_F h(x, \xi), \quad \mathbb{E}_G h(x, \xi), \\
\mathbb{E}_F g_0^1(x, \xi, \mathbb{E}_F h(x, \xi)), \quad \mathbb{E}_F g_0^1(x, \xi, \mathbb{E}_G h(x, \xi)), \quad \mathbb{E}_G g_0^1(x, \xi, \mathbb{E}_G h(x, \xi)),
\]

then for every \( x \in X \) it holds that

\[
| \mathbb{E}_F g_0^1(x, \xi, \mathbb{E}_F h(x, \xi)) - \mathbb{E}_G g_0^1(x, \xi, \mathbb{E}_G h(x, \xi)) | \leq \mathbb{E}_F L^y(x, \xi) || \mathbb{E}_F h(x, \xi) - \mathbb{E}_G h(x, \xi) ||_{m_1}^2 + \mathbb{E}_G g_0^1(x, \xi, \mathbb{E}_G h(x, \xi)) - \mathbb{E}_G g_0^1(x, \xi, \mathbb{E}_G h(x, \xi)) \|. \tag{12}
\]

\((\| \cdot \|_{m_1}^2 \) denotes the Euclidean norm in \( R^{m_1} \).

Consequently, the assumptions guaranteeing the relation (3) can be employed in this case to obtain new results (for the problem (11)). However, the classical Markowitz problem (7) is not covered by the problem (11). A special case (dealing with the Markowitz model) has been considered in [9]. To recall this assertion we denote by the symbol \( C(= C(n \times n)) \) the matrix with elements \( c_{k,j}, k, j = 1, \ldots, n \) defined by relation (5). Furthermore, we denote by the symbols \( x(F), x(F^N) \) the solutions of the problem (7) and the corresponding empirical problem.
Proposition 1. Let $Z_F$, $X$ be compact sets, $t > 0$, $\{\xi^i\}_{i=1}^{N}$ independent random sample, $N = 1, 2, \ldots, \beta \in (0, \frac{1}{2})$, then
\[
P\{\omega : N^\beta |\varphi^M(F) - \varphi^M(F^N)| > t\} \rightarrow (N \rightarrow \infty) 0.
\]
If, moreover $M > 0$ and the matrix $C$ is positive definite, then also
\[
P\{\omega : N^\beta \|x(F^N) - x(F)\|^2 > t\} \rightarrow (N \rightarrow \infty) 0.
\]
Proof. The assertion of Proposition 1 follows immediately from Theorem 3 [9].

3. SOME AUXILIARY Assertions

In this section we prove some auxiliary assertion. To this end we consider $s = 2$ and set $\xi := \xi, \xi_2 := \eta$, where $\xi := \xi(\omega), \eta := \eta(\omega)$ are random valuables defined on $(\Omega, \mathcal{S}, P)$ with finite second moments. If we denote by the symbols $F(= F(\xi, \eta)), F_\xi, F_\eta$ the distribution functions of the random vector $(\xi, \eta)$ and marginal distribution functions of the random values $\xi$ and $\eta$, then
\[
|E_F[(\xi - E_F\xi)(\eta - E_F\eta)]| - E_F[N((\xi - E_F\xi)(\eta - E_F\eta))]| \leq |E_F\xi - E_FN\xi| + |E_FN\xi - E_F\xi| + |E_FN\xi - E_F\xi| = \frac{1}{\sqrt{\beta}}
\]
\[
|E_F\xi - E_FN\xi| + |E_F\xi - E_FN\xi| + |E_F\xi - E_FN\xi| = \frac{1}{\sqrt{\beta}}
\]
Consequently for $t > 0$
\[
P\{\omega : N^\beta |E_F[(\xi - E_F\xi)(\eta - E_F\eta)]| - E_FN[(\xi - E_F\xi)(\eta - E_F\eta)]| \geq t\} \leq P\{\omega : N^\beta |E_F\xi - E_FN\xi| > \frac{1}{\sqrt{\beta}}\} + P\{\omega : N^\beta |E_F\xi - E_FN\xi| > \frac{1}{\sqrt{\beta}}\}
\]
\[
P\{\omega : N^\beta |E_F\xi - E_FN\xi| > \frac{1}{\sqrt{\beta}}\}
\]
If we set $\Omega_1^N(t) = \{\omega : |E_F\eta - E_F\eta| < \sqrt{t}\}, \Omega_1^N(t) = \Omega - \Omega_1^N$ and assume (without loss of generality) that $E,F\eta > 0$, then
\[
P\{\omega : N^\beta |E_F\eta - E_F\eta| > \frac{1}{\sqrt{\beta}}\} = P\{\omega : N^\beta |E_F\eta - E_F\eta| > \frac{1}{\sqrt{\beta}}\}
\]
\[
P\{\omega : N^\beta |E_F\eta - E_F\eta| > \frac{1}{\sqrt{\beta}}\}
\]
Lemma 2. Let $\tilde{\xi} = \tilde{\xi}\eta := \tilde{\xi}(\omega)\eta(\omega)$. Let moreover $F_{\tilde{\xi}}$ denote the distribution function of $\tilde{\xi}$. If
1. $P_{\tilde{\xi}}, P_{\tilde{\eta}}$ are absolutely continuous with respect to the Lebesgue measure on $R^1$ (we denote by $f_{\tilde{\xi}}, f_{\tilde{\eta}}$ the probability densities corresponding to $F_{\tilde{\xi}}, F_{\tilde{\eta}}$,
2. there exist constants $C_1^\xi, C_2^\xi, C_1^\eta, C_2^\eta > 0$ and $T > 0$ such that
\[
f_\xi(z) \leq C_1^\xi \exp\{-C_2^\xi |z|\} \quad \text{for} \quad z \in (-\infty, -T) \cup (T, \infty),
\]
\[
f_\eta(z) \leq C_1^\eta \exp\{-C_2^\eta |z|\} \quad \text{for} \quad z \in (-\infty, -T) \cup (T, \infty),
\]
then, there exist constants $\bar{C}_1^\xi, \bar{C}_2^\xi > 0, \bar{T} > 1$ such that for $z > \bar{T}$
\[
F_\xi(-z) \leq \frac{C_1^\xi}{\bar{C}_2^\xi} \exp\{-C_2^\xi \sqrt{z}\}, \quad (1 - F_\xi) \leq \frac{C_1^\xi}{\bar{C}_2^\xi} \exp\{-C_2^\xi \sqrt{z}\}.
\]

Proof. First, evidently, for $z > T$
\[
F_\xi(-z) \leq \frac{C_1^\xi}{\bar{C}_2^\xi} \exp\{-C_2^\xi z\}, \quad 1 - F_\xi(z) \leq \frac{C_1^\xi}{\bar{C}_2^\xi} \exp\{-C_2^\xi z\},
\]
\[
F_\eta(-z) \leq \frac{C_1^\eta}{\bar{C}_2^\eta} \exp\{-C_2^\eta z\}, \quad 1 - F_\eta(z) \leq \frac{C_1^\eta}{\bar{C}_2^\eta} \exp\{-C_2^\eta z\}.
\]
Consequently, if $\bar{\xi}(\omega) = \bar{\eta}(\omega)$ a.s., then
\[
P\{\omega : \bar{\xi} < -z\} = 0, \quad P\{\omega : \bar{\xi} > z\} = P\{\omega : |\bar{\xi}| > \sqrt{z}\} = 2 \frac{C_1^\xi}{\bar{C}_2^\xi} \exp\{-C_2^\xi \sqrt{z}\},
\]
if $\bar{\xi}(\omega) \neq \bar{\eta}(\omega)$, then evidently for $z > 1$
\[
P\{\omega : \bar{\xi} > z\} = P\{\omega : \bar{\xi} \bar{\eta} > z\} \leq
P\{\omega : \bar{\xi} \bar{\eta} > z; \ |\bar{\xi}| > \sqrt{z}\} + P\{\omega : \bar{\xi} \bar{\eta} > z; \ |\bar{\eta}| > \sqrt{z}\} \leq
P\{\omega : |\bar{\xi}| > \sqrt{z}\} + P\{\omega : |\bar{\eta}| > \sqrt{z}\} \leq 2 \frac{C_1^\xi}{\bar{C}_2^\xi} \exp\{-C_2^\xi \sqrt{z}\} + 2 \frac{C_1^\eta}{\bar{C}_2^\eta} \exp\{-C_2^\eta \sqrt{z}\}.
\]
Evidently, the assertion of Lemma 2 is valid. \(\square\)

Furthermore, it follows from Lemma 2 that for $z > \bar{T}$ there exist $D_1, D_2, D_3$ such that
\[
\int_{z}^{+\infty} (1 - F_\xi(u))du \leq D_1 \sqrt{z} e^{-D_2 \sqrt{z}} + D_3 e^{-D_2 \sqrt{z}}. \quad (16)
\]

4. CONVERGENCE RATE

To introduce more general assertion dealing with the Markowitz problem we employ an approach employing in [10].

Theorem. Let $X$ be a compact set. Let, moreover, $\{\xi_i\}_{i=1}^{N}$ be an independent random sample corresponding to the distribution function $F$, $N = 1, 2, \ldots$. If
1. $P_{F_i}$, $i = 1, \ldots, n$ are absolutely continuous with respect to the Lebesgue measure on $R^1$ (we denote by $f_i$ probability densities corresponding to $P_{F_i}$).

2. there exist constants $C_1^i, C_2^i > 0$ and $T > 0$ such that

$$f_i(z) \leq C_1^i \exp\{-C_2^isz\} \quad \text{for} \quad z \in (\infty, -T) \cup (T, \infty), \quad i = 1, \ldots, n,$$

then for $t > 0$, $\beta \in (0, \frac{1}{2})$

$$P\{\omega: N^{\beta}||\varphi^M(F) - \varphi^M(F_N)|| > t\} \longrightarrow_{(N \to \infty)} 0.$$

If, moreover $M > 0$ and the matrix $C$ is positive definite, then also

$$P\{\omega: N^{\beta}||x(F_N) - x(F)||^2 > t\} \longrightarrow_{(N \to \infty)} 0.$$

**Proof.** Employing the relations (14), (15), (16), the assertion of Lemma 2 and the technique employed in [10] we obtain the assertion of Theorem.

**Acknowledgment** The research was supported by the Czech Science Foundation under Grants 402/07/1113, 402/08/0107 and 402/06/0990.

**References**


RNDr. Vlasta Kaňková, CSc.,
Department of Econometrics, Institute of Information Theory and Automation of AS ČR, Pod vodárenskou věží 4, 18208 Praha 8, Czech Republic.