

Role of Dependence in Chance-constrained and Robust Programming[#]

Michal Houda^{*}

Introduction

Mathematical programming, a discipline covering methods that try to select the “best” solution is widely used by economic and engineering researchers and practitioners. The breakthrough was certainly the development of the simplex method that served a very useful purpose on solving linear optimization problems, by G. B. Dantzig (1947) – he provided an instrument for automated searching of solutions of such problems and also its applications to solve real-world problems. Methods of linear and (later developed) nonlinear programming deal with *deterministic* problems. But *uncertainty* is a natural property of all the real-world models. In the early paper, Dantzig (1955) himself realized very large importance of uncertainty in optimization. Sometimes, unknown parameters can be replaced without worries (approximated with some “good” estimates), but otherwise such simplification is not possible and could lead to false conclusions.

Origin of the uncertainty in optimization models

Consider an optimization problem of the form;

$$\text{minimize } c(x; \xi) \text{ subject to } x \in X, f(x; \xi) \leq 0, \quad (1)$$

where $\xi \in \mathbf{R}^s$ is a data element of the problem, $x \in X \subset \mathbf{R}^n$ is a decision vector, the dimensions n, s, M and the mappings $c : \mathbf{R}^n \times \mathbf{R}^s \rightarrow \mathbf{R}$ and $f : \mathbf{R}^n \times \mathbf{R}^s \rightarrow \mathbf{R}^M$ are structural elements of the problem.

Such a frame can be followed for a great variety of decision-making problems but different optimization methods have to be used according to the nature of structural and data elements of the problem. We extend the description of the problem by the following characterization:

1. the knowledge of the data is insufficient, all that is known about ξ is that it belongs to a given *uncertainty set* $\Xi \subset \mathbf{R}^s$;
2. the objective is required to be the best possible given an actual realization of ξ ;
3. the constraints of the problem (1) are required to be satisfied as much as possible given the actual realization of ξ .

[#] The article is processed as an output of the research projects *Nonlinear Dynamical Economic Systems: Theory and Applications* and *Economic Systems Under Uncertainty: Optimization and Approximation*, registered by the Czech Science Foundation under the registration numbers 402/03/H057 and 402/07/1113.

^{*} Mgr. Michal Houda – student of doctoral study; Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University in Prague, Ke Karlovu 3, 121 16 Prague 2; <houda@karlin.mff.cuni.cz>.

If a realization of ξ is known and fixed in advance (before the decision has to be taken), and the problem elements have suitable properties, standard deterministic optimization algorithms can be used to solve the problem (1). This is rarely the case; in practice, uncertainty of the data is to be taken into account. For example,

- the data ξ is not known at the time when a decision has to be taken and it is observed later in the future;
- the data ξ cannot be measured or estimated exactly even if it is realized before a decision is taken;
- the solution of the problem cannot be implemented exactly due to its physical characterization – this can be transformed to the model as uncertainty of the data (not uncertainty of the solution);
- the model itself is an approximation of a real problem.

Dealing with uncertainty of the data

During the history of mathematical programming, several approaches were developed to deal with uncertainty in optimization problems. First, the uncertainty is simply ignored, a fixed (nominal) value of the data is chosen and the problem is solved. The accuracy of the optimal solution is examined ex-post by the traditional *sensitivity analysis*. The usability of this kind of stability analysis is limited to a single, already generated solution; and, unfortunately, these solutions often happen to be very sensitive to perturbations of ξ .

If we declare ξ to be a random vector with known distribution, we can incorporate such information into the model – this is what *stochastic programming* does. The idea is to replace uncertainty of the data by known probability distribution and solve the newly created, deterministic optimization problem. There are many ways how to do that; among them, the most popular are

- penalization/recourse problems: the potential loss due to the uncertainty is compensated ex-post by solving another optimization program that depends on the first-stage decision x and the (realized) data ξ ; in the first-stage program, the expected value of the (uncertain) second-stage optimal value is taken into account;
- chance-constrained programming problems: the constraints of the original problem are not required to be satisfied in all cases but “only” with enough high probability.

There is a huge amount of literature devoted to the stochastic programming and its applications, starting with Dantzig’s paper (1955); we refer the reader to the books (Kall – Wallace, 1994), (Prékopa, 1995), (Birge – Louveaux, 1997), (Ruszczyński – Shapiro, 2003).

Completely different approach to solve the problem (1) is the “worst-case” analysis, also known as *robust programming*. The key is to find a solution that satisfies the constraints of (1) for all possible realization of ξ belonging to some given set Ξ . (In real applications, this set has not to be identical with the uncertainty set in (1), but the distinction is not important here and we shall use the same symbol Ξ for both

meanings.) The framework of the worst-case analysis was introduced by Ben-Tal and Nemirovski (1998) and developed by other authors in various directions, see e. g. (El Ghaoui – Oustry – Lebret, 1998).

Uncertain, stochastic and robust programming problems

Let us start with the following family of optimization programs (1) parameterized by $\xi \in \Xi$:

$$\text{minimize } c'x \text{ subject to } x \in X, f(x; \xi) \leq 0, \quad (2)$$

where $X \subset \mathbf{R}^n$ is convex and closed set. Without loss of generality, we assume that the objective is linear (non-linear objective can be transferred to the constraint part of the problem by minimizing t subject to $c(x; \xi) \geq t$) and that the function $f : X \times \Xi \rightarrow \mathbf{R}$ is scalar-valued (converted from the original multiple-valued form by $f(x; \xi) := \max_i f_i(x; \xi)$). Such parameterized program is known as *uncertain convex program*, and if the realization of ξ is known and fixed, we can use standard convex programming tools to solve the problem.

Assume that ξ is a random vector defined on some probability space $(\Omega, \mathcal{A}, \mathbf{P})$ with a known probability distribution P and allow a solution of the problem (2) violate the constraints with a small level of the “risk”, $\varepsilon \in [0; 1]$. The *chance-constrained program* (as a particular case of stochastic programming problems) reads

$$\text{minimize } c'x \text{ subject to } x \in X_\varepsilon := \{x \in X, P\{f(x; \xi) \geq 0\} \leq \varepsilon\}. \quad (3)$$

Conditions on P and f under which the set X_ε is convex are treated in the literature (Prékopa, 1995). Computational issues arise when we want to compute the probability measure in the definition of X_ε because it involves multidimensional integrals. For that reason, we are interested in approximations of the program (3), in particular by a sampled version (see next parts of the paper).

In the *robust convex program* we look for a solution of (2) that is feasible for all instances of $\xi \in \Xi$. The problem reads

$$\text{minimize } c'x \text{ subject to } x \in X, f(x; \xi) \leq 0 \text{ for all } \xi \in \Xi.$$

The program (4) has infinite number of constraints and some relaxation techniques are needed in order to solve the problem. In the next section, we concentrate on the “randomized” approach to approximate the problem (4).

Empirical approximations of optimization problems

Consider a set of independent samples ξ_1, \dots, ξ_N distributed according to P and define the *empirical distribution function* – a discrete random variable of the form

$$P_N := \frac{1}{N} \sum_{i=1}^N \delta_{\xi_i}, \quad (5)$$

where δ_ξ denotes the Dirac measure placing the unit mass at ξ . For a given realization of samples, the empirical distribution function possesses the properties of distribution functions and, under rather general conditions, converges almost surely to the distribution function of the original distribution P . We are now ready to use the empirical distribution function as an approximation of P in the problem (3); we obtain the following “sampled” chance-constrained problem:

$$\text{minimize } c'x \text{ subject to } x \in X, \frac{1}{N} \text{card} \{i; f(x; \xi_i) > 0\} \leq \varepsilon. \quad (6)$$

The essential idea of such approximation is that the relative frequency of constraint violations in (2) is approximately the desired level of infeasibility in (3). Under suitable assumptions, convergence properties of the program (6) are assured; see (Römisch, 2003).

The robust programming problem does not involve any information about stochastic nature of the uncertainty parameter ξ in its original form (4). Calafiore and Campi (2005) and de Farias and Van Roy (2004) proposed so-called randomized approach considering a set of independent samples ξ_1, \dots, ξ_N from a chosen distribution P , and so put back the stochastic nature of the uncertainty. The robust sampled convex problem reads:

$$\text{minimize } c'x \text{ subject to } x \in X, f(x; \xi_i) \leq 0 \text{ for all } i = 1, \dots, N. \quad (7)$$

The idea of the problem (7) is relaxing the original problem (4) and requiring the constraints to be satisfied not for all the possible realizations of ξ , but for a sufficiently high number of samples which are in addition most probably to happen.

Properties of the problems (6) and (7) are deeply examined in (Houda, 2006a), (Houda, 2006b); we demonstrated that both the programs (6) and (7) (and/or their original non-sampled versions (3) and (4)) can lead to very different solutions although they are solving mathematically the same optimization problem (1). In this paper we continue in this direction and take into account dependence properties that sample sets may acquire.

Financial application of stochastic and robust programming

Stochastic programming has a long history and therefore a great number of applications in many scientific areas. Typical applications of chance-constrained programming include financial problems; constraint inequalities involve uncertainties in the model parameters like product demand, currency exchange rates, etc. The other typical property of chance-constrained models is that infeasibility in constraints may not be expressed in terms of “compensation costs”, as is typical for models from another area of stochastic programming (problems with recourse). To illustrate one of possible applications, we give an example of the portfolio selection (cash matching) model (Dentcheva – Lai – Ruszczyński, 2004): the objective is to find an optimal portfolio of financial instruments that maximizes the final amount of money and ensures that all requested payments during the considered period is made. The uncertainty of the model (random yields and/or payments) is contained in liquidity constraints and we require that the constraints are satisfied with a sufficiently high probability.

Robust programming is recently used in many application contexts (truss design, reliability systems, filtering, etc.). As an example of a finance application we mention the robust portfolio selection problem (Goldfarb – Iyengar, 2003). In this framework, we consider perturbations in the market parameters. The uncertain parameters are considered bounded and the optimization problem is solved assuming worst-case behavior of the perturbations.

Example: simple optimization program and dependent samples

We adopt the settings of (Houda, 2006a) and consider the following simple uncertain convex program:

$$\text{minimize } x \text{ subject to } x \in \mathbf{R}, x \geq \xi, \quad (8)$$

where ξ is a random variable distributed according to standard normal distribution $N(0;1)$. The distribution function will be denoted by F . Following the direction outlined in previous sections, the chance-constrained program (9) and the approximated versions of chance-constrained (10) and robust program (11) of the case problem (8) are defined by:

$$\text{minimize } x \text{ subject to } x \in \mathbf{R}, x \geq F^{-1}(1 - \varepsilon), \quad (9)$$

$$\text{minimize } x \text{ subject to } x \in \mathbf{R}, \frac{1}{N} \text{card} \{i; x < \xi_i\} \leq \varepsilon, \quad (10)$$

$$\text{minimize } x \text{ subject to } x \in \mathbf{R}, x \geq \max \xi_i. \quad (11)$$

The robust program is not well defined here – the support $\Xi = \mathbf{R}$ is unbounded and there is no solution x that is feasible for all instances of ξ .

The solution of (9) is easy to calculate; with $\varepsilon = 0.05$ it is the 0.95-quantile of the standard normal distribution, which is approximately $x = 1.64$. As shown in (Houda, 2006a) the optimal solution of (10) approximates this value (more the number N of samples grows, better is the approximation) rather than the optimal solution of (11) which tends to “beat” this value (more the number N of samples grows, more the probability not to exceed the value is smaller). We will now take into account two types of dependence and compare the same results obtained with dependent and with independent sets of samples.

In the sequel we replace in (10) and (11) the set of independent samples by another one of samples that are weakly dependent. We take into consideration two particular series of samples:

1. moving-average process MA(2), defined by

$$\xi_k := 0.5\zeta_k + 0.3\zeta_{k-1} + 0.2\zeta_{k-2}; \quad (12)$$

2. autoregressive process AR(1), defined by

$$\xi_k := \zeta_k + 0.5\zeta_{k-1}. \quad (13)$$

The series ζ_1, ζ_2, \dots in (12) has independent samples with the same normal distribution with zero mean and variance 1/0.38 so the variance of ξ_k is equal 1 by convolution. In the case (13), the members of series ζ_1, ζ_2, \dots have the same normal distribution with zero mean and variance 0.75 so the variance of ξ_k is equal 1 again.

Moving-average process MA(2) is a special case of so-called m -dependent processes, here for $m=3$. A process $\{\xi_k\}$ is m -dependent if for any k the σ -algebra generated by the events $\dots, \xi_{k-2}, \xi_{k-1}, \xi_k$ is independent of the σ -algebra generated by the events $\xi_{k+m}, \xi_{k+m+1}, \xi_{k+m+2}, \dots$. AR(1) is not m -dependent but it has the strong mixing property leastwise. A sequence $\{\xi_k\}$ has the strong mixing property if the (supremum of) α -coefficient of dependence between the two σ -algebras defined above converges to zero with $m \rightarrow +\infty$, α -coefficient is defined as $\sup |P(AB) - P(A)P(B)|$ where A, B are events belonging respectively to the two σ -algebras defined above. m -dependent sequences as well as strong mixing sequences have suitable convergence properties and are of best interest to deal with them along the independent sequences. For theoretical background of weakly dependent sequences see for example (Yoshihara, 1992).

Numerical results are shown in the array of graphics (Figure 1 and 2) for various types of dependence and sample sizes. In each of the graphs, a short tick mark on the x -axis represents the optimal solution of the chance-constrained problem (9), dotted lines are estimated densities of the optimal solutions of the approximated chance-constrained problem (10), and solid lines represent optimal solutions of the robust sampled problem (11). The sampling procedure is repeated 200 times in order to estimate densities of optimal solutions in (10) and (11).

Convergence properties and differences between chance-constrained and robust programming were discussed in (Houda, 2006a) from a mathematical and in (Houda, 2006b) from an application point of view. We concentrate here on comparison concerning influence of weak dependence of sample sets.

Based on our expectation and conclusions of the theory of weakly dependent processes, such kind of dependence does not introduce any remarkable issue. With small sample sizes (30 samples, less with 300 samples) the only observable fact is the slightly augmented variance of approximated solutions, but importance of this weakness diminishes with growing number of samples at disposition. The important fact is that basic convergence properties of approximated stochastic as well as robust programs remain valid even if we consider some kind of weak dependence of sample sets instead of independent samples. This has a very large impact on practical applications in operation research because data series that we usually have at hand are rarely pure independent but some kind of weak dependence is frequent. The theoretical confirmation of this experimental fact remains still an open question.

Figure 1: Dependent and independent samples in optimization

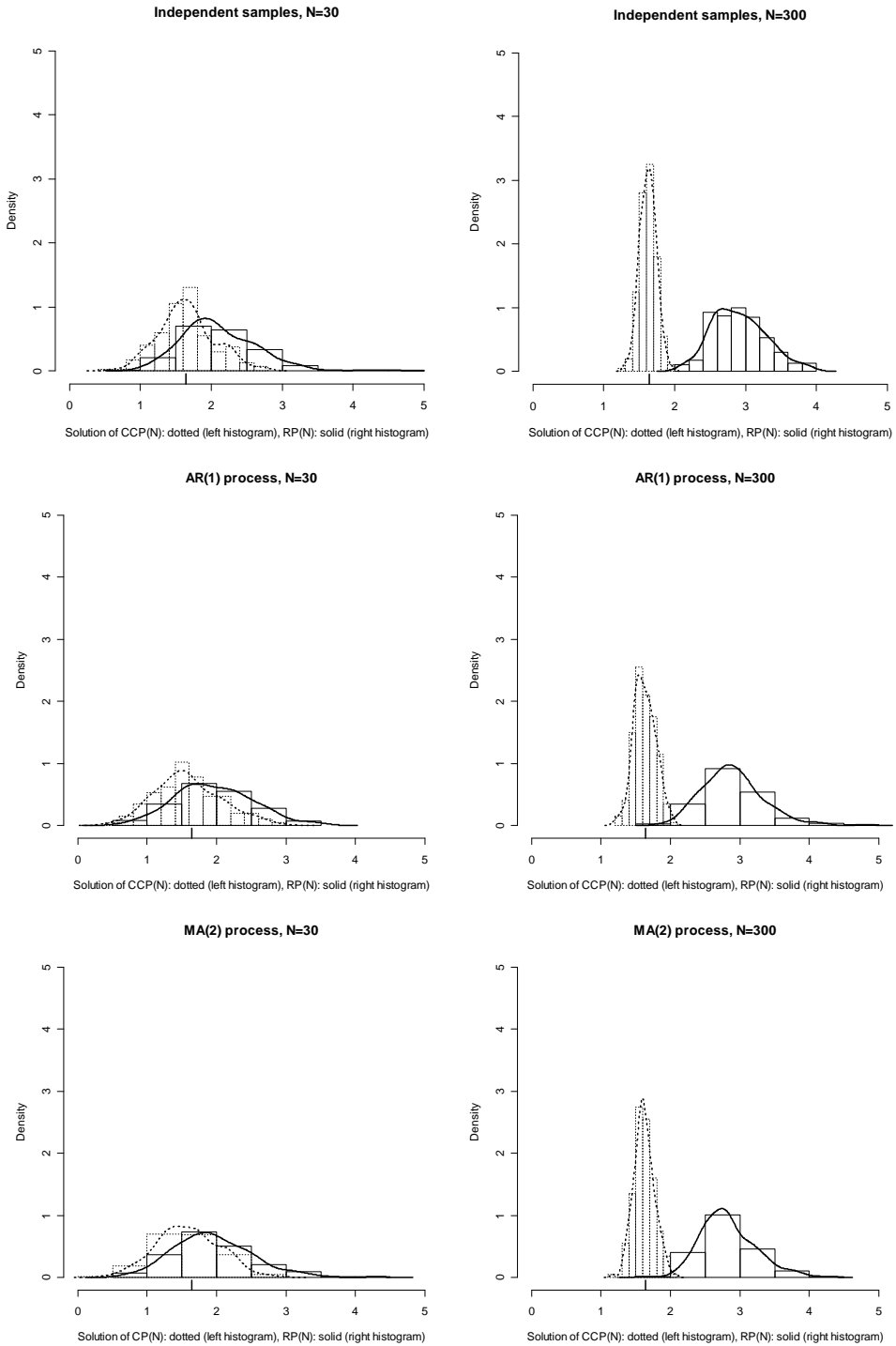
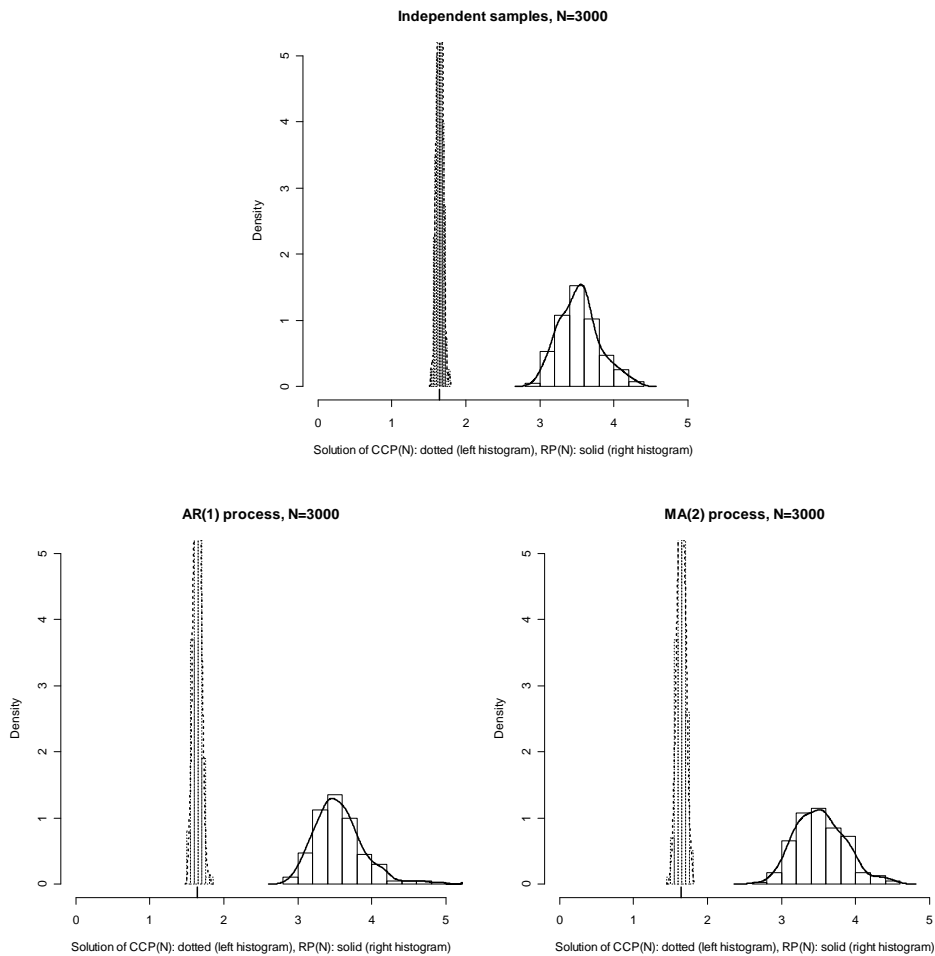


Figure 2: Dependent and independent samples in optimization (continued)



Conclusion

In the paper we have concentrated on special cases of approximations of uncertain convex program based on stochastic programming on one hand and on robust programming on the other hand. We took into consideration possible dependence of samples that one has at hand and demonstrated that this fact should not become an obstacle, especially if the number of samples is high. More consideration should be devoted to the correct choice of the optimization method because stochastic and robust programming lead to very different solutions although they are both solving the same original optimization program.

References

- [1] BEN-TAL, A.; NEMIROVSKI, A. (1998). *Robust convex optimization. Mathematics of Operations Research*, 1998, vol. 23, no. 4, pp. 769–805.
- [2] BIRGE, J. R.; LOUVEAUX, F. (1997). *Introduction To Stochastic Programming*. New York, Springer-Verlag, 1997.
- [3] CALAFIORE, G.; CAMPI, M. C. (2005). Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 2005, vol. 102, no. 1 (ser. A), pp. 25–46.
- [4] DANTZIG, G. B. (1947). Programming in a linear structure. *Econometrica*, 1949, vol. 17, pp. 73–74.
- [5] DANTZIG, G. B. (1955). Linear programming under uncertainty. *Management Science*, 1955, vol. 1, pp. 197–206.
- [6] DENTCHEVA, D.; LAI, B.; RUSZCZYŃSKI, A. (2004). Dual methods for probabilistic optimization problems. *Mathematical methods of operations research*, 2004, vol. 60, no. 2, pp. 331–346.
- [7] de FARIAS, D. P.; VAN ROY, B. (2004). On constraint sampling in the linear programming approach to approximate dynamic programming. *Mathematics of Operations Research*, 2004, vol. 29, no. 3, pp. 462–478.
- [8] EL GHAOUI, L.; OUSTRY, F.; LEBRET, H. (1998). Robust solutions to uncertain semidefinite programs. *SIAM Journal of Optimization*, 1998, vol. 9, no. 1, pp. 33–52.
- [9] GOLDFARB, D.; IYENGAR, G. (2003). Robust portfolio selection problems. *Mathematics of Operations Research*, 2003, vol. 28, no. 1, pp. 1–38.
- [10] HOUDA, M. (2006a). Comparison of approximations in stochastic and robust optimization programs. In Hušková, M.; Janžura, M. (eds.) *Prague Stochastics 2006*. Prague, Matfyzpress, 2006, pp. 418–425.
- [11] HOUDA, M. (2006b). Approximations in stochastic and robust programming problems. In Lukáš, L. (ed.). *Proceedings of the 24th International Conference Mathematical Methods in Economics 2006*. Plzeň, University of West Bohemia in Pilsen, 2006, pp. 249–254.
- [12] KALL, P.; WALLACE, S. W. (1994). *Stochastic Programming*. Chichester : Wiley, 1994.
- [13] PRÉKOPA, A. (1995). *Stochastic Programming*. Budapest : Akadémiai Kiadó, 1995.
- [14] RÖMISCH, W. (2003). Stability of stochastic programming problems. In Ruszczyński, A.; Shapiro, A. (eds.) *Stochastic Programming. Handbooks in Operations Research and Management Science*, Vol. 10. Amsterdam : Elsevier, 2003, pp. 483–554.
- [15] RUSZCZYŃSKI, A.; SHAPIRO, A. (eds.) (2003). *Stochastic Programming. Handbooks in Operations Research and Management Science*, vol. 10, Amsterdam : Elsevier, 2003.
- [16] YOSHIHARA, K. (1992). *Summation theory for weakly dependent sequences. Weakly dependent stochastic sequences and their applications* (Vol. 1). Tokyo : Sanseido, 1992.

Role závislosti v úlohách s pravděpodobnostními omezeními a v úlohách robustního programování

Michal Houda

Abstrakt

Článek se zabývá dvěma metodami pro řešení optimalizačních úloh, ve kterých se objevuje nejistota: stochastickým programováním (speciálně úlohami s pravděpodobnostními omezeními) a robustním programováním. Krátce představujeme způsoby, jakými obě metody pracují s nejistotou a jaké aproximace jsou obvykle při řešení používány. Dále se koncentrujeme na aproximace, které jsou založeny na množinách vzorků, ve kterých se objevuje určitá forma slabé závislosti. Ukazuje se, že takové typy závislostí nepředstavují významné riziko při užití zmíněných optimalizačních metod a přikládáme numerickou ilustraci na jednoduché optimalizační úloze.

Klíčová slova: stochastické programování; robustní programování; slabá závislost.

Role of Dependence in Chance-constrained and Robust Programming

Abstract

The paper deals with two methods of solving optimization programs where uncertainties occur: stochastic (in particular chance-constrained) programming and robust programming. We review briefly how these two methods deal with uncertainty and what approximations are commonly used. Furthermore, we are concentrated on approximations based on sample sets where some type of weak dependence occurs. We demonstrate that such kind of dependence does not imply any important malfunction of optimization methods used there. Numerical illustration on simple optimization program is given.

Key words: stochastic programming; robust programming; weak dependence.

JEL classification: C44, C61.