Georg Pflug's work and some personal memories

W. Römisch

Humboldt-University Berlin
Institute of Mathematics

http://www.math.hu-berlin.de/~romisch

Universität Wien, September 27, 2019
Emeritus Ceremony for Prof. Georg Ch. Pflug

Introduction

- Georg has much more than 100 publications in wide-ranging fields of interest covering and connecting theory and practical applications.
- To review parts of his work I had to select some main topics without intending to cover a high percentage of his research topics.
- My selection is
  - Statistical inference of stochastic programs
  - Risk measures
  - Scenario trees for multistage stochastic programs
- Georg organized a number of International Conferences (including WC Bernoulli Society1996, ICSP 2007 and OR 2015) and of Workshops.
- Georg's talks at some (early) SP Conferences are passed in revue, too.

INTERNATIONAL CONFERENCE ON STOCHASTIC PROGRAMMING

August 24 - 28, 1981

Kőszeg, Hungary

P R O G R A M

János Bolyai Mathematical Society

WEDNESDAY, August 26

SESSION V

Chairman: W.T. Ziemba

9.00 - 9.30 a.m.	×	Y. Ermoliev: ~~title will be announced later~~ Optimization problems with unknown distribution
9.30 - 10.00 a.m.	*	~~J.G. Kallberg: Approximate algorithms and computational complexity~~
10.00 - 10.30 a.m.	×	K. Marti: On the construction of descent directions in stochastic programs by means of stochastic dominance relations
10.30 - 10.45 a.m.		c o f f e e b r e a k
10.45 - 11.15 a.m.	×	G. Pflug: The penalization method in convex stochastic programming
11.15 - 11.45 a.m.	×	J. Pintér: Hybrid algorithms in stochastic programming

1.00 p.m.	Excursion by bus /Nagycenk, Fertőrákos, Sopron, dinner in Sopron/ /gathering in the hall of Hotel Irottkő/

* V. Arkin: Stochastic maximum principle in the problem of optimal control

ABSTRACTS

From Papers to be Presented
at the International Conference on

STOCHASTIC OPTIMIZATION

Kiev, USSR
9–16 September 1984

Sponsored by
The Committee for Systems Analysis of the USSR Academy of Sciences, Moscow,
The International Institute for Applied Systems Analysis, Laxenburg,
The Ukrainian Academy of Sciences, Kiev,
and The V. Glushkov Institute of Cybernetics, Kiev

Stochastic Minimization with constant Step - size - Asymptotic laws

G. Pflug (Gießen)

A multidimensional markovian process of the form

$$X_{n+1}^{a} = \Pi_S \, (X_n^a - a \, Y_n^a) \tag{1}$$

is considered. Here Π_S denotes the projection operator onto the closed convex set S and Y_n^a is a sequence of stochastic gradients. We are interested in the behavior of the invariant distribution of (1) as a tends to zero. Two different cases have to be distinguished:

(a) the tangent cone C_o at the minimal point x_o (with respect to S) contains a subspace of dimension greater than 0. In that case the invariant law, normalized by $a^{-1/2}$ tends to a degenerated normal distribution, if the curvature of S at x_o is quadratic.

(b) if S is "pointed" at x_o, then the invariant law, normalized by a^{-1} tends to a non-normal limiting distribution.

Faculty of Mathematics and Physics
Charles University

International Conference on Stochastic Programming

Abstracts
Instructions for the authors

Prague, September 15 – 19, 1986

PENALTY FUNCTIONS TECHNIQUE IN TWO-STAGE STOCHASTIC PROGRAMMING

Novikova N.M.

Let us consider the problem to search the value and a realization of the following stochastic minimum

$$(1) \quad \min_{x\in X} \int_Y \min_{z\in Z(x,y)} f(x,y,z)\,\sigma(dy),$$

where $Z(x,y)=\{z\in Z \mid g(x,y,z)\le 0\}$, functions f and g are convex with respect to z, x, sets X, Y, Z are compact in Euclid space.

We introduce the function

$$(2) \quad u(y)=\min_{z\in Z(x,y)} f(x,y,z)=\lim_{\alpha\to\infty}\min_{z\in Z}\{f(x,y,z)+\alpha g^{+}(x,y,z)\}.$$

It is continuous if $Z(x,y): Y\to Z$ is continuous by Housdorf. The sense of this construction is in better analytical properties of $u(\cdot)$ accordingly to the function $\arg\min_{Z(x,y)} f(x,y,z)$, used with traditional transformation the problem (1) to

$$\min_{x\in X}\ \min_{z(y)\in Z(x,y)\ \forall y\in Y} \int_Y f(x,y,z(y))\,\sigma(dy).$$

With the help of (2) the problem (1) transformes to the search of $\min\int_Y u(y)\sigma(dy)$ over $x, u(\cdot)$, under the constraint $u(y)\ge f(x,y,z)+\alpha g^{+}(x,y,z)$ $\forall z\in Z, \alpha\in R^{+}$, equivalent to $[f(x,y,z)+\alpha g^{+}(x,y,z)-u(y)]^{+}=0$ $\forall z\in Z, \alpha\in R^{+}$, or $\int_Z [f(x,y,z)+\alpha g^{+}(x,y,z)-u(y)]^{+}dz=0$ $\forall \alpha\in R^{+}$ $(\alpha\to+\infty)$.

Now on the base of integral penalty function of the problem (1) became

$$(3) \quad \lim_{C\uparrow\infty}\ \lim_{\alpha\uparrow\infty}\ \min_{x\in X, u(\cdot)} \int_Y \{u(y)+C\int_Z [f(x,y,z)+\alpha g^{+}(x,y,z)-u(y)]^{+}dz\}\,\sigma(dy).$$

Minimizing function in (3) is convex, functions u are continuous, uniformly boundered:

$$\forall y\in Y \quad \min_{X,Y,Z} f(x,y,z)\le u(y)\le \max_{X,Y,Z} f(x,y,z).$$

Therefore to solve the problem (3) let's apply stochastic gradients technique.

Optimization of simulated Markov Processes

G. Pflug, JLU Gießen

We consider a discrete-time continuous-state Markovian process on $\mathbb{R}^m$ with transition

$$Z_{n+1} = K(Z_n, x, \xi_n). \qquad (1)$$

x is a control vector and ξ_n are i.i.d. random variables. Suppose that μ_x is the unique invariant measure of (1). The performance of this stochastic system is described by a real-valued function $H(Z_n,x)$. We are interested in optimizing the expected value of H under the steady state law μ_x with respect to the control x, i.e.

$$\left\| \int H(z,x)\, d\mu_x = \max! \quad (\text{or } \min!) \right. \qquad (2)$$
$$x \in S$$

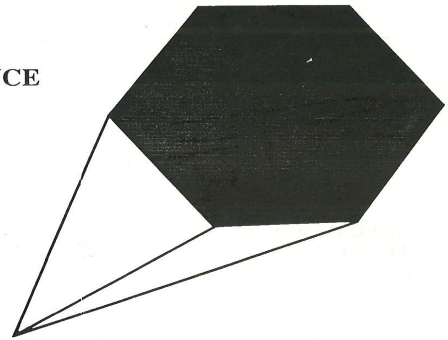
where S is a convex set of constraints. If μ_x is known, and the integral (2) may be evaluated, the problem is a deterministic constrained optimization problem. If the integral is complicated, but a random generator for μ_x is available, a Monte-Carlo like technique (the Stochastic Quasigradient Technique) may be used. Very often however, even μ_x is unknown and the only thing we can do is to simulate the process Z_n.

A method of on-line optimization is presented, which consists in a certain interplay of simulation (to get an approximation of the steady-state law μ_x) and a gradient technique (to approach the optimal control x*). The convergence properties of this new kind of procedure are studied and the nice behavior is demonstrated by some practical examples.

FIFTH INTERNATIONAL

CONFERENCE

ON



STOCHASTIC PROGRAMMING

* * * * * * * * * * * * * * *

FINAL PROGRAM

* * * * * * * * * * * * * * *

Ann Arbor, Michigan, USA
August 13–18, 1989

Sponsors:

The National Science Foundation

IBM Corporation

AT&T

Operations Research Society of America

The Institute of Management Sciences

Department of Industrial and Operations Engineering and

College of Engineering, The University of Michigan

The Committee for Stochastic Programming, Mathematical Programming Society

Technical Committee 7 (System Modeling and Optimization), IFIP covariance between objective function and RHS coefficients, ignored by the SQP model.

Stochastic Minimal Time Vessel Weather Routing

A.N. Perakis
University of Michigan

The stochastic minimal time vessel weather routing problem is defined. A procedure calculating the isochrone line, summarizing the exact knowledge of the sea state patterns dominating the initial part of the routing space is described.

This line is the first stage in a DP network. The rest of the routing space is characterized by stochastic sea states, which are estimated at the beginning of the voyage, utilizing past observations, current forecasts and seasonal statistical propeties.

The DP formulation in a Markovian context, the associated recursive equation, and their relation to the deterministic problem previously analyzed and published by the authors is next presented.

Bounds on the optimal state evolution, resulting in a serious reduction of the dimensionality of the problem, are finally derived.

Stochastic Optimization by Stochastic Approximation: Computer Intensive Methods Combining Simulation and Optimization

Georg Pflug
University of Vienna

Consider the problem of finding the solution of the stochastic program

$$F(x) = \int H(x,\xi)d\mu_x(\xi) = min!$$

$$x \in S$$

Remark that in our problem formulation, not only the integrand H, but also the probability measure μ may depend on x. We discuss Stochastic - Approximation type algorithms of the general form

$$X_{n+1} = \pi_S(X_n = a_n Y_n)$$

where π_S is the projection onto S, a_n are stepsize constants and Y_n is a stochastic quasigradient, i.e.,

$$E(Y_n) = \nabla F(X_n) + o(1)$$

The first part of the lecture is concerned with convergence properties of the recursion $\{X_n\}$ (speed of convergence, asymptotic laws, large deviations, stopping times and stepsize choices). The second part deals with the problem of finding appropriate quasigradients for stationary distribution of stochastic systems. We review the Kiefer-Wolfowitz method, the IPA (infinitesimal perturbation analysis), the notion of weak derivatives and the scores method. Finally, implementation details, including random number techniques and techinques which are appropriate for parallel computer architectures are discussed.

Use of Chance-Constrained Programming to Account for Stochastic Variation in the A-Matrix of Large-Scale Mathematical Programs

James B. Pickens and John Hof/Brian Kent
Michigan Technological University/USDA Forest Service

Linear programming in widely used to select the manner in which forest lands are managed. This application has several unique characteristics. Because of the nature of forestry, the models consider many different management actions over many years, resulting in very large problems with diverse data. In addition, almost none of the data are known with certainty. The most pervasive occurrence of stochastic information is in the production coefficients, which indicate the uncertain response of the managed forest ecosystem to various management options. A "chance- constrained" approach to handling this uncertainty would often be appropriate in forestry applications–managers and decision makers would like to specify a probability with which uncertain constraints are met. Unfortunately, chance-constrained procedures for linear programming are currently available only for random right hand sides, and the random production coefficients encountered in forestry applications are almost invariably entered as A-matrix coefficients. This paper will utilize a Monte Carlo simulation approach (a linear program will be repeatedly solved with randomly perturbed A-matrix coefficients) to describe the distribution of total output when the individual production coefficients are random. And, an iterative procedure will be developed and demonstrated for "chance-constraining" feasibility with this sort of random A-matrix. An iterative approach is required because the mean and variance of total output is an unknown function of the random A-matrix coefficients. This paper will be oriented toward forestry linear programs, but may have applications in other fields as well.

Deterministic Approximations of Probability Inequalities

János Pintér
Institute for Transport Sciences, Budapest, Hungary

A simple general framework is presented for deriving various deterministic approximations of probability inequalities of the form $P\{\xi \geq a\} \leq \alpha$. These approximations are based on limited parametric information about the involved random variables (such as their mean, variance, range or upper bound values). as examples of possible applications, stochastic extensions of the "knapsak problem" and of the stochastic linear programming problem are investigated: we provide approximate deterministic surrogates for these problems.

An Inexpensive Basis Recovery Procedure for Karmarkar's Dual Affine Method

Kumaraswamy Ponnambalam and Anthony Vannelli
University of Waterloo

The basis recovery procedure is useful for the following reasons:
(i) accurate primal and dual solutions are available (an accurate primal solution is difficult to determine if approximate methods such as the pre-conditioned conjugate gradient method is used in Karmarkar's Dual Affine (DA) method

ABSTRACTS

SIXTH INTERNATIONAL CONFERENCE

on

STOCHASTIC PROGRAMMING

September 14 - 18, 1992
CISM, Udine
Italy

Asymptotic stochastic programs

Georg Ch. Pflug
Univ. of Vienna

Consider the stochastic program

$$(P) \quad \| \ \mathbb{E}(H(x,\xi)) + \psi_C(x) = \min!$$

where ψ_C is the indicator of the set C and suppose that it has a unique solution

$$x^* = \arg\min_{x \in C} \mathbb{E}(H(x,\xi)).$$

For practical solution, the program (P) is approximated by the "empirical program"

$$(P^n) \quad \| F_n(x) = \frac{1}{n}\sum_{i=1}^{n} H(x,\xi_i) + \psi_C(x) = \min!$$

where (ξ_i) is a sequence of i.i.d random variables with the same distribution as ξ. Any measurable selection

$$\hat{X}_n \in \arg\min_x F_n(x)$$

is called an "empirical" solution and the quality of $\hat{X}_n$ (which is a random variable) may be judged by looking at the asymptotic distribution of

$$T_n = \Gamma_n^{-1}(\hat{X}_n - x^*),$$

where Γ_n is a sequence of regular matrices with nonnegative entries which converge to zero. By a simple change of coordinates

$$T_n \in \arg\min_t Z_n(t),$$

where $Z_n(t)$ is the stochastic process

$$Z_n(t) = \rho_n \sum_{i=1}^{n} [H(x^* + \Gamma_n t, \xi) - H(x^*, \xi_i)] + \psi_C(x^* + \Gamma_n t).$$

with (ρ_n) being an appropriately chosen sequence of positive constants. Under general assumptions $Z_n(\cdot)$ epi-converges in distribution to a limiting process

$$Z(\cdot) = D(\cdot) + S(\cdot) + \psi_K(\cdot),$$

where $D(t)$ is a regularly varying deterministic function and $S(t)$ is an infinitly divisible mean zero stochastic process on $\mathbb{R}^m$. The stochastic program

$$(P^\infty) \quad \left\| \begin{array}{c} D(t) + S(t) = \min! \\ t \in K \subseteq \mathbb{R}^m \end{array} \right.$$

is called the *asymptotic stochastic progam* associated to (P).

The aim of the talk is to identify classes of stochastic processes $Z(\cdot)$ which may occur as limits. Under some regularity assumptions, the limiting process $S(\cdot)$ has one of the following structures

(i) $S(t) = t \cdot Y$ where $Y \sim N(0, \Sigma)$,

(ii) $S(t) = W(t)$, the m–dimensional Wiener process,

(iii) $S(t)$ is a Poisson–Hyperplane process

MATHEMATICS OF OPERATIONS RESEARCH

(Continued on inside back cover)

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 20, No. 4, November 1995
Printed in U.S.A.

ASYMPTOTIC STOCHASTIC PROGRAMS

GEORG CH. PFLUG

Consider a stochastic program with unique solution. By the notion of epi-convergence in distribution in local coordinates, we define the asymptotic stochastic program associated to it. Stochastic programs may be classified according to the type of the pertaining asymptotic program. In particular, three such types are studied in detail: the normal shift program, the Wiener-type program and the Poisson-hyperplane program. Conditions for the convergence to each of the three types are given.

1. Introduction. We consider the stochastic program

$$(1) \qquad (P) \quad \left\| \begin{array}{l} \text{Minimize } \mathbb{E}(H(x,\xi)), \\ x \in C \subseteq \mathbb{R}^d, \end{array} \right.$$

where ξ is a random variable and C is a closed convex set. Suppose that (P) has a unique solution

$$x^* = \arg\min_{x \in C} \mathbb{E}(H(x,\xi)).$$

By introducing the indicator function

$$\psi_C(x) = \begin{cases} 0 & \text{if } x \in C, \\ \infty & \text{if } x \notin C, \end{cases}$$

we may write

$$(2) \qquad x^* = \arg\min_x F(x),$$

where

$$F(x) = \mathbb{E}(H(x,\xi)) + \psi_C(x).$$

For a practical solution, the program (1) is approximated by the "empirical program"

$$(3) \qquad (P^n) \quad \| \text{Minimize } F_n(x) = \frac{1}{n}\sum_{i=1}^{n} H(x,\xi_i) + \psi_C(x),$$

where (ξ_i) is a sequence of i.i.d. random variables with the same distribution as ξ.

Received July 5, 1993; revised June 28, 1994.
AMS 1991 subject classification. Primary 90C15. Secondary 62E20.
OR/MS Index 1978 subject classification. Primary: 663 Programming/Stochastic.
Key words. Stochastic programs, epi-convergence, asymptotic distribution.

0364-765X/95/2004/0769/$01.25

We consider the stochastic program

$$\min\left\{\int_{\Xi} f(x,\xi)P(d\xi) : x \in X\right\},$$

where f is a normal integrand on $\mathbb{R}^m \times \mathbb{R}^d$, X a closed subset of $\mathbb{R}^m$ and P a probability distribution on $\mathbb{R}^d$.

The empirical approximation of the stochastic program is

$$\min\left\{\frac{1}{n}\sum_{i=1}^{n} f(x,\xi^i) + \delta_X(x) : x \in \mathbb{R}^m\right\} \quad (n \in \mathbb{N}),$$

where δ_X denotes the indicator function of X and ξ^i, $i \in \mathbb{N}$ are i.i.d. random variables in $\mathbb{R}^d$ with common distribution P.

Assume that the original SP has a unique solution $x^\star \in X$ and let (ρ_n) be a positive sequence and (Γ_n) be a sequence of regular $m \times m$ matrices converging to zero. Consider the stochastic process

$$Z_n(t) = \rho_n \sum_{i=1}^{n} [f(x^\star + \Gamma_n t, \xi^i) - f(x^\star, \xi^i)] + \delta_X(x^\star + \Gamma_n t)) \quad (t \in \mathbb{R}^m).$$

Pflug 95 derives conditions under which the sequence (Z_n) epi-converges in distribution to some stochastic process Z which is explicitly characterized.

הטכניון – מכון טכנולוגי לישראל
Technion – Israel Institute of Technology

Faculty of Industrial Engineering and Management
Technion-City, Haifa 32000, Israel

הפקולטה להנדסת תעשיה וניהול
קרית-הטכניון, חיפה 32000

7th INTERNATIONAL CONFERENCE ON STOCHASTIC PROGRAMMING

June 26–29, 1995
Nahariya, ISRAEL

ABSTRACTS

STOCHASTIC PROGRAMS AND EMPIRICAL PROCESS THEORY

G. Ch. Pflug, Vienna

If μ is some probability measure and (X_i) is a sequence of independent random variables distributed according to μ, then

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$$

is the emprical measure (δ_x denotes the point mass at point x). Notice that $(\hat{\mu}_n)$ is a sequence of measure-valued random variables. Under mild regularity conditions, $\hat{\mu}_n$ converges almost surely weakly to μ.

This general fact is often used in stochastic programming, where the original program containing the measure μ is approximated by the emprirical program, which contains the emipirical measure $\hat{\mu}_n$. The empirical program is easier to solve and the quality of this approximation is an important issue.

Probabilists have developed a variety of fundamental results, which deal with the emprical process and its analytic properties. One group of results concern uniformity properties: What are the conditions for sets of functions $\mathcal{F}$ or families of sets $\mathcal{C}$ to guarantee that

$$\sup_{f\in\mathcal{F}} |\int f\, d\hat{\mu}_n - \int f\, d\mu| \longrightarrow 0$$

or

$$\sup_{C\in\mathcal{C}} |\hat{\mu}_n(C) - \mu(C)| \longrightarrow 0$$

as $n \longrightarrow \infty$?

Other results concern the asymptotic normality of

$$\sqrt{n}\left[\int f\, d\hat{\mu}_n - \int f\, d\mu\right].$$

What is the speed of convergence? Here, Banach-space-valued stochastic processes come into play. Also, strong approximations and the KMT-construction are relevant.

The laws of iterated logarithm deal with the pointwise asymptotic behavior of

$$\frac{\sqrt{n}}{\sqrt{2\log\log n}}\left[\int f\, d\hat{\mu}_n - \int f\, d\mu\right].$$

Finally, the large deviations results give informations about

$$\mathbb{P}\{|\int f\, d\hat{\mu}_n - \int f\, d\mu| \geq \epsilon\}.$$

All these results can be applied to yield interesting results about the approximation quality of emprirical stochastic programs. Some results will be presented, some not yet investigated areas will be pointed out.

OPTIMIZATION OF STOCHASTIC MODELS

The Interface Between Simulation and Optimization

Georg Ch. Pflug

Kluwer Academic Publishers

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 23, No. 1, February 1998
Printed in U.S.A.

ON THE GLIVENKO-CANTELLI PROBLEM IN STOCHASTIC PROGRAMMING: LINEAR RECOURSE AND EXTENSIONS

GEORG CH. PFLUG, ANDRZEJ RUSZCZYŃSKI AND RÜDIGER SCHULTZ

Integrals of optimal values of random optimization problems depending on a finite dimensional parameter are approximated by using empirical distributions instead of the original measure. Under fairly broad conditions, it is proved that uniform convergence of empirical approximations of the right hand sides of the constraints implies uniform convergence of the optimal values in the linear and convex case.

1. Introduction. Real-world decision problems are usually associated with high uncertainty due to unavailability or inaccuracy of some data, forecasting errors, changing environment, etc. There are many ways to deal with uncertainty; one that proved successful in practice is to describe uncertain quantities by random variables.

Using the probabilistic description of uncertainty within optimization problems leads to *stochastic programming models*. There is a large variety of such models, depending on the nature of information about the random quantities and on the form of objective and constraints. One of the most popular models, which found numerous applications in operations research practice, is the *two-stage problem*. In its simplest linear form, it can be formulated as follows:

$$\min_{x\in X}\left[c^T x + \int f(x,\omega)P(d\omega)\right], \tag{1.1}$$

where $X \subset \mathbb{R}^{n_x}$ is the first stage feasible set and $f: \mathbb{R}^{n_x} \times \Omega \mapsto \mathbb{R}$ denotes the *recourse function* dependent on x and on an elementary event in some probability space (Ω, Σ, P). The recourse function is defined as the optimal value of the *second stage problem*

$$f(x,\omega) = \min\{q(\omega)^T y \mid W(\omega)y = b(x,\omega),\ y \geq 0\}. \tag{1.2}$$

Here, the vector $y \in \mathbb{R}^{n_y}$ is the second stage decision (which may, in general, depend on x and ω), $q(\omega)$ is a random vector in $\mathbb{R}^{n_y}$, $W(\omega)$ is a random matrix of dimension $m_y \times n_y$ and $b: \mathbb{R}^{n_x} \times \Omega \mapsto \mathbb{R}^{m_y}$ is a measurable function.

There is a vast literature devoted to properties of the two-stage problem (1.1)–(1.2) and to solution methods (see Ermoliev/Wets (1988) or Kall/Wallace (1994) and the references therein). It is usually assumed that W is a deterministic matrix and

$$b(x,\omega) = h(\omega) - T(\omega)x. \tag{1.3}$$

For example, $h(\omega)$ may be interpreted as a random demand/supply and $T(\omega)$ as a certain "technology matrix" associated with the first stage decisions. Then $b(x,\omega)$ is the dis-

Received April 30, 1995; revised March 4, 1996, and May 13, 1997.
AMS 1991 subject classification. Primary: 90C15.
OR/MS subject classification: Primary: Programming/stochastic.
Key words: Stochastic programming, empirical measures, uniform convergence.

0364-765X/98/2301/0204/$05.00

crepancy between the technology input/output requirements and the demand/supply observed, and some corrective action y has to be undertaken to account for this discrepancy.

However, in some long-term planning problems in a highly uncertain environment, it is the data referring to the future that are random. For example, in long-term investment planning, where x denotes the investment decisions to be made now, while y represents future actions, the costs q and the technological characteristics W of the future investments are usually uncertain. Moreover, new technologies may appear that may increase our recourse capabilities. Therefore we focus on the *random recourse* case in a generalized sense, i.e. a situation when besides W and q also the number of columns of W is random.

Next, our model allows much more general relations between the first stage variables and the second stage problem than the linear relation (1.3). In (1.2) we allow, for example, nonlinear and random technologies $T(x,\omega)$; moreover, the supply/demand vector may be dependent on both x and ω. Apart from a broader class of potential applications, such a model appears to be interesting in its own right. In §6, we shall show how to apply results for (1.2) to some more general convex problems.

The fundamental question that will be analysed in this paper is the problem of approximation. Namely, given an i.i.d. sample $s = \{s_i\}_{i=1}^{\infty} \in \Omega^\infty = \Omega^{\mathbb{N}}$, we consider for $n \in \mathbb{N}$ the *empirical measures*

$$P_n(s) = \frac{1}{n}\sum_{i=1}^{n} \delta_{s_i}, \tag{1.4}$$

where δ_{s_i} denotes point mass at s_i. An empirical measure can be employed to approximate the expected recourse function

$$F(x) = \int f(x,\omega)P(d\omega) \tag{1.5}$$

by the empirical mean

$$F_n(x) = \int f(x,\omega)P_n(s)(d\omega) = \frac{1}{n}\sum_{i=1}^{n} f(x, s_i). \tag{1.6}$$

The main question is the following: *can uniform convergence of F_n to F take place for almost all s (with respect to the product probability P^∞ on Ω^∞)?* We shall show that a positive answer to this question can be given for a very broad class of functions $b(x,\omega)$ in (1.2). To this end we shall use some results on the Glivenko-Cantelli problem developed in Gine/Zinn (1984), Talagrand (1987), Vapnik/Cervonenkis (1981).

Compared with related contributions to the stability of two-stage stochastic programs, the scope of the present paper is novel in two respects: we allow recourse matrices with random entries and random size, and we are able to treat discontinuous and non-convex integrands in the expected recourse function. The tools from probability theory that we use here lead to uniform convergence. The approaches in Dupacova/Wets (1988), Kall (1987), Robinson/Wets (1987) utilize milder types of convergence (such as epigraphical convergence), and hence they can handle extended-real-valued functions. As in the present paper, the accent in King/Wets (1991) is on convergence of expected recourse functions in the context of empirical measures. The authors obtain consisitency results that cover convex stochastic programs with a fixed recourse matrix W. Perturbations going beyond empirical measures are studied in Kall (1987), Robinson/Wets (1987) for fixed-recourse problems with continuous integrands. Further related work is contained in Vogel (1992, 1994), where random approximations to random optimization problems are con-

Stochastic programs and statistical data

Georg Ch. Pflug
Department of Statistics and OR, University of Vienna, Universitätsstrasse 5, A-1010 Vienna, Austria
E-mail: pflug@smc.univie.ac.at

Using statistical data instead of true underlying distributions in a stochastic optimization problem leads to an approximation error. We discuss how bounds for this error can be derived from results on uniformity in the law of large numbers.

Keywords: risk functionals, confidence regions, metric entropy

AMS subject classification: 90C15, 62G15

1. Risk functionals

A stochastic optimization problem is characterized by the fact that the costs associated with the decision x are uncertain in the sense that they depend on a random variable (or random vector) ξ. We denote these costs by $f(x, \xi)$. The function f is known to the decision maker. The distribution of ξ is typically unknown, but a sample $\xi_1, \ldots, \xi_n$ is available, which follows this distribution. Therefore, a statistical estimation problem is associated with the optimization problem.

The cost variable $Y_x = f(x, \xi)$ is a real-valued random variable. Let G_x be its distribution function

$$G_x(u) = P\{Y_x \leq u\}.$$

For decision making, preference relations for cost distributions must be defined: Let Ψ be a set of measurable functions and define

$$G_{x_1} \overset{\prec}{\Psi} G_{x_2} \text{ iff } \int \psi(u) dG_{x_1}(u) \leq \int \psi(u) dG_{x_2}(u) \text{ for all } \psi \in \Psi.$$

Examples are:

- Stochastic dominance of order 1 (monotonic dominance).
  Here, Ψ is the set of all bounded, monotonic functions.

Let $v(P)$ and $S(P)$ denote the optimal value and solution set of the original stochastic program. It holds

$$|v(P) - v(Q)| \leq \sup_{x \in X} \left| \int_\Xi f(x,\xi)P(d\xi) - \int_\Xi f(x,\xi)Q(d\xi) \right|$$

$$\emptyset \neq S(Q) \subseteq S(P) + \Psi_P^{-1}\left(\sup_{x \in X} \left| \int_\Xi f(x,\xi)P(d\xi) - \int_\Xi f(x,\xi)Q(d\xi) \right| \right)\mathbb{B},$$

where X is assumed to be compact, $\mathbb{B}$ is the unit ball in $\mathbb{R}^m$, Q is a probability distribution approximating P and Ψ_P is the growth function of the objective near the solution set, i.e.,

$$\Psi_P(t) := \inf \left\{ \int_\Xi f(x,\xi)P(d\xi) - v(P) : x \in X, d(x, S(P)) \leq t \right\}.$$

Hence, the uniform distance $d_{\mathcal{F}}$ with $\mathcal{F} := \{f(x,\cdot) : x \in X\}$ becomes important

$$d_{\mathcal{F}}(P,Q) := \sup_{f \in \mathcal{F}} \left| \int_\Xi f(\xi)P(d\xi) - \int_\Xi f(\xi)Q(d\xi) \right|$$

when studying approximations of the original stochastic program.

With P_n denoting the empirical measure to P, it is, hence, of interest whether the function class $\mathcal{F}$ is a P-Glivenko-Cantelli class, i.e.,

$$\lim_{n\to\infty} d_{\mathcal{F}}(P, P_n) = 0 \quad \text{a.s.}$$

Pflug-Ruszczyński-Schultz 1998 provide sufficient conditions such that integrands from linear two-stage stochastic programming and its extensions are Glivenko-Cantelli-classes using tools from the work by Talagrand 1987.

The empirical process $\{n^{\frac{1}{2}}(P_n - P)f\}_{f\in\mathcal{F}}$ is called uniformly bounded in probability with tail $C_{\mathcal{F}}(\cdot)$ if the function $C_{\mathcal{F}}$ is decreasing on $(0, \infty)$ and the estimate

$$\mathbb{P}(n^{\frac{1}{2}} d_{\mathcal{F}}(P, P_n) \geq \varepsilon) \leq C_{\mathcal{F}}(\varepsilon)$$

holds for all $\varepsilon > 0$ and $n \in \mathbb{N}$. In his seminal work Talagrand 1994 proved that the tail can be chosen as

$$C_{\mathcal{F}}(\varepsilon) = p(\varepsilon) \exp\left(-2\varepsilon^2\right) \quad (\varepsilon > 0)$$

with a polynomial p if the class $\mathcal{F}$ is uniformly bounded and satisfies some metric entropy condition. Pflug 1999 showed that integrands from linear two-stage stochastic programming satisfy the metric entropy conditions.

Linear two-stage stochastic programs:

$$\min\left\{\langle c,x\rangle+\int_{\Xi}\Phi(q(\xi),h(\xi)-T(\xi)x)P(d\xi):x\in X\right\},$$

where $c\in\mathbb{R}^m$, Ξ and X are polyhedral subsets of $\mathbb{R}^d$ and $\mathbb{R}^m$, respectively, P is a probability measure on Ξ and the $s\times m$-matrix $T(\cdot)$, the vectors $q(\cdot)\in\mathbb{R}^{\overline{m}}$ and $h(\cdot)\in\mathbb{R}^s$ are affine functions of ξ.

The function Φ denotes the parametric infimum function of the linear second-stage program

$$\Phi(u,t)=\inf\{\langle u,y\rangle:Wy=t,y\in Y\},$$

which is finite and continuous on $\mathcal{D}\times W(Y)$, where $\mathcal{D}$ is the dual feasibility set

$$\mathcal{D}=\left\{u\in\mathbb{R}^{\overline{m}}:\left\{z\in\mathbb{R}^s:W^\top z-u\in Y^\star\right\}\neq\emptyset\right\},$$

where W is the $s\times\overline{m}$ recourse matrix, $W^\top$ the transposed of W and $Y^\star$ the polar cone to the polyhedral cone Y in $\mathbb{R}^{\overline{m}}$.

The function Φ is concave-convex polyhedral and finite and locally Lipschitz continuous with linearly growing local Lipschitz moduli on $\mathcal{D}\times W(Y)$.

A. Ruszczyński and A. Shapiro, Eds., *Handbooks in OR & MS, Vol. 10*

Chapter 7

Stochastic Optimization and Statistical Inference

G.Ch. Pflug

Department of Statistics and Decision Support Systems, University of Vienna, Universitätsstrasse 5, A-1090 Vienna, Austria

Abstract

If the distribution of the random parameters of a stochastic program is unknown, the empirical distribution based on a sample may be used as a proxy. This empirical approximation is related to the "true" stochastic program in the same way as a statistical estimate is related to the true parameter value. Properties of statistical estimators, like consistency, asymptotical distributions and the construction of confidence regions are reviewed in the realm of stochastic optimization. The entropic size of a stochastic program determines the quality of the approximation. In case that random constraints are present, the notion of epiconvergence replaces in a natural way the notion of uniform convergence of functions. The asymptotic structures are described by the asymptotic stochastic program associated to the sequence of empirical programs.

Key words: Empirical program, statistical estimates, asymptotic statistics, risk functionals, entropic size, epiconvergence, asymptotic stochastic programs.

1 Uncertain and ambiguous optimization problems

In deterministic optimization, a decision x must be found, which minimizes a known cost function $f(x)$ among all possible candidates x lying in the *feasible set* $\mathcal{X} \subseteq \mathbb{R}^d$, a closed subset of the euclidean d-dimensional space

$$\operatorname*{Min}_{x \in \mathcal{X}} f(x).$$

Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk

Georg Ch. Pflug (georg.pflug@univie.ac.at)
Department of Statistics and Decision Support Systems
University of Vienna

Abstract

The value-at-risk (VaR) and the conditional value-at-risk (CVaR) are two commonly used risk measures. We state some of their properties and make a comparison. Moreover, the structure of the portfolio optimization problem using the VaR and CVaR objective is studied.

Keywords: Risk measures, Value-at-Risk, Conditional Value-at-Risk, Portfolio optimization

1 Introduction

Let Y be a random cost variable and let F_Y be its distribution function, i.e. $F_Y(u) = \mathbb{P}\{Y \le u\}$. Let $F_Y^{-1}(v)$ be its right continuous inverse, i.e. $F_Y^{-1}(v) = \inf\{u : F_Y(u)) > v\}$. When no confusion may occur, we write simply F instead of F_Y.

For a fixed level α, we define (as usual) the *value-at-risk* VaR_α as the α-quantile, i.e.

$$\mathrm{VaR}_\alpha(Y) = F^{-1}(\alpha). \tag{1}$$

The *conditional value-at-risk* $CVaR_\alpha$ is defined as the solution of an optimization problem

$$\mathrm{CVaR}_\alpha(Y) := \inf\{a + \frac{1}{1-\alpha}\mathbb{E}[Y-a]^+ : a \in \mathbb{R}\}. \tag{2}$$

Here $[z]^+ = \max(z, 0)$. Uryasev and Rockafellar (1999) have shown that CVaR equals the conditional expectation of Y, given that $Y \ge \mathrm{VaR}_\alpha$, i.e.

$$\mathrm{CVaR}_\alpha(Y) \triangleq \mathbb{E}(Y|Y \ge \mathrm{VaR}_\alpha(Y)). \tag{3}$$

S.P. Uryasev (ed.), Probabilistic Constrained Optimization, 272–281.

In fact, (3) is the usual definition of CVaR_α.

We will prove some properties of CVaR and VaR and study the relation between these two measures of risk. To begin with, we show that the minimizer in (2) is VaR_α, even if F is not differentiable.

Proposition 1. Suppose that $F(b) \ge \alpha$ and $F(b-) \le \alpha$. Then

$$b + \frac{1}{1-\alpha}\mathbb{E}[Y-b]^+ \le a + \frac{1}{1-\alpha}\mathbb{E}[Y-a]^+$$

for all a.

Proof. Suppose first that $b \le a$. Then

$$\begin{aligned}\mathbb{E}[Y\mathbb{1}_{\{b<Y\}}] - \mathbb{E}[Y\mathbb{1}_{\{a<Y\}}] &= \mathbb{E}[Y\mathbb{1}_{\{b<Y\le a\}}]\\ &\le a[F(a)-F(b)] \le a[F(a)-\alpha] - b[F(b)-\alpha].\end{aligned}$$

Therefore

$$\begin{aligned}b[1-\alpha] - b[1-F(b)] + \mathbb{E}[Y\mathbb{1}_{\{b<Y\}}] &\le a[1-\alpha] - a[1-F(a)] + \mathbb{E}[Y\mathbb{1}_{\{a<Y\}}]\\ b[1-\alpha] + \mathbb{E}[(Y-b)\mathbb{1}_{\{b<Y\}}] &\le a[1-\alpha] + \mathbb{E}[(Y-a)\mathbb{1}_{\{a<Y\}}]\\ b + \frac{1}{1-\alpha}\mathbb{E}[(Y-b)\mathbb{1}_{\{b<Y\}}] &\le a + \frac{1}{1-\alpha}\mathbb{E}[(Y-a)\mathbb{1}_{\{a<Y\}}].\end{aligned}$$

Let now $a \le b$. Then

$$\begin{aligned}\mathbb{E}[Y\mathbb{1}_{\{a<Y\}}] - \mathbb{E}[Y\mathbb{1}_{\{b\le Y\}}] &= \mathbb{E}[Y\mathbb{1}_{\{a<Y<b\}}]\\ &\ge a[F(b-)-F(a)] \ge b[F(b-)-\alpha] - a[F(a)-\alpha].\end{aligned}$$

Therefore

$$\begin{aligned}a[1-\alpha] - a[1-F(a)] + \mathbb{E}[Y\mathbb{1}_{\{a<Y\}}] &\ge b[1-\alpha] - b[1-F(b-)] + \mathbb{E}[Y\mathbb{1}_{\{b\le Y\}}]\\ a[1-\alpha] + \mathbb{E}[(Y-a)\mathbb{1}_{\{a<Y\}}] &\ge b[1-\alpha] + \mathbb{E}[(Y-b)\mathbb{1}_{\{b\le Y\}}] - b[1-F(b-)]\\ b + \frac{1}{1-\alpha}\mathbb{E}[(Y-b)\mathbb{1}_{\{b<Y\}}] &\le a + \frac{1}{1-\alpha}\mathbb{E}[(Y-a)\mathbb{1}_{\{a<Y\}}].\end{aligned}$$

□

As a consequence, one sees that

$$\mathrm{VaR}_\alpha(Y) = F^{-1}(\alpha) \in \operatorname{argmin}\ \{a + \frac{1}{1-\alpha}\mathbb{E}[Y-a]^+\}.$$

Alternative, equivalent representations of CVaR are therefore

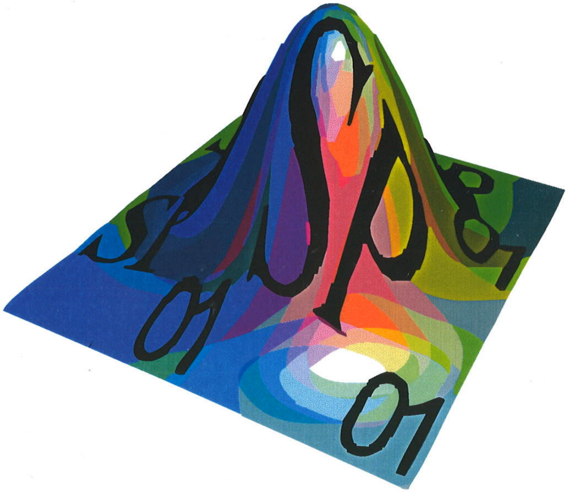
$$\begin{aligned}\mathrm{CVaR}_\alpha(Y) &= \mathbb{E}[Y|Y \ge F^{-1}(\alpha)]\\ &= \frac{1}{1-\alpha}\int_\alpha^1 F^{-1}(v)\,dv\\ &= \frac{1}{1-\alpha}\int_{F^{-1}(\alpha)}^\infty u\,dF(u).\end{aligned}$$

9th International Conference on

Stochastic Programming

Berlin, Germany

August 25-31, 2001



Program and Abstracts

Risk measures, convexity and the associated risk processes

Georg Pflug
Institut für Statistik und Decision Support Systems, Universität Wien, Universitätsstrasse 5, A-1090 Wien
e-mail: georg.pflug@univie.ac.at

Abstract

Risk measures map distribution functions of random variables to the real line. We introduce some concepts of monotonicity and convexity for such measures. Three notions of monotonicity (first order stochastic dominance, second order stochastic dominance and concave dominance) are discussed as well as three notions of convexity (compound convexity, convolution convexity and comonotone convexity).

The second half of the talk deals with tree processes and optimization problems defined on these processes. The optimal solution process and the clairvoyant process are identified. Their properties depend on the convexity structures of the considered risk measure. Convolution convexity translates into convexity of the associated multistage stochastic program whereas compound convexity translates into supermartingale structures for the optimal value process and the clairvoyant process. This generalizes earlier work of Dempster on EVPI (expected value of perfect information)- processes.

The Tenth International Conference on Stochastic Programming

Tucson, Arizona, USA ◆ October 9 - 15, 2004

Sponsoring Society: The Committee on Stochastic Programming, The Mathematical Programming Society

SP X 2004

Organizers:
Suvrajeet Sen (U. Arizona), General Chair
Julia L. Higle (U. Arizona), Program Chair
David P. Morton (U. Texas), Tutorials Chair
Kevin Wood (NPS), Multi-Disciplinary Outreach Chair

Hosted By:
The MORE Institute, SIE Department, THE UNIVERSITY OF ARIZONA®

often all natural candidates for a true model are infinite-dimensional optimization problems involving mathematical concepts that practitioners are not always familiar with. Indeed, in many problems, there are essential (not just technical) features that can be captured only in terms of infinite-dimensional spaces. Such general stochastic programming models were well-developed already in the 1970's, but it seems that, as stochastic programming gained popularity among practitioners, they were partly forgotten. Nowadays, stochastic programming models of real-life problems are often formulated in terms of scenario trees constructed in an ad-hoc manner. This has resulted in vague formulations of stochastic programs that lack interpretation.

Our aim is to describe an analytical version of the stochastic programming approach for practical decision making. In our approach, both the modeling and solution phases are broken down into two sub-phases:

1.1 Modeling the decision problem as an optimization problem,

1.2 Modeling the uncertainty,

2.1 Discretization of the optimization problem,

2.2 Numerical solution of the discretized problem.

The first step consists of modeling the decision problem as a stochastic optimization problem over a general probability space. The second step consists of specifying the probability distribution of the uncertain data. The purpose of the third step is to construct finite-dimensional, numerically solvable, consistent approximations of the optimization model specified in the first two steps. The discretized model is then solved in the fourth step using appropriate techniques for stochastic programs over finite scenario trees.

This kind of approaches to problem solving are familiar from other fields of applied mathematics such as ordinary or partial differential equations. Indeed, there also one models real phenomena by infinite-dimensional models, after which solutions are sought through discretization and numerical computation. Our approach has several advantages. First, it facilitates the solution process by decomposing it into more easily manageable pieces. Second, having a well-defined model allows for rigorous analysis of the problem and solution techniques. Third, it allows one to use well-developed models from various fields of stochastics where stochastic processes are not restricted to finite scenario trees. Fourth, this approach relates closely with other disciplines, making stochastic programming more attractive to a wider range of researchers and practitioners.
(ThE, Grand Ballroom)

Pflug, Georg. University of Vienna
Risk measures as solutions of stochastic programs
Practically all risk measures proposed in the literature can be seen as solutions of linear stochastic programs. This means that they have primal and dual representations, are monotone w.r.t. information and exhibit convexity properties. Also monotonicity properties may often be easily deduced from these representations.

We review representations of one- and multi-period risk measures and, following Kusuoka (2000), investigate special representations of risk measures which depend only on the distribution of the involved stochastic processes.
(MF, South Ballroom)

Philpott, Andy. University of Auckland
On unit commitment in electricity pool markets
We consider an electricity generator making offers of energy into an electricity pool market. The generator runs a set of generating units with given start-up costs and operating ranges. For a given time period, it must submit to the pool system operator a supply function, typically consisting of a fixed number of quantities of energy and prices at which it wants these dispatched. The amount of dispatch depends on the stack offered along with the offers of the other generators and market demand, both of which are random, but are assumed have known market distribution functions. After dispatch the generator determines which

Computational Optimization and Applications, 32, 161–178, 2005
2005 Springer Science + Business Media, Inc. Manufactured in The Netherlands.

Measuring Risk for Income Streams

GEORG CH. PFLUG georg.pflug@univie.ac.at
Department of Statistics and Decision Support Systems, Universitaetsstrasse 5, University of Vienna, 1090 Wien-Vienna, Austria

ANDRZEJ RUSZCZYŃSKI rusz@rutcor.rutgers.edu
Department of Management Science and Information Systems and RUTCOR, Rutgers University, 94 Rockefeller Rd, Piscataway, NJ 08854, USA

Abstract. A measure of risk is introduced for a sequence of random incomes adapted to some filtration. This measure is formulated as the optimal net present value of a stream of adaptively planned commitments for consumption. The new measure is calculated by solving a stochastic dynamic linear optimization problem which, for finite filtrations, reduces to a deterministic linear programming problem.

We analyze properties of the new measure by exploiting the convexity and duality structure of the stochastic dynamic linear problem. The measure depends on the full distribution of the income process (not only on its marginal distributions) as well as on the filtration, which is interpreted as the available information about the future. The features of the new approach are illustrated by a numerical example.

Keywords: dynamic risk measure, multistage stochastic programming, multiperiod mean-risk models, value of perfect information, conditional value at risk

1. Motivation

Since the seminal work of Markowitz it is well understood that consequences of economic activity with uncertain success must be judged in two different and well distinguished dimensions. The *mean* refers to the average result among a set of possible scenarios, while the *risk* dimension describes the possible variation of the results under varying scenarios. In the Markowitz model the risk is measured by the variance of the outcome (cf. [11, 12]). In the mean–risk setting the decision maker is faced with a two-objective situation: he/she wants to maximize the mean return and to minimize the risk at the same time. As for all multi-objective situations, there is in general no uniquely defined best decision, which is optimal in both dimensions and one has to seek for compromises. The set of solutions which are Pareto-efficient in the sense of these two objectives is called the mean–risk efficient frontier. In some models for optimal decision making the two dimensions are often mixed by introducing a nondecreasing concave utility function. Risk aversion, i.e. the degree of taking the risk dimension into account, can be modeled by the negative curvature of the utility function.

However, it is highly desirable to clearly separate the two dimensions and to make the compromising strategy as transparent as possible, and the efficient frontier approach provides such a transparency. In the first step, the efficient frontier is calculated for a given

Math. Program., Ser. B 108, 339–354 (2006)

Digital Object Identifier (DOI) 10.1007/s10107-006-0714-8

Georg Ch. Pflug

Subdifferential representations of risk measures

Received: August 23, 2004 / Accepted: October 5, 2005
Published online: June 2, 2006 – © Springer-Verlag 2006

Abstract. Measures of risk appear in two categories: Risk capital measures serve to determine the necessary amount of risk capital in order to avoid ruin if the outcomes of an economic activity are uncertain and their negative values may be interpreted as acceptability measures (safety measures). Pure risk measures (risk deviation measures) are natural generalizations of the standard deviation. While pure risk measures are typically convex, acceptability measures are typically concave. In both cases, the convexity (concavity) implies under mild conditions the existence of subgradients (supergradients). The present paper investigates the relation between the subgradient (supergradient) representation and the properties of the corresponding risk measures. In particular, we show how monotonicity properties are reflected by the subgradient representation. Once the subgradient (supergradient) representation has been established, it is extremely easy to derive these monotonicity properties. We give a list of Examples.

Key words. Risk measures – Duality – Stochastic dominance

1. Introduction

In recent years, starting from the seminal paper by Artzner et al. [1], axiomatic approaches to the definition of appropriate measures of risk for random variables and stochastic processes have been in the center of interest of many authors ([5, 8, 7, 9, 15, 17]). It is common sense, that convexity (concavity) plays a key role among the required properties for risk measures. A convex lower semicontinuous function is characterized by the fact that it is the dual of its own dual, hence completely characterized by its dual function. A concave function is characterized by the fact that its negative is convex.

Concavity (convexity) of risk functionals has been recently investigated by Ruszczyński and Shapiro [19]. They show the continuity and super(sub)-differentiability of risk functionals under mild conditions, i.e. the existence of dual representations. Moreover they investigate the dual structure of optimization problems involving super(sub)differentiable risk functionals. For positive homogeneous risk deviation measures (see below), Rockafellar et al. [18] have shown the existence dual representations and characterize the subgradient set, calling it the risk envelope.

In this paper, we show how for super(sub)differentiable risk functionals, the dual representation can be used to derive some properties (in particular monotonicity) of the risk functional in a very simple manner. Moreover we give many examples of dual representations of well known risk functionals.

G. Ch. Pflug: Department of Statistics and Decision Support Systems, Universitätsstrasse 5, University of Vienna, 1090 Wien-Vienna, Austria.
e-mail: georg.pflug@univie.ac.at

Statistics & Decisions 24, 45–60 (2006) / **DOI** 10.1524/stnd.2006.24.1.45

On distortion functionals

Georg Ch. Pflug

Received: November 11, 2005; Accepted: March 19, 2006

Summary: Distorted measures have been used in pricing of insurance contracts for a long time. This paper reviews properties of related acceptability functionals in risk management, called distortion functionals. These functionals may be characterized by being mixtures of average values-at-risk. We give a dual representation of these functionals and show how they may be used in portfolio optimization. An iterative numerical procedure for the solution of these portfolio problems is given which is based on duality.

1 Introduction: Distortion functionals as insurance premia

Let L be a random variable describing the (nonnegative) loss distribution of an insurance contract. Let G_L be the pertaining distribution function $G_L(u) = \mathbb{P}\{L \le u\}$. How much premium should the insurance company ask for coverage of L? Obviously, the premium should be greater than $\mathbb{E}[L]$ otherwise the insurance company will go bankrupt for sure.

Based on the well known formula

$$\mathbb{E}[L] = \int_0^\infty (1 - G_L(u))\, du$$

a safe insurance premium can be defined by

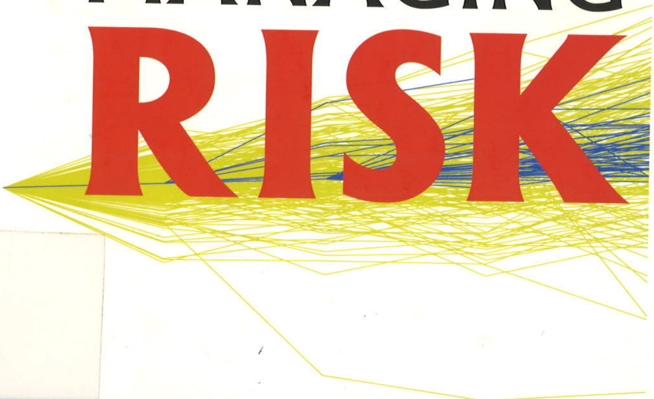
$$\pi_\psi[L] = \int_0^\infty \psi(1 - G_L(u))\, du, \tag{1.1}$$

where ψ is function, mapping $[0,1]$ to $[0,1]$, such that

$$\psi(p) \ge p \qquad \text{for } p \in [0,1]. \tag{1.2}$$

The condition (1.2) guarantees that the premium is not smaller than the expectation. However one usually considers more specific functions ψ.

MODELING, MEASURING AND MANAGING RISK



Georg Ch Pflug • Werner Römisch

Finance Stoch (2010) 14: 397–418
DOI 10.1007/s00780-009-0121-0

Asymptotic distribution of law-invariant risk functionals

Georg Pflug · Nancy Wozabal

Received: 16 December 2007 / Accepted: 11 December 2008 / Published online: 21 January 2010

Abstract Law-invariant or version-independent coherent risk or acceptability functionals do not explicitly depend on the underlying probability space and can be considered as functionals of the distribution function. In this paper, we consider estimates of these functionals based on the empirical distribution function and investigate their asymptotic properties.

Keywords Risk functionals · Law-invariance · Asymptotic normality · M-theorems

Mathematics Subject Classification (2000) 91B30 · 62E20 · 60F05 · 91B28

JEL Classification D81 · G32

1 Introduction

In this paper, we consider the asymptotic properties of coherent version-independent risk functionals. Such functionals have wide ranging applications in the financial industry, e.g. in portfolio optimization, asset pricing, capital allocation problems, performancé analysis and evaluation (see [2, 6, 7, 13, 18, 27]).

The property of version independence, also known as law-invariance, states that the risk functional assigns the same value to random variables following the same distribution, i.e., for random variables X and Y following the same distribution F, $\mathcal{A}(X) = \mathcal{A}(Y)$. In this case since the risk measure does not explicitly depend on the

G. Pflug · N. Wozabal (✉)
Department of Statistics and Decision Support Systems, University of Vienna, Universitaetsstrasse 5, 1010 Vienna, Austria
e-mail: nancy.wozabal@univie.ac.at

G. Pflug
e-mail: georg.pflug@univie.ac.at

International Journal of Theoretical and Applied Finance
Vol. 17, No. 1 (2014) 1450003 (33 pages)

DOI: 10.1142/S0219024914500034

ARE TIME CONSISTENT VALUATIONS INFORMATION MONOTONE?

RAIMUND M. KOVACEVIC*
Department of Statistics and Operations Research (ISOR)
University of Vienna, Universitaetsstrasse 5
A-1010 Wien-Vienna, Austria
raimund.kovacevic@univie.ac.at

GEORG CH PFLUG
ISOR and IIASA
Laxenburg, Austria
georg.pflug@univie.ac.at

Received 7 February 2012
Accepted 10 February 2013
Published 30 January 2014

Multi-period risk functionals assign a risk value to discrete-time stochastic processes. While convexity and monotonicity extend in straightforward manner from the single-period case, the role of information is more problematic in the multi-period situation. In this paper, we define multi-period functionals in such a way that the development of available information over time (expressed as a filtration) enters explicitly the definition of the functional. This allows to define and study the property of information monotonicity, i.e. monotonicity w.r.t. increasing filtrations. On the other hand, time consistency of valuations is a favorable property and it is well-known that this requirement essentially leads to compositions of conditional mappings. We demonstrate that generally spoken the intersection of time consistent and information monotone valuation functionals is rather sparse, although both classes alone are quite rich. In particular, the paper gives a necessary and sufficient condition for information monotonicity of additive compositions of positively homogeneous risk/acceptability mappings. Within the class of distortion functionals only compositions of expectation or essential infima are information monotone. Furthermore, we give a sufficient condition and examples for compositions of nonhomogeneous mappings exhibiting information monotonicity.

Keywords: Risk functional; acceptability functional; multi-period, conditional risk mapping; average value-at-risk; dual representation; information monotonicity; value of information.

AMS 2000 Classification: 91G99, 91B06

JEL Classification: G32

*This author was funded by FWF.

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_T$ of σ-fields and let $\mathfrak{F} = (\mathcal{F}_1, \ldots, \mathcal{F}_T)$ represent the observable relevant information. Let $\rho(\cdot, \mathfrak{F})$ denote a multi-period risk functional defined on some normed space $\mathcal{Y} = \times_{t=1}^{T} \mathcal{Y}_t$ of random vectors Y, where Y_t is $\mathcal{F}_t$-measurable, with values in the extended reals $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$, i.e., $\rho(\cdot, \mathfrak{F})$ is convex on $\mathcal{Y}$ and satisfies the monotonicity property $\rho(Y, \mathfrak{F}) \leq \rho(X, \mathfrak{F})$ if $X_t \leq Y_t$, $\mathbb{P}$-a.s., $t = 1, \ldots, T$.

The multi-period risk functional $\rho(\cdot, \cdot)$ is called information monotone if

$$\rho(Y, \mathfrak{F}') \leq \rho(Y, \mathfrak{F}) \quad \text{holds for all} \quad Y \in \mathcal{Y} \quad \text{if} \quad \mathcal{F}_t \subseteq \mathcal{F}'_t, t = 1, \ldots, T,$$

where $\mathfrak{F} = (\mathcal{F}_1, \ldots, \mathcal{F}_T)$ and $\mathfrak{F}' = (\mathcal{F}'_1, \ldots, \mathcal{F}'_T)$.

Now, consider a sequence of (risk) mappings $\rho^{(t)}(\,\cdot\,, \mathfrak{F}^{(t)})$ from $\times_{i=t+1}^{T} \mathcal{Y}_i$ to $\mathcal{Y}_t$ for $t = 0, \ldots, T-1$, where $\mathfrak{F}^{(t)} = (\mathcal{F}_t, \ldots, \mathcal{F}_T)$ and $Y^{(t)} = (Y_t, \ldots, Y_T)$. Such a sequence is called time consistent if

$$\begin{aligned} \rho^{(t)}(X^{(t+1)}, \mathfrak{F}^{(t)}) &\leq \rho^{(t)}(Y^{(t+1)}, \mathfrak{F}^{(t)}) \quad \text{and} \quad X_t \geq Y_t \quad \text{implies} \\ \rho^{(t-1)}(X^{(t)}, \mathfrak{F}^{(t-1)}) &\leq \rho^{(t-1)}(Y^{(t)}, \mathfrak{F}^{(t-1)}) \end{aligned}$$

for all $X, Y \in \mathcal{Y}$ and $t = 1, \ldots, T-1$ (Kovacevic-Pflug 14).

Multi-period risk functionals $\rho(\cdot, \mathfrak{F})$ are typically constructed by composing (risk) mappings in a suitable way.

Examples:

(a) SEC risk functionals: $\rho(Y, \mathfrak{F}) = \sum_{t=0}^{T-1} \mathbb{E}[\rho_t(Y_{t+1}|\mathcal{F}_t)]$, where the $\rho_t(\cdot|\mathcal{F}_t) : \mathcal{Y}_{t+1} \to \mathcal{Y}_t$ are conditional risk mappings. Such functionals are composed by a sequence of time consistent risk mappings and is information monotone if each $\rho_t(\cdot|\mathcal{F}_t)$ is information monotone.

(b) Additive conditional risk functional compositions:

$$\rho^{(t)}(Y^{(t+1)}, \mathcal{F}^{(t)}) \; = \; \rho_t(\cdot|\mathcal{F}_t) \circ \cdots \circ \rho_{T-1}\Big(\sum_{i=t+1}^{T} Y_i \Big| \mathcal{F}_{T-1}\Big) \; (t = 0, \ldots, T-1)$$

$$\rho(Y, \mathfrak{F}) \; = \; \rho^{(0)}(Y^{(1)}, \mathfrak{F}^{(0)}).$$

Additive risk functional compositions are time consistent, but lead to information monotone multi-period risk measures only in a few cases.

(c) Dynamic programming recursions are time consistent and may be used to obtain information monotone multi-period risk measures.

Math. Program., Ser. B 89: 251–271 (2001)

Digital Object Identifier (DOI) 10.1007/s101070000202

G.Ch. Pflug

Scenario tree generation for multiperiod financial optimization by optimal discretization*

Received: December 15, 1998 / Accepted: October 1, 2000
Published online December 15, 2000 –

Abstract. Multiperiod financial optimization is usually based on a stochastic model for the possible market situations. There is a rich literature about modeling and estimation of continuous-state financial processes, but little attention has been paid how to approximate such a process by a discrete-state scenario process and how to measure the pertaining approximation error.

In this paper we show how a scenario tree may be constructed in an optimal manner on the basis of a simulation model of the underlying financial process by using a stochastic approximation technique. Consistency relations for the tree may also be taken into account.

1. Introduction – the approximation problem

A (continuous-state) multistage financial optimization problem with decision periods $1, 2, \ldots T$ is based on

- a stochastic model of the future development of the economic environment (prices, interests, cash-flows, etc.). This scenario process is expressed as a (possibly vector-valued) stochastic process $\xi_1, \xi_2, \ldots, \xi_T$;
- a decision model for the actions to be taken. The decisions at time stage t, which may depend on the past observations $\xi_1, \ldots, \xi_{t-1}$ are $x_1, x_2(\xi_1), x_2(\xi_1, \xi_2)$, $\ldots x_T(\xi_1, \ldots, \xi_{T-1})$;
- the objective function, which expresses the long-term goals of the decision maker.

Except for extremely simple and unrealistic cases, continuous-state multiperiod financial optimization problems can only be formulated, but not solved.

The reason for practical unsolvability is the fact that the decisions are functions, making the problem a functional optimization problem, which cannot be numerically solved as it is.

The usual way of reducing the problem to a solvable one is to restrict to discrete-state multiperiod financial optimization problems, i.e. these cases in which the random vector $\xi_1, \ldots, \xi_T$ takes only finitely many values. In this case, the decision functions reduce to large decision vectors.

G.Ch. Pflug: Department of Statistics and Decision Support Systems, University of Vienna, Universitaets-strasse 5, A-1090 Vienna, Austria

Mathematics Subject Classification (1991): 90C15

* The work described in this paper was supported by the Special Research Program SFB F011 "AURORA" of the Austrian Science Fund

Ann Oper Res (2007) 152:257–272
DOI 10.1007/s10479-006-0140-6

Financial scenario generation for stochastic multi-stage decision processes as facility location problems

Ronald Hochreiter · Georg Ch. Pflug

Published online: 10 November 2006

Abstract The quality of multi-stage stochastic optimization models as they appear in asset liability management, energy planning, transportation, supply chain management, and other applications depends heavily on the quality of the underlying scenario model, describing the uncertain processes influencing the profit/cost function, such as asset prices and liabilities, the energy demand process, demand for transportation, and the like. A common approach to generate scenarios is based on estimating an unknown distribution and matching its moments with moments of a discrete scenario model. This paper demonstrates that the problem of finding valuable scenario approximations can be viewed as the problem of optimally approximating a given distribution with some distance function. We show that for Lipschitz continuous cost/profit functions it is best to employ the Wasserstein distance. The resulting optimization problem can be viewed as a multi-dimensional facility location problem, for which at least good heuristic algorithms exist. For multi-stage problems, a scenario tree is constructed as a nested facility location problem. Numerical convergence results for financial mean-risk portfolio selection conclude the paper.

Keywords Stochastic programming · Multi-stage financial scenario generation

1 Introduction

A large class of decision problems involves decision stages and uncertainty. Examples are multi-stage portfolio optimization or asset liability management problems, energy production models, as well as models in telecommunication, transportation, supply chain management. For a recent overview see Ruszczynski and Shapiro (2003) and Wallace and Ziemba (2005). A common feature of these models is the fact that a stochastic process describing the uncertain

R. Hochreiter (✉) · G. Ch. Pflug
Department of Statistics and Decision Support Systems, University of Vienna, Universitätsstraße 5/9, 1010 Vienna, Austria
e-mail: ronald.hochreiter@univie.ac.at

G. Ch. Pflug
e-mail: georg.pflug@univie.ac.at

SIAM J. OPTIM.
Vol. 22, No. 1, pp. 1–23

A DISTANCE FOR MULTISTAGE STOCHASTIC OPTIMIZATION MODELS*

GEORG CH. PFLUG† AND ALOIS PICHLER‡

Abstract. We describe multistage stochastic programs in a purely in-distribution setting, i.e., without any reference to a concrete probability space. The concept is based on the notion of nested distributions, which encompass in one mathematical object the scenario values as well as the information structure under which decisions have to be made. The nested distance between these distributions is introduced and turns out to be a generalization of the Wasserstein distance for stochastic two-stage problems. We give characterizations of this distance and show its usefulness in examples. The main result states that the difference of the optimal values of two multistage stochastic programs, which are Lipschitz and differ only in the nested distribution of the stochastic parameters, can be bounded by the nested distance of these distributions. This theorem generalizes the well-known Kantorovich–Rubinstein theorem, which is applicable only in two-stage situations, to multistage. Moreover, a dual characterization for the nested distance is established. The setup is applicable both for general stochastic processes and for finite scenario trees. In particular, the nested distance between general processes and scenario trees is well defined and becomes the important tool for judging the quality of the scenario tree generation. Minimizing—at least heuristically—this distance is what good scenario tree generation is all about.

Key words. stochastic optimization, quantitative stability, transportation distance, scenario approximation

AMS subject classifications. 90C15, 90C31, 90C08

DOI. 10.1137/110825054

1. Introduction.

Multistage stochastic programming models have been successfully developed for the financial sector (banking [9], insurance [5], pension fund management [18]), the energy sector (electricity production and trading of electricity [15] and gas [1]), the transportation [6] and communication sectors [10], and airline revenue management [22], among others. In general, the observable data for a multistage stochastic optimization problem are modeled as a stochastic process $\xi = (\xi_0, \ldots, \xi_T)$ (the scenario process) and the decisions may depend on its observed values, making the problem an optimization problem in function spaces. The general problem is only in rare cases solvable in an analytic way and for numerical solution the stochastic process is replaced by a *finite valued* stochastic scenario process $\tilde{\xi} = (\tilde{\xi}_0, \ldots, \tilde{\xi}_T)$. By this discretization, the decisions become high dimensional vectors, i.e., are themselves discretizations of the general decision functions. An extension function is then needed to transform optimal solutions of the approximate problem to feasible solutions of the basic underlying problem.

There are several results about the approximation of the discretized problem to the original problem, for instance, [24, 19, 21, 12]. All these authors assume that both processes, the original ξ and the approximate $\tilde{\xi}$, are defined on the same probability space. This assumption is quite *un*natural, since the approximate processes are finite

*Received by the editors February 18, 2011; accepted for publication (in revised form) October 5, 2011; published electronically January 5, 2012.
http://www.siam.org/journals/siopt/22-1/82505.html
†Department of Statistics and Operations Research, University of Vienna, 1010 Vienna, Austria (georg.pflug@univie.ac.at); International Institute for Applied Systems Analysis, Laxenburg, Austria.
‡Corresponding author. Department of Statistics and Operations Research, University of Vienna, 1010 Vienna, Austria (alois.pichler@univie.ac.at).

Springer Series in Operations Research and Financial Engineering

Georg Ch. Pflug
Alois Pichler

Multistage Stochastic Optimization

Springer

Thank you Georg for many years of successful cooperation, inspiring discussions and friendship!