SOLVE-OPERATOR METHODS FOR OPTIMIZATION OF RISK CONTROLLED STOCHASTIC PROCESSES

© Ivan Beyko, Petr Zinko

National Technical University of Ukraine "Kiev Politechnic Institute"
Taras Shevchenko National University of Kyiv. The Faculty of Cybernetics
e-mail: *ivan.beyko@gmail.com, petro.zinko@gmail.com*

Abstract. *In the paper we develop solve-operator methods for high order modelling, simulation and optimization of risk controlled stochastic processes described by general graph-operator control systems with incomplete data.*

The risk management includes increasing of the likelihood and impact of favorable events and reducing of the likelihood and impact of adverse processes. Development of new information technologies and computer based systems for solving risk minimization problems are based on optimization of adequate simulators of risk processes. A simulator is said to be adequate if it's practical implementation meets practical requirements to the allowable time $T(p)$ and error $E(p)$ of the calculations, where p is a vector-parameter of the simulator.

To design the optimal simulator, that minimizes the criterion function $J(p) = KT(p) + E(p)$ one uses available sets of mathematical models (with different aggregation levels and different resolving power) and available sets of sources of useful information. The optimal information sources are evaluated by "functions of information evaluation" (FIE) and the simulators are optimized by their iterative decomposition into optimal subsystems to perform substantiated prediction of risk processes in limited time in uncertain environment [1, 2]. Using FIE the iterative optimization procedures detect (on each iterative step) those of the subsystems that ought to be decomposed and those to be aggregated.

Risk optimization problems belong to the most difficult problems of controlled stochastic processes optimization. Their solution requires either simulation-based stochastic quasi-gradient methods [3] dealing with a general distribution of the random parameters, or special decomposition methods [4, 5] dealing with the distribution approximated by finitely many scenarios. Most of the existing computational methods are applicable only to convex problems and converge to a local minimum of multi-extremal problems [6].

To solve global stochastic non-convex optimization problems one may use the stochastic branch and bound algorithm based on the idea of global deterministic branch and bound algorithms [8]. The branch and bound algorithms are designed to solve those global stochastic non-convex problems, for which one can calculate (within a reasonable

time) a grate amount of alternative values of the objective function on allowable control sets.

To make it possible two optimization problems should be solved: the problem of mathematical models/simulators optimization and the problem of decision strategies optimization. In this way we implement solve-operator methods to design stochastic processes simulators and risk processes optimization under parametric uncertainties. In a rather general form the solution u^* of a risk optimization problem may be defined as the minimizer

$$u^* = \arg\min_{u\in\Omega} \bar{F}(u) \tag{1}$$

of a risk function

$$\bar{F}(u) = E \max_{q\in Q} \bar{f}(u,q,\theta) \tag{2}$$

where u is the control input, q is an uncertainty parameter, θ is a random variable defined on a probability space $(\Theta, \Sigma, \mathrm{P})$, $\bar{f}(u,q,\theta)$ is a random performance function, $\bar{F}(u)$ is the expected performance indicator, Q is a set of uncertainty, and Ω is a feasible control set.

We will consider time and space multidimensional interdependent risk processes where the random performance function

$$\bar{f}(u,q,\theta) \triangleq \tilde{f}(x(u,q,\theta),u,q,\theta)$$

depends on the stochastic process $x(u,q,\theta)$ simulated by the graph-operator system

$$A(x,u,q,\theta) \triangleq (A_1(x,u,q,\theta),\dots,A_{N_k}(x,u,q,\theta)) = 0, \tag{3}$$

$$A_k(x,u,q,\theta) \triangleq (A_{k1}(x_{k1},z_{k1},u_{k1},q_{k1},\theta_{k1}),\dots$$

$$\dots,A_{kN_{ks}}(x_{kN_{ks}},z_{kN_{ks}},u_{kN_{ks}},q_{kN_{ks}},\theta_{kN_{ks}})),$$

$$(x,u,q,\theta) \triangleq \{(x_k,u_k,q_k,\theta_k)\}_{k=1}^{N_k},\ (x_k,u_k,q_k,\theta_k) \triangleq \{(x_{ks},u_{ks},q_{ks},\theta_{ks})\}_{s=1}^{N_{ks}}.$$

where the ks-th subsystem

$$A_{ks}(x_{ks},z_{ks},u_{ks},q_{ks},\theta_{ks}) = 0 \tag{4}$$

of the graph's k-th knot describes interdependences between the ks-th subsystem states x_{ks}, subsystem controls u_{ks}, uncertainty parameters q_{ks}, random parameters θ_{ks}, and influences z_{ks} with the subsystem of environment,

$$z_{ks} = \varphi_{ks}(x,u,q,\theta), k = \overline{1,N_k}, s = \overline{1,N_{ks}}. \tag{5}$$

The designing of adequate computational procedures for calculating $x_{ks}(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})$, $\tilde{f}(x(u,q,\theta),u,q,\theta)$, and for calculating optimal solutions

$$u^* = \arg\min_{u\in\Omega} E \max_{q\in Q} \tilde{f}(x(u,q,\theta),u,q,\theta)$$

depends on types and dimensions reducing of all the algebraic, differential, and algebraic-integral-differential equations, that are being implemented to describe the ks-th subsystem.

Main difficulties of the optimal solution calculation arise in cases of non-convex multi-extremal performance function $\bar{F}$. There are different numerical algorithms designed for non-convex stochastic optimization. In simple cases, where calculations of $\bar{F}(u)$ may be done for many different alternative u, the branch and bound algorithm for stochastic global optimization may be used, capable of solving within a reasonable time small problems with highly non-convex functions and with a large number of local minima. The idea of deterministic branch and bound algorithm is to subdivide the set Ω into smaller subsets and to estimate from above and from below the optimal value of the objective within these subsets and to delete non perspective subsets from the Ω partition by using current lower and upper bounds of the optimal value within the subsets. In the stochastic deletion rule they do not delete subsets at each iteration, but only after carrying out a sufficiently large number of iterations, and after deriving an independent estimate of the objective value at the current approximate solution.

To simplify calculation difficulties we may replace too complicated subsystems models (4), (5) by simplified subsystems for which while there is some loss of accuracy using the simplified models, the results actually match fairly closely with the full solution. In this way there ware many successful attempts in searching for adequate approximations of stochastic subsystems $A_{ks}(x_{ks}, z_{ks}, u_{ks}, q_{ks}, \theta_{ks}) = 0$ by some simplified stochastic differential equation subsystems (SDE), that allow simplification of computation procedures for calculating $x_{ks}(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})$. For example, in many cases the adequate simplified approximation models may be described by simple SDE:

$$dx^1_{ks}(t) = a^1_{ks}(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})dt + b^1_{ks}(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})dw(t),$$

$$dx^2_{ks}(t) = x^2_{ks}(t)(a^2_{ks}(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})dt + b^2_{ks}(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})dw(t)),$$

$$dx^3_{ks}(t) = a^3_{ks}(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})x^3_{ks}(t)dt + b^3_{ks}(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})dw(t),$$

or by more general SDE linear systems

$$dx^4_{ks}(t) = (A(u(t), q(t), \theta(t), t)x^4_{ks}(t) + c(u(t), q(t), \theta(t), t))dt + +B(u(t), q(t), \theta(t), t)dw(t)$$

with Brownian movements $w(t) \triangleq (w_1(t), ..., w_m(t))$,

$$dw_i(t) \triangleq w_i(t+dt) - w_i(t), E(dw_i^2(t)) = \sigma_i^2 dt,$$

$E(dw_i(t)dw_j(t)) = 0$ for $i \neq j$.

The trajectories of these models are known to be:

$$x_{ks}^1(t) = x_{ks}^1(t_0) + a_{ks}^1(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})(t-t_0) + b_{ks}^1(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})w(t-t_0),$$

$$x_{ks}^2(t) = \exp((a_{ks}^2(u_{ks}, q_{ks}, \theta_{ks}, z_{ks}) - (b_{ks}^2)^2(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})/2)(t-t_0) + \\ + b_{ks}^2(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})w(t-t_0)),$$

$$x_{ks}^3(t) = x_{ks}^3(t_0)\exp(a_{ks}^3(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})(t-t_0)) + \\ + b_{ks}^3(u_{ks}, q_{ks}, \theta_{ks}, z_{ks}) \int_{t_0}^{t} \exp(a_{ks}^3(u_{ks}, q_{ks}, \theta_{ks}, z_{ks})(t-\tau))dw(\tau),$$

$$x_{ks}^4(t) = \Phi(t)x_{ks}^4(t_0) + \Phi(t) \int_{t_0}^{t} \Psi(\tau)(c(u(\tau), q(\tau), \theta(t), \tau) + + B(u(\tau), q(\tau), \theta(\tau), \tau)dw(\tau),$$

where $\Phi(\cdot)$ and $\Psi(\cdot)$ are the fundamental matrices of the associated homogeneous linear system and it's conjugate system.

In case of the nonlinear SDE

$$dx(t) = a(x(t), u, q, \theta)dt + b(x(t), u, q, \theta)dw(t),$$

$$x(t) \in R^n, a(\cdot) = \{a_i(\cdot), i = \overline{1,n}\}, b(\cdot) = \{b_{ij}(\cdot), i = \overline{1,n}, i = \overline{1,m}\},$$

the appropriate subsystem's risk increments of smooth random risk performance functions $\tilde{\tilde{f}}(t, x(t))$ satisfy the Ito formula

$$d\tilde{\tilde{f}}(t, x(t)) = [\partial_t \tilde{\tilde{f}}(t, x(t)) + a(x(t), u, q, \theta)\partial_x \tilde{\tilde{f}}(t, x(t))dt + \\ + 0,5b^2(x(t), u, q, \theta)\partial_{xx}^2 \tilde{\tilde{f}}(t, x(t))]dt + b(x(t), u, q, \theta)\partial_x \tilde{\tilde{f}}(t, x(t))dw(t)$$

and the probability density $p \triangleq p(x, t|x_0, t_0, u, q, \theta)$ may be calculated as the solution of the Fokker-Planck equations

$$\frac{\partial p(x,t)}{\partial t} = -\sum_{i=1}^{n} \frac{\partial}{\partial x_i}[a_i(x(t), u, q, \theta)p(x,t)] + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}[\bar{b}_{ij}(x(t), u, q, \theta)p(x,t)].$$

Using the calculated probability density p we calculate u^* as the solution of the significant simplified optimization problem

$$u^* = \arg\min_{u\in\Omega}\max_{q\in Q}\int \tilde{f}(x,u,q,\theta)dp(x,u,q,\theta).$$

In cases of convex optimization problems the global optimal solution u^* may be estimated by stochastic quasi-gradient methods using numerical SDE simulators. For example, the iterative Euler-Maruyama simulator

$$x(t_{i+1}) = x(t_i) + a(x(t_i),u,q,\theta)(t_{i+1}-t_i) + b(x(t_i),u,q,\theta)(w(t_{i+1})-w(t_i))$$

or more accurate Milstein simulator

$$x(t_{i+1}) = x(t_i) + a(x(t_i),u,q,\theta)(t_{i+1}-t_i) + b(x(t_i),u,q,\theta)(w(t_{i+1})-w(t_i)) +$$
$$+\frac{1}{2}b(x(t_i),u,q,\theta)b^T(x(t_i),u,q,\theta)((w(t_{i+1})-w(t_i))^2 + t_i - t_{i+1}).$$

Using the Ito Formula and the stochastic Taylor expansions of functionals of SDEs many other convergent, consistent, and strictly or marginally stable simulators are developed and may be implemented.

We develop higher order solve operator methods to calculate trajectories $x_{ksi}(t,p,q,\theta)$ of stochastic control processes

$$t_0 = \bar{t}(u_{ks0},q_{ks0},\theta_{ks0},z_{ks0}) \in R, x_{ks0}(t_0) = \bar{x}_{ks0}(t_0,u_{ks0},q_{ks0},\theta_{ks0},z_{ks0}) \in R^{n_{ksx}},$$
$$dx_{ksi}(t) = a_{ksi}(x_{ksi}(t),u_{ksi},q_{ksi},\theta_{ksi},z_{ksi},t,\omega_{ksi}(x_{ksi}(t_i),t_i,q_{ksi},\theta_{ksi},z_{ksi},t))dt, \quad (6)$$
$$t \in [t_i,t_{i+1}],$$
$$t_{i+1} = \tau(x_{ks(i-1)}(t_i),t_i,u_{ksi},q_{ksi},\theta_{ksi},z_{ksi}) > t_i,$$
$$x_{ks(i+1)}(t_{i+1}) = \psi(x_{ksi}(t_{i+1}),t_{i+1},u_{ks(i+1)},\theta_{ks(i+1)},z_{ks(i+1)}),$$

were $\theta_{ksi} \in \mathbb{R}^{n_{ksi}}$ are random vectors defined by adequate evaluated distribution functions $\tilde{F}_{ksi}(\tilde{\theta}_{ksi}|(x_{ks(i-1)}(t_i),t_i,q_{ks(i-1)},\theta_{ks(i-1)},z_{ks(i-1)})$.

For the given $\bar{q}$, $\bar{u}$ and for the given realization $\bar{\theta}$ of θ the trajectory $x = x(\tau) \triangleq x(\tau,\bar{u},\bar{q},\bar{\theta})$ of the system (6) in the neighbourhood $O(t) \in \prod_i(t_i,t_{i+1})$ of $t \in \prod_i(t_i,t_{i+1})$ may be described by the system (7)

$$dx(\tau)/d\tau = f(x(\tau),\tau), \quad (7)$$

$$f(x(\tau),\tau) \triangleq a(x(\tau),\bar{u},\bar{q},\bar{\theta},\tau,\omega(x(t_i),t_i,\bar{u},\bar{q},\bar{\theta},\tau)) \quad (8)$$

The operator F is said to be an asymptotic solve operator on the interval $\tau \in [t,t+H] \subset O(t)$ for the given function $v\,(x(t+H))$ with respect to the continue function $Z\,(Q(\tau))$ on the trajectory x of the system (7) if for continue functions p from

the neighbourhood of x holds the asymptotic neighbourhood

$$F(t,p,H,Z,Q) = v\left(x(t+H)\right) \ + \ \left(O(||Z(Q)||) + O(||p-x||)\right) H||p-x||.$$

And the operator $G(\tau)$ is said to be an s-asymptotic solve operator with respect to the parameter h if for the function $v\left(x(t+h)\right)$ holds the asymptotic equality

$$G(h) = v\left(x(t+h)\right) + O(h^s).G(h) = v\left(x(t+h)\right) + O(h^s).$$

Theorem 1. [8] *If* $v\left(x(t+H)\right) \triangleq Q(t+H)x(t+H), \ \ Z\left(Q(\tau)\right) \triangleq dQ(\tau)/d\tau + Q(\tau)A(\tau)$, *on the interval the functions* $Q(\tau), \ A(\tau) \triangleq f'_x\left(p(\tau),\tau\right)$ *and* $Z\left(Q(\tau)\right)$ *are continuous and* $f'_x\left(p(\tau),\tau\right)$ *is a Lipschitz matrices with respect to* $p(\tau)$, *then the asymptotic solve operator* F *is defined by the equality*

$$F(t,p,H,Z,Q) = Q(t+H)p(t+H) + \int\limits_t^{t+H} Q(\tau)\left(f(p(\tau),\tau) - dp(\tau)/d\tau\right)d\tau.$$

Theorem 2. *If in the conditions of the theorem 1 the functions* $Q(\tau), \ A(\tau) = f'_x\left(p(\tau),\tau\right), \ p(\tau)$ *and* $x(\tau)$ *satisfy on the interval* $\tau \in [t, t+h]$ *the asymptotic equality*

$$dQ\left(\tau\right)/d\tau = -Q\left(\tau\right)A\left(\tau\right) + \mathrm{O}(h^k), \ p\left(\tau\right) = x\left(\tau\right) + \mathrm{O}(h^l), \ p(t) = x(t),$$

then s-asymptotic solve operator $G(h)$, $s = k+l+1$, $l \le k$ *is defined by the equality*

$$G(h) = Q(t+h)p(t+h) + \int\limits_t^{t+H} Q(\tau)\left(f\left(p(\tau),\tau\right) - dp(\tau)/d\tau\right)d\tau. \tag{9}$$

The theorem statement follows from the given equalities

$$F(t,p,h,Z,Q) = Q(t+h)p(t+h) + \textstyle\int_t^{t+H} Q(\tau)\left(f\left(p(\tau),\tau\right) - dp(\tau)/d\tau\right)d\tau,$$
$$F(t,p,h,Z,Q) = v\left(x(t+h)\right) + \left(O\left(||Z(Q)||\right) + O\left(||p-x||\right)\right)||p-x||h.$$

Really, it follows

$$G(h) = F\left(t,p,h,Z,Q\right) = v\left(x(t+h)\right) + \left(O\left(||Z(Q)||\right) + O\left(||p-x||\right)\right)||p-x||h.$$

And taking into account

$$dQ(\tau)/d\tau = -Q(\tau)A(\tau) + O(h^k), \ \ p(\tau) = x(\tau) + O(h^l),$$

we obtain

$$G(h) = v\left(x(t+h)\right) + h\left(O\left(||Z(Q)||\right) + O\left(||p-x||\right)\right)||p-x|| =$$
$$= v\left(x(t+h)\right) + h\left(O(h^k) + O(h^l)\right)O(h^l),$$

and for $l \leq k$ we obtain the required equality

$$G(h) = v\left(x(t+h)\right) + O(h^{k+l+1}) = v\left(x(t+h)\right) + O(h^s).$$

From the theorem 2 it follows that for any given $\bar{q}$, $\bar{u}$, and for given realization $\bar{\theta}$ of θ, the s-order approximation $\bar{x}(t+h), \bar{x}(t+h) = x(t+h) + O(h^s), s = k+l+1$, the trajectory $x(\tau) \triangleq x(\tau, \bar{u}, \bar{q}, \bar{\theta})$ of the differential equation (7) may be calculated by the asymptotic solver-operator formula (10)

$$\bar{x}(t+h) = p(t+h) + \int_t^{t+h} Q(\tau)\left(f\left(p(\tau), \tau\right) - dp(\tau)/d\tau\right) d\tau, \tag{10}$$

using $p(\cdot)$ and $Q(\cdot)$ that satisfy (11), (12)

$$p(\tau) = x(\tau) + O(h^l), \tag{11}$$

$$dQ(\tau)/d\tau = -Q(\tau)A(\tau) + O(h^k),\ Q(t+h) = I. \tag{12}$$

Using asymptotic solve-operators (10)–(12) we construct many of the following high-order simulators to calculate trajectories of stochastic processes realization (6) and (7). For example, using Lagrange polynomials

$$\begin{aligned} p_{n+1}(\tau) = \tfrac{1}{h^n}\Big[& x(t)\tfrac{(\tau-t-h)\ldots(\tau-t-nh)}{(-1)\cdot(-2)\ldots(-n)} + \\ & + x(t+h)\tfrac{(\tau-t)(\tau-t-2h)\ldots(\tau-t-nh)}{1\cdot(-1)\cdot(-2)\ldots(-(n-1))} + \ldots \\ & + x(t+nh)\tfrac{(\tau-t)(\tau-t-h)\ldots(\tau-t-(n-1)h)}{n\cdot(n-1)\ldots2\cdot1}\Big]. \end{aligned}$$

for the given values $x(t+ih),\ \ i = \overline{0,n}$ we obtain the high-order simulators

$$\begin{aligned} x\left(t+(n+1)h\right) = {} & p_{n+1}\left(t+(n+1)h\right) + \textstyle\int_t^{t+(n+1)h} \left[E - (\tau - t - (n+1)h) \times\right. \\ & \left. \times f'_x\left(p_{n+1}\left(t+(n+1)h\right), t+(n+1)h\right)\right]\left[f\left(p_{n+1}(\tau), \tau\right) - \dot{p}_{n+1}(\tau)\right] d\tau, \end{aligned}$$

with simulators error $O(h^{n+3})$. And using the Newton-Cotes formula we obtain a number of numerical simulators

$$\begin{aligned} x\left(t+(n+1)h\right) = {} & p_{n+1}\left(t+(n+1)h\right) + \\ & + (n+1)h\textstyle\sum_{i=0}^{n+1} c_{i,n+1}\left[f\left(p_{n+1}(t+ih), t+ih\right) - \dot{p}(t+ih)\right] - \\ & - (n+1)h^2 f'_x\left(p_{n+1}\left(t+(n+1)h\right), t+(n+1)h\right)\textstyle\sum_{i=0}^{n+1} c_{i,n+1}(i-n-1)\times \\ & \times\left[f\left(p_{n+1}(t+ih), t+ih\right) - \dot{p}_{n+1}(t+ih)\right]. \end{aligned}$$

with estimated errors $O(h^{n+3})$. Using the Tylor's formula

$$\begin{aligned} x(t+mh) = {} & \int_t^{t+mh} \left[E - (\tau - t - mh) f'_x\left(p(t+mh), t+mh\right)\right] \times \\ & \times\left[f\left(p(\tau), \tau\right) - \dot{p}(\tau)\right] d\tau + p(t+mh), \end{aligned}$$

we obtain numerical simulators with the error estimate $O(h^{s+2})$.

Similar high-order simulators are constructed to calculate the probability densities using Fokker-Planck equations. Numerical experiments proved the practical efficiency of the designed high-order simulators implementation to calculate u minimizing the risk function

$$\mathrm{E} \max_{q \in Q} \tilde{f}(x(u, q, \theta), u, q, \theta)$$

using stochastic generalized gradient methods [9] and stochastic minimax algorihms [10].

Conclusions

The developed high order solve-operator methods may be implemented to solve problems of the general graph-operator stochastic control systems modelling, simulation and optimization under incomplete data.

References

1. Beyko, I. 1988. Extremal models of complex systems and decomposition methodology in numerical experiments. *Vestnik Kyiv University, Modeling and optimization of complex systems*, pp. 72–81.
2. Beyko, I. 1996. Functions for evaluation of information in the theory of optimal aggregate models. *Cybernetics and System Analysis*, 3, pp. 43–54.
3. Ermoliev, Yu. M. 1983. Stochastic quasi-gradient methods and their application to systems optimization, *Stochastics*, 4, pp. 1–37.
4. Mulvey, J. M., and Ruszczyn'ski, A. 1995. A new scenario decomposition method for large-scale stochastic optimization. *Operations Research*, 43, pp. 477–490.
5. Rockafellar, R. T., and Wets R. J.-B. 1991. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16, pp. 1–23.
6. Mikhalevich, V. S., Gupal, A. M. and Norkin, V. I. 1996. *Methods of Non-Convex Optimization.* Moscow: Nauka.
7. Norkin, V. I., Pug, G. Ch. and Ruszczyn'ski, A. 1996. A Branch and Bound Method for Stochastic Global Optimization. *IIASA WP-96-xxx, April 1996*, pp. 25.
8. Beyko, I. V. and Zinko, P. M. 2009. High-order methods for solving the Cauchy problems and the multidimensionalvalue problems by using an asymptotically-solving operators.*Mathematical and computer modeling* (Kamenets-Podolsky National University), 1, pp. 18–25.
9. Ermoliev, Y. M. and Norkin, V. I. 1997. Stochastic generalized gradient method with application to insurance risk management. IR-97-021, Int. Inst. for Appl. Syst. Anal., Laxenburg, Austria.
10. Beyko, I. V., Zinko, P. M., and Nakonechny, O. H. 2011. *Problems, methods and algorithms of optimization.* National University of Water Management and Nature Resources, Rivne.