# Chance constrained zero sum stochastic games

**Lucas OSMANI**, L2S - Paris **Abdel LISSER**, L2S - Paris
**Vikas SINGH**, IIT - New Delhi

We consider a two-person zero-sum discounted stochastic game [5] with random rewards and known transition probabilities. The players have opposite objectives and are interested in optimizing the expected discounted reward which they can obtain with a given confidence level when both the players play the worst possible move against each other.
Following the approach in [2], we model such a game problem by defining the chance-constrained optimization problem of each player, denoted by (P1) and (P2), and call this new formulation a chance-constrained stochastic game (CCSG). When the reward vector follows a multivariate elliptically symmetric distribution [4], the CCSG is equivalent to a minimax formulation. We consider CCSG with risk seeking and risk averse players separately.
We show that the risk-seeking problem is equivalent to a constrained optimization of a parameterized zero-sum stochastic game and the optimal payoff of player 1 and optimal cost of player 2 can be computed using algorithms based on a fixed point iteration. Later we can use the fixed point solutions to compute players' optimal strategies by solving linear programming problems.
We reformulate the risk averse problem as a discrete minimax problem. We propose an algorithm based on a linearization method [1] and discuss its convergence properties. Alternatively, we reformulate the risk averse problem as a second-order cone programming problem with bilinear constraints. This reformulation is an extension of [3]. We illustrate the theoretical results with numerical experiments.
The CCSG takes the following form :

$$\begin{aligned} \delta^*(p_1) &:= \max_{f \in F_S, \delta \in R} \delta \\ \text{s.t.} &\quad \min_{g \in G_S} P(\tilde{V}(m,f,g) \geq \delta) \geq p_1. \end{aligned} \tag{P1}$$

$$\begin{aligned} \eta^*(p_2) &:= \min_{g \in G_S, \eta \in R} \eta \\ \text{s.t.} &\quad \min_{f \in F_S} P(\tilde{V}(m,f,g) \leq \eta) \geq p_2. \end{aligned} \tag{P2}$$

Here, $f, g$ are stationary strategies, respectively, for players 1 and 2. $\tilde{V}(m, f, g)$ is a random discounted value function corresponding to strategies $f, g$, initial measure $m$, and a random reward function.

[1] M. S. Bazaraa, J. J. Goode. *An algorithm for solving linearly constrained minimax problems.* European Journal of Operational Research, **11(2)**, 158–166, 1982. doi :https ://doi.org/10.1016/0377-2217(82)90110-2.

[2] R. A. Blau. *Random-payoff two-person zero-sum games.* Operations Research, **22(6)**, 1243–1251, 1974.

[3] E. Delage, S. Mannor. *Percentile optimization for markov decision processes with parameter uncertainty.* Operations Research, **58(1)**, 203–213, 2010.

[4] K. Fang, S. Kotz, K. Ng. *Symmetric Multivariate and Related Distributions.* Monographs on Statistics and Applied Probability. Springer US, 2013.

[5] L. S. Shapley. *Stochastic games*.* Proceedings of the National Academy of Sciences, **39(10)**, 1095–1100, 1953. doi :10.1073/pnas.39.10.1095.

Contact : `lucas.osmani@centralesupelec.fr`