Novel Knowledge Gradient Policies

Boris Defourny (Department of Operations Research and Financial Engineering, Princeton University)

January 6, 2012

We consider control problems where the function of optimizing a performance measure (the implementation problem) can be separated from the function of estimating the parameters of the performance measure (the learning problem).

For the implementation problem, we consider linear programs (LP) with uncertain objective coefficients.

For the learning problem, we consider knowledge gradient policies (KG), that select a next measurement by solving a one-step optimization problem into the future, as if the solution to the implementation problem had to be implemented after having observed the outcome of the measurement.

We introduce a knowledge gradient policy which maximizes the expected value of information of the next measurement when the implementation problem takes itself the residual uncertainty into account by a worst-case performance measure, formulated using robust optimization techniques.

Under some assumptions, the learning problem becomes a nonconvex two-stage stochastic program, where the first stage is the choice of the measurement, and the second stage is the solution of to a convex robust problem with an updated uncertainty set. We develop convex relaxations for optimizing over continuous sets of measurements.

This is joint work with Ilya O. Ryzhov (U of Maryland) and Warren B. Powell (Princeton).