Quantization in discretization of stochastic programs.

Kengy Barty

EDF R&D

March 14, 2007

Introduction

[Rockafellar and Wets, 1991] addressed the following stochastic optimization problem:

$$\min_{u \in \mathcal{N}} J(u) = \sum_{s \in S} p_s j(u(s), s)$$

$$\mathcal{N} = \{u \in \mathcal{E} \mid u_t \text{ is constant on each bundle } A \in \mathcal{A}_t\}$$

S is a finite set of scenarios. $\mathcal{A}_t$ is the "information structure", $A \in \mathcal{A}_t$ is called a "bundle". The scenarios in any one bundle are regarded as observationally indistinguishable at time t. "For most purpose it is reasonable to suppose that the partition $\mathcal{A}_{t+1}$ is a refinement of the partition $\mathcal{A}_t$.".

Introduction

The same problem has been addressed under a continuous form:

$$\begin{array}{l} \min \mathbb{E}\left[j(u(X), X)\right] \\ u \in \mathcal{N}(\mathcal{F}) \end{array}$$

$$\mathcal{N}(\mathcal{F}) = \left\{ u \in L^2(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d}, \mathbb{P}_X) \mid u_t \text{ is } \mathcal{F}_t \text{ adapted function} \right\}$$

This kind of problem can be embedded in the set of problems with measurability constraints !

Introduction

Consider the following stochastic optimization problem:

$$V(\mathcal{B}) = \min\left\{\|v - u\|_{L^2} \mid u \text{ is } \mathcal{B} \text{ measurable}\right\}$$

clearly we have:

$$V(\mathcal{B}) = \|v - \mathbb{E}_{\mathbb{P}}[v \mid \mathcal{B}]\|_{L^2}$$

V has the 1-lipschitz-continuous property:

$$\left|V(\mathcal{B}) - V(\mathcal{B}')\right| \leq \underbrace{\left\|\mathbb{E}_{\mathbb{P}}[v \mid \mathcal{B}] - \mathbb{E}_{\mathbb{P}}\left[v \mid \mathcal{B}'\right]\right\|_{L^2}}_{\text{pseudo-metric on the space of } \sigma\text{-algebra}}$$

Now we can provide a way to compute the conditional expectation, but at the beginning the problem already satisfies a Lipsichitz property according to specific metric on the space of σ-algebra.

If now we want to compute the conditional expectation arising in the problem then we replace the underlying probability measure by the empirical probability measure denoted $\mathbb{P}_n$:

$$\begin{aligned} |V(\mathcal{B}) - V_n(\mathcal{B}')| &\leq |V(\mathcal{B}) - V(\mathcal{B}')| + |V(\mathcal{B}') - V_n(\mathcal{B}')| \\ &\leq \|\mathbb{E}_{\mathbb{P}}[v \mid \mathcal{B}] - \mathbb{E}_{\mathbb{P}}[v \mid \mathcal{B}']\|_{L^2} + \|\mathbb{E}_{\mathbb{P}}[v \mid \mathcal{B}] - \mathbb{E}_{\mathbb{P}_n}[v \mid \mathcal{B}']\|_{L^2} \end{aligned}$$

Theoretical framework

- Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a measurable probability space.
- Let $\mathcal{B}_{\mathbb{R}^d}$ denote the Borel σ-algebra. The smallest σ-algebra of subset of $\mathbb{R}^d$ which contains all open subsets of $\mathbb{R}^d$;
- Let $X : \Omega \to \mathbb{R}^d$ be a random variable, $X^{-1}(\mathcal{B}_{\mathbb{R}^d}) \subset \mathcal{F}$;
- $\mathbb{P}_X$ be the distribution of the random variable X defined by:

$$\forall A \in \mathcal{B}_{\mathbb{R}^d}, \quad \mathbb{P}_X(A) = \mathbb{P}(X^{-1}(A))$$

- If $u : \mathbb{R}^d \to \mathbb{R}^m$ we define $\sigma(u) = u^{-1}(\mathcal{B}_{\mathbb{R}^m})$, the smallest σ-algebra for which u is measurable;
- Let $j : \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}$ be a random functional (normal integrand)
- Let C denote a closed convex subset of $L^2(\mathbb{R}^p, \mathbb{P}_X; \mathbb{R}^m)$

For any Borel function $u : \mathbb{R}^d \rightarrow \mathbb{R}^m$ we define $J(u)$ by:

$$J(u) = \mathbb{E}\left[j(u(X), X)\right] = \int_{\mathbb{R}^d} j(u(x), x)\mathbb{P}_X(dx)$$

J is the objective function. Let S be defined by:

$$S = \left\{u \in L^2(\mathbb{R}^d, \mathbb{P}_X; \mathbb{R}^m) \mid u(x) = u(-x), \quad \mathbb{P}_X a.s.\right\}$$

Let $\mathcal{S}$ denote the smallest σ-algebra for which any $u \in S$ is measurable. $\mathcal{S}$ is called the symmetric σ-algebra. With that definition:

$$S = \left\{u \in L^2(\mathbb{R}^d, \mathbb{P}_X; \mathbb{R}^m) \mid \sigma(u) \subset \mathcal{S}\right\}$$

Consider now the following problem:

$$\mathcal{P} \begin{cases} \min \mathbb{E}\left[j(u(X), X)\right] \\ u \in S \end{cases}$$

or equivalently:

$$\mathcal{P} \begin{cases} \min \mathbb{E}\left[j(u(X), X)\right] \\ u \in L^2(\mathbb{R}^d, \mathbb{P}_X; \mathbb{R}^m) \\ u = \mathbb{E}\left[u \mid \mathcal{S}\right] \end{cases}$$

where $\mathbb{E}\left[u \mid \mathcal{S}\right]$ is the conditional expectation with respect to $\mathcal{S}$:

$$\left\| u - \mathbb{E}\left[u \mid \mathcal{S}\right] \right\|^2 = \min_{v \in S} \left\| u - v \right\|^2$$

The problem

We want to compute the value of $\mathcal{P}$ defined by:

$$V = \inf \left\{ \mathbb{E}\left[j(u(X), X) \right] \mid u \in S \right\}$$

Approximation

We could formulate directly a discrete problem:

$$\min \frac{1}{N} \sum_{i=0}^{N} j(u^i, X^i) \tag{1}$$

$$X^i = -X^j \Rightarrow u^i = u^j. \tag{2}$$

Whenever the distribution $\mathbb{P}_X$ absolutely continuous with respect to the Lebesgue measure the event:

$$\{X^i = -X^j\}$$

is null mass.

Transformation

The conditional expectation with respect to $\mathcal{S}$ is quite easy to calculate:

$$\mathbb{E}\left[u \mid \mathcal{S}\right] = \frac{u(x) + u(-x)}{2} \tag{3}$$

then we can transform the initial problem to the unconstraint problem:

$$\mathcal{P}_S \left\{ \begin{array}{l} \min \mathbb{E}\left[j(\mathbb{E}\left[u \mid \mathcal{S}\right](X), X)\right] \\ u \in L^2(\mathbb{R}^d, \mathbb{P}_X; \mathbb{R}^m) \end{array} \right.$$

if u^* is optimal for $\mathcal{P}_S$, $\mathbb{E}\left[u^* \mid \mathcal{S}\right]$ is obviously optimal for $\mathcal{P}$.

Two questions

1. How to transform:

$$J_S(u) = \mathbb{E}\left[j\left(\frac{u(X)+u(-X)}{2}, X\right)\right] \tag{4}$$

to :

$$J_S(u) = \mathbb{E}\left[j_s(u(X), X)\right] \tag{5}$$

2. How to link the optimal solution of the constraint problem $\mathcal{P}$ with the solution of the unconstraint problem $\mathcal{P}_S$?

Answer

[Dynkin and Evstigneev, 1976] have proved that:

$$\mathbb{E}\left[j(\mathbb{E}\left[u \mid \mathcal{S}\right](X), X) \mid \mathcal{S}\right] = j_s(\mathbb{E}\left[u \mid \mathcal{S}\right](X), X)$$

where j_s is the regular conditional expectation of the normal integrand j:

$$j_s(u, \cdot) = \mathbb{E}\left[j(u, \cdot) \mid \mathcal{S}\right]$$

That result has been also studied by J.M. Bismut.

The problem without information constraint

Let j_s denote the following normal integrand:

$$j_s(u, x) = \mathbb{E}\left[j(u, \cdot) \mid \mathcal{S}\right](x) = \frac{j(u, x) + j(u, -x)}{2}$$

consider now the following problem:

$$\mathcal{P}'_S \begin{cases} \min \mathbb{E}\left[j_s(u(X), X)\right] \\ u \in L^2(\mathbb{R}^d, \mathbb{P}_X; \mathbb{R}^m) \end{cases}$$

there is not information constraint in that new formulation.

A first result

- $\forall u \in S, \quad j_s(u(x), x) = \mathbb{E}\left[j(u(\cdot), \cdot) \mid \mathcal{S}\right](x)$ (Dynkin Evstigneev);
- Moreover if $u \to j(u, x)$ is convex then if $u^* \in L^2(\mathbb{R}^d, \mathbb{P}_X; \mathbb{R}^m)$ is solution of $\mathcal{P}'_S$ then:
  1. u^* is solution of $\mathcal{P}_S$;
  2. v^* defined by:

$$v^*(x) = \frac{u^*(x) + u^*(-x)}{2}$$

     is solution of $\mathcal{P}$

We can state this result without convexity assumption, but we couldn't establish a practical expression for the optimal decision rule v^* just its existence

The proof

$$\mathbb{E}\left[j_s(u(\cdot,\cdot)) \mid \mathcal{S}\right] = \frac{j(u(x),x)+j(u(-x),-x)}{2} \quad (6)$$

$$= \frac{j(u(x),x)+j(u(x),-x)}{2} \quad (7)$$

$$= j_s(u(x),x) \quad (8)$$

That result holds true in more general cases, see E.B. Dynkin and I.V. Evstigneev, 1976.

The proof

$$\begin{aligned}
\mathbb{E}\left[j(\mathbb{E}\left[u^* \mid \mathcal{S}\right](x), x)\right] &= \mathbb{E}\left[\mathbb{E}\left[j(\mathbb{E}\left[u^* \mid \mathcal{S}\right](x), x) \mid \mathcal{S}\right]\right] \text{(Cond Exp)} \\
&= \mathbb{E}\left[j_s(\mathbb{E}\left[u^* \mid \mathcal{S}\right](x), x)\right] \text{(Dyn Evs)} \\
&\leq \mathbb{E}\left[j_s(u^*(X), X)\right] \text{(Cond Jensen In for Integ)} \\
&\leq \mathbb{E}\left[j_s(\mathbb{E}\left[v \mid \mathcal{S}\right](x), x)\right] \text{(Optimality)} \\
&\leq \mathbb{E}\left[j(\mathbb{E}\left[v \mid \mathcal{S}\right](x), x)\right] \text{(Dyn Evs)}
\end{aligned}$$

Open loop problems

The problem $\mathcal{P}'_S$ turns out to be a Monte-Carlo simulation:

$$\mathcal{P}' \quad \Leftrightarrow \quad \mathbb{E}\left[\min_{u \in \mathbb{R}^m} j_s(u, X)\right]$$

In lack of information constraint we can use for $\mathcal{P}'$ a Monte-Carlo approximation technique.

Numerical scheme

We propose the following numerical scheme:

- Draw $(X^1, \dots, X^N)$ N iid of X;
- compute : $J_S^N(u^1, \dots, u^N) = \frac{1}{N}\sum_{i=1}^N j_s(u^i, X^i)$

Hence:

$$\lim_{N\to\infty}\left(\inf_{(u^i)_{i=1,\dots,N}} J_S^N(u^1, \dots, u^N)\right) = \text{Optimal value of } \mathcal{P}$$

But at this stage we suppose we are able to compute exactly the normal integrand j_s!

Conditional expectation approximation

The simplest idea is to replace the σ-algebra $\mathcal{S}$ by a σ-algebra denoted $\mathcal{S}^p$ generated by a finite partition of $\mathbb{R}^d$. Hence instead of handle j_s we will handle j_p defined by:

$$j_p(u,x) = \mathbb{E}\left[j(u,\cdot) \mid \mathcal{S}^p\right](x)$$

$$j_p(u,x) = \sum_{i=1}^{l_p} \frac{1}{\mathbb{P}_X(A_i^p)} \mathbb{E}\left[j(u,X)\mathbb{I}_{A_i^p}(X)\right] \mathbb{I}_{A_i^p}(x)$$

where $(A_i^p)_{i=1,\ldots,l_p}$ is a partition such that $\mathcal{S}^p = \sigma((A_i^p)_{i=1,\ldots,l_p})$. Is it consistent to replacing the continuous σ-algebra by a discrete one ?

σ-algebra metric

It well known that the space of sub-σ-algebra of $\mathcal{S}$ is metrizable.

Let $\mathcal{S}', \mathcal{S}'' \subset \mathcal{S}$ then:

$$d(\mathcal{S}', \mathcal{S}'') = \sum_{i=0}^{\infty} \frac{1}{2^i} \min\left(\left\| \mathbb{E}\left[f_i \mid \mathcal{S}' \right] - \mathbb{E}\left[f_i \mid \mathcal{S}'' \right] \right\|, 1 \right)$$

where $(f_i)_{i \in \mathbb{N}}$ is a separable set of $L^2(\mathbb{R}^d, \mathbb{P}_X; \mathbb{R})$ is a metric introduced first by Neveu in 1975. d is consistent with the weakest topology with respect to which all functions defined by:

$$L^2(\mathbb{R}^d, \mathbb{P}_X; \mathbb{R}) \to L^2(\mathbb{R}^d, \mathbb{P}_X; \mathbb{R}), \quad f \mapsto \mathbb{E}\left[f \mid \mathcal{S}' \right] \text{ with } \mathcal{S}' \subset \mathcal{S}$$

is continuous with respect to the norm topology.

Other metrics

$$d_{Boy}(\mathcal{S}', \mathcal{S}'') = \sup_{A \in \mathcal{S}'} \inf_{B \in \mathcal{S}''} \mathbb{P}(A \Delta B) + \sup_{B \in \mathcal{S}''} \inf_{A \in \mathcal{S}'} \mathbb{P}(A \Delta B)$$

$$d_{Bar}(\mathcal{S}', \mathcal{S}'') = \inf \left\{ 1 - \mathbb{P}(A) \mid \mathcal{S}' \cap A = \mathcal{S}'' \cap A, \quad A \in \mathcal{B} \right\}$$

Unfortunately discrete σ-algebras fail to be a discrete approximation of continuous σ-algebra with respect to those metrics.

Asymptotic result

it is easy to show from Neveu's metric definition that the following holds:

$$\lim_{p\to\infty} d(\mathcal{S}^p, \mathcal{S}) = 0 \Rightarrow \forall u \in \mathbb{R}^m \quad \lim_{p\to\infty} \|j_p(u,\cdot) - j_s(u,\cdot)\|_{L^2} = 0$$

Theoretical results

We consider now the following discrete problem:

$$\mathcal{P}_{n,p} \begin{cases} \min \frac{1}{n} \displaystyle\sum_{i=1}^{l_p} \sum_{k=1}^{n} j(u_i, X^k) \mathbb{I}_{A_i^p}(X^k) \\ u_i \in \mathbb{R}^m \end{cases}$$

- We denote $V_{n,p}$ the optimal value of $\mathcal{P}_{n,p}$ with respect to the partition $\mathcal{S}^p$ and to the empirical distribution $\mathbb{P}_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{X^k}$;
- we denote V_p the optimal value defined by:

$$V_p = \inf \left\{ \sum_{i=1}^{l_p} \mathbb{E}\left[j(u_i, X) \mathbb{I}_{A_i^p}(X) \right] \mid u_i \in \mathbb{R}^m \right\}$$

- V is the optimal value of the problem denoted $\mathcal{P}$

Theoretical results

We assume that $\mathbb{P}_X(\partial A_i^p) = 0$, [Dupačová and Wets, 1988]:

$$\lim_{n \to \infty} |V_{n,p} - V_p| = 0 \text{ J. Dupacova, A. Shapiro etc} \ldots$$

[Artstein, 1991] for a convex (J is convex) stochastic allocation problem and Barty (PhD, 2004) for a stochastic optimization program under continuity assumption (J is continuous) that:

$$\lim_{p \to \infty} d(\mathcal{S}^p, \mathcal{S}) = 0 \Rightarrow \lim_{p \to \infty} |V_p - V| = 0.$$

The information is given by a signal

Let $h : \mathbb{R}^p \to \mathbb{R}^\ell$ be a Borel measurable function. We suppose now the constraint is given by:

$$S = \left\{ u \in L^2(\mathbb{R}^d, \mathbb{P}_X; \mathbb{R}^m) \mid \sigma(u) \subset \sigma(h) \right\}$$

General result

$$\mathcal{P}(h) \begin{cases} \min J(u) = \mathbb{E}\left[j(u(X), X)\right] \\ u \in L^2(\mathbb{R}^d, \mathbb{P}_X; \mathbb{R}^m) \\ \sigma(u) \subset \sigma(h) \end{cases}$$

Let $(h_n)_{n\in\mathbb{N}}$ be a sequence of random variable.

- J is continuous with respect to the norm topology;
- $\sigma(h_n) \subset \sigma(h)$;
- $(h_n)_{n\in\mathbb{N}}$ converge in probability to h;

then $\lim_{n\to\infty} |V(h_n) - V(h)| = 0$.

$$\mathcal{P}(h_n) \begin{cases} \min \mathbb{E}\left[j(u(X), X)\right] \\ u \text{ is } \sigma(h_n) \end{cases}$$

General result with quantization

Let $(Q_p : \mathbb{R}^\ell \to \mathbb{R}^\ell)_{p\in\mathbb{N}}$ be a sequence of random variable.

- J is continuous with respect to the norm topology;
- the sequence Q_n converges pointwise to identity;

then $\lim_{p\to\infty} |V(Q_p(h)) - V(h)| = 0$.

$$\mathcal{P}(Q_p(h)) \begin{cases} \min \mathbb{E}\left[j(u(X), X)\right] \\ u \text{ is } \sigma(Q_p(h)) \end{cases}$$

The discrete problem

It is easy now to formulate a discrete problem:

$$j_p(u,\cdot) = \mathbb{E}\left[j(u,\cdot) \mid Q_p(h)\right]$$

$$\mathcal{P}_n(Q_p(h)) \Leftrightarrow \begin{cases} \min \dfrac{1}{n} \displaystyle\sum_{i=1}^{I_p} \sum_{k=1}^{n} j(u_i, X^k)\mathbb{I}_{\{Q_p(h)=y_i\}}(X^k) \\ (u_1, \ldots, u_{I_p}) \in (\mathbb{R}^m)^{I_p} \end{cases}$$

Conclusion

The class of problems for which quantization technique can be applied (or scenarios tree technique for multistage program with non-anticipativity constraint) contains the space of continuous objective functions J;

Comments

1. Due to the need of separable set of functions it doesn't seem relevant to consider the Neveu's metric for numerical applications. It would be better to consider a more intrinsic (eigentlich) metric like the following one:

$$d(\mathcal{B}', \mathcal{B}'') = \inf \left\{ 1 - \mathbb{P}_X(A) \mid \mathcal{B}' \cap A = \mathcal{B}'' \cap A, A \in \mathcal{B}(\mathbb{R}^d) \right\}$$

   (introduced by KB 2004).
2. All these results can be written for multistage problems (see C.Strugarek PhD.)

Artstein, Z. (1991).
Sensitivity to σ-fields of information in stochastic allocation.
Stochastic Stochastics Rep., (36):41–63.

Dupačová, J. and Wets, R.-B. (1988).
Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems.
Ann. Stat., 16(4):1517–1549.

Dynkin, E. and Evstigneev, I. (1976).
Regular conditional expectations of correspondences.
Theory Probab. Appl., 21(2):325–338.

Rockafellar, R. and Wets, R.-B. (1991).
Scenarios and policy aggregation in optimization under uncertainty.
Math. Oper. Res., 16(1):119–147.

Danke für Ihre Beachtung