Epiconvergence of relaxed stochastic optimization problem

V. Leclère (CERMICS, ENPC).

October 4, 2013
1. Dual Approximate Dynamic Programming (DADP) algorithm
   - Problem statement
   - Solving the decomposed problem

2. Epiconvergence of relaxed Problems
   - Dual approximation as constraint relaxation
   - Some useful convergences
   - Convergence result

3. Applications: some continuous $J$ and $\Theta$
Model presentation

Parameters:
- storage level $X^i_t$,
- hydroturbine outflows $U^i_t$,
- external inflows $W^i_t$.

Objective function:

$$
E \left( \sum_{i=1}^{N} \sum_{t=0}^{T-1} L^i_t(X^i_t, U^i_t, W_t) + K_i(X^i_T) \right)
$$

Figure: The river chain model
Decomposition Principle
Contents

1 Dual Approximate Dynamic Programming (DADP) algorithm
   - Problem statement
   - Solving the decomposed problem

2 Epiconvergence of relaxed Problems
   - Dual approximation as constraint relaxation
   - Some useful convergences
   - Convergence result

3 Applications: some continuous $J$ and $\Theta$
Abstract formulation of the problem

\[
\min_{X, U} \mathbb{E} \left( \sum_{t=0}^{T-1} \sum_{i=1}^{N} L_t^i(X_t^i, U_t^i, W_t) + K^i(X_T^i) \right)
\]

(dynamic) \quad \forall i, \quad X_{t+1}^i = f_t^i(X_t^i, U_t^i, W_t)

(bounds constraint) \quad \forall i, \quad X_0^i = x_0^i

(information constraint) \quad \forall i, \quad U_t^i \in \mathcal{U}_t^{ad}

\quad \forall i, \quad U_t^i \leq \mathcal{F}_t \quad \text{i.e. } U_t^i \text{ is } \mathcal{F}_t \text{ meas.}

(coupling constraint) \quad \sum_{i=1}^{N} \theta_t^i(U_t^i) = 0 \quad \text{a.s.}
Primal problem

\[
\min_{X, U} \sum_{i=1}^{N} \mathbb{E} \left( \sum_{t=0}^{T} L_t^i (X_t^i, U_t^i, W_t) + K_t^i (X_T^i) \right)
\]

\[
\forall i, \quad X_{t+1}^i = f_t^i (X_t^i, U_t^i, W_t)
\]

\[
\forall i, \quad X_0^i = x_0^i, \quad U_t^i \in U_{t, i}^{ad}, \quad U_t^i \leq F_t
\]

\[
\sum_{i=1}^{N} \theta_t^i (U_t^i) = 0
\]
Primal problem

\[
\begin{align*}
\min_{\mathbf{X}, \mathbf{U}} \quad & \max_{\lambda} \quad \sum_{i=1}^{N} \mathbb{E}\left( \sum_{t=0}^{T} L_{t}^{i}(\mathbf{X}_{t}^{i}, \mathbf{U}_{t}^{i}, \mathbf{W}_{t}) + \langle \lambda_{t}, \theta_{t}^{i}(\mathbf{U}_{t}^{i}) \rangle + K^{i}(\mathbf{X}_{T}^{i}) \right) \\
\forall i, \quad & \mathbf{X}_{t+1}^{i} = f_{t}^{i}(\mathbf{X}_{t}^{i}, \mathbf{U}_{t}^{i}, \mathbf{W}_{t}) \\
\forall i, \quad & \mathbf{X}_{0}^{i} = x_{0}^{i}, \quad \mathbf{U}_{t}^{i} \in \mathcal{U}_{t,i}^{ad}, \quad \mathbf{U}_{t}^{i} \preceq \mathcal{F}_{t}.
\end{align*}
\]
Dual problem

\[
\begin{align*}
\max_{\lambda} \min_{X,U} & \sum_{i=1}^{N} \mathbb{E} \left( \sum_{t=0}^{T} L^i_t(X^i_t, U^i_t, W_t) + \langle \lambda_t, \theta^i_t(U^i_t) \rangle + K^i(X^i_T) \right) \\
\forall i, & \quad X^i_{t+1} = f^i_t(X^i_t, U^i_t, W_t) \\
\forall i, & \quad X^i_0 = x^i_0, \quad U^i_t \in U^i_{ad_t}, \quad U^i_t \preceq F_t.
\end{align*}
\]
Decomposed problem

\[
\max_{\lambda} \sum_{i=1}^{N} \min_{x_i, u_i} \mathbb{E} \left( \sum_{t=0}^{T} L^i_t(x^i_t, u^i_t, w_t) + \langle \lambda_t, \theta^i_t(u^i_t) \rangle + K^i(x^i_T) \right)
\]

\[
x^i_{t+1} = f^i_t(x^i_t, u^i_t, w_t)
\]

\[
x^i_0 = x^i_0, \quad u^i_t \in U^ad_{t,i}, \quad u^i_t \preceq F_t.
\]
The price decomposition algorithm is done as follows, given a multiplier process \((\lambda_t)_{t \in [0, T-1]}\):

1. solve \(N\) problems with only one dam,
2. update the multiplier by a gradient step.

We need to specify:

- How to solve the one-dam problem?
- How to update the multiplier?
1 Dual Approximate Dynamic Programming (DADP) algorithm
   - Problem statement
   - Solving the decomposed problem

2 Epiconvergence of relaxed Problems
   - Dual approximation as constraint relaxation
   - Some useful convergences
   - Convergence result

3 Applications: some continuous $J$ and $\Theta$
Recalling the one-dam problem

Recall that the inner problem reads:

$$\min_{X^i_t, U^i_t} \mathbb{E}\left( \sum_{t=0}^{T} L_t^i(X^i_t, U^i_t, W_t) + \langle \lambda_t^{(k)}, \theta_t^i(U^i_t) \rangle + K^i(X^i_T) \right)$$

$$X^i_{t+1} = f_t^i(X^i_t, U^i_t, W_t)$$

$$X^i_0 = x^i_0, \quad U^i_t \in U^{ad}_{t,i}, \quad U^i_t \leq F_t$$

This problem can be solved by Dynamic Programming with the extended state $$(X^i_t, \lambda^{(k)}_{[t]})$$, where $$\lambda^{(k)}_{[t]} = (\lambda_0^{(k)}, \ldots, \lambda_t^{(k)})$$. 
Using a conditional expectation: \( \lambda \rightsquigarrow \mathbb{E}(\lambda \mid Y_t) \)

Idea behind the Dual Approximate Dynamic Programming algorithm:

\[
\lambda_t \rightsquigarrow \mathbb{E}(\lambda_t \mid Y_t).
\]

We will see that it is equivalent to

\[
\sum_{i=1}^{N} \theta^i_t(U^i_t) = 0 \quad \mathbb{P} - a.s. \rightsquigarrow \mathbb{E}\left( \sum_{i=1}^{N} \theta^i_t(U^i_t) \mid Y_t \right) = 0 \quad \mathbb{P} - a.s.
\]

If \((Y_t)_{t \in [0,T-1]}\) is a Markovian process, then the problem can be solved by Dynamic Programming with the state \((X_t^i, Y_t)\) which is numerically tractable if \(Y\) lives in a reasonably "small" state space.
Using a conditional expectation: $\lambda \rightsquigarrow \mathbb{E}(\lambda \mid Y_t)$

Idea behind the Dual Approximate Dynamic Programming algorithm:

$$\lambda_t \rightsquigarrow \mathbb{E}(\lambda_t \mid Y_t).$$

We will see that it is equivalent to

$$\sum_{i=1}^{N} \theta^i_t(U^i_t) = 0 \quad \mathbb{P} - a.s. \rightsquigarrow \mathbb{E}\left(\sum_{i=1}^{N} \theta^i_t(U^i_t) \mid Y_t\right) = 0 \quad \mathbb{P} - a.s.$$

If $(Y_t)_{t \in [0, T-1]}$ is a Markovian process, then the problem can be solved by Dynamic Programming with the state $(X^i_t, Y_t)$ which is numerically tractable if $Y$ lives in a reasonably “small” state space.
In an abstract point of view where $J$ and $\Theta$ incorporate the dynamics of the system the original problem is

$$\min_{U \in \mathcal{U}} J(U)$$

s.t. $\Theta(U) = 0$

where $J(U) = \mathbb{E}(j(U))$, which can be written as

$$\min_{U \in \mathcal{U}} \max_{\lambda} J(U) + \mathbb{E}(\langle \lambda, \Theta(U) \rangle)$$
Subsituting $\lambda$ by $E(\lambda|Y)$ gives

$$\min_{U \in U} \max_{\lambda} J(U) + E(\langle E(\lambda|Y), \Theta(U) \rangle)$$

which is equal to

$$\min_{U \in U} \max_{\lambda} J(U) + E(\langle \lambda, E(\Theta(U)|Y) \rangle)$$

and thus equivalent to

$$\min_{U \in U} J(U)$$

$$s.t. \ E(\Theta(U)|Y) = 0$$
Contents

1 Dual Approximate Dynamic Programming (DADP) algorithm
   - Problem statement
   - Solving the decomposed problem

2 Epiconvergence of relaxed Problems
   - Dual approximation as constraint relaxation
   - Some useful convergences
   - Convergence result

3 Applications: some continuous $J$ and $\Theta$
The approximation studied

We consider the stochastic optimization problem:

\[(\mathcal{P}) \quad \min_{\mathbf{U} \in \mathcal{U}} \quad J(\mathbf{U}) ,\]
\[s.t. \quad \Theta(\mathbf{U}) \in -C .\]

And its approximation

\[(\mathcal{P}_n) \quad \min_{\mathbf{U} \in \mathcal{U}} \quad J(\mathbf{U}) ,\]
\[s.t. \quad \mathbb{E}(\Theta(\mathbf{U})|\mathcal{F}_n) \in -C .\]

We give convergence results of the approximation \((\mathcal{P}_n)\) toward the original problem \((\mathcal{P})\).
1 Dual Approximate Dynamic Programming (DADP) algorithm
   - Problem statement
   - Solving the decomposed problem

2 Epiconvergence of relaxed Problems
   - Dual approximation as constraint relaxation
   - Some useful convergences
   - Convergence result

3 Applications: some continuous $J$ and $\Theta$
Kudo convergence of $\sigma$–algebras in a nutshell

A sequence $(\mathcal{F}_n)_{n \in \mathbb{N}}$ of $\sigma$–algebras Kudo-converges toward $\mathcal{F}_\infty$, iff

$$\forall X \in L^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}), \quad \mathbb{E}(X \mid \mathcal{F}_n) \rightarrow_{L^1} \mathbb{E}(X \mid \mathcal{F}_\infty).$$

The following result is shown by Piccinini (1998):

**Proposition**

Assume that the sequence of $\sigma$-algebras Kudo-converges and the sequence of random variable converges in $L^p$:

- $\mathcal{F}_n \rightarrow \mathcal{F}_\infty$,
- $X_n \rightarrow_{L^p} X$,

then

$$\mathbb{E}(X_n \mid \mathcal{F}_n) \rightarrow_{L^p} \mathbb{E}(X \mid \mathcal{F}_\infty).$$
Some properties of Kudo-convergence

A few properties on the Kudo-convergence of $\sigma$-algebras:

1. Kudo-convergence’s topology is metrizable;

2. the set of $\sigma$-fields generated by partition of $\Omega$ is dense in the set of all $\sigma$-algebras;

3. if a sequence of random variables $(X_n)_{n \in \mathbb{N}}$ converges in probability toward $X$ and for all $n \in \mathbb{N}$ we have $\sigma(X_n) \subseteq \sigma(X)$, then we have the Kudo-convergence of $(\sigma(X_n))_{n \in \mathbb{N}}$ toward $\sigma(X)$.
Painlevé-Kuratovski convergence of set

- $E$ is a topologic space,
- $A_n \subset E$,
- $J_n : E \to \mathbb{R} \cup \{+\infty\}$.

We define outer and inner limits of sequence of subset of a topological set $E$.

$$\lim_{n} A_n = \{ x \in E \mid \forall n \in \mathbb{N}, \ x_n \in A_n, \ \lim_{k \to \infty} x_n = x \} ,$$

$$\overline{\lim}_{n} A_n = \{ x \in E \mid \exists (n_k)_{k \in \mathbb{N}}, \ \forall k \in \mathbb{N}, \ x_{n_k} \in A_{n_k}, \ \lim_{k \to \infty} x_{n_k} = x \} .$$

And $(A_n)_{n \in \mathbb{N}}$ converges toward $A$ iff

$$\overline{\lim}_{n} A_n = \lim_{n} A_n = A .$$
Epi-convergence in a nutshell

The epiconvergence of the sequence of functions $J_n : E \rightarrow \mathbb{R} \cup \{+\infty\}$ toward $J$ is given as the convergence of their epigraphs:

$$J_n \rightarrow^e J \quad \text{iff} \quad \overline{\lim}_n \text{epi}(J_n) = \overline{\lim}_n \text{epi}(J_n) = \text{epi}(J).$$

Epiconvergence is the right notion of convergence in optimization as epiconvergence of $J_n$ toward $J$ almost implies:

- the convergence of $\min J_n$ toward $\min J$,
- the convergence of $\arg \min J_n$ toward $\arg \min J$.

More information can be found in the book by Rockafellar and Wets (1995).
1 Dual Approximate Dynamic Programming (DADP) algorithm
   - Problem statement
   - Solving the decomposed problem

2 Epiconvergence of relaxed Problems
   - Dual approximation as constraint relaxation
   - Some useful convergences
   - Convergence result

3 Applications: some continuous $J$ and $\Theta$
Convergence result

Theorem

Assume that

- the set of controls $\mathcal{U}$ is endowed with a topology $\tau$, 
- the image space of the constraint operator $\Theta$, is $\mathcal{V} = L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathcal{V})$, with $p \in [1, \infty)$, (strong or weak topo.),
- the cone of constraint $\mathcal{C}$ is such that $\mathbb{E}(\mathcal{C} \mid \mathcal{F}_n) \subset \mathcal{C}$,
- the sequence of $\sigma$-algebras $(\mathcal{F}_n)_{n \in \mathbb{N}}$ converges towards $\mathcal{F}$,
- the constraint operator $\Theta : (\mathcal{U}, \tau) \rightarrow (\mathcal{V}, \tau_{L^p})$ and the objective operator $J : (\mathcal{U}, \tau) \rightarrow \mathbb{R}$ are continuous.

Then $\tilde{\mathcal{J}}_n$ epi-converges toward $\tilde{\mathcal{J}}$, where

$$
\tilde{\mathcal{J}}_n(\mathcal{U}) = \begin{cases} 
J(\mathcal{U}) & \text{if } \mathcal{U} \text{ satisfies the constraint of } (\mathcal{P}_n), \\
+\infty & \text{otherwise}
\end{cases}
$$
Practical consequences

Consider a sequence of \( \epsilon_n \)-solution of \( (P_n) \) denoted \( U_n \), i.e.

\[
\tilde{J}_n(U_n) < \inf_{U \in \mathcal{U}} \tilde{J}_n(U) + \epsilon_n.
\]

Under the conditions of the preceding theorem for every converging sub-sequence \((U_{n_k})_{k \in \mathbb{N}}\), we have

\[
\tilde{J}(\lim_k U_{n_k}) = \min_{U \in \mathcal{U}} \tilde{J}(U) = \lim_k \tilde{J}_{n_k}(U_{n_k}).
\]

Which means that

- the optimal value of the relaxed problem converges toward the optimal value of the original problem,
- and if the optimal strategies \( U_n \) of the relaxed problems converges, then their limit is a solution of the original problem.
Contents

1 Dual Approximate Dynamic Programming (DADP) algorithm
   - Problem statement
   - Solving the decomposed problem

2 Epiconvergence of relaxed Problems
   - Dual approximation as constraint relaxation
   - Some useful convergences
   - Convergence result

3 Applications: some continuous $J$ and $\Theta$
Lemma

Let $\Theta : E \to F$, where $(E, \tau_P)$ is a space of random variables endowed with the topology of convergence in probability, and $(F, \tau)$ is a topological space. Assume that $\Theta$ is such that

$$U_n \xrightarrow{a.s.} U \implies \Theta(U_n) \to_{\tau} \Theta(U).$$

Then $\Theta$ is a continuous operator from $(E, \tau_P)$ into $(F, \tau)$.

Remarks:

- There is no topology $\tau_{a.s.}$.
- However $\tau_P$ is very close to what would be $\tau_{a.s.}$.
Collection of continuous operators

$j$ and $\theta$ are assumed to be continuous and bounded. Recall that $V = L^p(\Omega, \mathcal{F}, \mathbb{P})$.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Hypothesis</th>
<th>$(U, \tau)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J(U) = \mathbb{E}(j(U))$</td>
<td>$\rho$ c.r.m</td>
<td>$U = L^0(\Omega, \mathcal{F}, \mathbb{P}; \tau_L)$</td>
</tr>
<tr>
<td>$J(U) = \rho(j(U))$</td>
<td></td>
<td>$U = L^0(\Omega, \mathcal{F}, \mathbb{P}; \tau_P)$</td>
</tr>
<tr>
<td>$\Theta(U) = \theta(U)$</td>
<td>$\rho$ cond. c.r.m</td>
<td>$U = L^0(\Omega, \mathcal{F}, \mathbb{P}; \tau_P)$</td>
</tr>
<tr>
<td>$\Theta(U) = \mathbb{E}(U</td>
<td>B) - U$</td>
<td>$p' \geq p$</td>
</tr>
<tr>
<td>$\Theta(U) = \rho(U)$</td>
<td></td>
<td>$U = L^p(\Omega, \mathcal{F}, \mathbb{P}; \tau_P)$ cont. distribution, $\tau_L$</td>
</tr>
<tr>
<td>$\Theta(U) = \text{VaR}_\alpha(\theta(U))$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

c.r.m = convex risk measures
remark: there is no $\tau_{a.s.}$.
Conclusion

- We can apply price-decomposition methods in stochastic setting. However the subproblem have the same complexity that the original one.
- One idea is to approximate the stochastic multiplier process by its conditional expectation. This is equivalent to solve an approximate problem, where the almost sure constraint is relaxed in a conditional expectation constraint.
- We give epi-convergence results of the relaxation toward the original problem which relies on:
  - stability of cone of constraints w.r.t conditionning,
  - Kudo-convergence of the conditionning $\sigma$-algebras,
  - continuity of cost and constraint operators.
- The results are the same for a finite number of constraint operator (with different conditionning).
- Finally most cost and constraint operators found in stochastic optimization are continuous for the right topology (often $T_P$).
Conclusion

- We can apply price-decomposition methods in stochastic setting. However the subproblem have the same complexity that the original one.
- One idea is to **approximate the stochastic multiplier process** by its conditional expectation. This is equivalent to solve an approximate problem, where the almost sure constraint is relaxed in a conditional expectation constraint.
- We give **epi-convergence results** of the relaxation toward the original problem which relies on:
  - stability of cone of constraints w.r.t conditionning,
  - Kudo-convergence of the conditionning σ-algebras,
  - continuity of cost and constraint operators.
- The results are the same for a finite number of constraint operator (with different conditionning).
- Finally most cost and constraint operators found in stochastic optimization are continuous for the right topology (often $\tau^p$).