Stochastic Optimization - 2 hours

Vincent Leclère

16/01/2018

Name: ____________________

By "type of problem", we mean "Linear Program" denoted (LP), "Mixed Integer Linear Program" denoted (MILP), "Quadratic Program" denoted (QP) (linear constraint, convex quadratic objective, continuous variables), "Mixed Integer Quadratic Program" denoted (MIQP)... By size of a problem we mean the number of integer and linear variables.

1 A cutting plane algorithm for convex two stage programm (12 points)

We consider the following two stage stochastic programm

$$
\begin{aligned}
(SP) \qquad \min_{q,\boldsymbol{u}} \quad & J(q) + \mathbb{E}\left[c^T \boldsymbol{u}\right] \\
s.t. \quad & q \in Q \\
& A\boldsymbol{u} \leq \boldsymbol{b} + Bq \\
& \sigma(\boldsymbol{u}) \preceq \sigma(\boldsymbol{b})
\end{aligned}
$$

where $J : \mathbb{R}^{n_u} \to \mathbb{R}$ is a convex fonction, $c \in \mathbb{R}^{n_u}$ a deterministic cost vector, $Q \subset \mathbb{R}^{n_q}$ is a polytope (compact polyhedron), A and B deterministic matrices, and $\boldsymbol{b}$ a random vector with known discrete probability law : $\mathbb{P}(\boldsymbol{b} = b_i) = \pi_i$ for $i \in [\![1, n]\!]$.

1. (1 point) Justify that this problem is a two-stage stochastic programm by giving the first and second stage decision variable and the noise.

Solution: q is first stage decision variable, u is the second stage decision variable.

2. (2 points) Define $V(q) = \mathbb{E}\left[\hat{V}(q, \boldsymbol{b})\right]$ where

$$
\begin{aligned}
\hat{V}(q, b) = \min_{u \in \mathbb{R}^{n_u}} \quad & c^T u \\
s.t. \quad & Au \leq \boldsymbol{b} + Bq
\end{aligned}
$$

With this definition propose a simple formulation of (SP). Give an explicit (i.e. without domain of some function) relatively complete recourse condition

Solution:

$$\min_{q \in Q} J(q) + V(q)$$

RCR :

$$\forall q \in Q, \quad \forall i \in [\![i, n]\!], \quad \exists u_i \in \mathbb{R}^{n_u} \quad \text{such that} \quad Au_i \leq b_i + Bq$$

3. (3 points) For $i \in [\![1, n]\!]$, and $q \in Q$, write a linear program with value $\hat{V}(q, b_i)$ and show how to obain $\lambda_i \in \partial_q \hat{V}(q, b_i)$ from an optimal solution.

Solution: By RCR there is no duality gap, and the dual of $\hat{V}(q, b_i)$ is

$$\hat{V}(q, b_i) = \max_{\mu} \quad -\mu^T(b_i + Bq)$$
$$(c + A^T\mu) = 0$$

thus $-B^T\mu^\sharp \in \partial_q \hat{V}(q, b_i)$.

4. (2 points) For $q \in Q$, explain how to efficiently obtain $\alpha \in \mathbb{R}^{n_q}$ and $\beta \in \mathbb{R}$ such that, for all $q' \in Q$, $\alpha^T q' + \beta \leq V(q')$ and $\alpha^T q + \beta = V(q)$. Precise the number, type and size of problems solved.

Solution: From the previous question we have that $\alpha_i = \lambda_i$ and $\beta_i = \theta_i - \alpha_i^T q$. Then $\alpha = \sum_{i=1}^n \pi_i \alpha_i$ and $\beta = \sum_{i=1}^n \pi_i \beta_i$. Thus n LP of n_u variables are solved.

5. (3 points) Assuming that you are only allowed an LP solver, propose an algorithm solving (SP) and guarantee its convergence (using the results from the course). Precise the number, type and size of problems solved at each iteration.

Solution: For a given q a subgradient of $J + V$ is obtained as $\nabla J(q) + \alpha$. Hence we can apply Kelley's cutting plane algorithm with

Data: Initial point q^0
Set $J^{(0)} \equiv -\infty$;
for $k \in \mathbb{N}$ **do**
  for $i \in [\![1, n]\!]$ **do**
    solve the slave program to obtain α^i and θ^i
  end
  Compute α and β ;
  Define a cut $\mathcal{C}^{(k)} : q \mapsto J(q^{(k)}) + \theta + \langle \nabla J(q^{(k)}) + \alpha^{(k)}, q - q^{(k)} \rangle$;
  Update the lower approximation $J^{(k+1)} = \max\{J^{(k)}, \mathcal{C}^{(k)}\}$;
  Solve $(P^{(k)})$: $\min_{q \in Q} J^{(k+1)}(q)$;
  Set $\underline{v}^{(k)} = val(P^{(k)})$;
  Select $q^{(k+1)} \in sol(P^{(k)})$;
end

6. (2 points) Give the pseudo-code of a multicut version of your algorithm.

Solution:

Data: Initial point q^0
Set $J^{(0)} \equiv -\infty$;
for $k \in \mathbb{N}$ **do**
  for $i \in [\![1, n]\!]$ **do**
    solve the slave program to obtain α_k^i and θ_k^i ;
  end
  Solve

$$
\begin{aligned}
\min_{q \in Q} \quad & \eta + \sum_{i=1}^{n} \theta_i \\
s.t. \quad & \eta \geq \nabla(J)(q^\kappa) \cdot (q - q^{(\kappa)} + J(q^\kappa) && \forall \kappa \leq k \\
& \theta_i \geq \alpha_\kappa^i \cdot q + \beta_\kappa^i && \forall \kappa \leq k, \forall i
\end{aligned}
$$

  Select $q^{(k+1)} \in sol(P^{(k)})$;
end

2 A multi-stock problem (17 points)

Consider a chemical industry that use 10 different raw materials, to produce 5 different finished products. Denote by $a_{r,f}$ the quantity (in tons) of raw material r required to produce one ton of finished product f. At any point in time, the company cannot have more than 7 tons of each raw material. Each month, at the beginning of the month, the company discover the vector of cost $\boldsymbol{c}_t$ at which it can order any mix of raw material for a maximum of 20 tons in total, and the vector of prices $\boldsymbol{p}_t$ at which finished product will be sold (every finished product is sold). Material ordered at the beginning of the month arrive at the end of the month. The prices $\boldsymbol{p}_t$ and $\boldsymbol{c}_t$ are highly correlated, but independent in time and the law of $\boldsymbol{\xi}_t = (\boldsymbol{p}_t, \boldsymbol{c}_t)$ is finitely supported with a support Ξ of size 100 (we denote $\pi_\xi = \mathbb{P}(\boldsymbol{\xi} = \xi)$ for all $\xi \in \Xi$).

The company want to optimize its expected gain over one year, starting with 5 tons of each raw material. Stock at the end of the year are considered lost. We will denote by $b_t \in \mathbb{R}^{10}$ the quantity of raw material bought at month t, by $v_t \in \mathbb{R}^5$ the quantity of finished product sold during month t, and x_t the quantity of raw material available at the beginning of month t.

1. (0.5 points) Give a matrix A such that we have $\boldsymbol{x}_{t+1} = \boldsymbol{x}_t + \boldsymbol{b}_t - A\boldsymbol{v}_t$.

Solution: $A = (a_{r,f})$.

2. (2.5 points) Write the problem as a multistage stochastic optimization problem. Precise the information structure.

Solution:

$$
\begin{array}{lllr}
\min & \mathbb{E}\Big[\sum_{t=1}^{12} \boldsymbol{c}_t\boldsymbol{b}_t - \boldsymbol{p}_t\boldsymbol{v}_t\Big] & & \\
s.t. & \boldsymbol{x}_{t+1} = \boldsymbol{x}_t + \boldsymbol{b}_t - A\boldsymbol{v}_t, \quad \boldsymbol{x}_0 = x_0, & \forall t & \text{dynamic} \\
 & 0 \leq \boldsymbol{x}_t^r \leq 7, \qquad \forall r \in [\![1,5]\!], \quad \forall t & & \text{stock constraints} \\
 & A\boldsymbol{v}_t \leq \boldsymbol{x}_t, \qquad \forall t & & \text{use only available stock} \\
 & \sum_{r=1}^{10} \boldsymbol{b}_t^r \leq 20 \qquad \forall t & & \text{global buying capacity} \\
 & \boldsymbol{v}_t \geq 0, \boldsymbol{b}_t \geq 0 & & \\
 & \boldsymbol{b}_t, \boldsymbol{v}_t \preceq \sigma\big(\{\boldsymbol{c}_\tau, \boldsymbol{p}_\tau\}_{\tau \leq t}\big) & & \text{information constraint}
\end{array}
$$

We are in a hazard-decision framework.

3. (2 points) Justify that we can apply Dynamic Programming to this problem. Define adequate Bellman operators $\hat{\mathcal{B}}_t$ and $\mathcal{B}_t$ and give the associated Bellman Equation (beware of constraints on the stock).

Solution: By time-independence of noises we can apply Dynamic Programming. We define

$$
\begin{array}{rll}
\hat{\mathcal{B}}_t(R)(x,c,p) = \min\limits_{b,v,y} & cb - bv + R(y) & \\
s.t. & y = x + b - Av & \\
 & 0 \leq y \leq 7 & \\
 & Av \leq x & \\
 & \sum_r b_r \leq 20 & b \geq 0, v \geq 0
\end{array}
$$

and $\mathcal{B}_t(R)(x) = \mathbb{E}\big[\hat{\mathcal{B}}_t(x, \boldsymbol{c}_t, \boldsymbol{p}_t)\big]$. The Bellman equation is $V_t = \mathcal{B}_t(V_{t+1}) + \mathbb{I}_{0\leq x\leq 7}$, with $V_{13} = 0$.

4. Dynamic Programming for a discretized version. In this question we consider that the state x is constrained to be a vector of integer.

   (a) (2 points) We call $\Psi(x_t, x_{t+1}, \xi)$ the minimal cost of going from state x_t at the beginning of month t to state x_{t+1} knowing that the prices are given by $\xi = (p_t, c_t)$. Write a mathematical program computing $\Psi(x_t, x_{t+1}, \xi)$, and precise its type.

   Solution: It's a Linear Programm

$$\begin{aligned}
\Psi(x_t, x_{t+1}, (p, b)) = \quad & \min_{b,v} \quad cb - bv \\
& s.t. \quad x_{t+1} = x_t + b - Av \\
& \qquad\; Av \leq x_t \\
& \qquad\; \sum_r b_r \leq 20 \\
& \qquad\; b \geq 0, v \geq 0
\end{aligned}$$

   (b) (3 points) Using function Ψ, propose a Dynamic Programming algorithm on the discretized problem. How many time is the function Ψ called? Is it a reasonable approach?

   Solution:

```
V_13 ≡ 0 ;
for t = 12 → 1 do
    for x ∈ X do
        v = 0; for ξ ∈ Ξ do
            v_ξ = −∞ ;
            for y ∈ X do
                v_ξ = max(v_ξ, Ψ(x, y, ξ) + V_{t+1}(y));
            end
        end
        v = v + π_ξ v_ξ;
        V_t(x) = v ;
    end
end
```

   Ψ is called roughly $12 \times 8^{10} \times 100 \times 8^{10}$ times. Which is unreasonnable.

5. Stochastic Dual Dynamic Programming

   (a) (1 point) Justify that SDDP is adapted to the continuous problem (in particular the recourse assumption).

   Solution: Dynamic and cost are linear, constraints are polyhedral. Noise are finitely supported and time independent. We are in an extended relatively complete recourse assumption as we can always buy additional raw material required, or consume extra raw material (assuming that every raw material is used for at least one finished product).

   (b) (2 points) Consider the forward pass k of the SDDP algorithm at time t, give the LP problem solved ("one-stage-one-alea" problem).

Solution:

$$\begin{aligned}
\min_{b,v\geq 0} \quad & cb - bv + \theta \\
s.t. \quad & x_{t+1} = x_t + b - Av \\
& Av \leq x_t \\
& \sum_r b_r \leq 20 \\
& \theta \geq \alpha^\kappa \cdot x_{t+1} + \beta^\kappa & \forall \kappa \leq k
\end{aligned}$$

(c) (1 point) How many times is this (type of) problem solved during iteration k of SDDP?

Solution: The problem is solved 12 times in the forward pass, and 1200 times in the backward pass.

(d) (1 point) What is the "one-stage-one-alea" LP problem solved for a multicut version of the problem?

Solution:

$$\begin{aligned}
\min_{b,v\geq 0} \quad & cb - bv + \sum_{\xi\in\Xi} \pi_\xi \theta_\xi \\
s.t. \quad & x_{t+1} = x_t + b - Av \\
& Av \leq x_t \\
& \sum_r b_r \leq 20 \\
& \theta_\xi \geq \alpha_\xi^\kappa \cdot x_{t+1} + \beta_\xi^\kappa & \forall \kappa \leq k, \forall \xi
\end{aligned}$$