Data Driven Robust Optimization Exam

19/03/2018

The exam is made of two independant parts. If necessary, you can admit the results of previous questions. All documents authorized, all electronical device forbidden.

Some usefull recalls.

1. An SOCP constraint take the form $a^T x + b + \|c^T x + d\| \leq 0$.
2. We have, for any $\alpha \geq 0$, $(\alpha f)^\star(x) = \alpha f^\star(x/\alpha)$.
3. Function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is convex iff the perspective function $\varphi : \mathbb{R}^d \times \mathbb{R}^+_* \to \mathbb{R} \cup \{+\infty\}$ with $\varphi(x,t) = tf(x/t)$ is convex.
4. If $(g_i)_{i \in [\![1,d]\!]}$ are concave functions with $\cap_{i=1}^d ri(\text{dom}(g_i) \neq \emptyset$ we have

$$\Big(\sum_{i=1}^d g_i(\cdot)\Big)_\star(v) = \sup_{(v^i)_{i\in[\![1,d]\!]}} \Big\{ \sum_{i=1}^d (g_i)_\star(v^i) \;\Big|\; \sum_{i=1}^d v^i = v \Big\}$$

5. The value at risk of level ε is defined by

$$VaR_\varepsilon^{\mathbb{P}}(\boldsymbol{X}) := \inf \big\{ t \mid \mathbb{P}(\boldsymbol{X} \leq t) \geq 1 - \varepsilon \big\}$$

Entropy constrained optimization

1. Preliminary analysis

   We are interested in the following averaged entropy contraint $f(u,x) := -\sum_{i=1}^d x_i u_i ln(u_i) - C \leq 0$, were all matrices $C > 0$. Where f is defined on $\mathbb{R}^d \times \mathbb{R}^d_+$.

   (a) (1 point) Let $f_i(u_i) = -u_i ln(u_i)$. Compute $(f_i)_\star(v_i) := \inf_{u_i \in \mathbb{R}} v_i u_i - f_i(u_i)$

   > **Solution:** f_i is strictly concave. By differentiation $v_i + ln(u_i^\sharp) + 1 = 0$ thus $u_i^\sharp = e^{-1-v_i}$ and $(f_i)_\star(v_i) = -e^{-1-v_i}$.

   (b) (1 point) Show that

$$f_\star(v,x) := \inf_u v^T u - f(u,x) = \sup_{(v^i)_{i\in[\![1,d]\!]}} \Big\{ -e^{-1} \sum_{i=1}^d x_i e^{-v_i/x_i} + C \;\Big|\; \sum_{i=1}^d v^i = v \Big\}$$

   > **Solution:** By conjugation of the sum we have
   >
   > $$f_\star(v,x) = \sup_{(v^i)_{i\in[\![1,d]\!]}} \Big\{ (x_i f_i)_\star(v_i) + C \;\Big|\; \sum_{i=1}^d v^i = v \Big\}$$
   >
   > And, for any $x_i > 0$, $(x_i f_i)_\star(v^i) = x_i (f_i)_\star(\frac{v^i}{x_i})$.

2. Implementation

We are interested in the following problem

$$\min_{x\in\mathbb{R}^d_+} \quad c^T x \tag{1a}$$

$$s.t. \quad \mathbb{P}(-\sum_{i=1}^{d} x_i u_i ln(u_i) \leq C) \geq 0.95 \tag{1b}$$

$$Ax \leq b \tag{1c}$$

where $\tilde{u}$ is a random variable. We have a sample of 100 realizations of $\tilde{u}$, where $\|\tilde{u}\|_\infty \leq 10$ with an empirical mean $\bar{u}$ and variance Σ.

(a) (1 point) Show that $f_\star(v, x) \geq s$ is equivalent to

$$\begin{cases} \sum_{i=1}^{d} s_i \leq e(C - s) \\ x_i e^{-v_i/x_i} \leq s_i \qquad \forall i \\ \sum_{i=1}^{d} v_i = v \end{cases}$$

Solution: By the previous question we need v_i such that

$$\begin{cases} -e^{-1} \sum_{i=1}^{d} x_i e^{-v_i/x_i} \geq s - C \\ \sum_{i=1}^{d} v_i = v \end{cases}$$

(b) (2 points) Leveraging the CS test, explicit a convex optimisation problem (P) whose solution is a feasible solution for Problem 1 with 90% confidence (in the sampling). Is it an SOCP problem ?

Solution:

$$\begin{aligned} \min \quad & c^T x \\ s.t. \quad & Ax \leq b, x \geq 0 \\ & \sum_{i=1}^{d} s_i \leq e(C - s) \\ & x_i e^{-v_i/x_i} \leq s_i \qquad \forall i \\ & \sum_{i=1}^{d} v_i = v \\ & t - s \leq 0 \\ & \bar{u}^T v + \Gamma_1 \|v\|_2 + \sqrt{\frac{1-\varepsilon}{\varepsilon}} \sqrt{v^T(\Sigma + \Gamma_2 I)v} \leq t \end{aligned}$$

where $(\Gamma_1, \Gamma_2) = (2 + \sqrt{2\ln(10)}, 20(2 + \sqrt{2\ln(20)})) \approx (4.15, 89)$, $\varepsilon = 0.05$ and $\sqrt{\frac{1-\varepsilon}{\varepsilon}} \approx 4.36$. It is not an SOCP problem because of the exponential in the constraint.

3. Constraint generation

We would like to simplify the constraint

$$x_i e^{-v_i/x_i} \le s_i$$

through a constraint generation approach.

(a) (1 point) Show that $\varphi_i(v_i, x_i) = x_i e^{-v_i/x_i}$ (with domain $\mathbb{R} \times \mathbb{R}_*^+$) is convex, and compute its gradient.

> **Solution:** $x \mapsto e^{-x}$ is convex, thus φ_i is convex as a perspective function. We have
>
> $$\nabla\varphi(v_i^0, x_i^0) = e^{-v_i^0/x_i^0} \begin{pmatrix} -1 \\ 1 + \frac{v_i^0}{x_i^0} \end{pmatrix}$$

(b) (1 point) Construct, for $(v_i^0, x_i^0) \in \mathbb{R} \times \mathbb{R}_*^+$, an affine minorant of φ_i which is exact at (v_i^0, x_i^0).

> **Solution:** By convexity
>
> $$\varphi_i(v_i, x_i) \ge \nabla\varphi_i(v_i^0, x_i^0)^T \left(\begin{pmatrix} v_i \\ x_i \end{pmatrix} - \begin{pmatrix} v_i^0 \\ x_i^0 \end{pmatrix} \right) + \varphi_i(v_i^0, x_i^0)$$
>
> and the cut obtained is
>
> $$C_i(v_i^0, x_i^0) : (v_i, x_i) \mapsto e^{-v_i^0/x_i^0} \left[(1 + v_i^0/x_i^0)x - v \right]$$

(c) (2 points) Propose a constraint generation approach that (approximately) solves Problem (P) through a sequence of SOCP that you will explicit.

> **Solution:**
>
> 1. set $k = 0$
> 2. Choose $(v_i^0, x_i^0) \in \mathbb{R} \times \mathbb{R}_*^+$
> 3. Solve
>
> $$\begin{aligned} \min \quad & c^T x \\ s.t. \quad & Ax \le b, x \ge 0 \\ & \sum_{i=1}^d s_i \le e(C - s) \\ & C_i(v_i^\kappa, x_i^\kappa) \le s_i \qquad \forall i, \forall \kappa \le k \\ & \sum_{i=1}^d v_i = v \\ & t - s \le 0 \\ & \bar{u}^T v + \Gamma_1 \|v\|_2 + \sqrt{\frac{1-\varepsilon}{\varepsilon}} \sqrt{v^T(\Sigma + \Gamma_2 I)v} \le t \end{aligned}$$
>
>   index the solution by $k + 1$
> 4. Stop if $s_i^{k+1} \ge x_i^{k+1} e^{v i^{k+1}/x_i^{k+1}}$ otherwise increment k and go back to 3.

4. (1 point) How many sample are needed to ensure the same guarantee through a sampling approach ?

> **Solution:** $2/(0.1 \times 0.05) - 1 = 399$ samples.

Another data-driven approach

We are interested in the following optimization problem

$$\min_{x \in \mathbb{R}^d} \left\{ c^T x \quad | \quad \mathbb{P}(f(\tilde{u}, x) \leq 0) \geq 1 - \varepsilon \right\}$$

where $f(u, x)$ is a function concave in u, and convex in x.

We assume that $\tilde{u}$ is a gaussian variable of known variance. Let $S = u^1, \cdots, u^N$ be N independent realization of $\tilde{u}$. We define the empirical mean $M_N = 1/N \sum_{i=1}^N u^i$.

5. 1D case

   Assume that $\tilde{u}$ is a real valued Gaussian random variable following a law $\mathcal{N}(\mu^*, \sigma^2)$ where μ^* is unknown.

   (a) (1 point) Under the sampling probability, what is the law of M_N ? Deduce a confidence region $I(S)$ on μ such that $\mathbb{P}_S^*(\mu^* \in I(S)) = 0.95$ (choose the classical formulation - minimizing size of I). Is it an asymptotic or an exact confidence region ?

   > **Solution:** $M_N \sim \mathcal{N}(\mu^*, \sigma^2/N)$. Exact confidence region, $I(S) = [M_N \pm 1.96\sigma/\sqrt{N}]$.

   (b) (1 point) Compute $\sup_{\mathbb{P} \in I(S)} VaR_\varepsilon^{\mathbb{P}}(v^T \tilde{u})$

   > **Solution:** By symmetry of gaussian we have, $VaR_\varepsilon^{\mathbb{P}}(v^T \tilde{u}) = VaR_\varepsilon^{\mathbb{P}}(v\tilde{u}) = v\mu + \sigma|v|F_G^{-1}(1-\varepsilon)$ where F_G is the cdf of a centered reduced gaussian variable. Thus,
   >
   > $$\sup_{\mathbb{P} \in I(S)} VaR_\varepsilon^{\mathbb{P}}(v^T \tilde{u}) = vM_N + (1.96/\sqrt{N} + F_G^{-1}(1-\varepsilon))\sigma|v|.$$

   (c) (2 points) Give a set of linear constraints (in addition to the the constraint on $f_\star$) that imply, with confidence 95%, a probabilistic guarantee of level ε.

   > **Solution:**
   >
   > $$\begin{aligned} f_*(v, x) &\geq t \\ t - s &\leq 0 \\ vM_N + v(1.96/\sqrt{N} + F_G^{-1}(1-\varepsilon)) &\leq s \\ vM_N - v(1.96/\sqrt{N} + F_G^{-1}(1-\varepsilon)) &\leq s \end{aligned}$$

6. In dimension d

   $\tilde{u}$ is now Gaussian random vector of dimension d following a law $\mathcal{N}(\mu^*, \Sigma)$ where μ^* is unknown.

   (a) (2 points) Show that $\mathbb{P}_S^*(\mu^* \in M_N + z_\alpha/\sqrt{N}\Sigma^{1/2}B(0,1)) = 1 - \alpha$ for a z_α defined from the quantile of a well known law, and $B(0,1)$ being the ball in the euclidian norm of $\mathbb{R}^d$.

Solution: $M_N \sim \mathcal{N}(\mu, 1/N\Sigma)$. Thus $M_N = \mu^* + 1/\sqrt{N}\Sigma^{1/2}G$ where $G \sim \mathcal{N}(0, I)$, and $\|\sqrt{N}\Sigma^{-1/2}(M_N - \mu^*)\|_2^2$ follow a χ_2 of degree d. In particular, $\mathbb{P}_S^*(\|\sqrt{N}\Sigma^{-1/2}(M_N - \mu^*)\|_2^2 \leq z_\alpha^2) = 1 - \alpha$, where z_α^2 is the quantile $1 - \alpha$ of a $\chi_2(d)$ law.

(b) (3 points) For $d = 5$, deduce an SOCP formulation (in addition to the $f_\star$ constraint) that imply a probabilistic guarantee of level $\varepsilon = 0.1$ with confidence 95%.

Solution:

$$\begin{cases} f_\star(v, x) \geq s \\ t - s \leq 0 \\ \sup_{\mu \in \mathcal{E}(S)} VaR_\varepsilon^\mu(v^T \tilde{u}) \leq t \end{cases}$$

where $\mathcal{E}(s) := M_N + z_\alpha/\sqrt{N}\Sigma^{1/2}B(0,1))$. We have

$$VaR_\varepsilon^\mu(v^T\tilde{u}) = v^T\mu + \|\Sigma^{1/2}v\|_2 VaR_\varepsilon(G)$$

where $G \sim \mathcal{N}(0,1)$. Thus, F_G is the cdf of G we have

$$\begin{aligned} \sup_{\mu \in \mathcal{E}(S)} VaR_\varepsilon^\mu(v^T\tilde{u}) &= \sup_{\mu \in \mathcal{E}(S)} v^T\mu + F_G^{-1}(\varepsilon)\|\Sigma^{1/2}v\|_2 \\ &= F_G^{-1}(\varepsilon)\|\Sigma^{1/2}v\|_2 + v^T M_N + z_\alpha/\sqrt{N} \sup_{\mu' \in B(0,1)} (\Sigma^{1/2}v)^T\mu \\ &= F_G^{-1}(\varepsilon)\|\Sigma^{1/2}v\|_2 + v^T M_N + z_\alpha/\sqrt{N}\|\Sigma^{1/2}v\|_2 \end{aligned}$$

Thus the SOCP formulation reads

$$\begin{cases} f_\star(v, x) \geq s \\ t - s \leq 0 \\ (F_G^{-1}(\varepsilon) + z_\alpha/\sqrt{N})\|\Sigma^{1/2}v\|_2 + v^T M_N \leq t \end{cases}$$