Basics of Conic LP (Cont'd)

2020年8月18日 9:53

Let E be an Euclidean space with an inner product $\langle\cdot,\cdot\rangle$ defined on it

Let $K \subseteq E$ be a closed pointed cone with non-empty interior.

e.g: $\mathbb{R}^n_+$, $\mathcal{Q}^{n+1}$, $\mathcal{S}^n_+$

Recall the primal-dual pair of conic LP:

$$(P) \quad \begin{array}{ll} \inf & \langle c,x\rangle \\ \text{s.t.} & \langle a_i,x\rangle = b_i \quad (y_i) \quad i\in[m] \\ & x\in K \qquad (w) \end{array} \qquad \Bigg| \qquad (D) \quad \begin{array}{ll} \sup & b^T y \\ \text{s.t.} & c - \sum_{i=1}^m y_i a_i \in K^* \end{array}$$

* How to derive the dual of (P)? Write down the Lagrangian.

$$\mathcal{L}(x,y,w) = \langle c,x\rangle + \sum_{i=1}^m y_i(b_i - \langle a_i,x\rangle) + \underline{??}$$

– What should the term $\underline{??}$ be? It should penalize any violation of $x\in K$.

Warm-Up 1: $K=\mathbb{R}^n_+$

$$\mathcal{L}_{LP}(x,y,w) = \langle c,x\rangle + \sum_{i=1}^m y_i(b_i - \langle a_i,x\rangle) - w^T x \quad \text{with } w\geq 0$$

So that for each $x\in\mathbb{R}^n$,

$$\sup_{\substack{y\in\mathbb{R}^m \\ w\in\mathbb{R}^n_+}} \mathcal{L}_{LP}(x,y,w) = \begin{cases} \langle c,x\rangle & \text{if } \langle a_i,x\rangle = b_i \ \forall i \text{ and } x\geq 0, \\ +\infty & \text{o/w} \end{cases}$$

And hence

$$(P) \iff \inf_{x\in\mathbb{R}^n} \sup_{\substack{y\in\mathbb{R}^m \\ w\in\mathbb{R}^n_+}} \mathcal{L}_{LP}(x,y,w)$$

Warm-Up 2: $K=\mathcal{S}^n_+$

$$\mathcal{L}_{SDP}(X,y,W) = C\bullet X + \sum_{i=1}^m y_i(b_i - A_i\bullet X) - W\bullet X \quad \text{with } W\in\mathcal{S}^n_+$$

Why? Observe that since $X\in\mathcal{S}^n$, we can write

$X = U\Lambda U^T$, U orthogonal, Λ diagonal (eigen-decomposition)

Then, $W\bullet X = \operatorname{Tr}(WU\Lambda U^T) = (U^T W U)\bullet\Lambda = \sum_{i=1}^n \Lambda_{ii}(U^T W U)_{ii}$

From this, we deduce

1) If $W, X\in\mathcal{S}^n_+$, then $W\bullet X\geq 0$.

2) If $X\notin\mathcal{S}^n_+$, then we can find $W\in\mathcal{S}^n_+$ s.t. $W\bullet X<0$.

Thus, for each $X\in\mathcal{S}^n$,

$$\sup_{y\in\mathbb{R}^m} \mathcal{L}_{SDP}(X,y,W) = \begin{cases} C\bullet X & \text{if } A_i\bullet X = b_i \ \forall i \text{ and } X\in\mathcal{S}^n_+, \end{cases}$$

$$\sup_{\substack{y\in\mathbb{R}^m\\ W\in S^n_+}} \mathcal{L}_{SDP}(X,y,W) = \begin{cases} C\bullet X & \text{if } A_i\bullet X = b_i \ \forall i \text{ and } X\in S^n_+, \\ +\infty & \text{o/w.}\end{cases}$$

General Case. Claim: $\underline{??} = -\langle w,x\rangle$ with $w\in K^*$

In particular, if $x\notin K$, then there exists $w\in K^*$ s.t. $\langle w,x\rangle<0$.

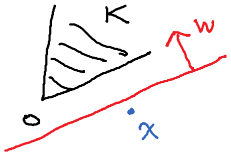
Proof:

1°: K is closed, and being a pointed cone, is convex

- K pointed cone: $\left\{\begin{array}{l} u,v\in K \Rightarrow u+v\in K \\ u\in K,\ \alpha>0 \Rightarrow \alpha u\in K \end{array}\right\} \Rightarrow \forall\alpha\in(0,1),\ \alpha u+(1-\alpha)v\in K$

2°: Since $x\notin K$, by separation theorem, there exists y s.t.

$$\langle w,x\rangle < \inf_{y\in K}\langle w,y\rangle$$



3°: Since $0\in K$, $\langle w,x\rangle < \langle w,0\rangle = 0$. We claim that $\inf_{y\in K}\langle w,y\rangle = 0$. Note that this implies $w\in K^*$.

(recall: $K^* = \{ z : \langle y,z\rangle \geq 0 \ \forall y\in K\}$)

Suppose not. Then, there exists $y'\in K$ s.t. $\langle w,y'\rangle<0$, because $\inf_{y\in K}\langle w,y\rangle \leq 0$ and $0\in K$. But $\alpha y'\in K$ $\forall\alpha>0$, so

$$\langle w,x\rangle < \inf_{y\in K}\langle w,y\rangle = -\infty.$$

a contradiction.

Hence, we conclude that

$$(P) \iff \inf_{x\in E}\ \sup_{\substack{y\in\mathbb{R}^m\\ w\in K^*}} \underbrace{\langle c,x\rangle + \sum_{i=1}^m y_i(b_i - \langle a_i,x\rangle) - \langle w,x\rangle}_{\mathcal{L}(x,y,w)}$$

The (Lagrangian) dual is simply

$$\sup_{\substack{y\in\mathbb{R}^m\\ w\in K^*}} \boxed{\inf_{x\in E} \left\langle C - \sum_{i=1}^m y_i a_i - w,\ x\right\rangle} + \sum_{i=1}^m b_i y_i$$

$$= \begin{cases} 0 & \text{if } C - \sum_{i=1}^m y_i a_i - w = 0 \\ & \end{cases}$$

$$L = \begin{cases} 0 & \text{if } c - \sum_{i=1} y_i a_i - w = 0 \\ -\infty & \text{o/w} \end{cases}$$

$$\Longleftrightarrow \quad \begin{aligned} \sup \quad & b^T y \\ \text{s.t.} \quad & c - \sum_{i=1}^{m} y_i a_i - w = 0, \\ & w \in K^* \end{aligned} \quad \Longleftrightarrow \quad (D)$$

Strong Duality for Conic LPs

Recall

$$(P) \quad \begin{aligned} v_p^* = \inf \quad & \langle c, x \rangle \\ \text{s.t.} \quad & \langle a_i, x \rangle = b_i, \quad i \in [m] \\ & x \in K \end{aligned} \qquad \Bigg| \qquad (D) \quad \begin{aligned} v_d^* = \sup \quad & b^T y \\ \text{s.t.} \quad & c - \sum_{i=1}^{m} y_i a_i \in K^* \end{aligned}$$

Theorem: (Strong Duality for Conic LPs)

Suppose that $\substack{\text{(P)} \\ \text{(D)}}$ is bounded $\substack{\text{below} \\ \text{above}}$ and is strictly feasible; i.e.,

there exists a feasible $\substack{\bar{x} \\ \bar{y}}$ to $\substack{\text{(P)} \\ \text{(D)}}$ such that $\substack{\bar{x} \in \text{int}(K) \\ c - \sum_{i=1}^{m} \bar{y}_i a_i \in \text{int}(K^*)}$.

Then, $v_p^* = v_d^*$ and there exists an optimal $\substack{\text{dual solution } y^* \\ \text{primal solution } x^*}$; i.e.,

$$\substack{b^T y^* \\ \langle c, x^* \rangle} = v_p^* = v_d^* .$$

It is instructive to compare this to the LP strong duality theorem ($K = \mathbb{R}_+^n$).

Theorem (Strong Duality for LPs)

Suppose that $\substack{\text{(P)} \\ \text{(D)}}$ is bounded $\substack{\text{below} \\ \text{above}}$ and is feasible. Then, $v_p^* = v_d^*$ and both (P) and (D) have optimal solutions.

** It is important to verify strict feasibility before applying the strong duality theorem for conic LPs.

Example: (Failure of Strong Duality)

Consider

$$(D) \quad \begin{aligned} v_d^* = \sup \quad & \overset{b}{\begin{bmatrix} -1 \\ 0 \end{bmatrix}}^T \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ \text{s.t.} \quad & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - y_1 \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix} - y_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \in \mathcal{Q}^3 \end{aligned} \qquad \Bigg| \qquad \begin{aligned} v_p^* = \inf \quad & x_2 \\ \text{s.t.} \quad & x_1 + x_3 = 1, \\ & x_1 - x_3 = 0, \\ & (x_1, x_2, x_3) \in \mathcal{Q}^3 \end{aligned} \quad (P)$$

(D) $\underset{c}{\begin{bmatrix}1\\0\end{bmatrix}} - y_1 \underset{a_1}{\begin{bmatrix}0\\-1\end{bmatrix}} - y_2 \underset{a_2}{\begin{bmatrix}0\\1\end{bmatrix}} \in Q$

(D) $\Longleftrightarrow$ $\sup\ -y_1$ s.t. $y_1 + y_2 \geq \sqrt{1 + (y_1 - y_2)^2}$

$$\Updownarrow$$

$4y_1y_2 \geq 1$, $y_1 + y_2 > 0$

$\therefore$ $V_d^* = 0$ but there is no optimal solution.

(P)

$x_1 - x_3 = 0$,

$(x_1, x_2, x_3) \in Q^3$

(P) only has one feasible solution $(x_1^*, x_2^*, x_3^*) = (\frac{1}{2}, 0, \frac{1}{2})$, which of course is optimal and $V_p^* = 0$.

Observe: (P) is not strictly feasible, but (D) is.

DR VaR Optimization

2020年8月18日 23:54

∗ Consider a one-period portfolio optimization problem. Let

$$\left.\begin{aligned} x_i &= \text{return of asset } i\text{, assumed random,} \\ w_i &= \text{allocation on asset } i\text{,} \end{aligned}\right. \quad i = 1, \cdots, n.$$

The return of the portfolio is $w^T x$.

∗ A commonly used measure of the portfolio's risk is the value-at-risk (VaR), defined as

$$\begin{aligned} V(w) = \inf \quad & \gamma \\ \text{s.t.} \quad & \Pr(\gamma \leq -w^T x) \leq \varepsilon \end{aligned} \qquad \text{— } (\Delta)$$

In words, this is the minimal level γ such that the portfolio's loss exceeds γ has probability at most ε.

∗ As previously mentioned, we do not usually know the distribution of x. On the other hand, we may have information about its mean $\bar{x}$ and covariance Γ. Hence, we may consider the DR counterpart of (Δ):

$$\begin{aligned} V_{DR}(w) = \inf \quad & \gamma \\ \text{s.t.} \quad & \sup_{\mathbb{P} \in \mathcal{P}} \Pr_{\mathbb{P}}(\gamma \leq -w^T x) \leq \varepsilon \end{aligned} \qquad \text{— (DR-VaR)}$$

Here, $\mathcal{P}$ is the set of probability measures on $(\mathbb{R}^n, \mathcal{F})$ with mean $\bar{x} \in \mathbb{R}^n$ and covariance $\Gamma \in S^n_{++}$; $\mathcal{F}$ is a Borel σ-algebra on subsets of $\mathbb{R}^n$.

— How to analyze and solve (DR-VaR)?

We begin by observing that

indicator of the event $\{\gamma \leq -w^T x\}$

$$\sup_{\mathbb{P} \in \mathcal{P}} \Pr_{\mathbb{P}}(\gamma \leq -w^T x) = \begin{aligned} \sup \quad & \int_{\mathbb{R}^n} \mathbb{1}_{\{\gamma \leq -w^T x\}}(x)\, d\mathbb{P}(x) \\ \text{s.t.} \quad & \int_{\mathbb{R}^n} d\mathbb{P}(x) = 1, \\ & \int_{\mathbb{R}^n} x\, d\mathbb{P}(x) = \bar{x}, \\ & \int_{\mathbb{R}^n} (x - \bar{x})(x - \bar{x})^T d\mathbb{P}(x) = \Gamma, \\ & \mathbb{P} \geq 0. \end{aligned}$$

Note that the RHS is linear in $\mathbb{P}$!

Intuitively, if $\mathbb{P}$ is supported on $\hat{x}_1, \cdots, \hat{x}_\ell$ with mass $P_1, \cdots, P_\ell$, resp, then the RHS is

$$\begin{aligned} \sup \quad & \sum_{j:\, \gamma \le -w^T \hat{x}_j} p_j \\ \text{s.t.} \quad & \sum_{j=1}^{\ell} p_j = 1, \\ & \sum_{j=1}^{\ell} p_j \hat{x}_j = \bar{x}, \\ & \sum_{j=1}^{\ell} p_j (\hat{x}_j - \bar{x})(\hat{x}_j - \bar{x})^T = \Gamma, \\ & p_j \ge 0 \quad \forall j. \end{aligned}$$

Now, observe that

$$\left.\begin{aligned} & \int_{\mathbb{R}^n} d\mathbb{P}(x) = 1, \\ & \int_{\mathbb{R}^n} x \, d\mathbb{P}(x) = \bar{x}, \\ & \int_{\mathbb{R}^n} (x-\bar{x})(x-\bar{x})^T d\mathbb{P}(x) = \Gamma \end{aligned}\right\} \iff \int_{\mathbb{R}^n} \begin{bmatrix} x \\ 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}^T d\mathbb{P}(x) = \underset{\substack{\shortparallel \\ \Sigma \succ 0}}{\begin{bmatrix} \Gamma + \bar{x}\bar{x}^T & \bar{x} \\ \bar{x}^T & 1 \end{bmatrix}}$$

So (DR-VaR) $\iff$

$$\begin{aligned} \sup \quad & \int_{\mathbb{R}^n} \mathbb{1}_{\{\gamma \le -w^T x\}}(x) \, d\mathbb{P}(x) \\ \text{s.t.} \quad & \int_{\mathbb{R}^n} \begin{bmatrix} x \\ 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}^T d\mathbb{P}(x) = \Sigma, \quad (M) \\ & \mathbb{P} \succcurlyeq 0. \end{aligned}$$

Consider the Lagrangian

$$\mathcal{L}(\mathbb{P}, M) = \int_{\mathbb{R}^n} \mathbb{1}_{\{\gamma \le -w^T x\}}(x) \, d\mathbb{P}(x) + \left\langle M, \Sigma - \int_{\mathbb{R}^n} \begin{bmatrix} x \\ 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}^T d\mathbb{P}(x) \right\rangle$$

with $M \in \mathcal{S}^{n+1}$. Then,

$$(\text{DR-VaR}) \iff \sup_{\mathbb{P} \ge 0} \inf_{M \in \mathcal{S}^n} \mathcal{L}(\mathbb{P}, M)$$

Now, consider

$$\inf_{M \in \mathcal{S}^n} \sup_{\mathbb{P} \ge 0} \mathcal{L}(\mathbb{P}, M).$$

For fixed $M \in \mathcal{S}^n$, let $q(x) = \begin{bmatrix} x \\ 1 \end{bmatrix}^T M \begin{bmatrix} x \\ 1 \end{bmatrix}$ and then

$$\sup_{\mathbb{P} \ge 0} \mathcal{L}(\mathbb{P}, M) = \langle M, \Sigma \rangle + \sup_{\mathbb{P} \ge 0} \int_{\mathbb{R}^n} \left[\mathbb{1}_{\{\gamma \le -w^T x\}}(x) - q(x) \right] d\mathbb{P}(x)$$

$$= \begin{cases} \langle M, \Sigma \rangle & \text{if } \mathbb{1}_{\{\gamma \le -w^T x\}}(x) \le q(x) \ \forall x \end{cases}$$

$$= \begin{cases} \langle M, \Sigma \rangle & \text{if } \mathbb{1}_{\{\gamma \le -w^T x\}}(x) \le g(x) \ \forall x \\ +\infty & \text{o/w.} \end{cases}$$

Note that

$$\mathbb{1}_{\{\gamma \le -w^T x\}}(x) \le g(x) \ \forall x \iff \begin{cases} g(x) \geqslant 0 \quad \forall x & (A) \\ g(x) \geqslant 1 \text{ whenever } \gamma + w^T x \le 0 & (B) \end{cases}$$

$\underline{1^o}$: $(A) \iff M \in S^n_+$

$\underline{2^o}$: $(B) \iff 1 \le \begin{array}{l} \inf \ g(x) \\ \text{s.t. } \gamma + w^T x \le 0 \ (\tau) \end{array} = \inf_{x \in \mathbb{R}^n} \sup_{\tau \geqslant 0} \{ g(x) + \tau(\gamma + w^T x) \}$

$\updownarrow$ strong duality: g convex, linear constraint

$= \sup_{\tau \geqslant 0} \inf_{x \in \mathbb{R}^n} \{ g(x) + \tau(\gamma + w^T x) \}$

Hence, there exists a $\tau \geqslant 0$ such that

$$\forall x \in \mathbb{R}^n: \quad g(x) \geqslant 1 - \tau(\gamma + w^T x) = \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} 0 & -\frac{1}{2}\tau w \\ -\frac{1}{2}\tau w^T & 1 - \tau\gamma \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}.$$

It follows that

$$\inf_{M \in S^n} \sup_{\mathbb{P} \geqslant 0} \mathcal{L}(\mathbb{P}, M) = \begin{array}{ll} \inf & \langle M, \Sigma \rangle \\ \text{s.t} & M + \begin{bmatrix} 0 & \frac{1}{2}\tau w \\ \frac{1}{2}\tau w^T & -1 + \tau\gamma \end{bmatrix} \in S^n_+, \\ & M \in S^n_+. \end{array}$$

It remains to argue why

$$\sup_{\mathbb{P} \geqslant 0} \inf_{M \in S^n} \mathcal{L}(\mathbb{P}, M) = \inf_{M \in S^n} \sup_{\mathbb{P} \geqslant 0} \mathcal{L}(\mathbb{P}, M). \quad \text{— } (+)$$

Consider the set of probability measures on $(\mathbb{R}^n, \mathcal{F})$:

$$\mathcal{M} = \{ \mathbb{P} : \mathbb{P} \geqslant 0, \ \mathbb{P}(\mathbb{R}^n) = 1 \}$$

Let $C = \text{cone}(\mathcal{M}) = \bigcup_{\lambda > 0} \lambda \mathcal{M}$ be the convex cone generated by $\mathcal{M}$.

Define the linear map A by

$$\mathbb{P} \mapsto \int_{\mathbb{R}^n} \begin{bmatrix} x \\ 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}^T d\mathbb{P}(x) \in S^n.$$

Then, the constraint of our problem can be written as

$$A(\mathbb{P}) = \Sigma \quad , \quad \mathbb{P} \in \mathcal{M} .$$

Since $\Sigma \in \mathbb{S}^n_{++}$, we have $\Sigma \in \text{int } A(C)$. Hence, strong duality (†) holds by invoking, e.g., Proposition 3.4 of

Shapiro: On Duality Theory of Conic Linear Problems

* The material discussed above comes from

El Ghaoui, Oks, Oustry: Worst-Case Value-at-Risk and Robust Portfolio Optimization: A Conic Programming Approach. Oper. Res. 51(4): 543–556, 2003.

Extension to "nonlinear" portfolios can be found at

Zymler, Kuhn, Rustem: Worst-Case Value at Risk of Nonlinear Portfolios. Manag. Sci. 59(1): 172–188, 2013.