Pure Mathematics 理论数学, 2020, 10(1), 17-22
Published Online January 2020 in Hans. http://www.hanspub.org/journal/pm
https://doi.org/10.12677/pm.2020.101004

Hans汉斯

Study on the Geometric Properties of Projection Operators on a Class of Variable Boxes

Dongxia Shen, Shiyun Wang, Guoqing Qi, Changlin Zhao, Jieyu Huang

College of Science, Shenyang Aerospace University, Shenyang Liaoning
Email: wsy0902@163.com

Received: Dec. 10th, 2019; accepted: Dec. 30th, 2019; published: Jan. 6th, 2020

Abstract

The geometric properties of closed convex cone projection operators are very important for the study of Lagrangian, the geometric properties of projection operators and the sensitivity analysis of optimization problems. In this paper, we study the dual cone, tangent cone, polar cone and critical cone of a class of variable box, which lays a foundation for the further study of the geometric properties of projection operators on this class of variable box and the sensitivity analysis of related optimization problems.

Keywords

Projection Operator, Dual Cones, Cutting Cones, Polar Cone, Critical Cone

一类可变盒子上投影算子的几何性质研究

沈东霞，王诗云，齐国庆，赵常霖，黄婕妤

沈阳航空航天大学理学院，辽宁 沈阳
Email: wsy0902@163.com

收稿日期：2019年12月10日；录用日期：2019年12月30日；发布日期：2020年1月6日

摘 要

闭凸锥投影算子几何性质对增广拉格朗日方法，投影算子几何性质的研究以及优化问题的灵敏性分析有着至关重要的作用。本文研究一类可变盒子的对偶锥、切锥、极锥与临界锥为进一步研究这一类可变盒子上投影算子的几何性质以及相关优化问题的灵敏度分析奠定了基础。

文章引用: 沈东霞, 王诗云, 齐国庆, 赵常霖, 黄婕妤. 一类可变盒子上投影算子的几何性质研究[J]. 理论数学, 2020, 10(1): 17-22. DOI: 10.12677/pm.2020.101004

关键词

投影算子，对偶锥，切锥，极锥，临界锥

1. 引言

闭凸锥的投影几何性质以及投影的几何性质在优化问题理论分析与算法设计中有非常重要的作用。凸集上的投影算子的插入[1] [2]在算法设计和理论研究中也有着非常关键的作用，例如最优化问题、增广拉格朗日方法的应用与收敛性分析、最优化问题的灵敏性分析均涉及到投影算子的性质[3]。我们研究其中一类可变盒子$\{(y,\tau)\in R^n\times R:0\le y\le \tau e\}$的几何性质，它是欧式空间$R^n\times R$的非空闭凸锥。文献[4] [5] [6]研究了由投影算子对应的凸二次优化的条件，给出的可变盒子集合投影算子显示解的计算方法，Han 等[7]研究了凸优化问题的求解，提出并详细阐述了某类凸锥投影算子显示表达式的计算方法。而对这些集合的刻画，离不开集合的几何性质。在这里我们对此研究的应用问题将在以后的研究中进行进一步的阐述。

2. 预备知识

由 Liu 等[4]研究的二阶锥上投影算子的 Clarke 广义 Jacobian 的具体表达式如下：

令 K 为 $n+1$ 维欧式空间 $R^n\times R$ 的闭凸锥。对 $\forall(x,t)\in R^n\times R$，我们用 $x^{\downarrow}$ 以表示一类向量组，x 按降序排列 $x^{\downarrow}:x_1^{\downarrow}\ge x_2^{\downarrow}\ge\cdots\ge x_n^{\downarrow}$，

规定 $\left[x_0^{\downarrow}\right]=+\infty$，$s_0=0$，$m=\left|\{x_i\,|\,x_i\ge 0\}\right|$，$s_k=\sum_{i=1}^{k}x_k^{\downarrow},k=1,\cdots,n$，

使 $\left[x_{k+1}^{\downarrow}\right]<\dfrac{s_k+t}{k+1}\le\left[x_k^{\downarrow}\right]$ 的最小整数 k 为 λ，

否则 $\lambda=m$。

令 $\theta=\dfrac{s_\lambda+t}{\lambda+1}$，则 $\begin{cases}\theta\ge 0, & \lambda<m\\ \theta\text{不变}, & \lambda=m\end{cases}$。

则 (x,t) 在 K 上的投影 $\prod_K(x,t)$ 可以计算为

$$\prod\nolimits_K(x,t)=(\bar{x},\bar{t}) \tag{1.1}$$

其中 $\bar{x_i}=\begin{cases}\bar{t}, & x_i\ge\bar{t}\\ x_i, & 0<x_i<\bar{t}\\ 0, & x_i\le 0\end{cases}$，$i=1,\cdots,n$，$\bar{t}\in R_+$，$\bar{x}\in R^n$。

令 $\alpha=\{i\,|\,x_i\le 0\}$，

则 $\theta=\dfrac{\sum_{i\in r}x_i+t}{|r|+1}$。

我们用 $\mathrm{int}(K)$ 表示 K 的内部。K 的对偶锥和极锥的定义分别为

$$K^*=\{(y,\tau)\in R^n\times R:\langle(x,t),(y,\tau)\rangle\ge 0,\forall(x,t)\in E\} \tag{1.2}$$

和

$$K^{o}=\left\{(y,\tau)\in R^{n}\times R:\left\langle (x,t),(y,\tau)\right\rangle \le 0,\forall (x,t)\in E\right\} \tag{1.3}$$

当 $K=\left\{(y,\tau)\in R^{n}\times R: f(y,\tau)\le 0\right\}$ 时，K 的切锥为

$$T_K=(\bar{x},\bar{t})=\left\{(d,\eta): f'\left((\bar{x},\bar{t});(d,\eta)\right)\le 0\right\} \tag{1.4}$$

K 的临界锥的定义为

$$\bar{K}(x,t)=T_K(\bar{x},\bar{t})\cap\left((x,t)-(\bar{x},\bar{t})\right)^{\perp} \tag{1.5}$$

3. 集合 K 的几何性质

对偶锥、切锥、极锥与临界锥在研究投影算子的方向导数以及在求最优化问题的临界锥时应用十分广泛。下面，主要从集合 K 的表达式以及对偶锥、切锥、极锥与临界锥定义出发，分别推导 K 的对偶锥、切锥、极锥与临界锥。

3.1. 集合 K 的对偶锥

取 $(x,t)\in K^{*}$，对 $\forall (y,\tau)\in K$，有

$$\langle x,y\rangle + t\tau \ge 0 \tag{2.1}$$

则 $t\ge 0$。

由(2.1)可知，$x^{\mathrm{T}}y+t\tau\ge 0$，进而可得 $-t\le x^{\mathrm{T}}\begin{pmatrix} y \\ \tau \end{pmatrix}$。

令 $z=\dfrac{y}{\tau}$，则 $z\in\Omega$，$\Omega=\left\{z\in R^{n}, 0\le z\le e\right\}$，则 $-t\le x^{\mathrm{T}}z$，对 $\forall z\in\Omega$，则只需 $-t\le \min\limits_{z\in\Omega} x^{\mathrm{T}}z$ 即可。

引理 1：对 $\forall x\in R^{n}$，$\forall z\in\Omega$，z^{*} 为 $\min\limits_{z\in\Omega} x^{\mathrm{T}}z$ 的最优解。则

$$\begin{cases} z_i^{*}=0, & x_i>0 \\ z_i^{*}=1, & x_i<0 \\ 0\le z_i^{*}\le 1, & x_i=0 \end{cases}$$

记当 $I_{<}=\left\{i: x_i<0\right\}$，$I_{>}=\left\{i: x_i>0\right\}$，$I_{=}=\left\{i: x_i=0\right\}$ 时，

$$\begin{aligned} -t &\le x^{\mathrm{T}}z^{*} \\ &=\sum_{i\in I_{>}} x_i z_i^{*}+\sum_{i\in I_{=}} x_i z_i^{*}+\sum_{i\in I_{<}} x_i z_i^{*} \\ &=0+0+\sum_{i\in I_{<}} x_i, \end{aligned}$$

即 $\sum\limits_{i\in I_{<}} x_i+t\ge 0$，$t\ge 0$。

引理 2：对 $\forall x\in R^{n}$，$\forall z\in\Omega$，z^{*} 为 $\max\limits_{z\in\Omega} x^{\mathrm{T}}z$ 的最优解。则

$$\begin{cases} z_i^{*}=0, & x_i>0 \\ z_i^{*}=1, & x_i<0 \\ 0\le z_i^{*}\le 1, & x_i=0 \end{cases}$$

记 $I_{<}=\left\{i: x_i>0\right\}$，$I_{>}=\left\{i: x_i<0\right\}$，$I_{=}=\left\{i: x_i=0\right\}$。

我们有如下定理。

定理 1：集合 K 的对偶锥为 $K^* = \left\{(x,t) \in R^n \times R : t + \sum_{i\in I_<} x_i \ge 0\right\}$

证明：令 $B = \left\{(x,t) \in R^n \times R : t \ge -\sum_{i\in I_<} x_i, t \ge 0\right\}$，要证 $B = K^*$；首先要证 $(x,t) \in K^*$ 对任意的 $(x,t) \in B$，由对偶锥的定义，对任意的 $(y,\tau) \in K$，都有下式成立

$$\langle x, y\rangle + t\tau = x^{\mathrm{T}} y + t\tau \tag{2.2}$$

若 $\tau > 0$，则(2.2)变为

$$\langle x, y\rangle + t\tau = x^{\mathrm{T}} y + t\tau = \tau\left(x^{\mathrm{T}} \frac{y}{\tau} + t\right) = \tau\left(x^{\mathrm{T}} z + t\right) \ge \tau\left(\min_{z\in\Omega}\left(x^{\mathrm{T}} z\right) + t\right) = \tau\left(x^{\mathrm{T}} z^* + t\right) = \tau\left(\min_{i\in I_<} x_i + t\right) \ge 0 \tag{2.3}$$

若 $\tau = 0$，则 $y = 0$，则 $\langle x, y\rangle + t\tau \ge 0$ 显然成立。

因此 $B \subset K^*$。

反过来，对 $\forall(x,t) \in K^*$，则对 $\forall(y,\tau)$ 有 $x^{\mathrm{T}} y + t\tau \ge 0$，若 $\tau > 0$，则

$$\tau\left(t + x^{\mathrm{T}} z\right) \ge \tau\left(t + \min_{z\in\Omega} x^{\mathrm{T}} z\right) = \tau\left(t + \sum_{i\in I_<} x_i\right) \ge 0,$$

由于 $\tau\left(t + \sum_{i\in I_<} x_i\right) \ge 0$ 且 $t \ge 0$ 则 $(x,t) \in B$，即 $K^* \subset B$。

由于对任意的 $(x,t) \in K^*$，$K^0 \subseteq K^*$，反过来对任意的 $(x,t) \in K^*$，要证 $(x,t) \in K^0$。

3.2. 集合 K 的切锥

定理 2：集合 K 的切锥为 $K^0 = \left\{(x,t) \in R^n \times R : t + \sum_{i\in I_<} x_i \le 0\right\}$

证明：由于 $K^0 = -K^*$，则该结论显然。

3.3. 集合 K 的极锥

定理 3：集合 K 的极锥为 $T_K(\bar{x},\bar{t}) = \begin{cases} R^n \times R; & (x,t) \in \operatorname{int} K \\ K; & (x,t) \in K^0 \\ \{(d,\eta) : f'((\bar{x},\bar{t}),(d,\eta)) \le 0\}; & \text{其他} \end{cases}$。

$f'((\bar{x},\bar{t}),(d,\eta)) \le 0$，则

$f'((\bar{x},\bar{t}),(d,\eta)) = \lim_{l\downarrow 0} \dfrac{f(\bar{x}+ld, \bar{t}+l\eta) - f(\bar{x},\bar{t})}{l}$，此时 $f(\bar{x},\bar{t}) = 0$

则

$$\begin{aligned} f'((\bar{x},\bar{t}),(d,\eta)) &= \lim_{l\downarrow 0} \frac{f(\bar{x}+ld, \bar{t}+l\eta)}{l} \\ &= \lim_{l\downarrow 0} \frac{\max\{-\bar{x}+ld, \bar{x}+ld-\bar{t}e+l\eta e\}}{l} \\ &= \lim_{l\downarrow 0} \frac{\max\{-ld_\alpha, ld_\gamma - l\eta e\}}{l} = \max\{-d_\alpha, d_\gamma - \eta e\} \end{aligned}$$

即 $T_K\left(\bar{x},\bar{t}\right)=\begin{cases}R^n\times R; & (x,t)\in \operatorname{int}K\\ 0; & (x,t)\in K^0\\ \left\{(d,\eta):-d_\alpha\le 0,d_\gamma\le \eta e\right\}; & \text{其他}\end{cases}$。

3.4. 集合 K 的临界锥

定理 4: 满足 $\min\frac{1}{2}\|y-x\|^2+\frac{1}{2}(\tau-t)^2$，使得 $0\le y\le \tau e$ 的临界锥为 $\bar{K}(x,t)=T_K\left(\bar{x},\bar{t}\right)\cap\left((x,t)-\left(\bar{x},\bar{t}\right)\right)^\perp$。

证明：i) 当 $(x,t)\in K$ 时，$\left(\bar{x},\bar{t}\right)=(x,t)$，则 $\bar{K}=T_K$ 即 $\begin{cases}R^n\times R; & (x,t)\in \operatorname{int}K\\ K; & (x,t)=(0,0)\\ \left\{(d,\eta):d_\alpha\ge 0,d_\gamma\le \eta e\right\}; & (x,t)\in bd(K)\end{cases}$；

ii) 当 $(x,t)\in \operatorname{int}K^0$ 时，$\left(\bar{x},\bar{t}\right)=(0,0)$，则 $T_K=K$，则 $\bar{K}=K\cap(x,t)^\perp$，而 $\sum_{i\in I_>}x_i+t<0$ 且 $\forall(d,\eta)\in K\cap(x,t)^\perp$，则 $0\le d\le \eta e$，$x^{\mathrm{T}}d+t\eta=0$，而

$$\begin{aligned}x^{\mathrm{T}}d+t\eta&=\eta\left(t+\left(\frac{d}{\eta}\right)^{\mathrm{T}}x\right)\le\eta\left(t+\max_{z\in\Omega}z^{\mathrm{T}}x\right)\\&=\eta\left(t+z_*^{\mathrm{T}}x\right)\\&=\eta\left(t+\sum_{i\in I_>}x_i\right)\\&\le 0\end{aligned}$$

则 $\eta=0$，即 $\bar{K}=\{(0,0)\}$；

iii) 当 $(x,t)\in bd\left(K^0\right)\setminus\{(0,0)\}$，$\left(\bar{x},\bar{t}\right)=(0,0)$ 时，$\bar{K}=K\cap(x,t)^\perp$，当 $\sum_{i\in I_>}x_i+t=0$ 时，对 $\forall(d,\eta)\in K\cap(x,t)^\perp$，有 $0\le d\le\eta e$，$0=x^{\mathrm{T}}d+t\eta=\eta\left(x^{\mathrm{T}}\left(\frac{d}{\eta}\right)+t\right)\le\eta\left(\max_{z\in\Omega}x^{\mathrm{T}}z+t\right)=0$，这说明 $\frac{d}{\eta}=\arg\max_{z\in\Omega}x^{\mathrm{T}}z$，即

$$\left(\frac{d}{\eta}\right)_i=\begin{cases}1, & x_i>0\\ 0, & x_i<0\\ [0,1], & x_i=0\end{cases},$$

$$\bar{K}=\left\{(d,\eta):\eta\ge 0,\left(\frac{d}{\eta}\right)_{T_>}=e,\left(\frac{d}{\eta}\right)_{T_<}=0,0\le\left(\frac{d}{\eta}\right)_{T_=}\le e\right\},$$

其中 $I_>=\gamma^{\neq}$，$I_<=\alpha^{\neq}$，$I_==\gamma_==\alpha_=$；

iv) 其他情况，$T_K\left(\bar{x},\bar{t}\right)=\left\{(d,\eta):-d_\alpha\le 0,d_\gamma\le\eta e\right\}$

下面求 $(x,t)-\left(\bar{x},\bar{t}\right)$，$\bar{t}>0$，$L(y,\tau,\upsilon,\lambda)=\frac{1}{2}\|y-x\|^2+\frac{1}{2}(\tau-t)^2-\mp\langle y,\lambda\rangle-\langle\tau e-y,\upsilon\rangle$

$$\begin{cases}y-x-\lambda+\upsilon=0 & (1)\\ \tau-t-e^{\mathrm{T}}\upsilon=0 & (2)\\ 0\le y\perp\lambda\ge 0 & (3)\\ 0\le(\tau e-y)\perp\upsilon\ge 0 & (4)\end{cases}$$

令$\left(y,\tau\right)=\left(\bar{x},\bar{t}\right)$代入，由于$\bar{x}_{\alpha}=0$，$\bar{x}_{\beta}=x_{\beta}$，$\bar{x}_{\gamma}=\bar{t}e$，则由(3)(4)得$\lambda_{\beta}=\lambda_{\gamma}=\upsilon_{\alpha}=\upsilon_{\beta}=0$，$\lambda_{\alpha}\geq 0$，$\upsilon_{\gamma}\geq 0$；

由(1)得$\begin{cases}\bar{x}_{\alpha}-x_{\alpha}-\lambda_{\alpha}+\upsilon_{\alpha}=0\Rightarrow\lambda_{\alpha}=-x_{\alpha} & (5)\\ \bar{x}_{\gamma}-x_{\gamma}-\lambda_{\gamma}+\upsilon_{\gamma}=0\Rightarrow\upsilon_{\gamma}=x_{\gamma}-\bar{t}e & (6)\end{cases}$；

由(2)得$t-\bar{t}=-e^{\mathrm{T}}\upsilon=-e^{\mathrm{T}}\upsilon_{\gamma}$。

下面求$\bar{K}$，对$\forall\left(d,\eta\right)\in T_{K}\left(\bar{x},\bar{t}\right)\cap\left(x-\bar{x},t-\bar{t}\right)$，有

$$\begin{aligned}0&=\left\langle d,x-\bar{x}\right\rangle+\eta\left(t-\bar{t}\right)\\&=\left\langle d_{\alpha},x_{\alpha}-\bar{x}_{\alpha}\right\rangle+\left\langle d_{\beta},x_{\beta}-\bar{x}_{\beta}\right\rangle+\left\langle d_{\gamma},x_{\gamma}-\bar{x}_{\gamma}\right\rangle+\eta\left(t-\bar{t}\right)\\&=\left\langle d_{\alpha},-\lambda_{\alpha}\right\rangle+\left\langle d_{\gamma},\upsilon_{\gamma}\right\rangle-\pm\left\langle\eta e,\upsilon_{\gamma}\right\rangle\\&=-\left\langle d_{\alpha},-\lambda_{\alpha}\right\rangle-\left\langle\eta e-d_{\gamma},\upsilon_{\gamma}\right\rangle\leq 0\end{aligned}$$

其中$\left(d,\eta\right)\in T_{K}\left(\bar{x},\bar{t}\right)$，由此可知$\left\langle d_{\alpha},\lambda_{\alpha}\right\rangle=0$，$\left\langle\eta e-d_{\gamma},\upsilon_{\gamma}\right\rangle=0$

由(5)得$\lambda_{\alpha}=0$，$\lambda_{\alpha^{\neq}}>0\Rightarrow d_{\alpha^{=}}\geq 0$，$d_{\alpha^{\neq}}=0$，由(6)得$\upsilon_{\gamma^{=}}=0$，$\upsilon_{\gamma^{\neq}}>0$，由此可知$\eta e-d_{\gamma^{=}}\geq 0$，$\eta e-d_{\gamma^{\neq}}=0$，即$d_{\gamma^{=}}\leq\eta e$，$d_{\gamma^{\neq}}=\eta e$。

则$$\bar{K}=\begin{cases}K; & \left(x,t\right)\in K\\ \left\{\left(0,0\right)\right\}; & \left(x,t\right)\in\operatorname{int}K^{0}\\ \left\{\left(d,\eta\right):\eta\geq 0,\left(\dfrac{d}{\eta}\right)_{T_{>}}=e,\left(\dfrac{d}{\eta}\right)_{T_{<}}=0,0\leq\left(\dfrac{d}{\eta}\right)_{T_{=}}\leq e\right\}; & \left(x,t\right)\in bd\left(K^{0}\right)\backslash\left\{\left(0,0\right)\right\}\\ \left\{\left(d,\eta\right):d_{\alpha^{=}}\geq 0,d_{\alpha^{\neq}}=0,d_{\gamma^{=}}\leq\eta e,d_{\lambda^{\neq}}=\eta e\right\}; & 其他\end{cases}$$。

4. 结论

本文研究了集合K的对偶锥、切锥、极锥和临界锥，为进一步研究增广拉格朗日方法，投影算子微分性质的研究以及灵敏性分析等相关优化问题奠定了一定的理论基础。

参考文献

[1] Sun, D.F. (2006) The Strong Second Order Sufficient Condition and Constraint Nondegeneracy in Nonlinear Semidefinite Programming and Their Implications. *Mathematics of Operations Research*, **31**, 649-848. https://doi.org/10.1287/moor.1060.0195

[2] Wu, B., Ding, C., Sun, D.F., *et al*. (2014) On the Moreau-Yoshida Regularization of the Vector K-Norm Related Functions. *SIAM Journal on Optimization*, **24**, 766-794. https://doi.org/10.1137/110827144

[3] Clarke, F.H. (1990) Optimization and Nonsmooth Analysis. John Wiley and Sons, New York, Chapter 2, Generalized Gradients, 50-62.

[4] Liu, Y.J., Wang, S.Y. and Sun, J.H. (2013) Finding the Projection onto the Intersection of a Closed Half-Space and a Variable Box. *Operations Research Letters*, **41**, 259-264. https://doi.org/10.1016/j.orl.2013.01.014

[5] Wang, S.Y. and Zhang, Q.Y. (2018) Variational Geometric Properties of Two Classes of Closed Convex Cones. *Journal of Shenyang Aerospace University*, **35**, 93-96.

[6] Hu, X., Liu, Y.J., Liu, M.J. and Gao, J.R. (2013) Study on the Projection Operator over the Set of Variable Box. *Journal of Shenyang Aerospace University*, **30**, 93-96.

[7] Han, N., Liu, Y.J. and Liu, M.J. (2013) Computation of the Metric Projection over a Class of Closed Convex Cones. *Journal of Shenyang Aerospace University*, **30**, 88-91.