بسم الله الرحمن الرحیم

دانشکده علوم ریاضی

گروه ریاضی کاربردی

(گرایش تحقیق در عملیات)

عنوان :

دسته‌بندی استوار و ماشین‌های بردار پشتیبان

از:

جواد نعیمی کلورزی

استاد راهنما:

دکتر مازیار صلاحی

اسفند ۱۳۹۱

تقدیم به

پدر و مادر عزیزم

آنان که زندگی را به من آموختند.

تقدیر و تشکر

بعد از شکر و سپاس خداوند را که بزرگترین امید و یاور در لحظه لحظه زندگیست، از خانواده عزیزم به خصوص پدر و مادر مهربانم که همواره یاوری دلسوز و پشتیبانی محکم برایم بوده‌اند، بسیار سپاسگزارم. وظیفه خود می‌دانم از استاد فرزانه و دلسوز، جناب آقای دکتر مازیار صلاحی که زحمت راهنمایی این پایان‌نامه را بر عهده گرفتند و بدون مساعدت ایشان، این پایان‌نامه به نتیجه مطلوب نمی‌رسید، کمال تشکر و قدردانی را نمایم.

از جناب آقایان دکتر سعید کتابچی و دکتر فرشید مهردوست که با نهایت دقت، زحمت داوری این پایان‌نامه را بر عهده گرفتند و همچنین از جناب آقای دکتر داوود خجسته سالکویه که به عنوان نماینده تحصیلات تکمیلی در جلسه دفاع حضور یافتند، سپاسگزارم. در پایان از اساتید دانشکده ریاضی و همه دوستان صمیمانه تشکر می‌نمایم و برایشان سلامتی و موفقیت را آرزومندم.

چکیده:

دسته‌بندی استوار و ماشین‌های بردار پشتیبان
جواد نعیمی کلورزی

دسته‌بندی از مسائل اصلی در یادگیری ماشین است به طوری که مسائل متعددی از دنیای واقعی را می‌توان به صورت آن مطرح و حل کرد. یکی از روش‌های قدرتمند که در حال حاضر به صورت گسترده برای مسئله دسته‌بندی مورد استفاده قرار می‌گیرد، روش ماشین‌های بردار پشتیبان است. یک فرض اساسی در این روش این است که داده‌ها قطعی هستند در حالی که در دنیای واقعی داده‌ها معمولاً دارای عدم‌قطعیت هستند. عدم‌قطعیت داده‌ها در مسائل دسته‌بندی بر جواب مسئله بسیار اثر گذار است، از این رو به مطالعه عدم‌قطعیت داده‌های ورودی در ماشین‌های بردار پشتیبان می‌پردازیم. با استفاده از مدل‌های عدم‌قطعیت مختلف، دسته‌بندی استوار را به صورت مسائل برنامه‌ریزی خطی یا برنامه‌ریزی مخروطی درجه دوم بیان می‌کنیم. در پایان نیز دسته‌بندی معمولی و استوار را پیاده‌سازی و روی چندین مثال واقعی با یکدیگر مقایسه می‌کنیم. آزمایشات عددی نشان از برتری مدل استوار دارند.

کلید واژه:
ماشین‌های بردار پشتیبان، دسته‌بندی استوار، داده‌های ناقطعی.

فهرست مطالب

لیست جداول

لیست تصاویر

پیشگفتار:

مسئله دسته‌بندی از مسائل مهم در حوزه علوم به خصوص علوم کامپیوتر است و بسیاری از مسائل مربوط به تشخیص و پیش‌بینی را می‌توان به صورت یک مسئله دسته‌بندی مطرح کرد. دسته‌بندی در زمینه‌های استخراج داده‌ها، تشخیص الگو، بینایی ماشین و بیوانفورماتیک نقش اساسی دارد [۱۰]. یکی از روش‌های کارا برای مسائل دسته‌بندی، روش ماشین‌های بردار پشتیبان است. این روش اولین بار توسط واپنیک[۱] ارائه شد که مدل خطی آن شامل یافتن ابرصفحه‌ای است که دو مجموعه را از هم جدا کند به طوری که فاصله بین ابرصفحه و نزدیک‌ترین نقاط به آن ماکزیمم شود. واپنیک و دستیارانش این روش را با استفاده از تابع هسته به حالت غیر خطی نیز تعمیم دادند و آن را به صورت یک مسئله برنامه‌ریزی درجه دوم با قیود خطی بیان کردند [۱۰، ۳۲، ۳۳]. در حال حاضر ماشین‌های بردار پشتیبان به ابزاری قدرتمند برای حل مسائل دسته‌بندی و رگرسیون تبدیل شده است. ماشین‌های بردار پشتیبان تعمیم پذیری بالایی دارند و بهینگی آن‌ها از نوع سراسری است، با این وجود فرض بر آن است که داده‌های بکار رفته در این روش، قطعی هستند، در حالی که در دنیای واقعی این طور نیست.

خطای اندازه‌گیری، خطای نمونه‌برداری و خطای مدلسازی ممکن است مانع شناخت دقیق داده‌های ورودی شود. بنابراین در اغلب مسائل با عدم‌قطعیت داده‌ها مواجه هستیم و این عدم‌قطعیت می‌تواند بر جواب مسئله اثرگذار باشد. بِن‌تال[۲] و نمیروفسکی[۳] مجموعه‌های محدب و کران‌دار عدم‌قطعیت را برای توصیف ضرایب ناقطعی در برنامه‌ریزی ریاضی معرفی کردند. آن‌ها با استفاده از این مجموعه‌های عدم‌قطعیت، یک روش بهینه‌سازی تحت داده‌های ناقطعی پیشنهاد دادند و آن را بهینه‌سازی استوار نامیدند [۶]. این مفهوم در زمینه‌های مختلف از جمله مسائل دسته‌بندی استفاده شده است [۳۰، ۳۱].

در این پایان‌نامه با بهره‌گیری از بهینه‌سازی استوار، ماشین‌های بردار پشتیبانی را به دست می‌آوریم که در برابر عدم‌قطعیت داده‌ها استوارند. ساختار این پایان‌نامه به صورت زیر است: در فصل اول مقدماتی از جبرخطی و بهینه‌سازی بیان می‌شود. فصل دوم به معرفی ماشین‌های بردار پشتیبان اختصاص دارد. در فصل سوم با در نظر گرفتن عدم‌قطعیت در داده‌های ورودی، مدل استوار ماشین‌های بردار پشتیبان بیان می‌شود. در فصل چهارم نیز به طور عملی به پیاده‌سازی و اجرای الگوریتم‌های دسته‌بندی و دسته‌بندی استوار مطرح شده در این پایان‌نامه می‌پردازیم و آن‌ها را برای حل چندین مسئله واقعی بکار می‌بریم. همچنین در این فصل از الگوریتم بهینه‌سازی

[۱] V.N. Vapnik [۲] A. Ben-Tal [۳] A. Nemirovski

مینیمال دنباله‌ای[۱] (SMO) استفاده شده است که روشی ابتکاری برای حل مسئله برنامه‌ریزی درجه دوم حاصل از دسته‌بندی معمولی است و در مقایسه با روش‌های کلاسیک حل مسائل درجه دوم بسیار سریع است.

[۱] Sequential minimal optimization

فصل ۱

پیش نیازهای جبر خطی و بهینه‌سازی

در این فصل برخی تعاریف و قضایای جبر خطی و بهینه‌سازی را که در فصل‌های بعدی مورد استفاده قرار می‌گیرند، بیان می‌کنیم.

۱-۱ مقدماتی از جبر خطی

۱-۱-۱ فضای برداری

تعریف ۱-۱-۱. یک فضای برداری (یا فضای خطی) روی میدان F مجموعه‌ای مانند V با دو عمل جمع و ضرب اسکالر است به طوری که به ازای هر $x, y \in V$ ، جمع x و y که با $(x+y)$ نشان می‌دهیم عضو یکتایی در V باشد، همچنین به ازای هر $x \in V$ و هر $\alpha \in F$، ضرب اسکالر α در x که با αx نشان می‌دهیم عضو یکتایی در V باشد، که در شرایط زیر صدق می‌کند:

۱. به ازای هر $x, y \in V$، $x + y = y + x$.

۲. به ازای هر $x, y, z \in V$، $(x + y) + z = x + (y + z)$.

۳. عضوی چون 0 در V وجود داشته باشد که به ازای هر $x \in V$، $x + 0 = x$.

۴. به ازای هر $x \in V$، $y \in V$ وجود داشته باشد به طوری که $x + y = 0$.

۵. به ازای هر $x \in V$، $1x = x$.

۶. به ازای هر $\alpha, \beta \in F$ و هر $x \in V$، $(\alpha\beta)x = \alpha(\beta x)$.

۷. به ازای هر $\alpha \in F$ و هر $x, y \in V$، $\alpha(x + y) = \alpha x + \alpha y$.

۸. به ازای هر $\alpha, \beta \in F$ و هر $x \in V$، $(\alpha + \beta)x = \alpha x + \beta x$.

مثال ۱-۱-۲. مجموعه $\mathbb{R}^n$ با دو عمل جمع و ضرب اسکالر مؤلفه‌ای تشکیل یک فضای برداری روی $\mathbb{R}$ می‌دهد. همچنین مجموعه همه ماتریس‌های $m \times n$ حقیقی با دو عمل جمع ماتریس‌ها و ضرب اسکالر یک فضای برداری روی $\mathbb{R}$ است که با $\mathbb{R}^{m\times n}$ یا $M_{m\times n}(\mathbb{R})$ نشان می‌دهند.

تعریف ۱-۱-۳. فرض کنید V یک فضای برداری روی F باشد. یک ضرب داخلی روی V تابعی است که برای هر $x, y \in V$، اسکالری چون $< x, y >$ در F نسبت می‌دهد به طوری که به ازای هر $x, y, z \in V$ و هر $c \in F$ داشته باشیم:

۱. $< x + y, z >=< x, z > + < y, z >$.

۲. $< cx, y >= c < x, y >$.

۳. $< \overline{x}, \overline{y} >=< y, x >$ ، که در آن علامت «$^{-}$» به معنی مزدوج مختلط است.

۴. اگر $x \neq 0$ آنگاه $< x, x > > 0$.

مثال ۱-۱-۴. به ازای $x = (x_1, \ldots, x_n)$ و $y = (y_1, \ldots, y_n)$ در $\mathbb{R}^n$، تعریف می‌کنیم
$$< x, y >= \sum_{i=1}^{n} x_i y_i.$$
به روشنی این یک ضرب داخلی در فضای $\mathbb{R}^n$ روی $\mathbb{R}$ است. این را ضرب داخلی استاندارد $\mathbb{R}^n$ می‌نامند و به صورت زیر نشان می‌دهند
$$< x, y >= x^T y.$$

تعریف ۱-۱-۵. یک فضای برداری که به یک ضرب داخلی مجهز شده باشد، فضای ضرب داخلی نام دارد.

مثال ۱-۱-۶. فضای برداری $\mathbb{R}^n$ با ضرب داخلی استاندارد، یکی از معروف‌ترین فضاهای ضرب داخلی است که آن را فضای اقلیدسی می‌نامند.

تعریف ۱-۱-۷. فرض کنید V یک فضای ضرب داخلی باشد. به ازای هر $x \in V$، نُرم یا طول x را به صورت زیر تعریف می‌شود
$$\|x\| = \sqrt{< x, x >}.$$

مثال ۱-۱-۸. برای بردار x در فضای اقلیدسی $\mathbb{R}^n$
$$\|x\| = \sqrt{< x, x >} = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}},$$
که آن را نرم اقلیدسی می‌نامند.

۱-۱-۲ فضاهای نرم دار

تعریف ۱-۱-۹. فرض کنید V یک فضای برداری روی $\mathbb{R}$ باشد، در اینصورت تابع $f : V \to \mathbb{R}$ یک نُرم در V نامیده می‌شود هرگاه دارای خواص زیر باشد:

۱. به ازای هر $\in V$، $f(x) \geq 0$.

۲. $f(x) = 0$ اگر و فقط اگر $x = 0$.

۳. به ازای هر $\alpha \in \mathbb{R}$ و $x \in V$، $f(\alpha x) =| \alpha | f(x)$.

۴. به ازای هر $x, y \in V$، $f(x+y) \leq f(x) + f(y)$.

به آسانی می‌توان ثابت کرد که تعاریف (۱-۱-۷) و (۱-۱-۹) هم‌ارزند [۱۳]. یک فضای برداری که دارای یک نرم باشد، فضای نرم دار می‌نامند.

۱-۱-۳ نرم برداری و ماتریسی

تعریف ۱-۱-۱۰. در فضای $\mathbb{R}^n$، یک نرم برداری به صورت تابع حقیقی و پیوسته $\|.\| : \mathbb{R}^n \to \mathbb{R}$ است که در شرایط زیر صدق کند:

۱. به ازای هر $x \in \mathbb{R}^n$، $\| x \| \geq ۰$ و $\| x \| = ۰$ اگر و فقط اگر $x = ۰$.

۲. به ازای هر $x \in \mathbb{R}^n$ و $\alpha \in \mathbb{R}$، $\| \alpha x \| = | \alpha | \| x \|$.

۳. به ازای هر $x, y \in \mathbb{R}^n$، $\| x + y \| \leq \| x \| + \| y \|$.

در ادامه چند نرم برداری مهم را برای فضای $\mathbb{R}^n$ معرفی می‌کنیم.

۱. نرم یک یا نرم مجموع:

$$\| x \|_۱ = \sum_{i=۱}^{n} | x_i | .$$

۲. نرم دو یا نرم اقلیدسی:

$$\| x \|_۲ = \left(\sum_{i=۱}^{n} x_i^۲ \right)^{\frac{۱}{۲}} .$$

۳. نرم بی‌نهایت یا نرم ماکزیمم:

$$\| x \|_\infty = \max_{۱ \leq i \leq n} | x_i | .$$

۴. $p-$نرم یا نرم هولدر برای $p \geqslant ۱$:

$$\| x \|_p = \left(\sum_{i=۱}^{n} x_i^p \right)^{\frac{۱}{p}} .$$

یکی از ویژگی‌های مهم نرم هولدر، نامساوی هولدر به صورت

$$|x^T y| \leqslant \|x\|_p \|y\|_q$$

است که در آن $۱ \leqslant p, q \leqslant \infty$ و $\frac{۱}{p} + \frac{۱}{q} = ۱$.

نامساوی کوشی-شوارتز حالت خاصی از نامساوی هولدر است

$$|x^T y| \leqslant \|x\|_۲ \|y\|_۲.$$

تعریف ۱-۱-۱۱. در فضای $\mathbb{R}^{m\times n}$، یک نرم ماتریسی به صورت تابع حقیقی و پیوسته $\|.\| : \mathbb{R}^{m\times n} \to \mathbb{R}$ است که در شرایط زیر صدق کند:

۱. به ازای هر $A \in \mathbb{R}^{m\times n}$، $\| A \|\geq \circ$ و $\| A \|= \circ$ اگر و فقط اگر A ماتریس صفر باشد.

۲. به ازای هر $A \in \mathbb{R}^{m\times n}$ و $\alpha \in \mathbb{R}$، $\| \alpha A \|=| \alpha | \| A \|$.

۳. به ازای هر $A, B \in \mathbb{R}^{m\times n}$، $\| A + B \|\leq \| A \| + \| B \|$.

ماتریس $A_{m\times n}$ و نرم برداری $\|.\|_p$ را در نظر بگیرید، تعریف می‌کنیم

$$\|A\|_p = \max_{x\neq \circ} \frac{\|Ax\|_p}{\|x\|_p}.$$

این تعریف تمام خواص نرم ماتریسی را دارد. این نرم را نرم ماتریسی سازگار با نرم برداری می‌نامند. با توجه به نرم‌های برداری به سادگی می‌توان نرم‌های ماتریسی زیر را به دست آورد، فرض کنید $A \in M_{m\times n}(\mathbb{R})$، آنگاه:

۱. $\| A \|_۱= \max_{۱\leq j\leq n} \sum_{i=۱}^{m} | a_{ij} |$ ، که نرم ماکزیمم مجموع ستونی نامیده می‌شود.

۲. $\| A \|_\infty= \max_{۱\leq i\leq m} \sum_{j=۱}^{n} | a_{ij} |$ ، که نرم ماکزیمم مجموع سطری نامیده می‌شود.

۳. $\| A \|_۲= \sqrt{\lambda_{\max}\left(A^TA\right)}$ ، که نرم طیفی نامیده می‌شود و در آن $\lambda_{max}\left(A^TA\right)$ نشان دهنده بزرگترین مقدار ویژه A^TA است.

یکی دیگر از نرم‌های ماتریسی معروف، نرم فروبنیوس[۱] است که به صورت زیر تعریف می‌شود

$$\| A \|_F= \left(\sum_{i=۱}^{m}\sum_{j=۱}^{n} a_{ij}^۲\right)^{\frac{۱}{۲}}.$$

اما این نرم با هیچ نرم برداری سازگار نیست.

۱-۱-۴ ماتریس

تعریف ۱-۱-۱۲. ماتریس متقارن A را معین مثبت گویند هرگاه برای هر بردار ناصفر x داشته باشیم

$$x^TAx > \circ.$$

همچنین ماتریس متقارن A نیمه معین مثبت است اگر به ازای هر x، $x^TAx \geqslant \circ$.

قضیه ۱-۱-۱۳ (*تجزیه چولسکی*[۲]). *فرض کنید $A \in \mathbb{R}^{n\times n}$ ماتریسی معین مثبت باشد، آنگاه ماتریس **بالا** مثلثی R با درایه‌های قطری اکیداً مثبت وجود دارد به طوری که*

$$A = R^TR.$$

[۱]Frobenuis [۲]Cholesky factorization

برهان. به [۱۵] رجوع شود. □

ماتریس مربعی A تجزیه چولسکی می‌شود اگر و تنها اگر متقارن و نیمه معین مثبت باشد. در حالت کلی تجزیه چولسکی ماتریس‌های نیمه معین مثبت یکتا نیست.

قضیه ۱-۱-۱۴ (*تجزیه مقدار تکین*[1] (*SVD*)). *فرض کنید* $A \in \mathbb{R}^{m\times n}$*، آنگاه ماتریس‌های متعامد*

$$U = [u_1, \dots, u_m] \in \mathbb{R}^{m\times m},\ V = [v_1, \dots, v_n] \in \mathbb{R}^{n\times n}$$

وجود دارند به طوری که

$$U^T AV = diag(\sigma_1, \sigma_2, \dots, \sigma_p) \in \mathbb{R}^{m\times n},$$

که در آن $p = \min\{m, n\}$ *و* $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$.

برهان. به [۱۵] رجوع شود. □

σ_i ها را مقادیر تکین ماتریس A می‌نامند. بردارهای u_i و v_i را به ترتیب i-امین بردار تکین چپ و i-امین بردار تکین راست گویند. اغلب تجزیه مقدار تکین ماتریس A را به صورت $A = U\Sigma V^T$ نشان می‌دهند که در آن

$$\Sigma = diag(\sigma_1, \sigma_2, \dots, \sigma_p).$$

هنگامی که ماتریس $A \in \mathbb{R}^{n\times n}$ معین مثبت است و $\lambda_1, \dots \lambda_n$ مقادیر ویژه آن و $q_1, \dots, q_n$ بردارهای ویژه وابسته آن‌ها باشند، تجزیه مقدار تکین را می‌توان به صورت

$$A = U\Lambda U^T$$

نوشت که در آن $U = [q_1, \dots, q_n]$ و $\Lambda = diag(\lambda_1, \dots \lambda_n)$. در این حالت مقادیر تکین با مقادیر ویژه منطبق می‌شوند.

۱-۲ توابع و مجموعه‌های محدب

تعریف ۱-۲-۱. مجموعه $S \subseteq \mathbb{R}^n$ محدب است اگر به ازای هر $x, y \in S$ و هر $\alpha \in [0, 1]$

$$\alpha x + (1 - \alpha)y \in S.$$

از لحاظ هندسی یک مجموعه محدب است اگر پاره‌خط واصل بین هر دو نقطه از آن، در آن مجموعه قرار گیرد.

[1] Singular value decomposition

مثال ۱-۲-۲. مجموعه‌های زیر نمونه‌هایی از مجموعه‌های محدب هستند.

۱. $\{(x_1, x_2): x_1^2 + x_2^2 \leq 1\}$.

۲. $\{x : Ax = b\}$، که در آن $A \in \mathbb{R}^{m\times n}$ و $b \in \mathbb{R}^m$.

۳. $\{x : Ax = b, x \geq 0\}$، که در آن $A \in \mathbb{R}^{m\times n}$ و $b \in \mathbb{R}^m$.

۴. $\{x : Ax \leq b, x \geq 0\}$، که در آن $A \in \mathbb{R}^{m\times n}$ و $b \in \mathbb{R}^m$.

تعریف ۱-۲-۳. تابع $f : S \to \mathbb{R}$ را در نظر بگیرید که در آن S زیرمجموعه محدب و ناتهی از $\mathbb{R}^n$ است. گوییم تابع f روی S محدب است اگر به ازای هر $x_1, x_2 \in S$ و هر $\alpha \in (0,1)$، داشته باشیم

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2).$$

تابع f را روی S اکیداً محدب گوییم هرگاه نامساوی فوق به صورت اکید برای هر x_1 و x_2 متمایز در S و برای هر $\lambda \in (0,1)$ برقرار باشد.

تابع $f : S \to \mathbb{R}$ را مقعر روی S گویند اگر $-f$ روی S محدب باشد.

قضیه ۱-۲-۴. *فرض کنید S یک مجموعه محدب باز ناتهی در $\mathbb{R}^n$ باشد و تابع $f : S \to \mathbb{R}$ دو بار مشتق‌پذیر روی S باشد. آنگاه f روی S محدب است اگر و فقط اگر ماتریس هسیان[1] f برای هر نقطه S نیمه معین مثبت[2] باشد.*

برهان. به [۵] رجوع شود. □

۱-۳ مقدماتی از بهینه‌سازی

عمده مسائل بهینه‌سازی در فصل‌های بعد از نوع برنامه‌ریزی درجه دوم و برنامه‌ریزی مخروطی درجه دوم است که در این بخش شکل کلی این دو بیان می‌شود.

۱-۳-۱ برنامه‌ریزی درجه دوم

یک مسئله بهینه‌سازی با تابع هدف درجه دوم و قیود خطی، مسئله برنامه‌ریزی درجه دوم[3] می‌نامند. شکل کلی آن به صورت زیر است

$$\begin{aligned} \min_{x} \quad & \frac{1}{2} x^T Q x + c^T x \\ \text{s.t.} \quad & Ax = b, \\ & Ex \leqslant d, \end{aligned} \tag{1-1}$$

[1] Hessian matrix [2] Positive semidefinite [3] Quadratic programming

که در آن $Q \in \mathbb{R}^{n\times n}$، $c \in \mathbb{R}^n$، $A \in \mathbb{R}^{m\times n}$، $b \in \mathbb{R}^m$، $E \in \mathbb{R}^{p\times n}$و $d \in \mathbb{R}^p$ معلوم و $x \in \mathbb{R}^n$ متغیر تصمیم است. حالت خاص مسئله (۱-۱) به صورت

$$\begin{aligned} \min_{x} \quad & \frac{1}{2}x^T Qx + c^T x \\ \text{s.t.} \quad & Ax = b, \\ & x \geqslant 0 \end{aligned} \tag{۱-۲}$$

است که در آن Q یک ماتریس معین مثبت است. مسئله (۱-۲)، یک مسئله محدب است و اگر ناحیه شدنی آن ناتهی باشد، آنگاه دارای مینیمم سراسری یکتایی است [۲۲]. دوگان لاگرانژ مسئله (۱-۲) به صورت

$$\begin{aligned} \max_{x,y} \quad & -\frac{1}{2}x^T Qx + b^T \alpha \\ \text{s.t.} \quad & -Qx + A^T y \leqslant c \end{aligned} \tag{۱-۳}$$

است. قضیه زیر شرایط بهینگی K.K.T[۱] را برای مسئله (۱-۲) بیان می‌کند.

قضیه ۱-۳-۱. *x و y جواب بهینه مسئله (۱-۲) و دوگان آن هستند اگر و فقط اگر*

۱. $Ax = b, x \geqslant 0$،

۲. $-Qx + A^T y \leqslant c$،

۳. $x^T(c + Qx - A^T y) = 0$.

برهان. به [۲۳] رجوع شود. □

۱-۳-۲ برنامه‌ریزی مخروطی درجه دوم

تعریف ۱-۳-۲. یک مخروط درجه دوم[۲] از بعد k به صورت زیر تعریف می‌شود

$$L^k = \left\{ \begin{bmatrix} u \\ t \end{bmatrix} \mid u \in \mathbb{R}^{k-1}, t \in \mathbb{R}, \|u\| \leqslant t \right\}.$$

همچنین L^k به مخروط بستنی[۳] یا مخروط لورنتز[۴] نیز معروف است. برای $k = 1$ مخروط درجه دوم واحد به صورت زیر است

$$L^1 = \{t \mid t \in \mathbb{R}, 0 \leqslant t\}.$$

[۱]Karush-Kuhn-Tucker [۲]Second-order cone [۳]Ice-cream cone [۴]Lorentz cone

شکل کلی یک مسئله برنامه‌ریزی مخروطی درجه دوم[1] (SOCP) به صورت

$$\begin{aligned} \min_{x} \quad & f^T x \\ \text{s.t.} \quad & \|A_i x + b_i\|_۲ \leqslant c_i^T x + d_i,\ i = ۱, \cdots, m \end{aligned} \tag{۱-۴}$$

است که در آن $x \in \mathbb{R}^n$ متغیر بهینه‌سازی است و $f \in \mathbb{R}^n$، $A_i \in \mathbb{R}^{k_i \times n}$، $b_i \in \mathbb{R}^{k_i}$، $c_i \in \mathbb{R}^n$ و $d_i \in \mathbb{R}$ داده‌های مسئله هستند. برای حل مسائل برنامه‌ریزی مخروطی درجه دوم از الگوریتم‌های نقطه درونی اولیه-دوگان استفاده می‌شود [۱، ۲۱، ۳۶].

[1] Second order cone programming