بسم الله الرحمن الرحیم

دانشگاه صنعتی شاهرود

دانشکده علوم ریاضی

رشته ریاضی محض، گرایش آنالیز ریاضی

پایان‌نامه کارشناسی ارشد

یافتن بهترین تقریب عنصری ازیک فضای ضرب داخلي نسبت به برد یک عملگر خطي روی یک چند وجهي

نگارنده: مجید میرزایی

استاد راهنما

دکتر مهدی ایرانمنش

شهریور ۱۳۹۶

شماره:
تاریخ:

باسمه‌تعالی

دانشگاه صنعتی شاهرود
مدیریت تحصیلات تکمیلی

فرم شماره (۳) صورتجلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم / آقای مجید میرزایی با شماره دانشجویی۹۴۱۷۴۶۴ رشته ریاضی محض گرایش آنالیز ریاضی تحت عنوان یافتن بهترین تقریب عنصری از یک فضای ضرب داخلی نسبت به برد یک عملگر خطی روی یک چند وجهی که در تاریخ ۱۳۹۶/۶/۲۲ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می گردد:

قبول (با درجه: خوب.......) ☑ مردود ☐

نوع تحقیق: نظری ☑ عملی ☐

عضو هیأت داوران	نام ونام خانوادگی	مرتبهٔ علمی	امضاء
۱ـ استادراهنمای اول	دکتر مهدی ایرانمنش	دانشیار	
۲- استادراهنمای دوم			
۳- استاد مشاور			
۴- نماینده تحصیلات تکمیلی	دکتر حسین باغیشنی	استادیار	
۵- استاد ممتحن اول	دکتر احمد معتمد نژاد	دانشیار	
۶ـ استاد ممتحن دوم	دکتر علیرضا خدامی	استادیار	علیرضا خدامی

نام و نام خانوادگی رئیس دانشکده: دکتر ابراهیم هاشمی

تاریخ و امضاء و مهر دانشکده:

تبصره: در صورتی که کسی مردود شود حداکثر یکبار دیگر (در مدت مجاز تحصیل) می تواند از پایان نامه خود دفاع نماید (دفاع مجدد نباید زودتر از ۴ ماه برگزار شود).

تقدیم به

پدر و مادر عزیز و مهربانم
که در سختی‌ها و دشواری‌های زندگی
همواره یاوری دلسوز و فداکار و پشتیبانی
محکم و مطمئن برایم بوده‌اند.

سپاس‌گزاری...

سپاس خداوندگار حکیم را که با لطف بی‌کران خود، آدمی را زیور عقل آراست.

در آغاز وظیفه خود می‌دانم از زحمات بی‌دریغ استاد راهنمای خود، جناب آقای دکتر مهدی ایرانمنش، صمیمانه تشکر و قدردانی کنم که قطعاً بدون راهنمایی‌های ارزنده ایشان، این مجموعه به انجام نمی‌رسید.

در پایان، بوسه می‌زنم بر دستان خداوندگاران مهر و مهربانی، پدر و مادر عزیزم و بعد از خدا، ستایش می‌کنم وجود مقدس‌شان را و تشکر می‌کنم از برادران عزیزم به پاس عاطفه سرشار و گرمای امیدبخش وجودشان، که در این سردترین روزگاران، بهترین پشتیبان من بودند.

مجید میرزایی

شهریور ۱۳۹۶

تعهد نامه

اینجانب **مجید میرزایی** دانشجوی کارشناسی ارشد رشته **ریاضی محض علوم ریاضی** دانشگاه شاهرود، نویسنده پایان‌نامه با عنوان **یافتن بهترین تقریب عنصری ازیک فضای ضرب داخلي نسبت به برد یک عملگر خطي روی یک چند وجهي** ، تحت راهنمایی **مهدی ایرانمنش** متعهد می‌شوم:

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های دیگر پژوهش‌گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان‌نامه، تا کنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ‌جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه صنعتی شاهرود تعلق دارد، و مقالات مستخرج با نام ‘‘ دانشگاه صنعتی شاهرود’’ یا “ Shahrood University of Technology ” به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به‌دست آوردن نتایج اصلی پایان‌نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافت‌های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده شده است)، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

مجید میرزایی
شهریور ۱۳۹۶

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه صنعتی شاهرود می‌باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان‌نامه بدون ذکر منبع مجاز نمی‌باشد.

چکیده

برد یک عملگر خطی روی یک جسم چند وجهی مطرح شده در این پایان‌نامه، زیر مجموعه محدب $\mathbb{H}$ از یک زیر فضای متناهی‌البعد $\mathbb{S}$ در فضای حاصلضرب داخلی $\mathbb{X}$ است. بر طبق اصل کاهش، بهترین تقریب یک نقطه از $\mathbb{X}$ روی $\mathbb{H}$ با بهترین تقریب $P_{\mathbb{S}}(x)$ روی $\mathbb{H}$ برابر است. پس طبق این اصل برای به‌دست آوردن بهترین تقریب x نسبت به $\mathbb{H}$، ابتدا باید بهترین تقریب x نسبت به $\mathbb{S}$ را به‌دست بیاوریم. برای این کار ممکن است $P_{\mathbb{S}}(x)$ موقعیت‌های مختلفی نسبت به $\mathbb{H}$ داشته باشد که این موقعیت‌ها را بررسی می‌کنیم و از آنها نتایجی ارائه می‌دهیم و با استفاده از قضیه بویل_ دیکسترا [1] بهترین تقریب x را نسبت به $\mathbb{H}$ با کمترین محاسبات به‌دست می‌آوریم. درنهایت کاربردی در داده‌های واقعی از صنعت نفت که یک راه حل برای یک مسأله خیلی مهم در عملیات تولید نفت و گاز که مسئله صلح نامیده می شود، ارائه می‌دهیم در جایی که سهم مشارکت تولیدات چاه‌های منحصر به فرد در تولیدات کل، ارزیابی و اندازه‌گیری می‌شود.

کلمات کلیدی: بهترین تقریب، اصل کاهش، زیرفضا، بویل دیکسترا

[1] Boyle-Dykstra

لیست مقالات مستخرج از پایان‌نامه

۱. بهترین تقریب روی مجموعه محدب در فضای ضرب داخلی، چهل وهشتمین کنفرانس ریاضی ایران، دانشگاه بوعلی سینا، همدان، ایران

فهرست مطالب

فهرست تصاویر

فصل ۱

تعاریف و مفاهیم اولیه

۱.۰.۱ مقدمه

در این فصل تعاریف مورد نیاز در فصل‌های بعدی به طور مختصر آورده شده است. که ابتدا برخی از مفاهیم ریاضی بصورت مختصر ارائه شده، سپس تعاریف و قضایای مرتبط به بهترین تقریب را بیان نموده و چند تعریف برای مخروط‌ها تعریف می‌کنیم. در ادامه ماتریس گرام اشمیت و فرایند متعامدسازی گرام اشمیت را بیان می‌کنیم و چند تعریف برای رسیدن به تعریف چندوجهی بیان می‌کنیم و در پایان فصل مفاهیم اولیه برای فرمولبندی کردن مسأله را بیان می‌کنیم.

۱.۱ مروری بر برخی مفاهیم ریاضی

تعریف ۱.۱.۱. فضای خطی $\mathbb{X}$ را فضای حاصلضرب داخلی گوییم اگر برای هر عضو x ، y و z در $\mathbb{X}$ و برای هر اسکالر α از $\mathbb{R}$ خصوصیات زیر را داشته باشیم:

(۱) $\langle x,x\rangle \geq 0$

(۲) $\langle x,x\rangle = 0 \Rightarrow x = 0$

(۳) $\langle x,y\rangle = \langle y,x\rangle$

(۴) $\langle \alpha x,y\rangle = \alpha\langle x,y\rangle$

(۵) $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle.$

تعریف ۲.۱.۱. اگر $\mathbb{X}$ و $\mathbb{Y}$ دو فضای حاصلضرب داخلی باشند آنگاه نگاشت $F:\mathbb{X}\to\mathbb{Y}.$ را خطی می‌گوییم اگر و فقط اگر برای هر x و y در $\mathbb{X}$ و α و β در $\mathbb{R}$ داشته باشیم:

$$F(\alpha x+\beta y) = \alpha F(x) + \beta F(y).$$

به نگاشت خطی، عملگر خطی نیز گفته می‌شود.

تعریف ۳.۱.۱. فرض کنید $\mathbb{X}$ یک فضای حاصلضرب داخلی و $x\in\mathbb{X}$ باشد، نرم x به صورت زیر تعریف می‌شود:

$$\|x\| = \sqrt{\langle x,x\rangle}.$$

و برای هر $x,y\in\mathbb{X}$ و هر $\alpha\in\mathbb{R}$ دارای خاصیت‌های زیر می‌باشد:

(۱) $\|x\| \geq 0.$

(۲) $\|x\| = 0 \Leftrightarrow x = 0.$

(۳) $\|\alpha x\| = |\alpha|\|x\|.$

(۴) $\|x+y\| \leq \|x\| + \|y\|.$

تعریف ۴.۱.۱. دو بردار x و y را متعامد می‌گوییم و با نماد $x\perp y$ نمایش می‌دهیم اگر $\langle x,y\rangle = 0$. بردار x را متعامد بر مجموعه $\mathbb{Y}$ می‌گوییم و با نماد $x\perp\mathbb{Y}$ نمایش می‌دهیم اگر برای $y\in\mathbb{Y}$ داشته باشیم $x\perp y$. مجموعه $\mathbb{A}$ را مجموعه متعامد می‌گوییم اگر هر جفت از بردارهای متمایز $\mathbb{A}$ متعامد باشند. پس مجموعه $\{x_i \quad | i\in I\}$ متعامد است اگر و فقط اگر

$$\langle x_i,x_j\rangle = \delta_{ij} := \begin{cases} 0 & \text{اگر} i\neq j, \\ 1 & \text{اگر} i=j. \end{cases}$$

و پایه‌های متعامدی که طول تک‌تک بردارهای آنها با هم برابر و برابر یک است پایه یکه متعامد می‌گوییم.

تعریف ۵.۱.۱. دنباله $\{x_n\}$ در فضای ضرب داخلی $\mathbb{X}$ را همگرا به $x\in\mathbb{X}$ گوییم و با $x_n\longrightarrow x$ نشان می‌دهیم هرگاه $\lim_{n\longrightarrow\infty}\|x_n-x\| = 0$ به عبارت دیگر هرگاه برای هر $\epsilon>0$ عدد صحیح N موجود باشد به‌طوری که

$$\forall n\geq N \quad , \|x_n - x\| < \epsilon.$$

تعریف ۶.۱.۱. دنباله $\{x_n\}$ در فضای ضرب داخلی $\mathbb{X}$ را کوشی گوییم اگر برای هر $\epsilon > ۰$ عدد صحیح N موجود باشد به طوری که

$$\forall n, m \geq N, \quad \|x_n - x_m\| < \epsilon.$$

و مجموعه ناتهی $\mathbb{B}$ را کراندار گوییم اگر

$$\sup\{\|x\| \quad | \quad x \in \mathbb{B}\} < \infty.$$

تعریف ۷.۱.۱. فضای ضرب داخلی $\mathbb{X}$ را کامل یا فضای هیلبرت گوییم اگر هر دنباله کوشی در $\mathbb{X}$ همگرا به نقطه‌ای در $\mathbb{X}$ باشد.

تعریف ۸.۱.۱. مجموعه $\mathbb{A}$ را بسته می گوییم اگر و فقط اگر به ازای هر دنباله مانند $\{x_n\}$ که $x_n \to x$ و به ازای هر $n \in \mathbb{N}$ داشته باشیم $x_n \in \mathbb{A}$ آنگاه نتیجه شود $x \in \mathbb{A}$.

۲.۱ بهترین تقریب

تعریف ۱.۲.۱. فرض کنید $\mathbb{A}$ زیر مجموعه ناتهی از فضای حاصلضرب داخلی $\mathbb{X}$ و $x \in \mathbb{X}$ باشد. عضو $a_\circ \in \mathbb{A}$ بهترین تقریب، یا نزدیکترین نقطه به x در مجموعه $\mathbb{A}$ می‌گوییم، اگر

$$\|x - a_\circ\| = d(x, \mathbb{A}).$$

که

$$d(x, \mathbb{A}) := \inf\{\|x - a\|; a \in \mathbb{A}\}.$$

عدد $d(x, \mathbb{A})$ فاصله x تا $\mathbb{A}$ ، یا خطای تقریب عنصر x به وسیله‌ی مجموعه $\mathbb{A}$ گفته می‌شود. مجموعه‌ی تمام تقریب‌های عنصر x از مجموعه‌ی $\mathbb{A}$ را با $P_{\mathbb{A}}(x)$ نمایش می‌دهیم. بنابراین

$$P_{\mathbb{A}}(x) = \{a \in \mathbb{A} : \quad \|x - a\| = d(x, \mathbb{A})\}.$$

اگر هر $x \in \mathbb{X}$ حداقل یک بهترین تقریب در مجموعه $\mathbb{A}$ داشته باشد، $\mathbb{A}$ یک مجموعه پروکسیمینال نامیده می‌شود. به عبارت دیگر $\mathbb{A}$ یک مجموعه پروکسیمینال است اگر و فقط اگر برای هر $x \in \mathbb{X}$، داشته باشیم $P_{\mathbb{A}}(x) \neq \emptyset$.

اگر هر $x \in \mathbb{X}$ دقیقا یک بهترین تقریب در مجموعه $\mathbb{A}$ داشته باشد، $\mathbb{A}$ یک مجموعه چبیشف نامیده می‌شود. به عبارت دیگر $\mathbb{A}$ یک مجموعه چبیشف است اگر و فقط اگر برای هر $x \in \mathbb{X}$، $P_{\mathbb{A}}(x)$ یک مجموعه‌ی تک عضوی باشد.

تعریف ۲.۲.۱. زیر مجموعه‌ی $\mathbb{A}$ از $\mathbb{X}$ را مجموعه محدب می‌گوییم اگر $x, y \in \mathbb{A}$ باشد، آنگاه به ازای هر $۰ \leq \lambda \leq ۱$ داشته باشیم:

$$\lambda x + (۱ - \lambda) y \in \mathbb{A}.$$

که $\{\lambda x+(1-\lambda)y\in\mathbb{A}: 0\leq\lambda\leq 1\}$ پاره‌خط واصل x به y است. پس مجموعه $\mathbb{A}$ محدب است اگر و فقط اگر به ازای هر $x,y\in\mathbb{A}$ تمام نقاط پاره خط واصل یعنی x و y نیز در $\mathbb{A}$ باشد.

تعریف ۳.۲.۱. زیر مجموعه $\mathbb{A}$ از فضای ضرب داخلی $\mathbb{X}$ را مخروط محدب می‌گوییم درصورتی که اگر برای هر $x,y\in\mathbb{X}$ و $\alpha,\beta\geq 0$ داشته باشیم:

$$\alpha x+\beta y\in\mathbb{A}.$$

برای مثال هر خط عبور کننده از مبدأ مختصات و مخروط دوگان و نیم‌فضا که در ادامه تعریف خواهیم کرد مخروط محدب هستند.

قضیه ۱.۲.۱. منحصربه فرد بودن بهترین تقریب: فرض کنید $\mathbb{H}$ زیر مجموعه محدب از $\mathbb{X}$ باشد آنگاه برای هر $x\in\mathbb{X}$ حداکثر یک بهترین تقریب در $\mathbb{H}$ داریم.

برهان. فرض کنید $x\in\mathbb{X}$ و فرض کنید y_1 و y_2 در $P_{\mathbb{H}}(x)$ باشد. آنگاه $\frac{y_1+y_2}{2}\in\mathbb{H}$ است و داریم:

$$d(x,\mathbb{H})=\|x-\frac{1}{2}(y_1+y_2)\|=\|\frac{1}{2}(x-y_1)+\frac{1}{2}(x-y_2)\|\leq\frac{1}{2}\|(x-y_1)\|+\frac{1}{2}\|(x-y_2)\|=d(x,\mathbb{H})$$

از این‌رو تساوی باید در این نابرابری صدق کند که شرط تساوی در نامساوی نامثلث $y_1=y_2$ و $x-y_1=\rho(x-y_2)$ برای هر $\rho\geq 0$ برقرار باشد. اما $\|(x-y_1)\|=d(x,\mathbb{H})=\|(x-y_2)\|$ پس نتیجه می‌گیریم $\rho=1$ و از این رو $y_1=y_2$. □

تعریف ۴.۲.۱. زیر مجموعه ناتهی $\mathbb{M}$ از $\mathbb{X}$ را زیرفضا می‌گوییم اگر برای هر $x,y\in\mathbb{X}$ و $\alpha,\beta\in\mathbb{R}$ داشته باشیم:

$$\alpha x+\beta y\in\mathbb{M}.$$

تعریف ۵.۲.۱. فرض کنید $\mathbb{A}$ زیر مجموعه ناتهی از فضای حاصلضرب داخلی $\mathbb{X}$ باشد، زیرفضای تولیدشده توسط $\mathbb{A}$ را با $span(\mathbb{A})$ نمایش می‌دهیم، که مجموعه همه ترکیبات خطی از عناصر $\mathbb{A}$ است. پس داریم:

$$span(\mathbb{A})=\{\sum_{i=1}^{n}\alpha_i x_i \quad | \quad x_i\in\mathbb{A},\alpha_i\in\mathbb{R},n\in\mathbb{N}\}.$$

تعریف ۶.۲.۱. فرض کنید $\mathbb{S}$ زیر مجموعه ناتهی از فضای حاصلضرب داخلی $\mathbb{X}$ باشد، مخروط دوگان $\mathbb{S}$ را به‌صورت زیر تعریف می‌کنیم:

$$\mathbb{S}^\circ:=\{x\in\mathbb{X} \quad | \quad \langle x,y\rangle\leq 0 \quad \forall y\in\mathbb{S}\}.$$

و متمم متعامد $\mathbb{S}$ مجموعه زیر است،

$$\mathbb{S}^\perp:=\mathbb{S}^\circ\cap(-\mathbb{S}^\circ)=\{x\in\mathbb{X} \quad | \quad \langle x,y\rangle=0 \quad \forall y\in\mathbb{S}\}.$$

تعریف ۷.۲.۱. پوسته مخروط $\mathbb{S}$ را با $con(\mathbb{S})$ نشان می‌دهیم، که اشتراک همه مخروط‌های محدب که شامل $\mathbb{S}$ است و پوسته مخروط به فرم زیر تعریف می‌شود:

$$con(\mathbb{S}) = \{\sum_{i=1}^{n} \rho_i x_i \quad | \quad x_i \in \mathbb{S}, \rho_i \geq 0, n \in \mathbb{N}\}.$$

تذکر: قضایا و تعاریف تا پایان این فصل از [۲] گرفته شده است.

۳.۱ ماتریس گرام اشمیت

قضیه ۱.۳.۱. (مشخص‌سازی بهترین تقریب از روی زیر فضا) فرض کنید $\mathbb{M}$ زیرفضایی از $\mathbb{X}$ باشد و فرض کنید $x \in \mathbb{X}$ و $y_0 \in \mathbb{M}$ باشد. آنگاه $y_0 = P_{\mathbb{M}}(x)$ اگر و فقط اگر $x - y_0 \in \mathbb{M}^{\perp}$.

برهان. برای اثبات به [۲] مراجعه می‌کنیم. □

قضیه ۲.۳.۱. (معادلات نرمال) فرض کنید $\{x_1, x_2, \dots, x_n\}$ یک پایه برای زیرفضای n بعدی $\mathbb{M}$ از $\mathbb{X}$ باشد. آنگاه $\mathbb{M}$ زیرفضای چیبشف است اگر برای هر اسکالر α_i و $x \in \mathbb{X}$ داشته باشیم:

$$P_{\mathbb{M}}(x) = \sum_{i=1}^{n} \alpha_i x_i \tag{۱.۱}$$

که در آن α_i به صورت منحصر به فردی از معادلات نرمال زیر بدست می‌آید:

$$\sum_{i=1}^{n} \alpha_i \langle x_i, x_j \rangle = \langle x, x_j \rangle \qquad (j = 1, 2, \dots, n). \tag{۲.۱}$$

به‌خصوص اگر $\{x_1, x_2, \dots x_n\}$ پایه متعامد برای $\mathbb{M}$ باشد، آنگاه

$$P_{\mathbb{M}}(x) = \sum_{i=1}^{n} \langle x, x_i \rangle x_i \qquad \forall x \in \mathbb{X}. \tag{۳.۱}$$

برهان. فرض کنید $x \in \mathbb{X}$ و $y_0 \in \mathbb{M}$ ثابت باشند. باتوجه به تعریف ۴.۲.۱ برای هر اسکالر α_i داریم $y_0 = \sum_{i=1}^{n} \alpha_i x_i$ با توجه به قضیه ۱.۳.۱، $y_0 = P_{\mathbb{M}}(x)$ که معادل است با $x - y_0 \in \mathbb{M}^{\perp}$ و برای هر $x_j \in \mathbb{M}$ داریم $\langle x - y_0, x_j \rangle = 0$ پس برای هر $j = 1, 2, \dots, n$ داریم:

$$\langle x - y_0, x_j \rangle = 0 \Rightarrow \langle x - \sum_{i=1}^{n} \alpha_i x_i, x_j \rangle = 0 \Rightarrow \langle x, x_j \rangle - \sum_{i=1}^{n} \alpha_i \langle x_i, x_j \rangle = 0$$

$$\Rightarrow \sum_{i=1}^{n} \alpha_i \langle x_i, x_j \rangle = \langle x, x_j \rangle$$

بنابراین معادله (۲.۱) برقرار است. اگر $\{x_1, x_2, \dots x_n\}$ پایه متعامد برای $\mathbb{M}$ باشد با توجه به تعریف ۴.۱.۱، $\langle x_i, x_j \rangle = \delta_{ij}$. بنابراین α_i در معادله (۲.۱) برابر $\langle x, x_i \rangle$ است بنابراین با توجه به معادله (۱.۱)، $P_{\mathbb{M}}(x) = \sum_{i=1}^{n} \langle x, x_i \rangle x_i$. درنتیجه برای هر i معادله (۳.۱) برقرار است. □

در نمادگذاری ماتریسی، معادلات نرمال (۲.۱) می‌تواند به صورت زیر بازنویسی شود:

$$G(x_1, x_2, \dots, x_n)\alpha = \beta \tag{۴.۱}$$

که به

$$G(x_1, x_2, \dots, x_n) := \begin{bmatrix} \langle x_1, x_1\rangle & \langle x_2, x_1\rangle & \cdots & \langle x_n, x_1\rangle \\ \langle x_1, x_2\rangle & \langle x_2, x_2\rangle & \cdots & \langle x_n, x_1\rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle x_1, x_n\rangle & \langle x_2, x_n\rangle & \cdots & \langle x_n, x_n\rangle \end{bmatrix},$$

ماتریس گرام $\{x_1, x_2, \dots, x_n\}$ می‌گویند و α و β بردارهایی ستونی هستند که $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$ و $\beta = (\langle x, x_1\rangle, \langle x, x_2\rangle, \dots, \langle x, x_n\rangle)^T$ است و دترمینان ماتریس گرام در نامساوی زیر صدق می‌کند:

$$0 \le \det G(x_1, x_2, \dots, x_n) \le \|x_1\|^2 \|x_2\|^2 \dots \|x_n\|^2.$$

و تساوی در سمت چپ(به‌ترتیب راست) برقرار است اگر و فقط اگر $\{x_1, x_2, \dots, x_n\}$ مجموعه‌ای مستقل خطی (به‌ترتیب متعامد) باشد. همچنین ماتریس گرام وارون‌پذیر است که در [۲] اثبات شده است.

۱.۳.۱ فرایند متعامد سازی گرام اشمیت

قضیه ۳.۳.۱. فرض کنید $\{x_1, x_2, \dots\}$ یک مجموعه متناهی یا شمارای نامتناهی از بردارهای مستقل خطی باشد و فرض کنید $M_0 = \{0\}$ باشد و برای $n \ge 1$، $M_n = span\{x_1, x_2, \dots, x_n\}$ اگر:

$$y_n = x_n - P_{M_{n-1}}(x_n), \qquad e_n = \frac{y_n}{\|y_n\|} \qquad \forall n \ge 1$$

آنگاه داریم:

(۱) $\{y_1, y_2, \dots\}$ مجموعه متعامد است.

(۲) $\{e_1, e_2, \dots\}$ مجموعه متعامد است.

(۳) $n \ge 1.$ $\quad span\{y_1, y_2, \dots, y_n\} = span\{e_1, e_2, \dots, e_n\} = M_n$

(۴) $y_1 = x_1$ و

$$y_n = x_n - \sum_{i=1}^{n-1} \langle x_n, e_n\rangle e_i \qquad \forall n \ge 2.$$

برهان. فرض کنید $\{x_1, x_2, \dots, x_n\}$ پایه‌ای برای $\mathbb{M}_n$ باشد. می‌خواهیم $\{x_1, x_2, \dots, x_n\}$ را به یک پایه متعامد تبدیل کنیم.

$$\mathbb{M}_1 = span(x_1), \quad e_1 = \frac{x_1}{\|x_1\|} = \|\frac{x_1}{\|x_1\|}\| = \frac{1}{\|x_1\|}\|x_1\| = 1.$$

پس e_1 یک پایه متعامد برای $\mathbb{M}_1$ است و $\|x_1\|e_1 = x_1$ پس داریم:

$$\mathbb{M}_1 = span(x_1) = span(e_1).$$

با توجه به شکل ۱.۱ $x_2 = x + y_2$ که $x \in \mathbb{M}_1$ و $y_2 \in \mathbb{M}_1^{\perp}$ پس $x = P_{\mathbb{M}_1}(x_2)$ پس داریم

$$x_2 = P_{\mathbb{M}_1}(x_2) + y_2 \Longrightarrow y_2 = x_2 - P_{\mathbb{M}_1}(x_2).$$

از طرفی داریم:

$$\mathbb{M}_2 = span(x_1, x_2) = span(e_1, x_2) = span(e_1, y_2).$$

چون e_1 پایه متعامد $\mathbb{M}_1$ است پس داریم:

$$P_{\mathbb{M}_1}(x_2) = \langle x_2, e_1\rangle e_1 \Longrightarrow y_2 = x_2 - \langle x_2, e_1\rangle e_1.$$

و چون $e_2 = \frac{y_2}{\|y_2\|}$ پس داریم:

$$\mathbb{M}_2 = span(e_1, y_2) = span(e_1, e_2) \Rightarrow x_2 = y_2 + \langle x_2, e_1\rangle e_1.$$

پس x_2 پایه متعامد $\mathbb{M}_2$ است.

$$\mathbb{M}_3 = span(e_1, e_2, x_3) \Rightarrow y_3 = x_3 - P_{\mathbb{M}_2}(x_3).$$

از طرفی

$$P_{\mathbb{M}_2}(x_3) = \langle x_3, e_1\rangle e_1 + \langle x_3, e_2\rangle e_2 \Rightarrow y_3 = x_3 - \langle x_3, e_1\rangle e_1 + \langle x_2, e_2\rangle e_2$$

$$\Rightarrow \mathbb{M}_3 = span(e_1, e_2, y_3).$$

و چون $e_3 = \frac{y_3}{\|y_3\|}$ است پس

$$\mathbb{M}_3 = span(e_1, e_2, x_3) = span(e_1, e_2, e_3).$$

پس x_3 یک پایه متعامد برای $\mathbb{M}_3$ و به همین ترتیب داریم:

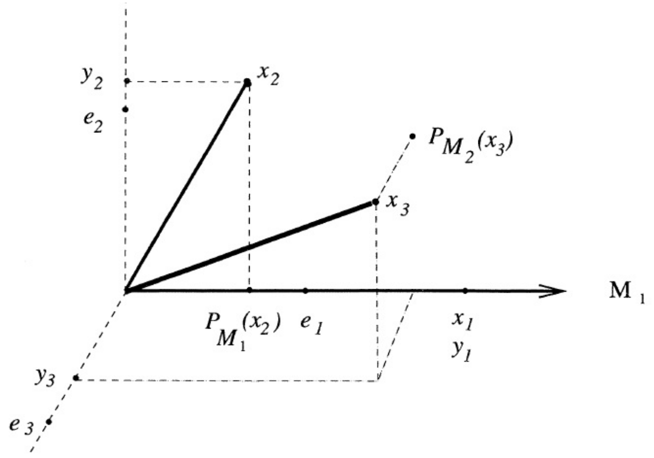
$$\mathbb{M}_4 = span(e_1, e_2, e_3, x_4) \Rightarrow y_4 = x_4 - P_{\mathbb{M}_3}(x_4) \Rightarrow x_4 = y_4 + P_{\mathbb{M}_3}(x_4).$$

پس داریم:

$$\mathbb{M}_3 = span(e_1, e_2, e_3, x_4) = span(e_1, e_2, e_3, e_4).$$

پس x_4 یک پایه متعامد برای $\mathbb{M}_4$ است و این فرایند ادامه دارد که به فرایند گرام اشمیت معروف است. □

در شکل ۱.۱ می‌توان فرایند گرام اشمیت را مشاهده کرد:



شکل ۱.۱: فرایند گرام اشمیت

قضیه ۴.۳.۱. فرض کنید $\{e_۱, e_۲, \ldots, e_n\}$ مجموعه متناهی متعامد در فضای حاصلضرب داخلی $\mathbb{X}$ و $M := span\{e_۱, e_۲, \ldots, e_n\}$ باشد.آنگاه برای هر $x \in \mathbb{X}$ داریم:

$$P_{\mathbb{M}}(x) = \sum_{i=۱}^{n} \langle x, e_i \rangle e_i$$

و

$$\sum_{i=۱}^{n} |\langle x, e_i \rangle|^۲ \leq \|x\|^۲.$$

تعریف ۱.۳.۱. سری فوریه فرض کنید ϑ مجموعه متعامد در فضای حاصلضرب داخلی $\mathbb{X}$ و $x \in \mathbb{X}$ باشد.اگر ϑ_x مجموعه شمارای تعریف شده در لم ؟؟ باشد، آنگاه سری فوریه x نسبت به ϑ به صورت زیر است:

$$\sum_{e \in \vartheta} \langle x, e \rangle e := \begin{cases} ۰ & \text{اگر}\ \vartheta_x = \emptyset \\ \sum_{i=۱}^{n} \langle x, e_i \rangle e_i & \text{اگر}\ \vartheta_x = \{e_۱, e_۲, \ldots, e_n\} \end{cases}$$

و اگر $\vartheta_x = \{e_۱, e_۲, \ldots\}$ شمارای نامتناهی باشد، داریم:

$$\sum_{e \in \vartheta} \langle x, e \rangle e := \lim_{n} \sum_{i=۱}^{n} \langle x, e_i \rangle e_i.$$

که این حد وجود دارد و به $\langle x, e \rangle$ ضرایب فوریه می‌گوییم.سری فوریه $\sum_{e \in \vartheta} \langle x, e \rangle e$ همیشه همگراست زمانی که زیرفضای بسته $\mathbb{M} := \overline{span\vartheta}$ در فضای حاصلضرب داخلی $\mathbb{X}$ کامل باشد.در حقیقت در هر یک از این موارد داریم:

$$\sum_{e \in \vartheta} \langle x, e \rangle e = P_{\mathbb{M}}(x) \qquad \forall x \in \mathbb{X}. \qquad (۵.۱)$$

مجموعه متعامد ϑ را با خاصیت $\mathbb{M} := \overline{span\vartheta}$ پایه متعامد $\mathbb{M}$ می‌گوییم.در حقیقت ϑ پایه متعامد $\mathbb{M}$ است اگر و فقط اگر برای هر $x \in \mathbb{M}$ داشته باشیم:

$$x = \sum_{e\in\vartheta} \langle x, e\rangle e \tag{۶.۱}$$

۴.۱ چندوجهی

تعریف ۱.۴.۱. فرض کنید $\mathbb{X}$ فضای حاصلضرب داخلی باشد. به تابع حقیقی f در $\mathbb{X}$ تابعک خطی می‌گوییم اگر برای هر $x, y \in \mathbb{X}$ و $\alpha, \beta \in \mathbb{R}$ داشته باشیم:

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y).$$

تابعک خطی f را کراندار می‌گوییم، اگر ثابت C وجود داشته باشد به طوری که برای هر $x \in \mathbb{X}$ داشته باشیم:

$$|f(x)| \leq C\|x\|.$$

کوچکترین C را نرم f می‌گوییم و داریم:

$$\|f\| = \sup_{x\neq ۰} \frac{|f(x)|}{\|x\|} = \sup_{\|x\|\leq ۱} |f(x)| = \sup_{\|x\|=۱} |f(x)|. \tag{۷.۱}$$

تعریف ۲.۴.۱. فضای دوگان فضای حاصلضرب داخلی $\mathbb{X}$ را با $\mathbb{X}^*$ نشان می‌دهیم.که مجموعه همه تابعک‌های کراندار در $\mathbb{X}$ است. به علاوه روی این فضا جمع و ضرب اسکالری برای هر $x \in \mathbb{X}$ و $\alpha \in \mathbb{R}$ به صورت زیر تعریف می‌شود:

$$(f + g)(x) := f(x) + g(x),$$

$$(\alpha f)(x) = \alpha f(x)$$

تعریف ۳.۴.۱. یک ابر صفحه در فضای حاصلضرب داخلی $\mathbb{X}$ برای هر $c \in \mathbb{R}$ و $x^* \in \mathbb{X}^* - \{۰\}$ مجموعه ای به فرم زیر است:

$$\mathrm{H} = \{y \in \mathbb{X} \quad | \quad x^*(y) = c\}.$$

کرنل یا فضای پوچ یک تابعک $x^* \in \mathbb{X}^*$ مجموعه‌ای به فرم زیر است:

$$\ker x^* := \{y \in \mathbb{X} \quad | \quad x^*(y) = ۰\}.$$

همچنین فاصله تا ابر صفحه به صورت زیر محاسبه می شود:

$$d(x, \mathrm{H}) = \frac{۱}{\|x^*\|}|x^*(x) - c| \qquad \forall x \in \mathbb{X}.$$

تعریف ۴.۴.۱. یک نیم فضا در فضای حاصلضرب داخلی $\mathbb{X}$ برای هر $c \in \mathbb{R}$ و $x^* \in \mathbb{X}^* - \{0\}$ مجموعه ای به فرم زیر است:

$$\mathcal{H} = \{y \in \mathbb{X} \quad | \quad x^*(y) \le c\} \qquad \text{یا} \qquad \mathcal{H} = \{y \in \mathbb{X} \quad | \quad x^*(y) \ge c\}.$$

به عبارت دیگر نیم فضا مجموعه همه نقاطی است که روی هم ابرصفحه یا در یک طرف ابرصفحه قرار دارند.

تذکر: هر نیم فضا به فرم $\mathcal{H} = \{y \in \mathbb{X} \quad | \quad x^*(y) \ge c\}$ را می توان به فرم :

$$\mathcal{H} = \{y \in \mathbb{X} \quad | \quad x_1^*(y) \le c_1\}$$

نوشت که $x_1^* = -x^*$ و $c_1 = -c$ است.
پس در تعریف نیم فضا تنها فرم $\mathcal{H} = \{y \in \mathbb{X} \quad | \quad x^*(y) \le c\}$ کافی است.

تعریف ۵.۴.۱. یک جسم چندوجهی یا مجموعه چندوجهی در فضای حاصلضرب داخلی مجموعه‌ای ناتهی است که اشتراک تعداد متناهی از نیم فضاها است بنابراین $\mathcal{Q}$ جسمی چند وجهی در فضای حاصلضرب داخلی است اگر و فقط اگر برای هر $c_i \in \mathbb{R}$ و $x^* \in \mathbb{X}^* - \{0\}$ داشته باشیم:

$$Q = \cap_{i=1}^{n} \{y \in \mathbb{X} \quad | \quad x_i^*(y) \le c_i\} \neq \emptyset \qquad i = 1, 2, \ldots, n$$

لم ۱.۴.۱. اگر بردار ناصفر x یک ترکیب خطی مثبت از بردارهای ناصفر $\{x_1, x_2, \ldots, x_n\}$ باشد آنگاه x ترکیب خطی مثبت از زیرمجموعه‌ای مستقل خطی از $\{x_1, x_2, \ldots, x_n\}$ است.

تعریف ۶.۴.۱. مخروط محدب $\mathbb{A}$ در فضای حاصلضرب داخلی $\mathbb{X}$ متناهیاً تولید شده گفته می‌شود اگر پوسته مخروط از تعداد متناهی بردار تشکیل شده باشد پس $\mathbb{A}$ متناهیا تولید شده است اگر برای هر زیر مجموعه متناهی $\{x_1, x_2, \ldots, x_n\}$ از $\mathbb{X}$ داشته باشیم:

$$\mathbb{A} = con\{x_1, x_2, \ldots, x_n\} = \{\sum_{i=1}^{n} \rho_i x_i \quad | \quad \rho_i \ge 0\}.$$

۵.۱ تعریف مسأله

این بخش را با فرموله کردن مسأله مطرح شده در این پایان‌نامه آغاز می‌کنیم:
فرض می‌کنیم فضای حاصلضرب داخلی $\mathbb{X}$ با حاصلضرب داخلی $\langle \cdot, \cdot \rangle : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ و یک مجموعه متناهی از عناصر مستقل خطی از $\mathbb{X}$ به صورت زیر داده شده است:

$$\mathbb{Y} = \{y_1, \ldots, y_n\}. \tag{۸.۱}$$

و نیز فرض می‌کنیم گردایه‌ای از اعداد حقیقی $\alpha_1, \cdots, \alpha_n$ و $\beta_1, \cdots, \beta_n$ که $\alpha_i < \beta_i$ و مجموعه زیر داده شده باشد :

$$\Gamma = \{\chi \in \mathbb{R}^n \mid \chi_i \in [\alpha_i, \beta_i] \quad, i = 1, \ldots, n\} \tag{۹.۱}$$

که $[\alpha_i, \beta_i]$ در $\mathbb{R}$ بسته است. نگاشت :

$$Q_{\mathbb{Y}} : \Gamma \to \mathbb{X}$$

که با ضابطه‌ی $\gamma \longmapsto \sum_{i=1}^{n} \gamma_i y_i$ تعریف می‌شود که $\gamma_i = (\gamma)_i$ به این معنی که مختص i ام γ در $\mathbb{R}^n$ است. این نگاشت یک عملگر خطی است زیرا:

$$\begin{aligned} Q_{\mathbb{Y}}(\alpha\gamma_1 + \beta\gamma_2) &= \sum_{i=1}^{2}(\alpha\gamma_1 + \beta\gamma_2)y_i \\ &= (\alpha\gamma_1)y_1 + (\beta\gamma_2)y_2 = \alpha(\gamma_1 y_1) + \beta(\gamma_2 y_2) = \alpha Q_{\mathbb{Y}}(\gamma_1) + \beta Q_{\mathbb{Y}}(\gamma_2) \end{aligned}$$

آنگاه برد $Q_{\mathbb{Y}}$ به صورت زیر تعریف می شود:

$$Q_{\mathbb{Y}}(\Gamma) := \mathbb{H} = \{y \in \mathbb{X} \mid y = \sum_{i=1}^{n} \gamma_i y_i, y_i \in \mathbb{Y}, \gamma \in \Gamma\} \qquad (۱۰.۱)$$

در این پایان‌نامه می‌خواهیم بهترین تقریب $x \in \mathbb{X}$ را نسبت به $\mathbb{H}$ به‌دست آوریم.
اینک به توضیح‌های زیرین درباره‌ی مسأله بالا توجه می‌کنیم:

- مجموعه Γ در معادله (۹.۱) در $\mathbb{R}^n$ یک جسم چند وجهی است. که از [۷] گرفته شده است.
- مجموعه $\mathbb{H}$ در معادله(۱۰.۱) مجموعه‌ی محدب بسته در فضای حاصلضرب داخلی است. ساختمان چند وجهی Γ به وسیله‌ی نگاشت $Q_{\mathbb{Y}}$ روی $\mathbb{H}$ که در $\mathbb{X}$ است. در $\mathbb{R}^n$ نشانده می‌شود که این نگاشت در نظریه قاب‌ها، عملگر پیش‌قاب گفته می‌شودکه از [۱] گرفته شده است.

تذکر:عنصر x در فضای ضرب داخلی $\mathbb{X}$ و مجموعه های $\mathbb{Y}$ ، $\mathbb{H}$ واعداد حقیقی $\alpha_1, \ldots, \alpha_n$ و $\beta_1, \ldots, \beta_n$ در این پایان‌نامه تماماً ثابت هستند، بنابراین ما از تکرار غیر ضروری تعاریف این اشیا در عبارات مختلف جلوگیری می‌کنیم. علاوه بر این از دو زیر مجموعه از اعداد طبیعی، مکرراً در این پایان‌نامه استفاده خواهد شد. به همین دلیل آنها را با نام‌های $\mathbb{I}$ و $\mathbb{J}$ به صورت زیر مشخص می‌کنیم:

$$\mathbb{I} = \{1, \ldots, n\} \qquad \mathbb{J} = \{1, \ldots, 2^n\}$$

فصل ۲

ساختار هندسی مسأله تصویر

۱.۲ مقدمه

در این فصل ابتدا به مشخصه سازی بهترین تقریب روی مجموعه محدب $\mathbb{H}$ می‌پردازیم و از آن نتیجه می‌گیریم که باید به تمام عناصر آن دسترسی داشته باشیم برای این ابتدا زیرفضای $\mathbb{S}$ را از فضای ضرب داخلی تعریف می‌کنیم که $\mathbb{H}$ را در آن قرار دهیم. سپس اصل کاهش را بیان می‌کنیم بر طبق این اصل برای محاسبه $P_{\mathbb{H}}(x)$ ابتدا تصویر متری x روی $\mathbb{S}$ را بدست می‌آوریم و سپس تصویر $P_{\mathbb{S}}(x)$ روی $\mathbb{H}$ را محاسبه می‌کنیم. پس برای محاسبه $P_{\mathbb{H}}(x)$ نیاز داریم $P_{\mathbb{S}}(x)$ را محاسبه کنیم که موقعیت‌های مختلفی برای $P_{\mathbb{S}}(x)$ وجود دارد. برای بیان این موقعیت‌ها ما نیاز به معرفی مفاهیمی داریم که آنها را تعریف کرده و از آنها برای شکل دادن $\mathbb{H}$ استفاده می‌کنیم و سپس به تجزیه $\mathbb{H}$ در شرایط متناهیاً تولید شده می‌پردازیم.

۲.۲ مشخصه‌سازی بهترین تقریب روی مجموعه محدب

تعریف ۱.۲.۲. نگاشت $P_{\mathbb{H}}$ از $\mathbb{X}$ به توی زیرمجموعه‌هایی از $\mathbb{H}$ را تصویر متری به توی $\mathbb{H}$ می‌گوییم.

قضیه ۱.۲.۲. نامساوی شوارتز: فرض کنید $\mathbb{X}$ فضای حاصلضرب داخلی باشد برای هر جفت

$x, y \in \mathbb{X}$ داریم:

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

که در آن تساوی برقرار است اگر و فقط اگر x و y مستقل خطی باشند.

حال در زیر قضیه مشخصه‌سازی بهترین تقریب روی مجموعه محدب $\mathbb{H}$ را بیان می‌کنیم

قضیه ۲.۲.۲. فرض کنید $\mathbb{H}$ زیرمجموعه محدبی از فضای حاصلضرب داخلی $\mathbb{X}$ و $\alpha_\circ \in \mathbb{H}$ و $x \in \mathbb{X}$ باشد. آنگاه $\alpha_\circ = P_{\mathbb{H}}(x)$ اگر و تنها اگر

$$\langle x - \alpha_\circ, a - \alpha_\circ\rangle \leq \circ \qquad \forall \alpha \in \mathbb{H}. \tag{۱.۲}$$

برهان. اگر معادله (۱.۲) برقرار باشد و $\alpha \in \mathbb{H}$ آنگاه داریم

$$\begin{aligned}\|x - \alpha_\circ\|^۲ &= \langle x - \alpha_\circ, x - \alpha_\circ\rangle = \langle x - \alpha_\circ - \alpha + \alpha, x - \alpha_\circ\rangle \\ &= \langle x - \alpha_\circ, x - \alpha\rangle + \langle x - \alpha_\circ, \alpha - \alpha_\circ\rangle \\ &\leq \langle x - \alpha_\circ, x - \alpha\rangle \leq \|x - \alpha_\circ\|\|x - \alpha\|\end{aligned}$$

که نامساوی اخیر با توجه به نامساوی شوارتز برقرار است پس $\|x - \alpha_\circ\| \leq \|x - \alpha\|$ و بنابرین $\alpha_\circ = P_{\mathbb{H}}(x)$

برای برهان برگشت قضیه با استفاده از برهان خلف فرض کنید معادله (۱.۲) رد شود. آنگاه

$$\langle x - \alpha_\circ, a - \alpha_\circ\rangle > \circ \qquad \forall \alpha \in \mathbb{H}$$

برای هر $\circ < \lambda < ۱$، عنصر $\alpha_\lambda = \lambda\alpha + (۱ - \lambda)\alpha_\circ = \alpha_\circ + \lambda(\alpha - \alpha_\circ)$ در مجموعه محدب $\mathbb{H}$ است و

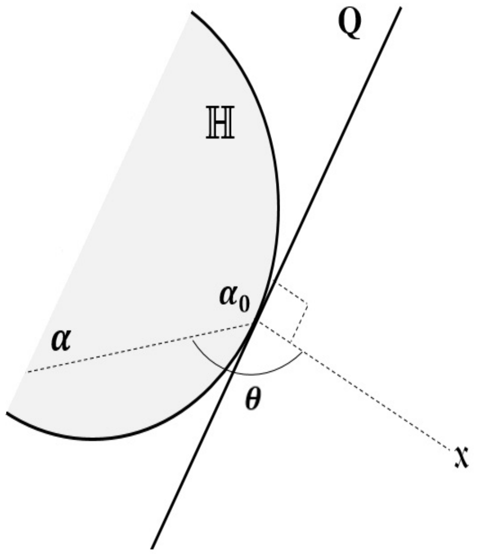
$$\begin{aligned}\|x - \alpha_\lambda\|^۲ &= \langle x - \alpha_\lambda, x - \alpha_\lambda\rangle = \langle x - \alpha_\circ - \lambda(\alpha - \alpha_\circ), x - \alpha_\circ - \lambda(\alpha - \alpha_\circ)\rangle \\ &= \|x - \alpha_\circ\|^۲ - ۲\lambda\langle x - \alpha_\circ, \alpha - \alpha_\circ\rangle + \lambda^۲\|\alpha - \alpha_\circ\|^۲ \\ &= \|x - \alpha_\circ\|^۲ - \lambda[۲\langle x - \alpha_\circ, \alpha - \alpha_\circ\rangle - \lambda\|\alpha - \alpha_\circ\|^۲].\end{aligned}$$

برای $\lambda > \circ$ به اندازه کافی کوچک ،عبارت داخل براکت مثبت است و در نتیجه

$$\|x - \alpha_\lambda\|^۲ < \|x - \alpha_\circ\|^۲ \qquad \Rightarrow \alpha_\circ \neq P_{\mathbb{H}}(x).$$

□

با توجه به شکل ۱.۲ دو تعبیر هندسی از قضیه ۲.۲.۲ وجود دارد اول این که زاویه θ بین دو بردار $x - \alpha_\circ$ و $\alpha - \alpha_\circ$ برای هر $\alpha \in \mathbb{H}$ کمتر از ۹۰ درجه است و تفسیر دوم این‌که مجموعه محدب $\mathbb{H}$ در یک طرف ابرصفحه Q قرار دارد که Q متعامد با $x - \alpha_\circ$ و از میان $\alpha_\circ$ می‌گذرد.



شکل ۱.۲: مشخصه‌سازی بهترین تقریب از روی مجموعه محدب

بنابراین این قضیه نشان می دهد که برای محاسبه بهترین تقریب x روی مجموعه $\mathbb{H}$ در هر صورت باید امکان رسیدگی به همه عناصر $\mathbb{H}$ را داشته باشیم. برای این منظور ممکن است نیاز داشته باشیم دستگاه مختصات مرجعی در $\mathbb{H}$ ایجاد کنیم، یا سریع‌تر $\mathbb{H}$ را در قاب مناسب جا بدهیم. بنابراین به نظر می‌رسد درست ترین راه برای رسیدن به این مقصود آن است زیرفضای تولیدشده‌ی خطی توسط $\mathbb{H}$ را در نظر بگیریم.

تعریف ۲.۲.۲. زیرفضای تولید شده توسط $\mathbb{H}$ را با $\mathbb{S}$ نشان می‌دهیم که $\mathbb{S}$ تمام ترکیبات خطی عناصر $\mathbb{Y}$ است و به صورت زیر آمده است:

$$\mathbb{S} := span(\mathbb{H}) = span(\mathbb{Y}) = \{y \in \mathbb{X} \mid y = \sum_{i=۱}^{n} \tau_i y_i, \quad y_i \in \mathbb{Y}, \tau \in \mathbb{R}, n \in \mathbb{R}\} \qquad (۲.۲)$$

که $\mathbb{Y}$ در معادله (۸.۱) آمده است. خواهیم دید که $\mathbb{S}$ واقعاً یک محیط مناسب برای یافتن بهترین تقریب x نسبت به مجموعه $\mathbb{H}$ است.

تعریف ۳.۲.۲. زیرمجموعه ناتهی $\mathbb{H}$ از $\mathbb{X}$ را فشرده می‌گوییم اگر هر دنباله کوشی در $\mathbb{H}$ همگرا به نقطه‌ای در $\mathbb{H}$ باشد و تقریبا فشرده گوییم اگر برای هر $x \in \mathbb{X}$ و هر دنباله $\{y_n\}$ در $\mathbb{H}$ که $\|x - y_n\| \to d(x, \mathbb{H})$ ، یک زیردنباله همگرا به یک نقطه در $\mathbb{H}$ داشته باشد. اگر دنباله $\{y_n\}$ در $\mathbb{H}$ در $\|x - y_n\| \to d(x, \mathbb{H})$ صدق کند می‌گوییم $\{y_n\}$ دنباله مینیمم‌کننده برای $\mathbb{H}$ است پس $\mathbb{H}$ تقریبا فشرده است اگر هر دنباله مینیمم‌کننده زیر دنباله‌ای همگرا به نقطه‌ای از $\mathbb{H}$ داشته باشد و این اظهار عقیده مفید است که هر دنباله مینیمم‌کننده کراندار است این نتیجه منطقی با عبارت زیرمعادل است:

$$\|y_n\| \leq \|y_n - x\| + \|x\| \longrightarrow d(x, \mathbb{H}) + \|x\|.$$

قضیه ۳.۲.۲. هر مجموعه تقریبا فشرده پروکسیمینال است.

برهان. فرض کنید $\mathbb{H}$ تقریبا فشرده و $x \in \mathbb{X}$ باشد دنباله مینیمم‌کننده $\{y_n\}$ در $\mathbb{H}$ را برای x انتخاب می‌کنیم. فرض کنید $\{y_{n_i}\}$ زیر دنباله‌ای باشد که همگرا به $y \in \mathbb{H}$ است آنگاه داریم:

$$\|x-y\| = \lim\|x-y_{n_i}\| = d(x,\mathbb{H})$$

پس y یک بهترین تقریب x روی $\mathbb{H}$ است پس $\mathbb{H}$ پروکسیمینال است. $\square$

قضیه ۴.۲.۲. فرض کنید $\mathbb{S}$ یک زیر فضای با بعد متناهی از فضای ضرب داخلی $\mathbb{X}$ باشد آنگاه

(۱) هر دنباله کراندار در $\mathbb{S}$ یک زیر دنباله دارد که به یک نقطه در $\mathbb{S}$ همگرا است.

(۲) $\mathbb{S}$ بسته است.

برهان. فرض کنید $\mathbb{S} = span\{x_1,\dots,x_n\}$ زیرفضای n بعدی باشد. با روش استقرا به اثبات (۱) و (۲) به طور همزمان می‌پردازیم. برای $n=1$ فرض می‌کنیم $\{y_k\}$ دنباله‌ای کراندار در $\mathbb{S}$ باشد. پس $y_k = \alpha_k x_1$ بنابراین $\{\alpha_k\}$ باید دنباله‌ای کراندار در $\mathbb{R}$ باشد پس زیردنباله‌ای $\{\alpha_{k_j}\}$ و $\alpha_0 \in \mathbb{R}$ وجود دارد به طوری که $\{\alpha_{k_j}\} \longrightarrow \alpha_0$ پس در نتیجه $y_0 = \alpha_0 x_1 \in \mathbb{S}$. پس داریم:

$$\|y_{k_j} - y_0\| = |\alpha_{k_j} - \alpha_0|\|x_1\| \longrightarrow 0$$

البته اگر $\{y_k\}$ واقعا دنباله همگرا باشد پس $\{y_k\}$ کراندار است و در بحث بالا نشان داده شد که حد آنها در $\mathbb{S}$ است.

حال فرض می‌کنیم وقتی $\mathbb{S}$، nبعدی است، (۱) و (۲) درست باشد. و حال درستی (۱) و (۲) را برای $n+1$ ثابت می‌کنیم فرض می‌کنیم $\mathbb{S} = span\{x_1,\dots,x_n,x_{n+1}\}$ و $n+1$ بعدی باشد.

مجموعه $\mathbb{S}_n = span\{x_1,\dots,x_n\}$ که $\mathbb{S}_n$، n بعدی است. بنابراین $\mathbb{S}_n$ با توجه به فرض استقرا بسته است و $x_{n+1} \notin \mathbb{S}_n$. که از آن نتیجه می‌شود $d(x_{n+1},\mathbb{S}_n) > 0$. فرض کنید $\{y_k\}$ دنباله‌ای کراندار در $\mathbb{S}$ باشد. آنگاه $y_k = z_k + \beta_k x_{n+1}$ که $z_k \in \mathbb{S}_n$ است. با توجه به اینکه $0 \in \mathbb{S}_n$ است پس داریم:

$$\|y_k\| \geq d(y_k,\mathbb{S}_n) = d(\beta_k x_{n+1},\mathbb{S}_n) = |\beta_k| d(x_{n+1},\mathbb{S}_n),$$

که از آن عبارت زیر نتیجه می‌شود:

$$|\beta_k| \leq \frac{\|y_k\|}{d(x_{n+1},\mathbb{S}_n)}$$

پس β_k دنباله‌ای کراندار است و

$$\|z_k\| = \|y_k + \beta_k x_{n+1}\| \leq \|y_k\| + |\beta_k|\|x_{n+1}\|$$

در نتیجه $\{z_k\}$ دنباله‌ای کراندار است. با استفاده از فرض استقرا زیر دنباله‌ی $\{z_{k_j}\}$ و $z_\circ \in \mathbb{S}_n$ وجود دارد به طوری که $z_{k_j} \longrightarrow z_\circ$. چون $\{\beta_{k_j}\}$ کراندار است پس زیر دنباله بیشتری لازم است فرض می‌کنیم $\beta_{k_j} \longrightarrow \beta_\circ \in \mathbb{R}$. از این رو $y_\circ = z_\circ + \beta_\circ x_{n+1}$ و نتیجه می‌دهد که $y_\circ \in \mathbb{S}$ و

$$\|y_{k_j} - y_\circ\| \leq \|z_{k_j} - z_\circ\| + |\beta_{k_j} - \beta_\circ|\|x_{n+1}\| \longrightarrow \circ$$

دوباره اگر $\{y_k\}$ واقعاً دنباله همگرا باشد این ثابت می‌کند که حد آنها در $\mathbb{S}$ است پس (۱) و (۲) برای $\mathbb{S}$ صدق می‌کند. □

قضیه ۵.۲.۲. فرض کنید $\mathbb{X}$ فضای حاصلضرب داخلی باشد آنگاه:

(۱) هر زیرمجموعه بسته از یک زیرفضای متناهی‌البعد از $\mathbb{X}$ پروکسیمینال است.

(۲) هر زیرمجموعه محدب بسته از یک زیرفضای متناهی‌البعد از $\mathbb{X}$ چیبشف است.

(۳) هر زیرمجموعه متناهی‌البعد چیبشف است.

برهان. (۱) فرض کنید $\mathbb{H}$ زیر مجموعه بسته از زیرفضای متناهی‌البعد $\mathbb{S}$ باشد با توجه به قضیه ۳.۲.۲ کافیست نشان دهیم $\mathbb{H}$ تقریباً فشرده است فرض کنید $x \in \mathbb{X}$ و فرض کنید $\{y_n\}$ دنباله مینمم‌کننده در $\mathbb{H}$ برای x باشد. پس $\{y_n\}$ کراندار و با استفاده از قضیه قبل دارای زیردنباله‌ای است که به نقطه $y_\circ \in \mathbb{S}$ همگرا است. چون $\mathbb{H}$ بسته و $y_\circ \in \mathbb{H}$ پس $\mathbb{H}$ تقریبا فشرده است.

(۲) با توجه به قسمت (۱) و اینکه هر مجموعه محدب پروکسیمینال چیبشف است برقرار است.

(۳) باتوجه به قسمت (۲) و قضیه ۴.۲.۲ قسمت دوم هر زیر مجموعه محدب بسته از یک زیرفضای متناهی‌البعد چیبشف است.

□

نتیجه ۱.۲.۲. هر زیرمجموعه محدب بسته از $\mathbb{S}$ چیبشف است.همچنین $\mathbb{S}$ خودش چیبشف است.

بنابراین بویژه $\mathbb{H}$ به عنوان یک زیر مجموعه محدب بسته از $\mathbb{S}$، چیببشف است. به این معنی که $P_{\mathbb{H}}(x)$ یکتاست. این مسیر بهترین تقریب $x \in \mathbb{X}$ روی $\mathbb{H}$ توسط $\mathbb{S}$ است. اما سوال این است که آیا در تمام روند به‌دست آوردن $P_{\mathbb{H}}(x)$، $\mathbb{S}$ تمام این مسیر را طی خواهد کرد؟
نتایج زیر نشان می دهد که جواب این سؤال مثبت است.

۳.۲ اصل کاهش

قضیه ۱.۳.۲. فرض کنید $\mathbb{H}$ زیرمجموعه محدب از فضای حاصلضرب داخلی $\mathbb{X}$ باشد و فرض کنید $\mathbb{S}$ یک زیرفضای چیبشف از $\mathbb{X}$ شامل $\mathbb{H}$ باشد آنگاه

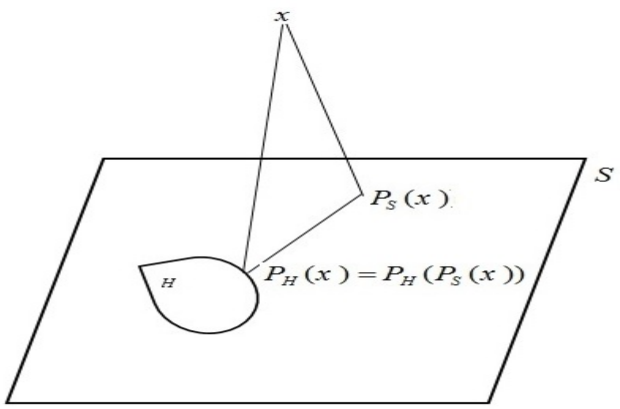
$$P_{\mathbb{H}}(x) = P_{\mathbb{H}}(P_{\mathbb{S}}(x)).$$

برهان. چون $\mathbb{H} \subset \mathbb{S}$ است.برای هر $y \in \mathbb{H}$ داریم که $y \in \mathbb{S}$ و با توجه به قضیه ۱.۳.۱ $x - P_{\mathbb{S}}(x) \in \mathbb{S}^T$ بنابراین داریم:

$$\|x-y\|^۲ = \|x - P_{\mathbb{S}}(x)\|^۲ + \|P_{\mathbb{S}}(x) - y\|^۲. \tag{۳.۲}$$

چون اولین شرط در سمت راست از معادله (۳.۲) مستقل از y است که از آن نتیجه می‌شود $y \in \mathbb{H}$ است $\|x-y\|^۲$ به حداقل می‌رسد اگر و فقط اگر $\|P_{\mathbb{S}}(x) - y\|^۲$ به حداقل برسد.این ثابت می‌کند که $P_{\mathbb{H}}(x)$ وجود دارد اگر و فقط اگر $P_{\mathbb{H}}(P_{\mathbb{S}}(x))$ وجود داشته باشد و $P_{\mathbb{H}}(x) = P_{\mathbb{H}}(P_{\mathbb{S}}(x))$ از این رو
$P_{\mathbb{H}}(x) = \emptyset$ است اگر و فقط اگر $P_{\mathbb{H}}(P_{\mathbb{S}}(x)) = \emptyset$ باشد. □

تعبیر هندسی اصل کاهش با توجه به شکل ۲.۲ به صورت زیر است:
برای پیدا کردن بهترین تقریب x از روی $\mathbb{H}$ ، ابتدا تصویر x روی $\mathbb{S}$ و سپس تصویر $P_{\mathbb{S}}(x)$ روی $\mathbb{H}$ را بدست می‌آوریم و در نتیجه $P_{\mathbb{H}}(x)$ بدست می‌آید.



شکل ۲.۲: اصل کاهش

بنابراین برای محاسبه بهترین تقریب روی $\mathbb{H}$ ما ابتدا به محاسبه تصویر متری روی $\mathbb{S}$ که نباید یک مسأله باشد می‌پردازیم. چون $\mathbb{S}$ یک زیر فضا و بر طبق نتیجه ۱.۲.۲ چبیشف است و $P_{\mathbb{S}}(x)$ منحصر به فرد است پس تصویر $P_{\mathbb{S}}(x)$ روی $\mathbb{H}$ را بدست می‌آوریم.

اکنون ما ظاهراً یک مختصات مرجع مناسب که $\mathbb{H}$ در آن قرار گرفته است را پیدا کردیم. اما یک مشکل وجود دارد برای مثال به همه عناصر $\mathbb{H}$ برای بررسی نابرابری در مشخصه‌سازی بهترین تقریب در قضیه ۲.۲.۲ دسترسی نداریم. از آنجایی که یک مرحله جدید برای محاسبه بهترین تقریب x روی $\mathbb{H}$ اضافه می‌شود یعنی ابتدا باید $P_{\mathbb{S}}(x)$ را بدست بیاوریم که $P_{\mathbb{S}}(x)$ موقعیت‌های متفاوتی نسبت به $\mathbb{H}$ دارد پس باید موقعیت‌های مختلفی که $P_{\mathbb{S}}(x)$ در آن قرار دارد را بررسی کنیم تا بتوانیم به محاسبه $P_{\mathbb{H}}(x)$ بر طبق اصل کاهش قضیه ۱.۳.۲ بپردازیم.

برای شروع به تعداد متناهی عنصر y_i از مجموعه $\mathbb{Y}$ که در شرط متناهیاً تولید شده صدق کند برای توصیف مجموعه محدب $\mathbb{H}$ نیاز داریم. باید نشان دهیم که یک توصیف هندسی از $\mathbb{H}$ در شرایط متناهیاً تولید شده ذکر شده در بالا را داریم. اگرچه ابتدا نیاز به معرفی تعدادی از اشیاء قبل از شروع این نتیجه داریم. تعاریف زیر از [۲] داده شده است که در ارتباط با شکل دادن $\mathbb{H}$ مهم است.

۴.۲ تعاریفی برای شکل دادن $\mathbb{H}$

تعریف ۱.۴.۲. زیرمجموعه $\mathbb{E}$ از مجموعه محدب $\mathbb{H}$ را زیرمجموعه اکسترمال $\mathbb{H}$ می‌گوییم اگر از $f, g \in \mathbb{H}$ و $\lambda f + (1-\lambda)g \in \mathbb{E}$ که $0 < \lambda < 1$ نتیجه شود $f, g \in \mathbb{E}$.

تعریف ۲.۴.۲. $h \in \mathbb{H}$ را یک نقطه اکسترمم $\mathbb{H}$ می‌گوییم اگر $f, g \in \mathbb{H}$ و $h = \lambda f + (1-\lambda)g$ به ازای هر $0 < \lambda < 1$ نتیجه دهد $f = g = h$.

برای مثال نقاط اکسترمم یک گوی واحد نقاط مرزی گوی می‌باشد و نقاط اکسترمم یک مکعب سه بعدی رئوس مکعب هستند.

با مراجعه به این تعریف مجموعه $\mathbb{E}$ از نقاط اکسترمم $\mathbb{H}$ را به صورت زیر نشان می‌دهند:

$$\mathbb{E} = \{r^1, \dots, r^{2^n} \mid r^j = \sum_{i=1}^{n} \psi_i^j y_i; \quad y_i \in \mathbb{Y}, \psi_i^j = (\alpha_i \vee \beta_i), j \in \mathbb{J}, \psi_i^j \neq \psi_i^k \quad (j \neq k)\} \tag{۴.۲}$$

که α_i و β_i و y_iها همان هستند که در معادله‌های ۸.۱ و ۹.۱ داده شده‌اند. جفت نقاط اکسترمم متقابل $\mathbb{H}$ و همچنین مفهوم نقطه اکسترمم مجاور یک نقطه اکسترمم از اهمیت ویژه‌ای برخوردار هستند که آنها در زیر تعریف می‌کنیم:

تعریف ۳.۴.۲. $\{r^j, r^k\}$ که $(r^j, r^k \in \mathbb{E}, j \neq k)$ را یک جفت نقاط اکسترمم متقابل از $\mathbb{H}$ می‌گویند اگر برای هر i، $\psi_i^j \neq \psi_i^k$ که $(j \neq k)$. همچنین بدون از دست دادن کلیت فرض می کنیم که مجموعه $\mathbb{E}$ مرتب باشد به طوری که $\{\{r^{2j-1}, r^{2j}\} \mid j = 1, \dots, 2^{n-1}\}$ جفت‌های اکسترمم متقابل از $\mathbb{E}$ هستند. برای مثال $\psi_i^1 = \alpha_i, \psi_i^2 = \beta_i, \dots$ مرتب ثابت هستند.

$r^k \in \mathbb{E}$ را یک نقطه اکسترمم مجاور r^j می‌گوییم اگر برای یک $i \in \mathbb{I}$ داشته باشیم $\psi_i^j \neq \psi_i^k$ و برای بقیه $i \in \mathbb{I}$ها داشته باشیم $\psi_i^j = \psi_i^k$.

در زیر اهمیت شیء هندسی می‌تواند وابسته به یک نقطه انتهایی $r^j \in \mathbb{E}$ باشد.

تعریف ۴.۴.۲. فرض می‌کنیم مخروط $\mathbb{Z}^j$ به صورت زیر باشد:

$$\mathbb{Z}^j = \{z_1^j, \dots, z_n^j \mid z_i^j = \begin{cases} y_i & \text{اگر } \alpha_i = \psi_i^j \\ -y_i & \text{اگر } \beta_i = \psi_i^j \end{cases} \} \tag{۵.۲}$$

پوسته مخروط $\mathbb{Z}^j$ را که α_i و β_i و y_iها همان هستند که در معادله‌های ۸.۱ و ۹.۱ داده شده‌اند، به صورت زیر تعریف می‌کنیم:

$$con(\mathbb{Z}^j) = \{y \in \mathbb{X} \mid y = \sum_{i=1}^{n} \tau_i z_i^j, z_i^j \in \mathbb{Z}^j, \tau \in \mathbb{R}^+ \cup \{0\} \quad \} \tag{۶.۲}$$

حال نقطه اکسترمم r^j را نسبت به مخروط $\mathbb{Z}^j$ بدست می‌آوریم. برای هر $k \in \mathbb{J}$ و $i \in \mathbb{I}$ داریم:

$$\sigma_i^k = sgn(\alpha_i - \psi_i^k) \tag{۷.۲}$$

و برای هر $\rho \in \mathbb{R}$، $sgn(\rho)$ از رابطه زیر بدست می‌آید:

$$sgn(\rho) = \begin{cases} 1 & \text{اگر } \rho \geq 0 \\ -1 & \text{اگر } \rho < 0 \end{cases} \tag{۸.۲}$$

پس با توجه به معادله ۴.۲ داریم:

$$r^j = \sum_{i=1}^{n} \psi_i^j y_i \quad \forall y_i \in \mathbb{Y}, \psi_i^j = (\alpha_i \vee \beta_i)$$

حال دو حالت $\psi_i^j = \alpha_i$ و $\psi_i^j = \beta_i$ را بررسی می‌کنیم:

$$\psi_i^j = \alpha_i \Rightarrow \sigma_i^j = sgn(\alpha_i - \alpha_i) = sgn(0) = 1 \Rightarrow r^j = \psi_i^j y_i$$

و

$$\psi_i^j = \beta_i \Rightarrow \sigma_i^j = sgn(\alpha_i - \beta_i) < 0 \Rightarrow \sigma_i^j = -1 \Rightarrow r^j = (-1)\psi_i^j(-y_i) \Rightarrow r^j = \psi_i^j y_i$$

پس داریم:

$$r^j = \sum_{i=1}^{n} \sigma_i^j \psi_i^j z_i^j = \sum_{i=1}^{n} \sigma_i^k \psi_i^j z_i^k \tag{۹.۲}$$

حال پوسته مخروط انتقال‌یافته را تعریف می‌کنیم:

تعریف ۵.۴.۲. پوسته مخروط انتقال‌یافته متناظر با نقطه اکسترمم $r^j \in \mathbb{E}$ به صورت زیر تعریف می‌شود:

$$\mathbb{K}^j := r^j + con(\mathbb{Z}^j) \tag{۱۰.۲}$$

نمایشی که در تعریف ۵.۴.۲ گفته شد را به جمع دو مجموعه $\{r^j\}$ و $con(\mathbb{Z}^j)$ تعبیر می کنیم:

$$\mathbb{K}^j := \{r^j\} + con(\mathbb{Z}^j)$$

که $\{r^j\}$ها نقاط اکسترمم است و $con(\mathbb{Z}^j)$ پوسته یک مخروط محدب متناهیا تولید شده است زیرا آن پوسته مخروط از تعداد متناهی بردار تشکیل شده است.

قضیه ۱.۴.۲. فرض کنید K یک زیر مجموعه ناتهی از فضای ضرب داخلی $\mathbb{X}$ باشد و $x, y \in \mathbb{X}$ باشد.آنگاه

(۱) $d(x+y, K+y) = d(x, K)$.

(۲) $P_{K+y}(x+y) = P_K(x) + y$.

(۳) K چبیشف است اگر و تنها اگر $K+y$ چبشف باشد.

برهان. (۱)

$$d(x+y, K+y) = \inf_{z\in K}\|(x+y)-(z+y)\| = \inf_{z\in K}\|(x-z)\| = d(x, K)$$

(۲)

$$y_\circ \in P_{K+y}(x+y) \Leftrightarrow y_\circ \in K+y, \|x+y-y_\circ\| = d(x+y, K+y) = d(x, K) \Leftrightarrow y_\circ - y \in K$$

$$, \|x-(y-y_\circ)\| = d(x, K) \Leftrightarrow y_\circ - y \in P_K(x) \Leftrightarrow y_\circ \in P_K(x) + y$$

(۳) با توجه به قسمت (۲) اثبات می‌شود.

□

اهمیت پوسته مخروط انتقال یافته از گزاره زیر نتیجه می‌شود:

گزاره ۱.۴.۲. $y + con(\mathbb{Z}^j)$ برای هر $y \in \mathbb{S}$ چبیشف است.

برهان. اگر $con(\mathbb{Z}^j)$ چبیشف باشد پس با توجه به قضیه ۱.۴.۲ قسمت (۳) $y+con(\mathbb{Z}^j)$ چبیشف است .پس کافیست نشان دهیم $con(\mathbb{Z}^j)$ چبیشف است. بنابراین با توجه به قسمت دوم قضیه ۵.۲.۲ نشان می‌دهیم که $con(\mathbb{Z}^j)$ بسته در $\mathbb{S}$ است. با برهان خلف ثابت می‌کنیم. فرض می‌کنیم $con(\mathbb{Z}^j)$ بسته نباشد پس دنباله‌ی $\{x_n\}$ در $con(\mathbb{Z}^j)$ وجود دارد به طوری که $x_n \longrightarrow x$ و $x \notin con(\mathbb{Z}^j)$ پس $d(x, con(\mathbb{Z}^j)) \leq \|x - x_n\| \longrightarrow \circ$ بنابراین $d(x, con(\mathbb{Z}^j)) = \circ$ اما برای هر $y \in con(\mathbb{Z}^j)$ چون $x \notin con(\mathbb{Z}^j)$ پس $\|x - y\| > \circ$ و این با $d(x, con(\mathbb{Z}^j)) = \circ$ در تناقض است پس $con(\mathbb{Z}^j)$ بسته است و حکم ثابت شد. □

۱.۴.۲ تجزیه $\mathbb{H}$ در شرایط متناهیاً تولید شده و مخروط‌های محدب انتقال‌یافته

گزاره ۲.۴.۲. فرض کنید $\{r^{۲j-۱}, r^{۲j}\}$ یک جفت نقاط انتهایی متقابل از $\mathbb{H}$ باشد که $r^{۲j-۱}, r^{۲j} \in \mathbb{E}$. آنگاه

$$r^{۲j} \in r^{۲j-۱} + con(\mathbb{Z}^{۲j-۱})$$

$$r^{۲j-۱} \in r^{۲j} + con(\mathbb{Z}^{۲j})$$

$$\mathbb{H} = (r^{۲j-۱} + con(\mathbb{Z}^{۲j-۱})) \cap (r^{۲j} + con(\mathbb{Z}^{۲j})).$$

برهان. با توجه به معادله‌های (۶.۲) و (۹.۲) داریم:

$$r^j = \sum_{i=۱}^{n} \sigma_i^j \psi_i^j z_i^j \qquad , con(\mathbb{Z}^j) = \sum_{i=۱}^{n} \tau_i z_i^j$$

چون $\alpha_i < \beta_i$ است فرض می‌کنیم $\alpha_i \in \beta_i - \tau_i$ باشد که با توجه به معادله (۶.۲)، $\tau \in \mathbb{R} \cup \{۰\}$ چون $\{r^{۲j-۱}, r^{۲j}\}$ جفت نقاط اکسترمم متقابل هستند پس با توجه به تعریف ۳.۴.۲، $\psi_i^{۲j} \neq \psi_i^{۲j-۱}$ پس فرض می‌کنیم $\psi_i^{۲j} = \alpha_i$ و $\psi_i^{۲j-۱} = \beta_i$ پس با توجه به معادله (۵.۲) داریم

$$r^{۲j} = \sum_{i=۱}^{n} \sigma_i^{۲j} \psi_i^{۲j} z_i^{۲j} = \sum_{i=۱}^{n} \alpha_i y_i$$

و

$$r^{۲j-۱} = \sum_{i=۱}^{n} \sigma_i^{۲j-۱} \psi_i^{۲j-۱} z_i^{۲j-۱} = \sum_{i=۱}^{n} \beta_i y_i$$

بنابراین با توجه به معادله (۶.۲) داریم:

$$r^{۲j-۱} + con(\mathbb{Z}^{۲j-۱}) == \sum_{i=۱}^{n} \beta_i y_i - \sum_{i=۱}^{n} \tau_i y_i = \sum_{i=۱}^{n} (\beta_i - \tau_i) y_i$$

$$r^{۲j} = \sum_{i=۱}^{n} \alpha_i y_i$$

چون $\alpha_i \in \beta_i - \tau_i$ بنابراین داریم:

$$\Rightarrow r^{۲j} \in r^{۲j-۱} + con(\mathbb{Z}^{۲j-۱})$$

برای رابطه دوم هم به همین ترتیب عمل می‌کنیم. حال آخرین عبارت را اثبات می‌کنیم با توجه به معادله‌های (۹.۱) و (۱۰.۱) داریم:

$$\mathbb{H} = \{y \in \mathbb{X} \mid y = \sum_{i=۱}^{n} \gamma_i y_i, \alpha_i \leq \gamma_i \leq \beta_i\} \qquad (۱۱.۲)$$

فرض کنید $\psi_i^j = \alpha_i$ با توجه به این‌که $\gamma_i \geq \alpha_i$ پس $\sigma_i^{۲j-۱}\gamma_i \geq \sigma_i^{۲j-۱}\psi_i^{۲j-۱}$ و چون $\psi_i^j = \alpha_i$ پس با توجه به معادله (۵.۲)، $z_i^j = y_i$ پس $z_i^{۲j-۱} = y_i$ و $z_i^{۲j} = y_i$ است پس داریم:

$$\mathbb{H} = \{y \in \mathbb{X} \mid y = \sum_{i=۱}^{n}(\sigma_i^{۲j-۱}\gamma_i)z_i^{۲j-۱}, \sigma_i^{۲j-۱}\gamma_i \geq \sigma_i^{۲j-۱}\psi_i^{۲j-۱}\} \cap$$

$$\{y \in \mathbb{X} \mid y = \sum_{i=۱}^{n}(\sigma_i^{۲j}\gamma_i)z_i^{۲j}, \sigma_i^{۲j}\gamma_i \geq \sigma_i^{۲j}\psi_i^{۲j}\}$$

که

$$\{y \in \mathbb{X} \mid y = \sum_{i=۱}^{n}(\sigma_i^{۲j-۱}\gamma_i)z_i^{۲j-۱}, \sigma_i^{۲j-۱}\gamma_i \geq \sigma_i^{۲j-۱}\psi_i^{۲j-۱}\} =$$

$$\{y \in \mathbb{X} \mid y = r^{۲j-۱} - r^{۲j-۱} + \sum_{i=۱}^{n}(\sigma_i^{۲j-۱}\gamma_i)z_i^{۲j-۱}, \sigma_i^{۲j-۱}\gamma_i \geq \sigma_i^{۲j-۱}\psi_i^{۲j-۱}\} =$$

$$\{y \in \mathbb{X} \mid y = r^{۲j-۱} - \sum_{i=۱}^{n}\sigma_i^{۲j-۱}\psi_i^{۲j-۱}z_i^{۲j-۱} + \sum_{i=۱}^{n}(\sigma_i^{۲j-۱}\gamma_i)z_i^{۲j-۱}, \sigma_i^{۲j-۱}\gamma_i \geq \sigma_i^{۲j-۱}\psi_i^{۲j-۱}\} =$$

$$\{y \in \mathbb{X} \mid y = r^{۲j-۱} + \sum_{i=۱}^{n}(\sigma_i^{۲j-۱}\gamma_i - \sigma_i^{۲j-۱}\psi_i^{۲j-۱})z_i^{۲j-۱},$$

$$(\sigma_i^{۲j-۱}\gamma_i - \sigma_i^{۲j-۱}\psi_i^{۲j-۱}) \geq ۰\} =$$

$$r^{۲j-۱} + \{y \in \mathbb{X} \mid y = \sum_{i=۱}^{n}\tau_i z_i^{۲j-۱}, \tau_i \in \mathbb{R}^+ \cup \{۰\}\}$$

$$= r^{۲j-۱} + con(\mathbb{Z}^{۲j-۱})$$

و

$$\{y \in \mathbb{X} \mid y = \sum_{i=۱}^{n}(\sigma_i^{۲j}\gamma_i)z_i^{۲j}, \sigma_i^{۲j}\gamma_i \geq \sigma_i^{۲j}\psi_i^{۲j}\} =$$

$$\{y \in \mathbb{X} \mid y = r^{۲j} - r^{۲j} + \sum_{i=۱}^{n}(\sigma_i^{۲j}\gamma_i)z_i^{۲j}, \sigma_i^{۲j}\gamma_i \geq \sigma_i^{۲j}\psi_i^{۲j}\} =$$

$$\{y \in \mathbb{X} \mid y = r^{۲j} - \sum_{i=۱}^{n}\sigma_i^{۲j}\psi_i^{۲j}z_i^{۲j} + \sum_{i=۱}^{n}(\sigma_i^{۲j}\gamma_i)z_i^{۲j}, \sigma_i^{۲j}\gamma_i \geq \sigma_i^{۲j}\psi_i^{۲j}\} =$$

$$\{y \in \mathbb{X} \mid y = r^{۲j} + \sum_{i=۱}^{n}(\sigma_i^{۲j}\gamma_i - \sigma_i^{۲j}\psi_i^{۲j})z_i^{۲j-۱},$$

$$(\sigma_i^{۲j-۱}\gamma_i - \sigma_i^{۲j}\psi_i^{۲j}) \geq ۰\} =$$

$$r^{۲j} + \{y \in \mathbb{X} \mid y = \sum_{i=۱}^{n}\tau_i z_i^{۲j}, \tau_i \in \mathbb{R}^+ \cup \{۰\}\}$$

$$= r^{۲j} + con(\mathbb{Z}^{۲j})$$

پس داریم:

$$\mathbb{H} = (r^{۲j-۱} + con(\mathbb{Z}^{۲j-۱})) \cap (r^{۲j} + con(\mathbb{Z}^{۲j})).$$

□

قضیه ۲.۴.۲. فرض کنید $\mathbb{H}$ مجموعه چیبشف و محدب باشد. دراین‌صورت $P_{\mathbb{H}}$ خودتوان است یعنی برای هر $x \in \mathbb{X}$ داریم: $P_{\mathbb{H}}(P_{\mathbb{H}}(x)) = P_{\mathbb{H}}(x)$. یعنی بهترین تقریبِ بهترین تقریب هر عضو در $\mathbb{H}$ با بهترین تقریب آن عضو در $\mathbb{H}$ برابر است.

ما اکنون مسأله را از مکان‌های مختلف در $P_{\mathbb{S}}(x)$ نسبت به $\mathbb{H}$ در نظر می‌گیریم. واضح است اگر $P_{\mathbb{S}}(x)$ در $\mathbb{H}$ اتفاق افتد جای هیچ نگرانی نیست زیرا در حقیقت تصویر متری عملگر در این مورد خودتوان است پس $P_{\mathbb{H}}(x) = P_{\mathbb{S}}(x)$. اگر $P_{\mathbb{S}}(x) \notin \mathbb{H}$ ما دو موقعیت مختلف را تشخیص می‌دهیم یعنی $P_{\mathbb{S}}(x)$ در محدوده نقطه انتهایی $r^j \in \mathbb{E}$ است که می‌توانیم این ایده دقیق را در راه‌های زیر بسازیم.

۵.۲ تعاریفی برای موقعیت $P_{\mathbb{S}}(x)$

تعریف ۱.۵.۲. مخروط قطبی انتقال یافته حاصل از انتقال پوسته مخروط $\mathbb{K}^j$ که در تعریف ۵.۴.۲ آمده است زیرمجموعه محدب از $\mathbb{S}$ است که به صورت تعریف می‌شود:

$$(\mathbb{K}^j)_\circ = \{s \in \mathbb{S} \mid \langle (s - r^j), z^j \rangle \leq \circ \quad \forall z^j \in con(\mathbb{Z}^j)\}. \tag{۱۲.۲}$$

همچنین $P_{\mathbb{S}}(x)$ را در محدوده نقطه اکسترمم $r^j \in \mathbb{E}$ گوییم اگر $P_{\mathbb{S}}(x) \in (\mathbb{K}^j)_\circ$ باشد.

اثبات این اصطلاحات برای محاسبه $P_{\mathbb{H}}(x)$ در این وضعیت را در فصل بعدی که نتیجه می‌شود ارائه می‌دهیم. موقعیت دیگری که $P_{\mathbb{S}}(x)$ نسبت به $\mathbb{H}$ می‌تواند داشته باشد موقعیت $P_{\mathbb{S}}(x) \in \mathbb{D}$ است که مجموعه $\mathbb{D}$ به صورت زیر نمایش داده می‌شود:

$$\mathbb{D} = \mathbb{S} \setminus ((\bigcup_{j=۱}^{۲^n} (\mathbb{K}^j)_\circ) \cup \mathbb{H}). \tag{۱۳.۲}$$

مجموعه‌ی $\mathbb{D}$ با مؤلفه‌های محدب ناهمبند است که در [۶] دیده می‌شود. بنابراین $P_{\mathbb{S}}(x) \in \mathbb{D}$ به این معنی است که $P_{\mathbb{S}}(x)$ در یکی از مؤلفه‌های $\mathbb{D}$ است. این در ارتباط با تصویر متری روی $\mathbb{H}$ اهمیت دارد. در حقیقت مجموعه مرزی $\mathbb{H}$ نسبت به یکی از آن مؤلفه‌ها شرایط $\mathbb{K}^j$ که $(j \in \mathbb{J})$ را جدا می‌کند که $P_{\mathbb{S}}(x)$ در قسمت جلویی مجموعه مرزی قرار دارد وقتی که آن در یکی از مؤلفه‌های $\mathbb{D}$ باشد. برای توصیف مجموعه مرزی $\mathbb{H}$ و این موقعیت $P_{\mathbb{S}}(x)$ ما به نقطه اکسترمم مجاور در تعریف ۳.۲.۲ برمی‌گردیم.

تعریف ۲.۵.۲. فرض کنید $j \in \mathbb{J}$ باشد دراین‌صورت گردایه‌ای از n نقطه اکسترمم مجاور با $r^j \in \mathbb{E}$ نمایش داده می‌شود که به صورت زیر آمده است:

$$\mathbb{E}^{r^j} = \{r^k \in \mathbb{E} \mid \text{نقطه اکسترمم مجاور } r^j \text{ است } r^k\}.$$

همچنین $\mathbb{E}^{r^j}_{r^k}$ زیر مجموعه اکسترمال $\mathbb{H}$ مربوط به $r^j \in \mathbb{E}$ است یکی از نقاط اکسترمم مجاور r^j، مانند $r^k \in \mathbb{E}^{r^j}$ است که این زیر مجموعه اکسترمال به صورت زیر بدست می‌آید:

$$\mathbb{E}^{r^j}_{r^k} = \{h \in \mathbb{H} \mid h = \lambda r^j + (1-\lambda) r^k, 0 \leq \lambda \leq 1\}.$$

ما بارها نیاز داریم که $(n-1)$ نقطه اکسترمم مجاور نقطه اکسترمم r^j را در نظر بگیریم n حالت در این رابطه وجود دارد که هر انتخاب با پارامتر $i \in \mathbb{I}$ در روش زیر داده شده است:

$$\mathbb{C}^j_i = \{r^{j_{i_l}} \in \mathbb{E}^{r^j} \mid l = 1, \ldots, n-1\}. \tag{۱۴.۲}$$

اگر $P_{\mathbb{S}}(x) \in \mathbb{D}$ باشد آنگاه برای هر $j \in \mathbb{J}$ و برای هر $i \in \mathbb{I}$ تصویر متری $P_{\mathbb{S}}(x)$ در بیرون یک زیر مجموعه از $\mathbb{K}^j$ با توجه به تعریف ظاهر می‌شود.

تعریف ۳.۵.۲. حاصل‌ضرب دکارتی مجموعه اندیس $\mathbb{I} * \mathbb{J}$ داده شده است و فرض کنید (i, j) یک عنصر از این مجموعه اندیس باشد مجموعه مرزی $\mathbb{H}$ مربوط به جفت (i, j) زیر مجموعه‌ای از $\mathbb{K}^j$ است که به صورت زیر تعریف می‌شود:

$$\mathbb{M}^j_i = \{(\mathbb{E}^{r^j}_{r^{j_{i_1}}} + \ldots + \mathbb{E}^{r^j}_{r^{j_{i_{n-1}}}}) \setminus ((\bigcup_{l=1}^{n-1} \{r^{j_{i_l}}\}) + \{r^j\}) \mid r^{j_{i_l}} \in \mathbb{C}^j_i\}. \tag{۱۵.۲}$$

در حقیقت مجموعه مرزی $\mathbb{M}^j_i$ مجموع زیرمجموعه‌های اکسترمال مجاور r^j منهای نقاط حدی مجموعه مرزی $\mathbb{M}^j_i$ است که این نقاط حدی عضوی از $\mathbb{M}^j_i$ نیستند. نقاط حدی $\mathbb{M}^j_i$ در $\mathbb{S}$ که در $\mathbb{M}^j_i$ نیستند که از [۶] گرفته شده به صورت زیر داده شده است:

$$\mathbb{L}^j_i = \mathbb{C}^j_i \cup \{r^j\}. \tag{۱۶.۲}$$

که $\mathbb{C}^j_i$ در تعریف ۲.۵.۲ آمده است که آن انتخاب $(n-1)$ نقطه اکسترمم مجاور r^j است به طوری که باقیمانده نقاط مجاور r^j مانند $r^{j_{i_n}}$ در $\mathbb{E}^{r^j} \setminus \mathbb{C}^j_i$ یک نقطه حدی از $\mathbb{M}^j_i$ نیست.

نتایجی که برای مجموعه $\{\mathbb{M}^j_i \mid j \in \mathbb{J}, i \in \mathbb{I}\}$ از تعاریفی که گفته شد را به صورت یک گزاره در زیر بیان می‌کنیم.

گزاره ۱.۵.۲. ۱. فرض کنید $2j-1, 2j \in \mathbb{J}$ است در این‌صورت

$$\mathbb{M}^{2j-1}_i \subset \mathbb{K}^{2j} \qquad \forall i \in \mathbb{I}$$

که $\{r^{2j-1}, r^{2j}\}$ یک جفت نقاط انتهایی متقابل از $\mathbb{H}$ است.

۲. فرض کنید $j \in \mathbb{J}$ و $i \in \mathbb{I}$ باشد در این‌صورت برای هر $l \in \mathbb{J}$ به طوری که $r^l \in \mathbb{E}^{r^j} \setminus \mathbb{C}^j_i$ باشد داریم $r^l \notin \mathbb{L}^j_i$.

۳. جفت از مجموعه های مرزی $\{\mathbb{M}^{2j}_l, \mathbb{M}^k_m\}$ را در نظر بگیرید در این‌صورت $r^k \in \mathbb{L}^{2j-1}_l$ و $\{r^{2j}, r^{2j-1}\}$ نقاط انتهایی مجاور هستنداگر و تنها اگر $l = m$.

۴. فرض کنید $j,k \in \mathbb{J}$ و $i,m \in \mathbb{I}$ آنگاه

$$(r^j \in \mathbb{L}_i^k \wedge r^k \in \mathbb{L}_m^j) \Leftrightarrow \mathbb{M}_i^k = \mathbb{M}_m^j \wedge i = m.$$

۵. مجموعه مرزی $\mathbb{M}_i^j$ را در نظر بگیرید.آنگاه

$$\mathbb{M}_i^j = \mathbb{M}_i^{j_{i_l}} \quad \forall j_{i_l} \in \mathbb{J} \quad s.t \quad r^{j_{i_l}} \in \mathbb{C}_i^j.$$

۶. فرض کنید $j \in \mathbb{J}$ و $i \in \mathbb{I}$ باشد و فرض کنید $\mathbb{M}_i^j$ از معادله ۱۶.۲ داده شده باشد و $\mathbb{E}_{r^{j_{i_n}}}^{r^j}$ از تعریف ۲.۵.۲ داده شده است با اندیس‌های i_n و j_{i_n} به طوری که $r^{jin} \in \mathbb{E}^{r^j} \setminus \mathbb{C}_i^j$ باشد.آنگاه

$$\mathbb{M}_i^j \subset r^j + span(\{z_{i_1}^j, \ldots, z_{i_{n-1}}^j\}) \subset \mathbb{S} \qquad (۱۷.۲)$$

$$\mathbb{E}_{r^{j_{i_n}}}^{r^j} = \{r^j + \lambda \sigma_{i_n}^j \psi_{i_n}^j z_{i_n}^j \mid ۰ \leq \lambda \leq ۱\}.$$

معادله ۱۷.۲ شرح دیگری از انتخاب مقدار مشخص i و به علاوه شرح داده شده تعریف ۳.۵.۲ را آشکار می‌کند، یعنی زیرفضاهای $(n-۱)$ بعدی شامل $\mathbb{S}$ را بررسی می‌کند و با نقطه اکسترمم r^j انتقال داده می‌شود آنگاه مجموعه مرزی $\mathbb{M}_i^j$ زیر مجموعه از این انتقال می‌شود.
مفهوم استفاده شده در معادله ۱۷.۲ در ارتباط با انتخاب دلخواه پارامتر i به طور مداوم در طول این پایان‌نامه مورد استفاده قرار می گیرد.
این نتیجه گیری شرح هندسی ما از اطراف $\mathbb{H}$ در زمینه بهترین تقریب است. در فصل بعدی موقعیت های که گفتیم $P_{\mathbb{S}}(x)$ می‌تواند در آنها باشد را بررسی و محاسبه می کنیم و نتایجی را بدست می‌آوریم که این نتایج حائز اهمیت می‌باشند.

فصل ۳

محاسبات صریح تصویر متری بر روی یک مخروط محدب انتقال‌یافته

۱.۳ مقدمه

در این فصل به بررسی تصویر $P_{\mathbb{S}}(x)$ روی عناصر مجموعه محدب $\mathbb{H}$ می‌پردازیم برای این کار مخروط دوگان را تعریف می‌کنیم و با استفاده از آن به فرموله کردن مشخصه‌سازی بهترین تقریب که در فصل قبل گفتیم می‌پردازیم. در این فصل ما موقعیتی که $P_{\mathbb{S}}(x)$ در مخروط قطبی انتقال‌یافته قرار دارد را بررسی می‌کنیم و قضیه‌ای برای مقایسه مخروط قطبی انتقال‌یافته و مخروط دوگان بیان می‌کنیم و سپس به محاسبه موقعیت $P_{\mathbb{S}}(x) \in \mathbb{D}$ می‌پردازیم و از آن نتایج مهمی بدست می‌آوریم و در پایان فصل نتیجه‌ای را بیان می‌کنیم که نشان می‌دهد بهترین تقریب روی مخروط قطبی انتقال‌یافته به عنوان یک مؤلفه از $\mathbb{H}$ را می‌توان به طور مستقیم محاسبه کرد.

۲.۳ یافتن بهترین تقریب x روی $\mathbb{S}$

اولین گام، یافتن بهترین تقریب $x \in \mathbb{X}$ روی $\mathbb{S}$ است که بر طبق قضیه ۲.۲.۲ منحصر به فرد است. برای این منظور مجموعه زیر را معرفی می‌کنیم:

$$\mathbb{B} = \{b_i \in \mathbb{X} \mid i = ۱, \dots, n\}, \tag{۱.۳}$$

که $\mathbb{B}$ یک پایه متعامد برای $\mathbb{S}$ است به این معنی که برای هر $j \in \mathbb{J}$، $\mathbb{S} = span(\mathbb{Z}^j) = span(\mathbb{B})$ پس با توجه به تعریف ۴.۱.۱، $\langle b_i, b_j\rangle = ۱$ اگر $i = j$ و در غیر این‌صورت $\langle b_i, b_j\rangle = ۰$. با توجه به معادله ۵.۱، بهترین تقریب x روی $\mathbb{S}$ برابر بسط فوریه $P_{\mathbb{S}}(x)$ نسبت به $\mathbb{B}$ است که به صورت زیر بدست می‌آید:

$$P_{\mathbb{S}}(x) = \sum_{i=۱}^{n} \langle b_i, x\rangle b_i \tag{۲.۳}$$

بنابراین بین پایه متعامد $\mathbb{B}$ در $\mathbb{S}$ حرکت می‌کنیم که در آن محاسبات با استفاده از ضرایب فوریه به‌عنوان مختصات آسان می‌شود و در پایه طبیعی $\mathbb{Z}^j$ برای هر $j \in \mathbb{J}$ نتایج می‌تواند مستقیما تعبیر شود. همچنین با توجه به معادله (۱.۱)، بسط $P_{\mathbb{S}}(x)$ نسبت به پایه $\mathbb{Z}^j$ به صورت زیر است:

$$P_{\mathbb{S}}(x) = \sum_{i=۱}^{n} \sigma_i^j \delta_i z_i^j \qquad (\delta_i \in \mathbb{R}, j \in \mathbb{J}) \tag{۳.۳}$$

که σ_i^j در معادله (۷.۲) فصل قبل داده شده است حال رابطه بین $\{\sigma_۱^j \delta_۱, \dots, \sigma_n^j \delta_n\}$ و $\{\langle b_۱, x\rangle, \dots, \langle b_n, x\rangle\}$ را تعیین می‌کنیم.

تبدیل خطی $T^j : \mathbb{S} \to \mathbb{S}$ را که با ضابطه $T^j z_i^j = b_i$ تعریف می‌شود را در نظر بگیرید. ماتریس این تبدیل نسبت به پایه $\mathbb{Z}^j$ با (t_{ik}^j) نشان داده می‌شود به این معنی که

$$b_k = T^j z_k^j = \sum_{i=۱}^{n} t_{ik}^j z_i^j$$

که در آن (t_{ik}^j) از رابطه زیر بدست می‌آید:

$$(t_{ik}^j) = G^{-۱}(z_۱^j, z_۲^j, \dots, z_n^j)(u_{ik}^j) \tag{۴.۳}$$

و

$$(u_{ik}^j) = (\langle b_k, z_i^j\rangle)$$

که $G(z_۱^j, z_۲^j, \dots, z_n^j)$ یک ماتریس گرام است که به صورت زیر تعریف می‌شود:

$$(G(z_۱^j, z_۲^j, \dots, z_n^j))_{i,k} = \langle z_k^j, z_i^j\rangle \qquad (i, k = ۱, \dots, n) \tag{۵.۳}$$

روابط اخیر از معادلات نرمال و ماتریس گرام اشمیت که در فصل اول گفته شد بدست آمده است . در نظر بگیرید که $G(z_1^j, z_2^j, \ldots, z_n^j)$ از [۲] و (t_{ik}^j) از [۳] وارون‌پذیر هستند بنابراین (u_{ik}^j) وارون‌پذیر است. ارتباط بین مجموعه‌های مختصات $P_{\mathbb{S}}(x)$ بترتیب در معادله (۲.۳) و (۳.۳) از روی محاسبات مستقیم بدست می‌آید.

$$P_{\mathbb{S}}(x) = \sum_{i=1}^{n} \langle b_i, x \rangle b_i \quad , \quad P_{\mathbb{S}}(x) = \sum_{i=1}^{n} \sigma_i^j \delta_i z_i^j$$

$$\Rightarrow \sum_{i=1}^{n} \langle b_i, x \rangle b_k = \sum_{i=1}^{n} \sigma_i^j \delta_i z_i^j \Rightarrow \sum_{i,k=1}^{n} t_{ik}^j z_i^j \langle b_k, x \rangle = \sum_{i=1}^{n} \sigma_i^j \delta_i z_i^j$$

$$\Rightarrow \sigma_i^j \delta_i = \sum_{k=1}^{n} t_{ik}^j \langle b_k, x \rangle. \tag{۶.۳}$$

توجه کنید که معادله (۶.۳) مساوی با حل کردن معادله نرمال است و همچنین از معکوس تبدیل $V^j : \mathbb{S} \to \mathbb{S}$ که با ضابطه $z_k^j = V^j b_k$ تعریف می‌شود استفاده می‌کنیم ماتریس این تبدیل نسبت به $\mathbb{B}$ با (v_{ik}^j) نشان داده می‌شود به این معنی که $z_k^j = V^j b_k = \sum_{i=1}^n v_{ik}^j b_i$ که

$$(v_{ik}^j) = (u_{ik}^j)^t, \tag{۷.۳}$$

اندیس بالای t ترانهاده را نشان می‌دهد و بنابراین ضرایب فوریه از $P_{\mathbb{S}}(x)$ در معادله (۲.۳) تابعی از ضرایب معادله (۳.۳)است که با محاسبات مستقیم زیر بدست می‌آید:

$$P_{\mathbb{S}}(x) = \sum_{k=1}^{n} \langle b_k, x \rangle b_k \quad , \quad P_{\mathbb{S}}(x) = \sum_{k=1}^{n} \sigma_k^j \delta_k z_k^j$$

$$\Rightarrow \sum_{k=1}^{n} \langle b_k, x \rangle b_k = \sum_{k=1}^{n} \sigma_k^j \delta_k z_k^j \Rightarrow \sum_{k=1}^{n} b_k \langle b_k, x \rangle = \sum_{i,k=1}^{n} v_{ik}^j b_k \sigma_k^j \delta_k$$

$$\langle b_k, x \rangle = \sum_{k=1}^{n} v_{ik}^j \sigma_k^j \delta_k. \tag{۸.۳}$$

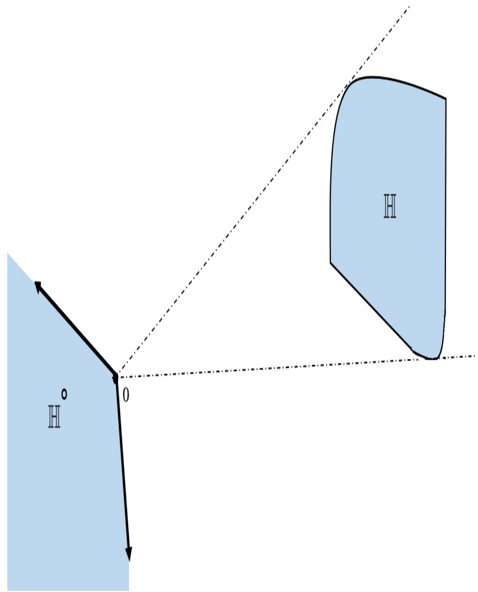
اکنون آماده هستیم تصویر $P_{\mathbb{S}}(x)$ روی مؤلفه‌های $\mathbb{H}$ را بررسی کنیم ما با یک مفهوم هندسی شروع می‌کنیم که این معادل است با فرموله کردن مشخصه‌سازی بهترین تقریب در قضیه ۲.۲.۲ است برای این منظور نیاز به معرفی شیء هندسی مهم داریم که در زیر بیان می‌کنیم:

مخروط دوگان

تعریف ۱.۲.۳. فرض کنید $\mathbb{H}$ زیر مجموعه ناتهی از فضای حاصلضرب داخلی $\mathbb{X}$ باشد، به مجموعه زیر مخروط دوگان می‌گوییم:

$$\mathbb{H}^\circ := \{x \in \mathbb{X} \mid \quad \langle x, y \rangle \leq 0 \quad \text{برای همه } y \in \mathbb{H}\}$$

به طور هندسی با توجه به شکل ۱.۳ مخروط دوگان $\mathbb{H}^\circ$ مجموعه همه بردارهای در $\mathbb{X}$ است که زاویه کمتر از ۹۰ درجه با هر بردار در $\mathbb{H}$ می‌سازد.



شکل ۱.۳: مخروط دوگان

قضیه ۱.۲.۳. فرض کنید $\mathbb{H}$ زیر مجموعه محدب در فضای حاصلضرب داخلی $\mathbb{X}$ باشد و $X \in \mathbb{X}$ و $h \in \mathbb{H}$ باشد. آنگاه $h = P_{\mathbb{H}}(x)$ اگر و تنها اگر

$$x - h \in (\mathbb{H} - h)^{\circ}$$

برهان. با توجه به قضیه ۲.۲.۲ داریم:

$$h = P_{\mathbb{H}}(x) \Leftrightarrow \langle x - h, y - h \rangle \leq 0 \quad \forall y \in \mathbb{H} \Leftrightarrow x - h \in (\mathbb{H} - h)^{\circ} \qquad \forall y - h \in \mathbb{H} - h$$

. $\square$

اکنون می‌توانیم شرح زیر را از قضیه مشخصه‌سازی بهترین تقریب روی مجموعه محدب در فصل قبل بدهیم. در اینجا موقعیت‌های که $P_{\mathbb{S}}(x)$ می‌تواند نسبت به $\mathbb{H}$ قرار بگیرد را بررسی می کنیم و با محاسبات صریح و روشن بدست می‌آوریم.

۱.۲.۳ محاسبه موقعیت $P_{\mathbb{S}}(x) \in (\mathbb{K}^j)_0$

گزاره ۱.۲.۳. فرض کنید $j \in \mathbb{J}$ و $y_0 \in \mathbb{K}^j$. آنگاه

$$y_0 = P_{\mathbb{K}^j}(P_{\mathbb{S}}(x)) \Leftrightarrow \langle P_{\mathbb{S}}(x) - y_0, z^j \rangle \leq 0 \quad \forall z^j \in con(\mathbb{Z}^j)$$

و همچنین داریم:

$$\langle P_{\mathbb{S}}(x) - y_0, y_0 - r^j \rangle = 0$$

به این معنی که $P_{\mathbb{S}}(x) - y_0$ و $y_0 - r^j$ با یکدیگر متعامد هستند.

برهان. با توجه به قضیه ۱.۲.۳ و تعریف ۱.۲.۳ داریم:

$$y_\circ = P_{\mathbb{K}^j}(P_{\mathbb{S}}(x)) \Leftrightarrow (P_{\mathbb{S}}(x) - y_\circ) \in (\mathbb{K}^j - y_\circ)^\circ \Leftrightarrow$$

$$\langle P_{\mathbb{S}}(x) - y_\circ, z^j + r^j - y_\circ\rangle \leq \circ \Leftrightarrow \langle P_{\mathbb{S}}(x) - y_\circ, z^j - (y_\circ - r^j)\rangle \leq \circ \qquad \forall z^j \in con(\mathbb{Z}^j).$$

شرایط نامثبت بودن حاصلضرب داخلی در رابطه اخیر برای همه $z^j \in con(\mathbb{Z}^j)$ برقرار است. پس برای دو انتخاب $z^j = \circ$ و $z^j = ۲(y_\circ - r^j)$ نیز برقرار است پس این دو حالت را در رابطه اخیر قرار می‌دهیم:

$$z^j = \circ \Rightarrow P_{\mathbb{S}}(x) - y_\circ, -(y_\circ - r^j)\rangle \leq \circ \Rightarrow \langle P_{\mathbb{S}}(x) - y_\circ, (y_\circ - r^j)\rangle \geq \circ$$

$$z^j = ۲(y_\circ - r^j) \Rightarrow \langle P_{\mathbb{S}}(x) - y_\circ, ۲(y_\circ - r^j) - (y_\circ - r^j)\rangle \leq \circ \Rightarrow \langle P_{\mathbb{S}}(x) - y_\circ, (y_\circ - r^j)\rangle \leq \circ.$$

$$\Longrightarrow \langle P_{\mathbb{S}}(x) - y_\circ, (y_\circ - r^j)\rangle = \circ.$$

پس $P_{\mathbb{S}}(x) - y_\circ$ و $y_\circ - r^j$ متعامد هستند. □

اکنون با مراجعه به بحث در فصل قبلی وقتی که $P_{\mathbb{S}}(x) \in (\mathbb{K}^j)_\circ$ با توجه به تعریف ۱.۵.۲، $P_{\mathbb{S}}(x)$ در محدوده برخی نقطه اکسترمم $r^j \in \mathbb{E}$ است پس نتیجه مفیدی که از گزاره ۱.۲.۳ می‌توانیم داشته باشیم را به صورت زیر بیان می‌کنیم:

نتیجه ۱.۲.۳. فرض کنید $P_{\mathbb{S}}(x) \in (\mathbb{K}^j)_\circ$ برای هر $j \in \mathbb{J}$ باشد آنگاه $P_{\mathbb{K}^j}(P_{\mathbb{S}}(x)) = r^j$.

برهان. با توجه به تعریف ۱.۵.۲ مخروط قطبی انتقال‌یافته داریم:

$$P_{\mathbb{S}}(x) \in (\mathbb{K}^j)_\circ \Rightarrow \langle P_{\mathbb{S}}(x) - r^j, z^j\rangle \leq \circ \qquad \forall z^j \in con(\mathbb{Z}^j)$$

پس با توجه به گزاره ۱.۲.۳ اگر به جای $y_\circ$، r^j در نظر بگیریم داریم:

$$\langle P_{\mathbb{S}}(x) - r^j, z\rangle \leq \circ \Leftrightarrow P_{\mathbb{K}^j}(P_{\mathbb{S}}(x)) = r^j.$$

□

در فصل بعدی با استفاده از الگوریتم بویل دیکسترا نشان می‌دهیم که در حالتی که $P_{\mathbb{S}}(x) \in (\mathbb{K}^j)_\circ$، بهترین تقریب x نسبت به مجموعه محدب $\mathbb{H}$ برابر نقاط اکسترمم $r^j \in \mathbb{E}$ می‌شود اکنون ما به برخی از نتایجی که در آن مخروط قطبی انتقال‌یافته با مخروط دوگان مقایسه می‌شود می‌پردازیم.

۳.۳ مقایسه مخروط قطبی انتقال‌یافته و مخروط دوگان

قضیه ۱.۳.۳. فرض کنید $j \in \mathbb{J}$ باشد و فرض کنید $\mathbb{N}_\epsilon$ همسایگی از $\circ \in \mathbb{S}$ باشد. آنگاه

.۱

$$(r^j + con(\mathbb{Z}^j))^\circ = (r^j + con(\mathbb{Z}^j))_\circ \Leftrightarrow r^j = \circ.$$

.۲

$$(r^j + con(\mathbb{Z}^j))_\circ \neq \emptyset.$$

.۳

$$\mathbb{N}_\epsilon \subset \mathbb{H} \Rightarrow \mathbb{H}^\circ = \emptyset \Rightarrow (r^j + con(\mathbb{Z}^j))^\circ = \emptyset.$$

.۴

$$\mathbb{N}_\epsilon \subset (r^j + con(\mathbb{Z}^j)) \Rightarrow (r^j + con(\mathbb{Z}^j))^\circ = \emptyset.$$

برهان. (۱) با توجه به تعریف ۱.۲.۳ و معادله (۱۲.۲) داریم:

$$(r^j + con(\mathbb{Z}^j))^\circ = \{x \in \mathbb{X} \mid \langle x, r^j + z^j \rangle \leq \circ \qquad z^j \in con(\mathbb{Z}^j)\},$$

$$(r^j + con(\mathbb{Z}^j))_\circ = \{x \in \mathbb{X} \mid \langle x - r^j, z^j \rangle \leq \circ \qquad z^j \in con(\mathbb{Z}^j)\}$$

$$\Longrightarrow r^j = \circ \Leftrightarrow (r^j + con(\mathbb{Z}^j))^\circ = (r^j + con(\mathbb{Z}^j))_\circ.$$

(۲) با توجه به معادله (۱۲.۲) داریم:

$$(r^j + con(\mathbb{Z}^j))_\circ = \{x \in \mathbb{X} \mid \langle x - r^j, z^j \rangle \leq \circ \qquad z^j \in con(\mathbb{Z}^j)\} \neq \emptyset.$$

پس $(r^j + con(\mathbb{Z}^j))_\circ \neq \emptyset$.

(۳) با توجه به به تعریف ۱.۲.۳ داریم:

$$\mathbb{N}_\epsilon \subset \mathbb{H} \Rightarrow \mathbb{H}^\circ = \{x \in \mathbb{X} \mid \langle x, \mathbb{N}_\epsilon \rangle > \circ\} \Rightarrow \mathbb{H}^\circ = \emptyset \Rightarrow (r^j + con(\mathbb{Z}^j))^\circ = \emptyset.$$

(۴) با توجه به معادله (۱۲.۲) داریم:

$$\mathbb{N}_\epsilon \subset (r^j + con(\mathbb{Z}^j)) \Rightarrow (r^j + con(\mathbb{Z}^j))^\circ = \{x \in \mathbb{X} \mid \langle x, \mathbb{N}_\epsilon \rangle > \circ\} \Rightarrow (r^j + con(\mathbb{Z}^j))^\circ = \emptyset.$$

□

۴.۳ محاسبه موقعیت $P_{\mathbb{S}}(x)$ در یکی از مؤلفه‌های $\mathbb{D}$

می‌خواهیم موقعیت دیگر $P_{\mathbb{S}}(x)$ را محاسبه کنیم. فرض می‌کنیم که $P_{\mathbb{S}}(x)$ در هر مؤلفه از مجموعه $\mathbb{D}$ که در معادله (۱۳.۲) داده شده، باشد. به این معنی است که برای هر $j \in \mathbb{J}$ و برای هر $i \in \mathbb{I}$ تصویر متری $P_{\mathbb{S}}(x)$ در قسمت جلویی مجموعه $\mathbb{M}_i^j \subset \mathbb{K}^j$ است. قبل از ارائه نتیجه برای محاسبه $P_{\mathbb{K}^j}(P_{\mathbb{S}}(x))$ که $P_{\mathbb{S}}(x) \in \mathbb{D}$ نیاز داریم زیر مجموعه از $\mathbb{R}^n$ را که به صورت زیر است معرفی کنیم:

$$\Xi^j = \Bigg\{ \overbrace{\begin{pmatrix} \langle b_1, P_{\mathbb{S}}(x) - r^j \rangle \\ \cdots \\ \cdots \\ \langle b_{n-1}, P_{\mathbb{S}}(x) - r^j \rangle \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ . \\ . \\ 0 \\ \langle b_n, P_{\mathbb{S}}(x) - r^j \rangle \end{pmatrix}}^{\text{عناصر } \sum_{i=1}^{n-1}\binom{n}{i} = 2^n - 2} \Bigg\}, \tag{۹.۳}$$

که بردارها در این مجموعه ترکیبی از تمام انتخاب‌های ممکن از عناصر روی $\mathbb{R}^n$ هستند که دو بردارهای $(\langle P_{\mathbb{S}}(x) - r^j, b_i \rangle)$ و 0 با حداقل یکی از این عناصر است و در آن $P_{\mathbb{S}}(x)$ در معادله‌های (۲.۳) و (۳.۳) و $b_i \in \mathbb{B}$ در معادله (۱.۳) آمده است.

قضیه ۱.۴.۳. فرض کنید $P_{\mathbb{S}}(x) \in \mathbb{D}$ به این معنی که $j \in \mathbb{J}$ و $i \in \mathbb{I}$ وجود دارد به طوری که تصویر متری $P_{\mathbb{S}}(x)$ در قسمت جلویی مجموعه $\mathbb{M}_i^j \subset \mathbb{K}^j$ است آنگاه عنصر منحصر به فرد $\varepsilon_0^j \in \Xi$ وجود دارد به طوری که

$$P_{\mathbb{K}^j}(P_{\mathbb{S}}(x)) = \sum_{i=1}^{n} \sigma_i^j \rho_{0_i} z_i^j + r^j,$$

که $(\sigma^j \rho_0) = (\sigma_i^j \rho_{0_i}) \in \mathbb{R}^n$ حاصل شده با

$$(\sigma^j \rho_0) = (t_{ik}^j)\varepsilon_0^j,$$

که (t_{ik}^j) ماتریس پایه تبدیل $T^j : \mathbb{S} \to \mathbb{S}$ است که در معادله (۴.۳) داده شده است در حقیقت این رابطه حالت کلی معادله (۶.۳) است. به خصوص بهترین تقریب $P_{\mathbb{S}}(x)$ روی $\mathbb{K}^j$ بیشتر از $(2^n - 2)$ آزمایش پیدا می‌شود.

برهان. از گزاره ۱.۴.۲ نتیجه می‌شود که یک تصویر منحصر به فرد روی $\mathbb{K}^j$ است. این راه حل y_0 را می‌تواند به روش زیر نشان دهد:

$$y_0 = \sum_{i=1}^{n} \sigma_i^j \rho_{0_i} z_i^j + r^j = \sum_{i=1}^{n} \sigma_i^j \rho_{0_i} z_i^j + \sum_{i=1}^{n} \sigma_i^j \psi_i^j z_i^j = \sum_{i=1}^{n} \sigma_i^j (\rho_{0_i} + \psi_i^j) z_i^j \quad (\rho_{0_i} \geq 0) \tag{۱۰.۳}$$

شرط متعامد بودن در گزاره ۱.۲.۳ می‌تواند برای $y_\circ$ یک راه حل باشد و بردارهای $P_{\mathbb{S}}(x)-y_\circ$ و $y_\circ - r^j$ با توجه به معادله‌های (۱۰.۳) و (۳.۳) نسبت به پایه $\mathbb{Z}^j$ در $\mathbb{S}$ به صورت زیر نمایش داده می‌شود:

$$P_{\mathbb{S}}(x)-y_\circ=\sum_{i=1}^{n}\sigma_i^j\delta_i z_i^j-\sum_{i=1}^{n}\sigma_i^j(\rho_{\circ_i}+\psi_i^j)z_i^j=\sum_{i=1}^{n}\sigma_i^j(\delta_i-\rho_{\circ_i}-\psi_i^j)z_i^j.$$

$$y_\circ-r^j=\sum_{i=1}^{n}\sigma_i^j\rho_{\circ_i}z_i^j+r^j-r^j=\sum_{i=1}^{n}\sigma_i^j\rho_{\circ_i}z_i^j.$$

پس داریم:

$$P_{\mathbb{S}}(x)-y_\circ=\sum_{i=1}^{n}\sigma_i^j(\delta_i-\rho_{\circ_i}-\psi_i^j)z_i^j, \tag{۱۱.۳}$$

$$y_\circ-r^j=\sum_{i=1}^{n}\sigma_i^j\rho_{\circ_i}z_i^j. \tag{۱۲.۳}$$

که با توجه به گزاره ۱.۲.۳ باید دو معادله (۱۱.۳) و (۱۲.۳) با یکدیگر متعامد باشند. استفاده پایه تبدیل $V^j:\mathbb{S}\to\mathbb{S}$ برای بردارهای $P_{\mathbb{S}}(x)-y_\circ$ و $y_\circ-r^j$ در زیر معادل با نمایش دادن آنها نسبت به پایه $\mathbb{B}$ از $\mathbb{S}$ است پس با توجه به معادله (۶.۱) داریم:

$$P_{\mathbb{S}}(x)-y_\circ=\sum_{i=1}^{n}\langle b_i,P_{\mathbb{S}}(x)-y_\circ\rangle b_i,$$

که با توجه به معادله (۸.۳) داریم:

$$\langle b_i,P_{\mathbb{S}}(x)-y_\circ\rangle=\sum_{i=1}^{n}v_{ik}^j\sigma_i^j(\delta_i-\rho_{\circ_i}-\psi_i^j) \tag{۱۳.۳}$$

و همچنین با توجه به معادله (۶.۱) داریم:

$$y_\circ-r^j=\sum_{i=1}^{n}\langle b_i,y_\circ-r^j\rangle b_i,$$

که با توجه به معادله (۸.۳) داریم:

$$\langle b_i,y_\circ-r^j\rangle=\sum_{i=1}^{n}v_{ik}^j\sigma_i^j\rho_{\circ_j}. \tag{۱۴.۳}$$

پس $P_{\mathbb{S}}(x)-y_\circ$ و $y_\circ-r^j$ با یکدیگر متعامد هستند اگر $P_{\mathbb{S}}(x)-y_\circ$ در زیر فضای $\mathbb{S}$ باشد که متمم متعامد زیرفضای $\mathbb{S}$، $y_\circ-r^j$ است که یک عنصر از$\mathbb{S}$ است. به عبارت دیگر ضرایب فوریه در نمایش $P_{\mathbb{S}}(x)-y_\circ$ نسبت به $\mathbb{B}$ باید ناصفر باشد و باید صفر در نمایش $y_\circ-r^j$ نسبت به $\mathbb{B}$ باشد. و روش دیگری از این قبیل است که $۲^n-۲=\sum_{i=1}^{n-1}\begin{pmatrix} n \\ i \end{pmatrix}$ روش با ضرایب دو

جمله‌ای $\begin{pmatrix} . \\ . \end{pmatrix}$ در $\mathbb{S}$ وجود دارد که می‌تواند متمم متعامد را تفکیک کند. به ویژه فرض کنید که

$$y_0 - r^j \in span\{\{b_1, \ldots, b_m\}\} \wedge P_{\mathbb{S}}(x) - y_0 \in span\{\{b_{m+1}, \ldots, b_n\}\}$$

یا به طور هم ارز

$$y_0 - r^j \notin span\{\{b_{m+1}, \ldots, b_n\}\} \wedge P_{\mathbb{S}}(x) - y_0 \notin span\{\{b_1, \ldots, b_m\}\}.$$

این معادل است با دستگاه معادلات خطی با ضرایب مجهول $\rho_0 = (\rho_{0_i})$ در نمایش y_0 نسبت به z^j است که با استفاده از معادله‌های (۱۳.۳) و (۱۴.۳) به صورت زیر بدست آمده است : :

$$(v_{ik}^j)(\sigma^j \rho_0) = \begin{pmatrix} \langle b_1, P_{\mathbb{S}}(x) - r^j \rangle \\ \cdots \\ \cdots \\ \langle b_m, P_{\mathbb{S}}(x) - r^j \rangle \\ 0 \\ . \\ 0 \end{pmatrix}, \tag{۱۵.۳}$$

که (v_{ik}^j) ماتریس $n \times n$ است که در معادله (۶.۳) تعریف شده است. توجه کنید که طرف راست معادله (۱۵.۳) عناصر از مجموعه Ξ^j تعریف شده در معادله (۸.۳) هستند. یادآوری می‌کنیم که (v_{ik}^j) ماتریس پایه تبدیل $V^j : \mathbb{S} \to \mathbb{S}$ است که وارون‌پذیر است که راه حل منحصر به فرد ρ_0 است. درواقع،

$$(\sigma^j \rho_0) = (t_{ik}^j) \begin{pmatrix} \langle b_1, P_{\mathbb{S}}(x) - r^j \rangle \\ \cdots \\ \cdots \\ \langle b_m, P_{\mathbb{S}}(x) - r^j \rangle \\ 0 \\ . \\ 0 \end{pmatrix}, \tag{۱۶.۳}$$

که (t_{ik}^j) در معادله (۳.۳) داده شده است البته باید بررسی کنیم که y_0 واقعا در $\mathbb{K}^j$ است

$$y_0 \in \mathbb{K}^j \Leftrightarrow \rho_0 \geq 0. \tag{۱۷.۳}$$

فرض کنید y_0 در $\mathbb{K}^j$ باشد برطبق گزاره ۱.۲.۳ برای y_0 پس ازگذشت آزمون غیر مثبت راه حلی ارائه می‌دهیم که به صورت زیر آمده است:

$$\begin{aligned} \langle P_{\mathbb{S}}(x) - y_\circ, z^j\rangle \leq \circ \qquad \forall z^j \in con(\mathbb{Z}^j) \quad &\Leftrightarrow \\ \langle P_{\mathbb{S}}(x) - y_\circ, z_i^j\rangle \leq \circ \qquad \forall i \in \mathbb{I} \quad &\Leftrightarrow \qquad (۱۸.۳) \\ G(z_1^j, \dots, z_n^j)(\sigma^j(\delta - \rho_\circ - \psi^j)) \leq \circ. & \end{aligned}$$

در اینجا اضافه می‌کنیم که آزمایش معادله (۱۷.۳) اضافی است زیرا $y_\circ$ شرط متعامد بودن را در گزاره ۱.۲.۳ ادا می‌کند و با قبول آزمایش (۱۸.۳) یک عنصر از $\mathbb{K}^j$ است. $\rho_\circ$ نقطه ای است که باید در هر صورت برای آزمون (۱۸.۳) در دسترس باشد و بنابراین اگر $y_\circ$ شرط متعامد بودن را بدون قبول آزمایش ۱۷.۳ ادا می‌کند آزمایش (۱۸.۳) می تواند حذف شود و این فرایند پی‌درپی از عناصر دیگر مجموعه‌ی Ξ^j که در سمت راست معادله (۱۵.۳) هستند انتخاب می‌شود. زیرا راه حل مجموعه‌ی Ξ^j جامع است. وجود راه حل تضمین می‌کند که یکی از عناصر Ξ^j هر دو شرط در گزاره‌ی ۱.۲.۳ را ادا می‌کند. از منحصر به فرد بودن راه حل نتیجه می‌شود که این فرایند می‌توان به زودی تمام شود که این عناصر از Ξ^j یافت می‌شود. □

قضیه ۱.۴.۳ وجود نتیجه را واقعاً خوش تعریف می‌کند که وجود آن را پس از ساخت قراردادی برای راه‌حل، ثابت می‌کند مگر آنکه n خیلی کوچک باشد و راه حل قضیه ۱.۴.۳ ممکن است خیلی مفید روش را ارائه ندهد چرا که ممکن است در $۲^n - ۲$ امتحان، راه حل را پیدا کند وقتی که مؤلفه‌ای از $\mathbb{D}$ داده شده که $P_{\mathbb{S}}(x)$ در آن جای‌گذاری می‌شود. اما این فرض می‌تواند به طور کلی ساخته نشود و محاسبه بهترین تقریب را تا حداکثر $۲^n(۲^n - ۲)$ امتحان را اضافه می‌کند. بنابراین سوال بدیهی رخ می‌دهد که اتفاق می‌افتد اگر بدانیم تصویر متری $P_{\mathbb{S}}(x)$ در مؤلفه‌ای از $\mathbb{D}$واقع شده است به این معنی اگر بدانیم به ویژه $P_{\mathbb{S}}(x)$ در قسمت جلویی مجموعه $\mathbb{M}_i^j$ که در معادله ۱۴.۳ تعریف شده باشد. از اثبات قضیه ۱.۴.۳ نتیجه می‌شود که دو انتقال متفاوت برای جستجو بهترین تقریب یعنی $y_\circ$ داریم که به ترتیب در عناصر تجزیه متعامد $\mathbb{S}$ واقع شده‌اند. برای مثال انتقال نسبت به $P_{\mathbb{S}}(x)$ و یکی دیگر نسبت به نقطه اکسترمم $r^j \in \mathbb{E}$ که مربوط به مجموعه $\mathbb{M}_i^j$ است. نتیجه بعدی قضیه ۱.۴.۳ را نشان می دهد که این اطلاعات داده می شود به ما که منجر به محاسبه مستقیم بهترین تقریب می شود.

نتیجه ۱.۴.۳. فرض کنید $j \in \mathbb{J}$ و $i \in \mathbb{I}$ و فرض کنید تصویر متری از $x \in \mathbb{X}$ روی $\mathbb{S}$ باشد و $P_{\mathbb{S}}(x)$ درقسمت جلویی مجموعه $\mathbb{M}_i^j \subset \mathbb{K}^j$ که در تعریف ۵.۴.۲ آمده است، باشد و در یکی از مؤلفه‌های $\mathbb{D}$ باشد آنگاه

$$P_{\mathbb{K}^j}(P_{\mathbb{S}}(x)) = \sum_{l=1}^{n} \sigma_{i_l}^j \rho_{\circ_{i_l}} z_{i_l}^j + r^j,$$

که $(\sigma^j\rho_0) = (\sigma^j_{i_l}\rho_{0\,i_l}) \in \mathbb{R}^n$ که به صورت زیر بدست می‌آید:

$$(\sigma^j\rho_0) = (t^j_{i_l i_k})\begin{pmatrix} \langle b_{i_l}, P_{\mathbb{S}}(x) - r^j\rangle \\ \cdots \\ \cdots \\ \langle b_{i_{n-1}}, P_{\mathbb{S}}(x) - r^j\rangle \\ 0 \end{pmatrix},$$

و $\mathbb{B}^j_{i_n} = \{b^j_{i_1}, b^j_{i_2}, \dots, b^j_{i_n}\}$ پایه متعامد گذرنده از گرام اشمیت است که $\mathbb{Z}^j = \{z^j_{i_1}, \dots, z^j_{i_{n-1}}, z^j_{i_n}\}$ را ایجاد می‌کند به طوری که $b^j_{i_n} = z^j_{i_n}/\|z^j_{i_n}\|$, و اندیس $i_n \in \mathbb{I}$ به طوری‌که $r^{j i_n} \in \mathbb{E}^{r^j} \setminus \mathbb{C}^j_i$ به این معنی

$$\mathbb{E}^{r^j}_{r^{j i_n}} = \{r^j + \lambda\sigma^j_{i_n}\psi^j_{i_n} z^j_{i_n} \mid 0 \leq \lambda \leq 1\} \nsubseteq \mathbb{M}^j_i$$

و در نهایت

$$(t^j_{i_l i_k}) = G^{-1}(z^j_{i_1}, z^j_{i_2}, \dots, z^j_{i_n})(u^j_{i_l i_k})$$

$$(u^j_{i_l i_k}) = (\langle z^j_{i_l}, b_{i_k}\rangle).$$

نتیجه ۱.۴.۳ نشان می دهد که بهترین تقریب روی مخروط محدب انتقال‌یافته به عنوان یک مؤلفه از مجموعه $\mathbb{H}$ را می‌تواند به طور مستقیم محاسبه کرد به ویژه بدون هیچ جستجو کننده یک مرتبه می‌دانیم که $P_{\mathbb{S}}(x)$ در مؤلفه‌ای از $\mathbb{D}$ قرار دارد. از این رو برای این که نتیجه مفید باشد ممکن است مکان $P_{\mathbb{S}}(x)$ را نسبت به $\mathbb{H}$ پیدا کنیم اکنون این یک موضوع ضروری است و البته این تعیین موقعیت با محاسباتی کم پیدا می‌شود در هر صورت خیلی کمتر از محاسبه $2^n(2^n-2)$ امتحان دستوری است.

فصل ۴

الگوریتم بویل دیکسترا

۱.۴ مقدمه

در این فصل قضیه بویل دیکسترا که همگرایی یک فرایند تکراری را برقرار می‌کند و با کمترین محاسبات به محاسبه بهترین تقریب روی مجموعه منحصر به فرد می‌پردازد را بیان می‌کنیم. مفروضات قضیه بویل دیکسترا متناظر با موقعیت مورد نظر ما است. ابتدا چند قضیه و لم برای اثبات این قضیه بیان می‌کنیم و سپس قضیه را اثبات می‌کنیم و از آن در اثبات قضیه اصلی که خصوصیات کامل بهترین تقریب یک عنصر از فضای ضرب داخلی نسبت به زیرمجموعه‌های محدب در آن فضا را نشان می‌دهد، در فصل بعد استفاده می‌کنیم.

۲.۴ همگرایی ضعیف

تعریف ۱.۲.۴. دنباله $\{x_n\}$ در فضای ضرب داخلی $\mathbb{X}$ را همگرای ضعیف به x گوییم و با $x_n \xrightarrow{w} x$ نمایش می‌دهیم و می‌گوییم x_n به طور ضعیف همگرا به xاست اگر برای هر $y \in \mathbb{X}$ آنگاه $\langle x_n, y\rangle \longrightarrow \langle x, y\rangle$ و این معادل است با x_n به طور ضعیف همگرا به x است اگر و تنها اگر برای هر $x^* \in \mathbb{X}^*$ داشته باشیم: $x^*(x_n) \longrightarrow x^*(x)$

قضیه ۱.۲.۴. (۱) اگر $x_n \longrightarrow x$ آنگاه $x_n \xrightarrow{w} x$

(۲) $x_n \longrightarrow x$ اگر و تنها اگر $x_n \xrightarrow{w} x$ و $\|x_n\| \longrightarrow \|x\|$

(۳) اگر X با بعد متناهی باشد آنگاه $x_n \xrightarrow{w} x$ اگر و فقط اگر $x_n \longrightarrow x$

برهان. (۱) فرض کنید $x_n \longrightarrow x$ آنگاه برای هر $y \in \mathbb{X}$ داریم

$$|\langle x_n, y\rangle - \langle x, y\rangle| = |\langle x_n - x, y\rangle| \leq \|x_n - x\|\|y\| \longrightarrow ۰$$

بنابراین $\langle x_n, y\rangle \longrightarrow \langle x, y\rangle$ پس $x_n \xrightarrow{w} x$.

(۲) اگر $x_n \longrightarrow x$ پس با توجه به قسمت (۱) $x_n \xrightarrow{w} x$ و برای برهان $\|x_n\| \longrightarrow \|x\|$ داریم:

$$\|x_n\| = \|(x_n - x) + x\| \leq \|x_n - x\| + \|x\|$$
$$\|x_n\| - \|x\| \leq \|x_n - x\| \Longrightarrow \|x_n\| \longrightarrow \|x\|$$

برعکس اگر $x_n \xrightarrow{w} x$ و $\|x_n\| \longrightarrow \|x\|$ آنگاه

$$\|x_n - x\|^۲ = \|x_n\|^۲ - ۲\langle x_n, x\rangle + \|x\|^۲ \Rightarrow \|x\|^۲ - ۲\langle x, x\rangle + \|x\|^۲ = ۰$$

پس $x_n \longrightarrow x$.

(۳) اگر $dim\mathbb{X} = N$ باشد فرض کنید که $\{e_۱, \ldots e_N\}$ پایه متعامد $\mathbb{X}$ باشد آنگاه با توجه به معادله (۵.۱) برای هر $x \in \mathbb{X}$ داریم:

$$x = \sum_{i=۱}^{N} \langle x, e_i\rangle e_i \qquad , \quad \|x\|^۲ = \sum_{i=۱}^{N} |\langle x, e_i\rangle|^۲$$

از این رو چون $x_n \xrightarrow{w} x$ آنگاه داریم:

$$\|x_n\|^۲ = \sum_{i=۱}^{N} |\langle x_n, e_i\rangle|^۲ \longrightarrow \sum_{i=۱}^{N} |\langle x, e_i\rangle|^۲ = \|x\|^۲$$

پس با توجه به قسمت (۲) $x_n \longrightarrow x$ و برعکس با توجه به قسمت (۱) اثبات می‌شود.
□

قضیه ۲.۲.۴. هر دنباله کراندار در فضای هیلبرت یک زیردنباله همگرای ضعیف دارد.

برهان. برای اثبات به [۲] مراجعه می‌کنیم. □

لم ۱.۲.۴. فرض کنید دنباله x_n همگرا ضعیف به x باشد آنگاه x_n کراندار است و داریم:

$$\|x\| \leq \lim_{n \longrightarrow \infty} inf \|x_n\|.$$

برهان. برای اثبات به [۲] مراجعه می‌کنیم. □

تعریف ۲.۲.۴. زیرمجموعه K از فضای ضرب داخلی $\mathbb{X}$ را به طور ضعیف بسته گوییم اگر $x_n \in K$ و $x_n \xrightarrow{w} x$ از آن نتیجه شود که $x \in K$.

قضیه ۳.۲.۴. (۱) هر مجموعه بطور ضعیف بسته، بسته است.

(۲) یک مجموعه محدب بسته است اگر و فقط اگر به طور ضعیف بسته باشد.

برهان. (۱) فرض کنید K بسته ضعیف باشد و $x_n \in K$ و $x_n \xrightarrow{w} x$ آنگاه نشان می‌دهیم K بسته است. کافیست نشان دهیم $x \in K$ است. با توجه به قضیه ۱.۲.۴ قسمت(۱) چون $x_n \longrightarrow x$ پس $x_n \xrightarrow{w} x$. چون K بسته ضعیف است پس با توجه به تعریف ضعیف بسته $x_n \in K$ است پس K بسته است.

(۲) کافیست نشان دهیم که اگر K بسته و محدب باشد آنگاه K بسته ضعیف است. فرض می‌کنیم $x_n \in K$ و $x_n \xrightarrow{w} x$ اگر $x \notin K$ پس $x^* \in \mathbb{X}^*$ وجود دارد که $\|x^*\| = ۱$ و $x^*(x) > \sup x^*(K)$ است پس نتیجه می‌شود که:

$$x^*(x) > \sup_{n \longrightarrow \infty} x^*(x_n) \geq \lim_{n \longrightarrow \infty} x^*(x_n) = x^*(x_n)$$

که این تناقض است پس $x \in K$ است و K ضعیف بسته است.

□

۳.۴ الگوریتم بویل دیکسترا

می‌خواهیم همگرایی را از یک الگوریتم تکراری که الگوریتم دیکسترا گفته می‌شود اثبات کنیم که اجازه می‌دهد به محاسبه بهترین تقریب اشتراک تعداد متناهی مجموعه محدب بسته $\mathbb{K}_i$ بپردازیم. با کمترین محاسبات بهترین تقریب را روی مجموعه منحصر به فرد پیدا کنیم. فرض کنیم $\mathbb{K}_i$ زیر مجموعه محدب بسته از فضای هیلبرت $\mathbb{X}$ که $i = ۱, \ldots, r$ باشد و فرض کنیم $\mathbb{H} = \cap_{i=۱}^{r} \mathbb{K}_i$ باشد که $\mathbb{H} \neq \emptyset$ است برای هر $x \in \mathbb{X}$ داده شده ما می‌خواهیم $P_{\mathbb{H}}(x)$ را محاسبه کنیم.

برای هر $n \in \mathbb{N}$ فرض می‌کنیم $[n]$ نشانندە'' n به پیمانه r'' باشد که به صورت زیر تعریف می‌شود:

$$[n] = \{۱, ۲, \ldots, r\} \cap \{n - kr \mid k = ۰, ۱, ۲, \ldots\}$$

برای مثال داریم:

$$[۱] = ۱, [۲] = ۲, \ldots [r] = r, [r + ۱] = ۱, [r + ۲] = ۲, \ldots, [۲r] = r, \ldots$$

مثلا برای $r = ۲$ داریم n به پیمانه ۲ به این معنی که:

$$[n] = \{۱, ۲\} \cap \{n - ۲k \mid k = ۰, ۱, ۲, \ldots\}$$

پس برای هر $x \in \mathbb{X}$ دنباله $\{x_n\}$ را به صورت زیر تعریف می‌کنیم:

$$x_۰ = P_{\mathbb{S}}(x) \qquad e_{-(r-۱)} = \ldots = e_{-۱} = e_۰ = ۰$$

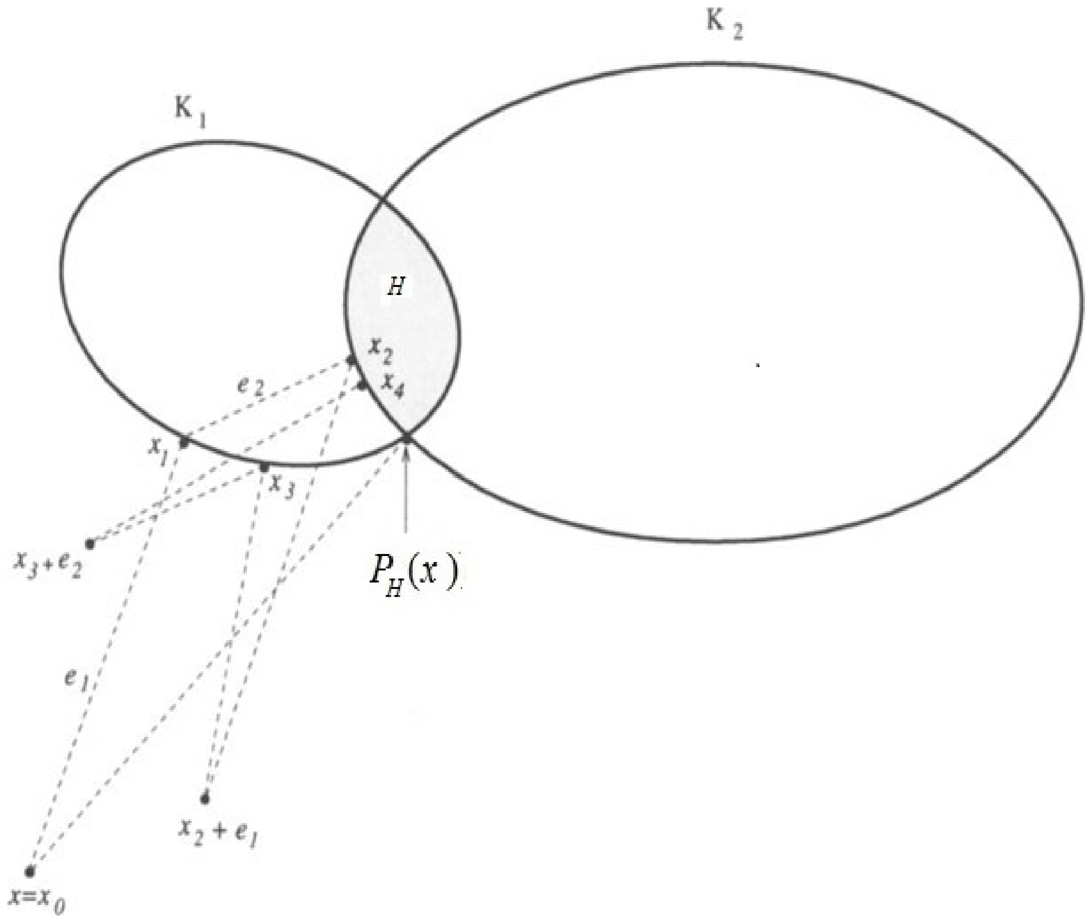
$$x_n = P_{\mathbb{K}[n]}(x_{n-۱} + e_{n-۲}) \tag{۱.۴}$$

$$e_n = x_{n-۱} + e_{n-۲} - x_n \qquad (n = ۱, ۲, \ldots)$$

که این دنباله همگرا به بهترین تقریب x روی $\mathbb{H}$ است که از روش زیر حاصل می‌شود:

$$\lim_{n \to \infty} \|x_n - P_{\mathbb{H}}(x)\| = ۰$$

برای اثبات این عبارات مناسب است چند لم را ارائه دهیم که در زیر آنها را بیان می‌کنیم. شکل ۱.۴ الگوریتم بویل دیکسترا را نشان می‌دهد که $K_۱$ و $K_۲$ دو مجموعه محدب بسته هستند که $\mathbb{H} = K_۱ \cap K_۲$.



شکل ۱.۴: الگوریتم بویل دیکسترا

لم ۱.۳.۴. برای هر n و برای همه $y \in \mathbb{K}^{[n]}$ داریم:

$$\langle x_n - y, e_n \rangle \geq ۰ \tag{۲.۴}$$

برهان. با توجه به مشخصه‌سازی بهترین تقریب روی مجموعه محدب قضیه ۲.۲.۲ اگر فرض کنیم $x = x_{n-۱} + e_{n-۱}$ ، $\alpha_۰ = P_{\mathbb{K}^{[n]}}(x_{n-۱} + e_{n-۲})$ و $\mathbb{H} = \mathbb{K}^{[n]}$ آنگاه برای هر $y \in \mathbb{K}^{[n]}$ با توجه به معادله (۱.۴) داریم:

$$\langle x_n - y, e_n \rangle = \langle P_{\mathbb{K}^{[n]}}(x_{n-۱} + e_{n-۲}) - y, x_{n-۱} + e_{n-۱} - P_{\mathbb{K}^{[n]}}(x_{n-۱} + e_{n-۲}) \rangle \geq ۰$$

□

لم ۲.۳.۴. برای هر $n \geq ۰$ داریم:

$$x - x_n = e_{n-(r-۱)} + e_{n-(r-۲)} + \ldots + e_{n-۱} + e_n \tag{۳.۴}$$

برهان. با استقرا ثابت می‌کنیم فرض می‌کنیم$n = ۰$ باشد آنگاه طرف چپ معادله ۳.۴ و طرف راست آن برابر می‌شود زیرا

$$x - x_۰ = x - x = ۰$$

و

$$e_{-(r-۱)} + e_{-(r-۲)} + \ldots + e-۱ + e_{\circ} = \circ$$

پس حکم برای $n = \circ$ برقرار است حال فرض میکنیم برای حالت n برقرار باشد برای حالت $n+۱$ ثابت می‌کنیم پس داریم:

$$\begin{aligned} x - x_{n+۱} &= (x - x_n) + (x_n - x_{n+۱}) \\ &= e_{n-(r-۱)} + e_{n-(r-۲)} + \ldots + e_{n-۱} + e_n + (e_{n+۱} - e_{n+۱-r}) \\ &= e_{n-(r-۲)} + \ldots + e_{n-۱} + e_n + e_{n+۱} \\ &= e_{n+۱-(r-۱)} + e_{n+۱-(r-۲)} + \ldots + e_{n-۱} + e_n + e_{n+۱}. \end{aligned}$$

پس برای $n+۱$ حکم ثابت شد پس حکم برقرار است. □

لم ۳.۳.۴. برای هر $n \in \mathbb{N}$ و $\circ \leq m \leq n$ و هر $y \in K$ داریم:

$$\begin{aligned} \|x_m - y\|^۲ = \|x_n - y\|^۲ + \sum_{k=m+۱}^{m} \|x_k - x_{k-۱}\|^۲ + ۲ \sum_{k=m+۱}^{n} \langle e_{k-r}, x_{k-r} - x_k \rangle \\ + ۲ \sum_{k=n-(r-۱)}^{n} \langle e_k, x_k - y \rangle - ۲ \sum_{k=m-(r-۱)}^{m} \langle e_k, x_k - y \rangle. \end{aligned}$$

برهان. برای اثبات به [۲] مراجعه می‌کنیم. □

لم ۴.۳.۴. $\{x_n\}$ یک دنباله کراندار است و $\sum_{k=۱}^{\infty} \|x_{k-۱} - x_k\|^۲ < \infty$ بخصوص اگر $n \longrightarrow \infty$ آنگاه $\|x_{n-۱} - x_n\| \longrightarrow \circ$.

برهان. اگر در لم ۱.۳.۴ فرض کنیم $m = \circ$. جمله سوم و چهارم در سمت راست معادله با توجه به لم ۱.۳.۴ نامنفی هستند در صورتی که با توجه به معادله (۱.۴) آخرین شرط برابر صفر است پس داریم:

$$\|x_{\circ} - y\| \geq \|x_n - y\|^۲ + \sum_{k=۱}^{n} \|x_k - x_{k-۱}\|^۲.$$

که نتیجه می‌شود $\{\|x_n - y\|\}$ کراندار است و از این رو $\{x_n\}$ کراندار است پس لم ثابت شد. □

لم ۵.۳.۴. برای هر $n \in \mathbb{N}$ داریم:

$$\|e_n\| = \sum_{k=۱}^{n} \|x_{k-۱} - x_k\|.$$

برهان. با استقرا ثابت می‌کنیم برای $n = ۱$ داریم:

$$\|e_{\circ}\| = \|x_{\circ} - x_۱ - e_{۱-r}\| = \|x_{\circ} - x_۱\|.$$

پس برای ۱ $n=1$ برقرار است فرض می‌کنیم برای همه $m \geq n$ برقرار باشد آنگاه با توجه به معادله (۱.۴) داریم:

$$\begin{aligned}\|e_{n+1}\| &= \|x_n - x_{n+1-r} + e_{n+1-r}\| \leq \|x_n - x_{n+1-r}\| + \|e_{n+1-r}\| \\ &\leq \|x_n - x_{n+1-r}\| + \sum_{k=1}^{n+1-r} \|x_{k-1} - x_k\| \leq \sum_{k=1}^{n+1} \|x_{k-1} - x_k\|.\end{aligned}$$

پس برای $n+1$ برقرار است پس برای n نیز برقرار است پس لم ثابت شد. □

لم ۶.۳.۴.

$$\liminf \sum_{k=n-(r-1)}^{n} |\langle x_k - x_n, e_k\rangle| = 0.$$

برهان. برای اثبات به [۲] مراجعه می‌کنیم. □

لم ۷.۳.۴. زیر دنباله $\{x_{n_j}\}$ از $\{x_n\}$ موجود است به طوری که

$$\limsup_{j\longrightarrow 0} \langle y - x_{n_j}, x - x_{n_j}\rangle \leq 0 \qquad \forall y \in K$$

و

$$\lim_{j\to\infty} \sum_{k=n_j-(r-1)}^{n_j} |\langle x_k - x_{n_j}, e_k\rangle| = 0.$$

برهان. با استفاده از لم ۲.۳.۴ برای همه $y \in k$ و $n \geq 0$ داریم:

$$\begin{aligned}\langle y - x_n, x - x_n\rangle &= \langle y - x_n, e_{n-(r-1)} + e_{n-(r-2)} + \ldots + e_{n-1} + e_n\rangle \\ &= \sum_{k=n-(r-1)}^{n} \langle y - x_n, e_k\rangle \\ &= \sum_{k=n-(r-1)}^{n} \langle y - x_k, e_k\rangle + \sum_{k=n-(r-1)}^{n} \langle x_k - x_n, e_k\rangle. \qquad (۴.۴)\end{aligned}$$

و با استفاده از لم ۱.۳.۴ اولین مجموع بیشتر از صفر نیست بنابراین

$$\langle y - x_n, x - x_n\rangle \leq \sum_{k=n-(r-1)}^{n} \langle x_k - x_n, e_k\rangle \qquad (۵.۴)$$

و با توجه به لم ۶.۳.۴ نتیجه می‌گیریم که زیر دنباله $\{n_j\} \in \mathbb{N}$ وجود دارد به طوری که

$$\lim_{j\to\infty} \sum_{k=n_j-(r-1)}^{n_j} |\langle x_k - x_{n_j}, e_k\rangle| = 0. \qquad (۶.۴)$$

□

۱.۳.۴ قضیه بویل دیکسترا

قضیه ۱.۳.۴. فرض کنید ساختمان دنباله $\{x_n\}$ در $\mathbb{S}$ به صورت زیر باشد:

$$x_۰ = P_{\mathbb{S}}(x), \qquad e_{-۱} = e_۰ = ۰,$$

$$x_n = P_{\mathbb{K}[n]}(x_{n-۱} + e_{n-۲})$$

$$e_n = x_{n-۱} + e_{n-۲} - x_n \qquad (n = ۱, ۲, \dots)$$

و فرض کنید $\mathbb{H}_۱, \mathbb{H}_۲, \dots, \mathbb{H}_r$ زیرمجموعه‌های بسته محدب از فضای هیلبرت $\mathbb{X}$ باشد به طوری که $\mathbb{H} = \cap_{i=۱}^{r} \mathbb{H}_i$ باشد که $\mathbb{H} \neq \emptyset$ آنگاه برای هر $x \in \mathbb{X}$ و دنباله $\{x_n\}$ داریم:

$$\lim_{n \to \infty} \|x_n - P_{\mathbb{H}}(x)\| = ۰$$

که $\|.\|$ نرمی در $\mathbb{X}$ است.

برهان. با توجه به لم ۷.۳.۴ زیر دنباله $\{x_{n_j}\}$ از $\{x_n\}$ موجود است به طوری که

$$\forall y \in \mathbb{H} \qquad \limsup_{j \longrightarrow \infty} \langle y - x_{n_j}, x - x_{n_j} \rangle \leq ۰. \tag{۷.۴}$$

چون با توجه به لم ۴.۳.۴ $\{x_n\}$ کراندار است و با توجه به قضیه ۳.۲.۴ هر دنباله کراندار دارای زیردنباله‌ی همگرای ضعیف است پس $y_۰ \in \mathbb{X}$ وجود دارد به طوری که

$$x_{n_j} \xrightarrow{w} y_۰ \tag{۸.۴}$$

و

$$\lim_{j \longrightarrow \infty} \|x_{n_j}\| \tag{۹.۴}$$

موجود است و با توجه به لم ۱.۲.۴ داریم:

$$\|y_۰\| \leq \liminf_{j \longrightarrow \infty} \|x_{n_j}\| = \lim_{j \longrightarrow \infty} \|x_{n_j}\|. \tag{۱۰.۴}$$

چون تنها تعدادی متناهی از مجموعه‌های $\mathbb{H}_i$ وجود دارد. تعدادی متناهی از x_{n_j} ها باید در مجموعه تنهای $\mathbb{H}_i$ قرار بگیرد. زیرا $\mathbb{H}_i$ بسته و محدب است که با توجه به قضیه ۳.۲.۴ آن به طور ضعیف بسته است و بنابراین $y_۰ \in \mathbb{H}_i$ است. با توجه به لم ۴.۳.۴ $x_n - x_{n-۱} \longrightarrow ۰$ است. با استفاده مکرر از این واقعیت می‌بینیم که همه دنباله‌های $\{x_{n_j+۱}\}$ ، $\{x_{n_j+۲}\}$ و $\{x_{n_j+۳}\}$ به طور ضعیف همگرا به $y_۰$ هستند بنابراین برای هر i ، $y_۰ \in \mathbb{H}_i$ است که

$$y_۰ \in \mathbb{H} \tag{۱۱.۴}$$

است. برای هر $y \in \mathbb{H}$ معادله (۷.۴) و (۱۰.۴) مفهوم زیر را می‌سازند:

$$\begin{aligned}\langle y-y_\circ, x-y_\circ\rangle &= \langle y,x\rangle - \langle y,y_\circ\rangle - \langle y_\circ,x\rangle + \|y_\circ\|^{۲}\\ &\le \lim_{j\longrightarrow\infty}[\langle y,x\rangle - \langle y,x_{n_j}\rangle - \langle x_{n_j},x\rangle + \|x_{n_j}\|^{۲}]\\ &= \lim_{j\longrightarrow\infty}\langle y-x_{n_j}, x-x_{n_j}\rangle \le \circ.\end{aligned}$$

با توجه به قضیه ۲.۲.۲

$$y_\circ = P_{\mathbb{H}}(x) \tag{۱۲.۴}$$

به علاوه در نابرابری بالا $y = y_\circ$ را قرار دهید. ما یک تساوی در زنجیر نابرابری بدست می‌آوریم از این رو

$$\lim_{j\longrightarrow\infty}\|x_{n_j}\|^{۲} = \|y_\circ\|^{۲} \tag{۱۳.۴}$$

و

$$\lim_{j\longrightarrow\infty}\langle y_\circ - x_{n_j}, x-x_{n_j}\rangle \le \circ. \tag{۱۴.۴}$$

با توجه به معادله‌های ۱۳.۴ و ۱۴.۴ و قضیه ۱.۲.۴ قسمت (۲) نتیجه زیر به‌دست می‌آید:

$$\|x_{n_j} - y_\circ\| \longrightarrow \circ. \tag{۱۵.۴}$$

از این رو

$$\|x_{n_j} - P_{\mathbb{H}}(x)\| = \|x_{n_j} - y_\circ\| \longrightarrow \circ. \tag{۱۶.۴}$$

برای کامل شدن اثبات باید نشان دهیم که دنباله‌ی $\{x_n\}$ همگرا به $y_\circ$ است از معادله (۴.۴) با قرار دادن $y = y_\circ$ و $n = n_j$ داریم

$$\langle y - x_{n_j}, x - x_{n_j}\rangle = \sum_{k=n_j-(r-۱)}^{n_j}\langle y_\circ - x_k, e_k\rangle + \sum_{k=n_j-(r-۱)}^{n_j}\langle x_k - x_{n_j}, e_k\rangle. \tag{۱۷.۴}$$

طرف چپ معادله (۱۷.۴) وقتی $j \longrightarrow \infty$ با توجه به معادله (۱۴.۴) به صفر میل می‌کند و مجموع دوم در طرف راست با توجه به معادله (۵.۴) به صفر میل می‌کند از این رو داریم:

$$\lim_{j\longrightarrow\infty}\sum_{k=n_j-(r-۱)}^{n_j}\langle y_\circ - x_k, e_k\rangle = \circ \tag{۱۸.۴}$$

با استفاده از لم ۱.۳.۴ و لم ۳.۳.۴ با قرار دادن $y = y_\circ$ و $m = n_j$ برای هر $n \ge n_j$ داریم:

$$\|x_{n_j} - y_\circ\|^{۲} \ge \|x_n - y_\circ\|^{۲} - ۲\sum_{k=n_j-(r-۱)}^{n_j}\langle e_k, x_k - y_\circ\rangle$$

یا

$$\|x_n - y_\circ\|^۲ \leq \|x_{n_j} - y_\circ\|^۲ + ۲ \sum_{k=n_j-(r-۱)}^{n_j} \langle e_k, x_k - y_\circ \rangle. \qquad (۱۹.۴)$$

اما دو شرط سمت راست نابرابری (۱۹.۴) برای هر $j \longrightarrow \infty$ با توجه به معادله‌های (۱۸.۴) و (۱۵.۴) به صفر میل می‌کند پس نتیجه می‌شود که

$$\lim_{n \longrightarrow \infty} \|x_n - y_\circ\| = \circ$$

و اثبات کامل می‌شود. $\square$

اکنون آماده به ارائه تصویر متری روی مجموعه $\mathbb{H}$ هستیم که در فصل بعدی و نهایی آن را تحلیل می‌کنیم.

فصل ۵

نتیجه گیری

۱.۵ تصویر متری $P_{\mathbb{S}}(x)$ روی مجموعه $\mathbb{H}$

قضیه ۱.۱.۵. فرض کنید $i \in \mathbb{I}$ و $j \in \mathbb{J}$ باشد و فرض کنید با توجه به معادله (۳.۳)
$P_{\mathbb{S}}(x) = \sum_{k=1}^{n} \sigma_k^j \delta_k z_k^j$. دراین‌صورت بهترین تقریب $x \in \mathbb{X}$ روی $\mathbb{H}$ منحصر به فرد است و از یکی از ۳ عبارت زیر به‌دست می‌آید:

(۱) اگر $P_{\mathbb{S}}(x) \in \mathbb{H}$ باشد آنگاه $P_{\mathbb{H}}(x) = P_{\mathbb{S}}(x)$.

(۲) اگر $P_{\mathbb{S}}(x) \in (\mathbb{K}^j)_\circ$ باشد آنگاه $P_{\mathbb{H}}(x) = r^j$.

(۳) اگر $P_{\mathbb{S}}(x)$ در قسمت جلویی مجموعه $\mathbb{M}_i^j \subset \mathbb{K}^j$ در یکی از مؤلفه‌های $\mathbb{D}$ باشد آنگاه $P_{\mathbb{H}}(x) = P_{\mathbb{K}^j}(P_{\mathbb{S}}(x))$.

برهان. (۱) این ادعا ناشی از این واقعیت است که تصویر عملگر $P_{\mathbb{H}}$ خودتوان است. به این معنی که با توجه به قضیه ۲.۴.۲ بهترین تقریبِ بهترین تقریب هر عضو $\mathbb{H}$ نسبت به $\mathbb{H}$ برابر با بهترین تقریب آن عضو نسبت به $\mathbb{H}$ می‌باشد.

(۲) با توجه به نتیجه ۱.۲.۳ داریم:

$$P_{\mathbb{S}}(x) \in (P_{\mathbb{K}^j})_\circ \Rightarrow P_{\mathbb{K}^j}(P_{\mathbb{S}}(x)) = r^j.$$

حال با استفاده از قضیه بویل دیکسترا نشان می‌دهیم که $P_{\mathbb{H}}(x) = r^j$ بدون از دست دادن کلیت فرض می‌کنیم که j فرد است به این معنی که $j = ۲k - ۱$ برای هر $k \in \{۱, \dots, ۲^{n-۱}\}$ با توجه به گزاره ۲.۴.۲ داریم:

$$r^j = r^{۲k-۱} \in r^{۲k} + con(\mathbb{Z}^{۲k}) = \mathbb{K}^{۲k}.$$

و به همین ترتیب اگر فرض کنیم که j زوج باشد به این معنی که $j = ۲k$ برای هر $k \in \{۱, \dots, ۲^n\}$ با توجه به گزاره ۲.۴.۲ داریم:

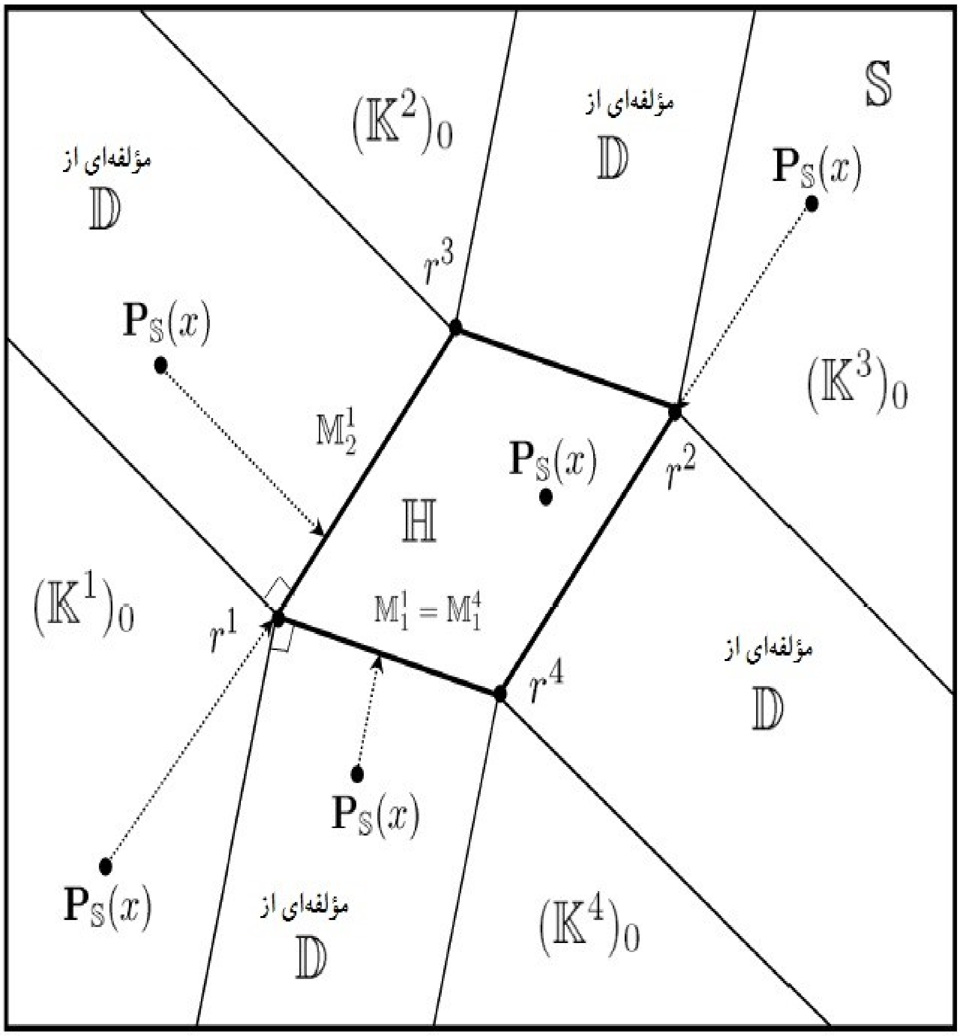
$$r^j = r^{۲k} \in r^{۲k-۱} + con(\mathbb{Z}^{۲k-۱}) = \mathbb{K}^{۲k-۱}.$$

به این ترتیب با توجه به گزاره ۲.۴.۲ $\mathbb{H} = \mathbb{K}^{۲k-۱} \cap \mathbb{K}^{۲k}$ پس قضیه بویل دیکسترا نتیجه می‌شود پس دنباله $\{x_n\}$ که در قضیه بویل دیکسترا گفته شد ساخته می‌شود که همگرا به دنباله ثابت $\{r^j\}$ است. پس $P_{\mathbb{H}}(x) = r^j$.

(۳) بدون از دست دادن کلیت دوباره فرض می‌کنیم که j فرد باشد. به این معنی که $j = ۲k - ۱$ که در آن $k \in \{۱, \dots, ۲^{n-۱}\}$ است. با استفاده از نتیجه ۱.۴.۳ $P_{\mathbb{K}^{۲k-۱}}(P_{\mathbb{S}}(x)) \in \mathbb{M}_i^{۲k-۱} \subset \mathbb{K}^{۲k-۱}$ محاسبه می‌شود. اما چون با توجه به گزاره ۱.۵.۲ $\mathbb{M}_i^{۲k-۱} \subset \mathbb{K}^{۲k}$ پس $P_{\mathbb{K}^{۲k-۱}}(P_{\mathbb{S}}(x)) \in \mathbb{K}^{۲k}$ پس با توجه به گزاره ۲.۴.۲ داریم $\mathbb{H} = \mathbb{K}^{۲k-۱} \cap \mathbb{K}^{۲k}$. بنابراین قضیه بویل دیکسترا نتیجه می‌شود. پس دنباله $\{x_n\}$ که در قضیه بویل دیکسترا گفته شد ساخته می‌شود که همگرا به دنباله ثابت $\{P_{\mathbb{K}^{۲k-۱}}(P_{\mathbb{S}}(x))\}$ است. بنابراین $P_{\mathbb{H}}(x) = P_{\mathbb{K}^j}(P_{\mathbb{S}}(x))$.

□

شکل ۱.۵ به معنی دادن یک تصویر ذهنی برای نشان دادن قضیه ۱.۱.۵ است که اگرچه این نمایش دقیقا درست نیست این شکل نشان می‌دهد که کار با قضیه یک روش شهودی خوشایند است همچنین توجه شود که ساختار هندسی زیبا مسأله را تنظیم می‌کند.توسط این شکل نشان داده می‌شود ایده شکل این است که تعدادی از موقعیت‌های ممکن برای $P_{\mathbb{S}}(x)$ در زیرفضای $\mathbb{S}$ را نشان می‌دهد.تصویری که روی $\mathbb{H}$ نتیجه می‌دهد توسط پیکانهایی در شکل مشخص شده البته به جز موقعیتی که $P_{\mathbb{S}}(x)$ در $\mathbb{H}$ است.



شکل ۱.۵: نشان دهنده تصویر متری روی $\mathbb{H}$

قضیه ۱.۱.۵ خصوصیات کامل بهترین تقریب یک عنصر از فضای ضرب داخلی نسبت به زیر مجموعه‌های محدب در آن فضا را نشان می‌دهد. در بخش بعدی و نهایی ما به توصیف یک برنامه زندگی واقعی از این قضیه می‌پردازیم.

۲.۵ برنامه‌ها:تولید منابع در صنعت نفت

یک وضعیت مشترک در عملیات تولید در صنعت نفت وجود دارد که تعداد چاه‌ها n چاه است و تولیدات روی یک لوله با قطعات بزرگ که به آن لوله اصلی گروهی می‌گویند انجام می‌شود. و تولیدات مشترک از لوله اصلی به گروهی جداکننده که در فازهای مختلف به صورت نفت و آب و گاز که مجزا هستند انتقال داده می‌شوند و مقدار تولید در هر فاز به صورت جداگانه اندازه گیری می‌شود. فازها جداگانه به مناطق پایین دست منتقل داده می‌شوند.

برآورد یا پیش‌بینی تولیدات روی چاه‌های منحصر به فرد از طریق یک مدل ریاضی برای تولیدات یک چاه بدست می‌آید. این مدل از طریق یک چاه آزمایش پیدا می‌شود. برای آزمایش یک چاه به‌ویژه چاه i ، گروهی از چاه‌ها آزمایش می‌شوند که در آن چاه از لوله اصلی گروهی جدا شده و به لوله اصلی آزمایش متصل شده است. که به آزمایش جدا کننده متصل است. در اینجا دوباره فاز تولیدات اندازه‌گیری می‌شود. اما این بار فقط چاه i اندازه‌گیری می‌شود. فاز تولیدات روی چاه آزمایش در پایین دست، ترکیبی از آزمایش جدا کننده است و ترکیبی با تولیدات گروهی مشترک روی چاه‌های دیگر است. مدل چاه نگاشتی از مقادیری است که تولیدات را طی می‌کند. به‌ویژه انواع مختلف فشار روی فاز تولیدات از چاه است. در طول چاه آزمایش دو مقادیر دامنه این نگاشت و برد آن اندازه‌گیری می‌شود. و این اطلاعات برای پیدا کردن نگاشت از داده‌های اندازه‌گیری شده استفاده می‌شود. وقتی که چاه در وضعیت تولید باز می‌گردد مقادیر طی شده هنوز هم در دسترس هستند و با پردازش آنها توسط مدل چاه اعلام‌ها از تولیدات چاه مربوط به دست می‌آید. ایده تولید منابع مسأله در حال حاضر همکاری چاه منحصر به فرد با تولیدات کل را پیدا می‌کند. زیرا چاه‌های مؤثر دیگری در طول تولید هستند و تولید کلی به طور عمومی انجام می‌شود که با مجموع اعلام‌های تولیدات جداگانه برابر نیست. به عبارت دیگر ما باید بهترین تقریب تولیدات کل از روی اعلام‌های تولیدات چاه منحصر به فرد پیدا کنیم. شرایط نمادگذاری مسأله در فصل ۱ شرح داده شده است ما مکاتبات زیر خودمان را با توجه به یک نوع فاز تولیدات فقط محدود می‌کنیم:

$x \in \mathbb{X}$: تولیدات کل

$y_i \in \mathbb{Y}$: تولیدچاهi

n : تعداد چاه‌ها

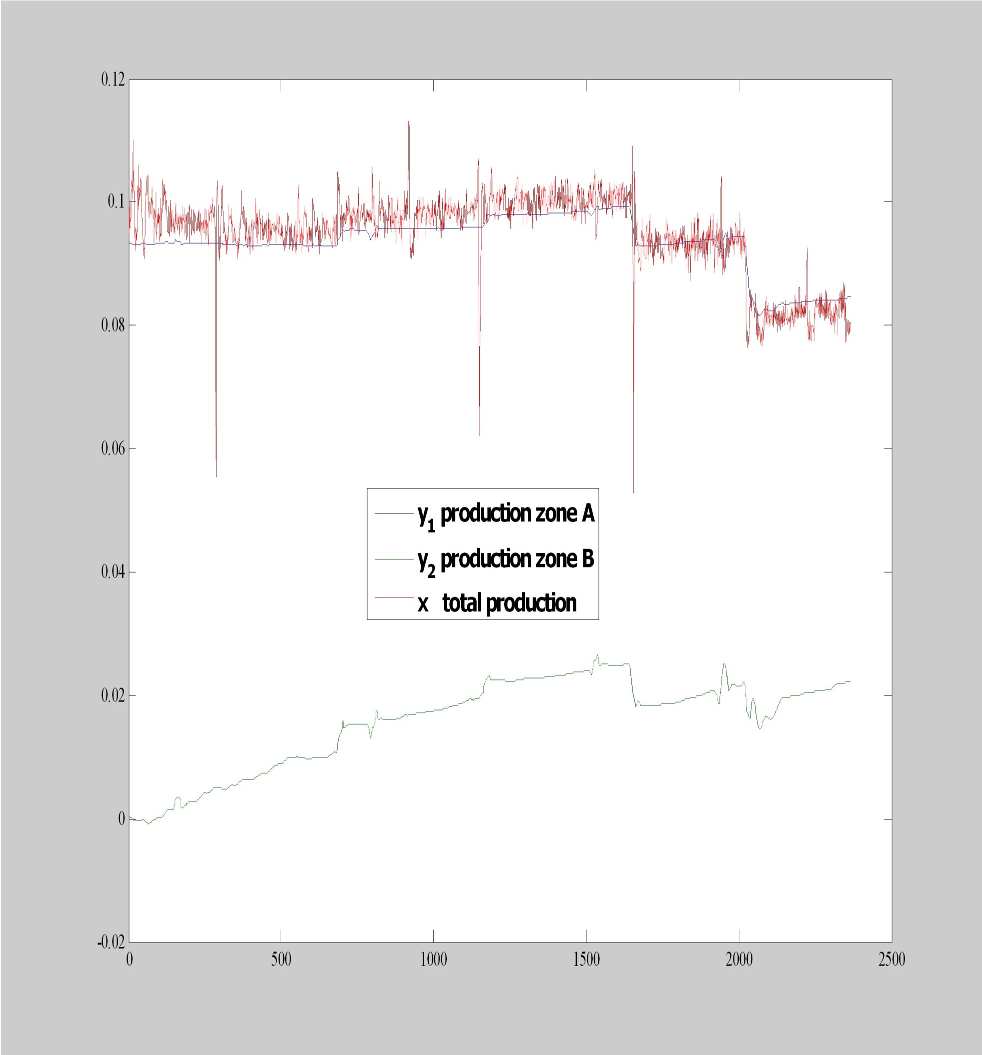
$\mathbb{H} \subset \mathbb{Y}$: ترکیبات قابل قبول

برای ترکیبات قابل قبول باید توجه کنیم که همکاری‌های منفی مجاز به استفاده نیستند. با این حال تا حدودی شمارنده شهودی ممکن است که در حقیقت سهم بزرگتری نیز هم داشته باشند. این به این معنا است که تولید روی چاه خاص ممکن است القا شود زمانی که در آن تولیدات در کم آبی توسط یک یا چند چاه دیگر انجام شود و ممکن است به طور

خاص اصطلاحا در چاه‌های چند منظوره واقع شده باشد که چاه‌های منحصر به فرد در واقع در منطقه‌های مختلف در مخزنی هستند که توسط لوله تولید از چند چاه چندمنظوره بررسی می‌شود. تولیدات در این چاه‌ها به این صورت به‌دست می‌آید که چاه‌هایی که تولیدات آنها به پایان رسیده است در نزدیکی آنها یک چاه دیگر احداث می‌شود و به وسیله چاه‌هایی که تولیدات آنها به پایان رسیده تزریق می‌شود و تولیدات چاه جدید بیشتر می‌شود که این در مدل ریاضی ما به عنوان " قابل قبول نیست" ارزیابی می‌شود.

شکل ۲.۵ در صفحه بعد تولیدات کل روی یک چاه چند منطقه را نشان می‌دهد که شامل ۲ منطقهA و B است. فرض کنید تولیدات موافق منطقه A را با y_1 و منطقه B را با y_2 نشان دهیم. تولیدات روی مناطق جداگانه با استفاده از مدل‌های چاه محاسبه می‌شوند و اعلام‌هایی از عملکرد خود وقتی که مناطق در حال تولید به طور جداگانه هستند اعلام می‌شود.

از شکل ۲.۵ در صفحه بعد نتیجه می‌شود که اگر منطقه B قصد داشته باشد که در تولیدات کل سهم داشته باشد عملکرد آن خیلی از منطقه A کمتر خواهد بود و وقتی که آن به تنهایی تولید می‌کند خیلی هزینه بر خواهد بود. از منظر دیگر انتظار می‌رود که منطقه A تقریبا به طور کامل در فاصله‌ی دورتری از منطقه B قرار دارد وقتی که همزمان تولید وجود دارد پس منطقه A بهتر از منطقه B است.



شکل ۲.۵: تولید کل x و تولید منحصر به فرد $y_۱$ و $y_۲$

حال موقعی است که $P_{\mathbb{S}}(x)$ را پیدا کنیم فرض کنید $P_{\mathbb{S}}(x)$ به صورت زیر باشد:

$$P_{\mathbb{S}}(x) = ۱/۰۵ y_۱ - ۰/۱۸ y_۲.$$

همکاری منفی $y_۲$ البته از نظر فیزیکی غیرممکن است. ارزیابی کامل‌تر این نتیجه پس از تعریف مجموعه $\mathbb{H}$ از ترکیبات پذیرفتنی ممکن می‌شود. در تمرین عملیات تولید موقعیت $P_{\mathbb{S}}(x)$ نسبت به $\mathbb{H}$ ممکن است ارزش تشخیص مفهوم را داشته باشد که عملیات قطعاً انجام شده است یا نتیجه‌گیری‌ها ممکن است براساس این موقعیت‌های نسبی $P_{\mathbb{S}}(x)$، کشیده شود. اما برای این ضروری است که یک انتخاب عاقلانه از پارامترهای قابل قبول برای پارامترهای شکل α و β در تعریف معادله (۱۰.۱)، $\mathbb{H}$ ساخته شود. در جدول زیر اثر $P_{\mathbb{H}}(x)$ ناشی از انتخاب‌های مختلف برای پارامترهای α و β را نشان می‌دهد. به منظور درک نتایج ارائه شده در اینجا نمایش‌هایی از نقاط اکسترمم که در محاسبات استفاده شده تعیین می‌کنیم:

$$r^۱ = \alpha_۱ y_۱ + \alpha_۲ y_۲.$$

$$r^{۲} = \beta_{۱} y_{۱} + \beta_{۲} y_{۲}.$$

$$r^{۳} = \alpha_{۱} y_{۱} + \beta_{۲} y_{۲}.$$

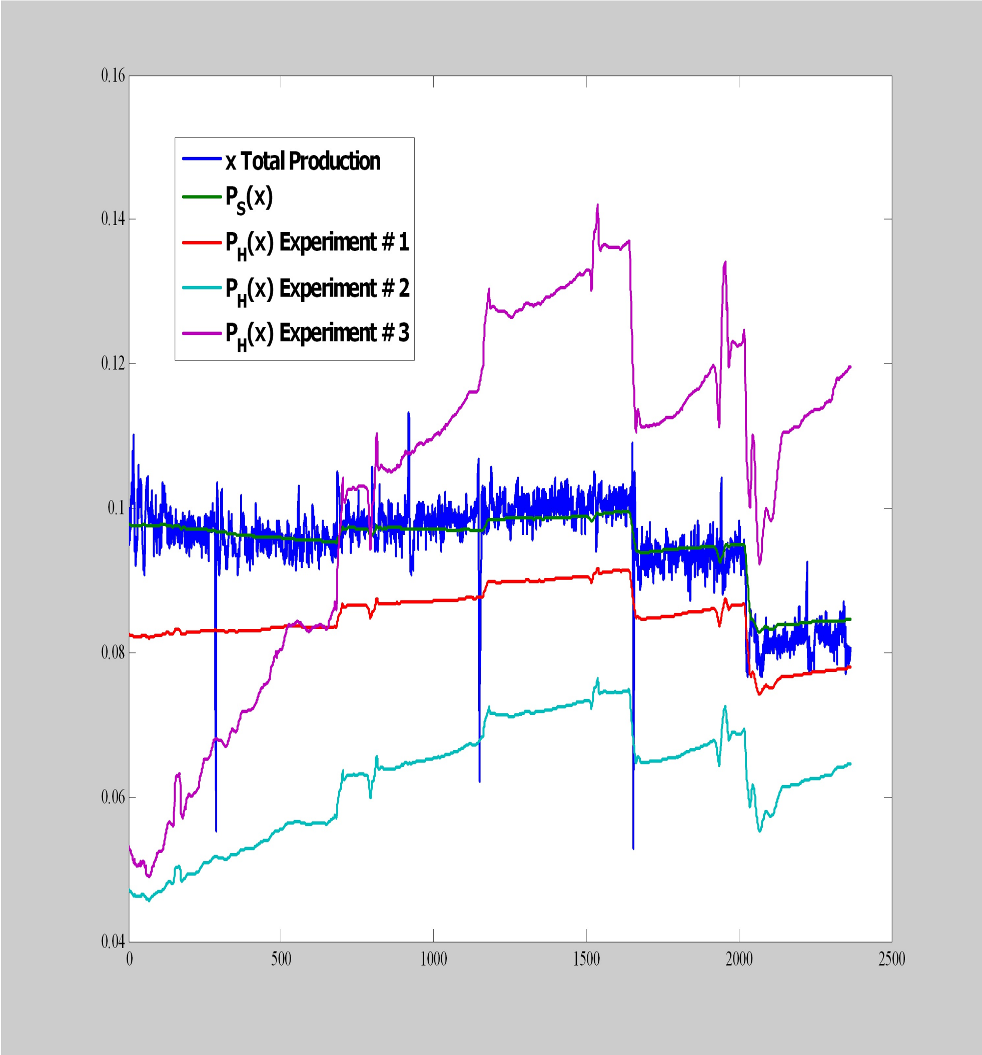
$$r^{۴} = \beta_{۱} y_{۱} + \alpha_{۲} y_{۲}.$$

شماره آزمایش	$[\alpha_{۱}, \alpha_{۲}]$	$[\beta_{۱}, \beta_{۲}]$	مشاهده $P_{\mathbb{S}}(x)$
۱	$[/۱; /۰۵]$	$[۲; ۱]$	$\mathbb{M}^{۱}_{۲} = \mathbb{M}^{۳}_{۲}$
۲	$[/۱; /۱]$	$[/۵; ۱]$	$(\mathbb{K}^{۲})_{\circ}$
۳	$[\circ; /\circ]$	$[۵; ۳]$	$\mathbb{M}^{۲}_{۱} = \mathbb{M}^{۳}_{۱}$

شکل ۳.۵ نمایش گرافیکی از موقعیت‌های مختلف در جدول ملاحظه شده است. انتخاب ما از پارامترهای شکل مسلما پیرو یک اثر نمایش است به‌خصوص انتخاب ساخته شده برای پارامترهای شکل در آزمایش سوم که پوچ است به این معنی که تولید کل تقریبی خیلی بزرگتر از تولید کل اندازه‌گیری شده است. این شرح از چنین وضعیتی که در بالا ذکر شده غیر قابل قبول است که اجازه می‌دهد سهم میزان تولید چندتا سهم واحد باشد. در هر صورت این نتایج با این حال ضرورت انتخاب معقول از پارامتر شکل را نشان می‌دهد. یک راه برای انجام طرح مسأله بالا یک محیط کاملا متفاوت ریاضی مخصوصا در شرایط نظریه ایده‌آل به عنوان بخشی از جبر جابجایی است. تولید کل پس از یک عضو از ایده آل تولید شده توسط تولید جداگانه است. با فرض این‌که نمایش‌های چندجمله‌ای برای مدل‌های چاه باشد این امر منجر به نمایش زیر برای تولید کل می‌شود:

$$x = \sum_{i=۱}^{n} s_i y_i,$$

که s_i چندجمله‌ها هستند. مدل‌سازی اثر متقابل چاه است. s_i پس از آن برای ساخت صحیح پارامترهای شکل ساخته شده است.



شکل ۳.۵: تولید کل x، $P_{\mathbb{S}}(x)$ و $P_{\mathbb{H}}(x)$ با انتخاب‌های مختلف از پارامترهای شکل.

مراجع

[1] Ole Christensen (2003), "**An Introduction to Frames and Riesz Bases**", Birk¨auser Verlag, Boston,

[2] Frank Deutsch(2001), "**Best Approximation in Inner Product Spaces**", Springer Verlag, New York,

[3] Paul R. Halmos (1987), "**Finite-Dimensional Vector Spaces**", Springer Verlag, New York, Reprinted second edition,

[4] Daniel Heldt, Martin Kreuzer, Sebastian Pokutta, and Hennie Poulisse (2006), "**Approximate Calculations of Zero-Dimensional Ideals**", submitted to Journal on Symbolic Computation,

[5] Jean-Baptiste Hiriart-UrrutyandClaudeLemar´echal (2001),"**Fundamentals of Convex Analysis** ", Springer Verlag, Berlin,

[6] John M. Lee (2000), "**Introduction to Topological Manifolds**", Springer Verlag, New York,

[7] Roger Webster (2002), "**Convexity** ", Oxford University Press, Reprinted edition,

Abstract

The range of a linear operator over a polyhedron, considered in this dissertation, is a convex subset H of a finite dimensional subspace S of the ambient inner product space X. According to the reduction principle, the best approximation to a point of X from H equals the best approximation $P_S(x)$ from H. From a geometric point of view different locations of $P_S(x)$ with respect to H, these differences may have their bearings on the computation best approximation to a point of X from H of according to the reduction principle, Using the Boyle-Dykstra theorem,by reducing the computation to a problem of finding best approximations x from H . The result is finally applied to real-life data from oil industry in that a solution is presented for a very important problem in oil - and gas production operations called the reconciliation problem, where the contribution of individual wells to a measured total production has to be assessed.

keywords: best approximation, reduction principle, subspace, Boyle-Dykstra

Shahrood University of Technology

Faculty Of Mathematical Sciences

MSc Thesis in: Mathematical analysis

Best approximation to an element of an inner product space from the range of a linear operator over a polyhedron

By: Majid Mirzayi

Supervisor

Mahdi Iranmanesh

september 2017