بسم الله الرحمن الرحیم

دانشگاه حکیم سبزواری

دانشکده ریاضی و علوم کامپیوتر

پایان‌نامه برای دریافت درجه کارشناسی ارشد در رشته ریاضی کاربردی
گرایش تحقیق در عملیات

شناسایی فریم های زمینه ویدیو با استفاده از تجزیه QR

استاد راهنما

دکترمحمود امین‌طوسی

استاد مشاور

دکتر امین رفیعی

پژوهشگر:

سمیه پورصدیق

بهمن ۱۳۹۴

دانشگاه حکیم سبزواری

فرم 113 - ت

بسمه تعالی

فرم ارزشیابی و صورتجلسه دفاع از پایان نامه کارشناسی ارشد*

جلسه دفاع از پایان نامه آقای/خانم سمیه پورصدیق دانشجوی رشته ریاضی کاربردی
گرایش: تحقیق در عملیات به شماره دانشجویی ۹۲۱۳۱۳۳۰۲۰ با عنوان: شناسایی فریم های زمینه ویدیو با استفاده از تجزیه QR
در مورخه ۹۴/۱۱/۱۹ در دانشکده ریاضی و علوم کامپیوتر تشکیل و توسط هیات داوران مورد ارزشیابی قرار گرفت و نمره......۱۹...... برابر درجه......عالی......... برای آن تعیین گردید.
به این ترتیب از این تاریخ آقای/ خانم.....سمیه پورصدیق................... به عنوان کارشناس ارشد در رشته مذکور شناخته می شود.

موارد ارزشیابی	موارد	حداکثر نمره	نمره کسب شده
1- کیفیت نگارش	رعایت اصول نگارش انسجام در تنظیم بخشهای مختلف ،کیفیت تصاویر ، جداول و اشکال ،تنظیم فهرست ها ، منابع و ماخذ	4	۳/۵
2- کیفیت علمی	بررسی تاریخچه و سابقه تجربی و نظری موضوع انسجام منطقی دربخش های مختلف پایان نامه ،ابتکار و نوآوری ،اهمیت و ارزش علمی پایان نامه ،استفاده از منابع معتبر و جدید ،کیفیت تجزیه و تحلیل یافته ها و نتیجه گیری ،روشن بودن روش کار ،هدف ها و فرضیه های تحقیق ،جدید بودن روش تحقیق	10	۱۰
3-کیفیت ارایه در جلسه دفاع	تسلط بر موضوع و بیان واضح وتفهیم آن ،توانایی در پاسخگویی به سوالات مطرح شده در جلسه ، رعایت زمان ارایه ، روش ارائه	4	۴
4-ارزشیابی گزارشات	گزارش های دوره ای پیشرفت کار(حداقل 4مورد)	1	۱
5- خروجی پایان نامه	مقاله مستخرج از پایان نامه: این نمره به صورت زیر اختصاص می یابد 1) چکیده کنفرانسی هر مورد 0/25نمره تا سقف0/5 نمره 2) مقاله کامل در مجموع مقالات همایشهای معتبر یا مقاله در مجلات علمی-ترویجی معتبر پذیرفته شده یا چاپ شده هر مورد 0/5 نمره تا سقف 1نمره 3)مقاله پذیرفته شده یا چاپ شده در مجلات علمی پژوهشی معتبر 1 نمره 4)مقاله ارسال شدهبه مجلات علمی پژوهشی معتبرهر مورد 25/. نمره تا سقف 0/5 نمره 5) دستگاه ساخته شده دارای گواهی ثبت اختراع یا به سفارش سازمان ها تا سقف 1 نمره 6) دستگاه ساخته شده کاربردی که به تایید رئیس دانشکده رسیده باشد تا سقف 0/5 نمره	1	۰/۵
جمع			۱۹

درجه معادل کسب شده: (از 19تا 20 عالی☑ از18تا 18/99بسیار خوب ☐ از 16تا 17/99خوب ☐ از 14 تا 15/99قابل قبول☐ کمتر از 14 غیر قابل قبول☐)

مشخصات هیات دوران

ردیف	نام ونام خانوادگی	سمت	مرتبه علمی	محل کار	امضاء
1	محمود امین طوسی	استاد راهنما	استادیار	دانشگاه حکیم سبزواری	
2	امین رفیعی	استاد مشاور	استادیار	دانشگاه حکیم سبزواری	
3	مسعود مشرقی	استاد داور	استادیار	دانشگاه حکیم سبزواری	
4	سید محمد صادق نبوی ثالث	نماینده تحصیلات تکمیلی	استادیار	دانشگاه حکیم سبزواری	

امضاء
مدیر گروه ریاضی کاربردی

امضاء
رئیس دانشکده

* این فرم الزاما باید به صورت تایپ شده تهیه، ارسال و در پایان نامه درج شود

سوگند نامه دانش آموختگان دانشگاه حکیم سبزواری

به نام خداوند جان و خرد   کزین برتر اندیشه بر نگذرد

اینک که به خواست آفریدگار پاک، کوشش خویش و بهره گیری از دانش استادان و سرمایه‌های مادی و معنوی این مرز و بوم، توشه‌ای از دانش و خرد گردآورده‌ام، در پیشگاه خداوند بزرگ سوگند یاد می‌کنم که در به کارگیری دانش خویش، همواره بر راه راست ودرست گام بردارم. خداوند بزرگ، شما شاهدان، دانشجویان و دیگر حاضران را به عنوان داورانی امین گواه می گیرم که ازهمه دانش و توان خود برای گسترش مرزهای دانش بهره گیرم و از هیچ کوششی برای تبدیل جهان به جایی بهتر برای زیستن، دریغ نورزم. پیمان می‌بندم که همواره کرامت انسانی را در نظر داشته باشم و همنوعان خود را در هر زمان و مکان تا سر حد امکان یاری دهم. سوگند می خورم که در به کارگیری دانش خویش به کاری که با راه و رسم انسانی، آیین پرهیزگاری، شرافت و اصول اخلاقی برخاسته از ادیان بزرگ الهی، به ویژه دین مبین اسلام، مباینت دارد دست نیازم. همچنین در سایه اصول جهان شمول انسانی و اسلامی، پیمان می‌بندم از هیچ کوششی برای آبادانی و سرافرازی میهن و هم میهنانم فروگذاری نکنم و خداوند بزرگ را به یاری طلبم تا همواره در پیشگاه او و در برابر وجدان بیدار خویش و ملت سرافراز، بر این پیمان تا ابد استوار بمانم.

نام و نام خانوادگی: سمیه پورصدیق

تاریخ و امضا:

تأییدیه‌ی صحت و اصالت نتایج

باسمه تعالی

اینجانب سمیه پورصدیق به شماره دانشجویی ۹۲۱۳۱۳۳۰۲۰ دانشجوی رشته ریاضی کاربردی مقطع تحصیلی کارشناسی ارشد تأیید می‌نمایم که کلیه‌ی نتایج این پایان‌نامه حاصل کار اینجانب و بدون هرگونه دخل و تصرف است و موارد نسخه برداری شده از آثار دیگران را با ذکر کامل مشخصات منبع ذکر کرده‌ام. درصورت اثبات خلاف مندرجات فوق، به تشخیص دانشگاه مطابق با ضوابط و مقررات حاکم (قانون حمایت از حقوق مؤلفان و مصنفان و قانون ترجمه و تکثیر کتب و نشریات و آثار صوتی، ضوابط و مقررات آموزشی، پژوهشی و انضباطی ...) با اینجانب رفتار خواهد شد و حق هرگونه اعتراض درخصوص احقاق حقوق مکتسب و تشخیص و تعیین تخلف و مجازات را از خویش سلب می‌نمایم. در ضمن، مسؤولیت هرگونه پاسخگویی به اشخاص اعم از حقیقی و حقوقی و مراجع ذی صلاح (اعم از اداری و قضایی) به عهده ی اینجانب خواهد بود و دانشگاه هیچ گونه مسؤولیتی در این خصوص نخواهد داشت.

نام و نام خانوادگی: سمیه پورصدیق

تاریخ و امضا:

تقدیم به:

همسر و فرزندم

و

پدر و مادرم

سپاس خداوندگار حکیم را که با لطف بی کران خود، آدمی را زیور عقل آراست.
در آغاز وظیفه خود می دانم از زحمات بی دریغ استاد راهنمای خود، جناب آقای دکتر محمود امین‌طوسی صمیمانه تشکر و قدردانی کنم که قطعاً بدون راهنمایی های ارزنده ایشان، این مجموعه به انجام نمی رسید.
از جناب آقای دکتر امین رفیعی که زحمت مطالعه و مشاوره این رساله را تقبل فرمودند و در آماده سازی این رساله، به نحو احسن اینجانب را مورد راهنمایی قرار دادند، کمال امتنان را دارم.
همچنین لازم می دانم از گروه پارسی‌لاتک در پاسخگویی به مشکلات کاربران کمال قدردانی را داشته باشم.
در پایان، بوسه می زنم بر دستان خداوندگاران مهر و مهربانی، پدر و مادر عزیزم و بعد از خدا، ستایش می‌کنم وجود مقدس شان را و تشکر می‌کنم از خانواده عزیزم به پاس عاطفه سرشار و گرمای امیدبخش وجودشان، که بهترین پشتیبان من بودند.

سمیه پورصدیق

بهمن ۱۳۹۴

فهرست مطالب

فهرست تصاویر

فرم چکیده ی پایان‌نامه ی دوره ی تحصیلات تکمیلی

مدیریت تحصیلات تکمیلی

دانشگاه حکیم سبزواری

نام خانوادگی دانشجو: پورصدیق	نام: سمیه	ش. دانشجویی: ۹۲۱۳۱۳۳۰۲۰
استاد راهنما: دکتر محمود امین‌طوسی		
استاد مشاور: دکتر امین رفیعی		
دانشکده ریاضی و علوم کامپیوتر	رشته: ریاضی کاربردی	گرایش: تحقیق در عملیات
مقطع: کارشناسی ارشد	تاریخ دفاع: بهمن ۱۳۹۴	تعداد صفحات: ۱۰۲
عنوان پایان‌نامه: شناسایی فریم های زمینه ویدیو با استفاده از تجزیه QR		
کلید واژه‌ها: الگوریتم گرام اشمیت، تجزیه QR، مدل‌سازی زمینه		

چکیده: شناسایی و استخراج زمینه یکی از مهم ترین مراحل در پردازش ویدیو و مراقبت بینایی می‌باشد. در برخی از روش‌های مورد استفاده در این حوزه از مقادیر قطر اصلی ماتریس R در تجزیه QR برای شناسایی فریم‌های زمینه استفاده شده است. یکی از مشکلات این شیوه، محاسبه مجدد تجزیه QR در هنگام ورود فریم جدید ویدئو است. در این پایان‌نامه چند روش بروزرسانی تجزیه QR و برخی کاربردهای آن در یادگیری ماشین مورد بررسی قرار گرفته است. به عنوان یک کار جدید، با تغییر چیدمان ماتریس ورودی و اصلاح یکی از روش‌های تجزیه‌ی ماتریس با نام «الگوریتم گرام-اشمیت»، تجزیه ماتریس بروزرسانی شده در زمان کمتری انجام شده است. با روش پیشنهادی، حجم محاسبات مدل‌سازی زمینه، نسبت به محاسبه مجدد تجزیه ماتریس کمتر شده و زمان پردازش از نظر تئوری به نصف کاهش پیدا کرده است.

پیش‌گفتار

شناسایی و استخراج زمینه یکی از مهم ترین مراحل در پردازش ویدیو و کاربرد های مراقبت بینایی می باشد. هدف مساله مدلسازی زمینه یک ویدیو، تولید مدلی از زمینه ویدیو به نحوی است که با در دست داشتن آن بتوان به نحو موثری اشیاء متحرک صحنه را استخراج نمود. در این پایان نامه از تجزیه QR برای مشخص کردن فریم هایی که به احتمال زیاد متعلق به زمینه هستند استفاده شده و سپس با یکی از روش های مدل سازی زمینه،زمینه مدل می‌شود . در یک کاربرد عملی نیاز است که مدل سازی زمینه هماهنگ با ورود فریمهای جدید انجام شود. لازمه این امر با روش فوق الذکر، بروزرسانی تجزیه QR مبتنی بر تجزیه قبلی می باشد. در این پایان نامه با استفاده از الگوریتم اصلاح شده گرام - اشمیت این کار انجام شده است. در کاربردهایی دیگری هم نیاز به بروزرسانی تجزیه QR می باشد که دو مثال مورد بررسی قرار گرفته اند.

این پایان‌نامه شامل ۴ فصل است:

در فصل ۱ تعاریف و مفاهیم مورد نیاز مطرح خواهند شد.

در فصل ۲ به بیان تعاریف و مقدمات نظری برای تجزیه ماتریس ها و تبدیل هوسهولدر و تجزیه QR و الگوریتم های آن پرداخته شده است.

در فصل ۳ به بیان بروزرسانی تجزیه QR و کاربردهای آن می پردازیم.

فصل آخر به بیان چگونگی استفاده از تجزیه QR در شناسایی فریم های زمینه و ذکر نتایج پیاده سازی آن با کمک الگوریتم گرام اشمیت اختصاص یافته است.

تذکر:تمام نرم های به کار رفته در پایان نامه نرم ۲ می باشند.

فصل ۱

تعاریف و مفاهیم اولیه

۱-۱ مقدمه

در این فصل ابتدا به مفاهیم و تعاریف مورد نیاز این پایان‌نامه از قبیل متعامد سازی، تصاویر متعامد و اهمیت آنها پرداخته و سپس مساله حداقل مربعات و کاربرد آن در حل دستگاه معادلات ناسازگار مطرح می شود. سپس روند حل مساله حداقل مربعات با استفاده از معادلات نرمال ،تجزیه چالسکی و تجزیه QR ارائه می شود. در پایان به موضوع کاربرد روش حداقل مربعات در برازش داده‌ها، که یکی از مباحث پایه‌ای در تخمین و شناسایی سیستم‌ها می‌باشد پرداخته شده است.

۱-۱-۱ فضا های برداری

در مطالعه مفاهیم جبر خطی و دستگاه معادلات جبری مفهوم میدان [۱]و فضای برداری [۲]از اهمیت ویژه ای برخوردار است و اساس کلیه تحلیل های جبر خطی را تشکیل می دهد.

تعریف ۱-۱-۱. میدان. یک میدان مجموعه ای از اسکالر ها است به طوریکه همراه دو عمل جمع وضرب شرایط زیر را برآورده می سازد [۱].

۱-برای هر اسکالر $\alpha, \beta \in F$، $\alpha + \beta \in F, \alpha\beta \in F$ باشد. (بسته بودن نسبت به اعمال جمع و ضرب)

۲-برای هر اسکالر $\alpha, \beta, \gamma \in F$ نتیجه می شود:

الف-قوانین جابجایی پذیری $\alpha + \beta = \beta + \alpha \;\;,\;\; \alpha\beta = \beta\alpha$

ب-قوانین شرکت پذیری $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma) \;\;,\;\; (\alpha\beta)\gamma = \alpha(\beta\gamma)$

---

[۱] field [۲] vector space

ج- قوانین توزیع پذیری $\alpha(\beta+\gamma) = (\alpha\beta) + (\alpha\gamma)$

۳-F دارای یک عنصر ۰ و یک عنصر ۱ است به گونه ای که به ازای هر $\alpha \in F$،

$$\alpha + 0 = 0 + \alpha = \alpha \ , \ \alpha \times 1 = 1 \times \alpha = \alpha$$

(عضو خنثی نسبت به عمل جمع و ضرب)

۴-برای هر عضو $\alpha \in F$ یک عنصر منحصر به فرد $(-\alpha) \in F$ موجود باشد به گونه ای که

$$\alpha + (-\alpha) = (-\alpha) + \alpha = 0$$

(عنصر قرینه نسبت به عمل جمع)

۵-برای هر $\alpha \in F$، عنصرغیر صفر و منحصر به فرد $\alpha^{-1} \in F$ موجود باشد به طوریکه:

$$\alpha \times \alpha^{-1} = \alpha^{-1} \times \alpha = 1$$

(عضو قرینه نسبت به عمل ضرب)

تعریف ۱-۱-۲. فضای برداری. یک فضای برداری مانند V بر روی میدان F ،مجموعه ای از بردارها است که با دو عمل جمع و ضرب شرایط زیر را برآورده می سازد.

۱-$\forall u, v \in V \to u + v \in V$

۲-$\forall u \in V, \forall c \in F \to cu \in V$

۳-$\forall u, v \in V \to u + v = v + u$

۴-$\forall u, v, w \in V \to u + (v + w) = (u + v) + w$

۵-$\forall u \in V, \exists 0 \in V \to u + 0 = 0 + u$

۶-$\forall u \in V, \exists (-u) \in V \to u + (-u) = (-u) + u = 0$

۷-$\forall u, v \in V, \forall a, b \in F \to (a + b)u = au + bu, a(u + v) = au + av$

۸-$\forall u \in V, \forall a, b \in F \to a(bu) = (ab)u$

۹-$\forall u \in V, \exists 1 \in F \to 1u = u$

تعریف ۱-۱-۳. ترکیب خطی. اگر $\alpha_1, \alpha_2, ..., \alpha_n \in F$, $v_1, v_2, ..., v_n \in V$ باشد،آنگاه

$$u = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n$$

را ترکیب خطی $v_1, v_2, ..., v_n$ می نامند.

تعریف ۱-۱-۴. وابستگی و استقلال خطی. فرض کنیم $v_1, v_2, ..., v_n$ ، n بردار در فضای برداری V باشند، این بردارها وابسته خطی اند اگر n اسکالر مانند $c_1, c_2, ..., c_n$ که همه صفر نیستند وجود داشته باشند به طوریکه

$$c_1 v_1 + c_2 v_2 + ... + c_n v_n = 0$$

اگر بردارها وابسته خطی نباشند ،آنها را مستقل خطی می نامند. به بیان دیگر ،$v_1, v_2, ..., v_n$ مستقل خطی اند اگر معادله

$$c_1 v_1 + c_2 v_2 + ... + c_n v_n = 0$$

فقط به ازای $c_1 = c_2 = ... = c_n = 0$ برقرار باشد.

قضیه ۱-۱-۵. اگر $n > m$ باشد دستگاه همگن بی نهایت جواب دارد.

برهان. [۲]. □

قضیه ۱-۱-۶. یک مجموعه از n بردار در $\mathbb{R}^m$ همیشه وابسته خطی است اگر $n > m$ باشد.

برهان. فرض کنیم $v_1, v_2, ..., v_n$ ، n بردار در $\mathbb{R}^m$ باشند،و سعی می کنیم ثابتهای $c_1, c_2, ..., c_n$ را که همه صفر نیستند چنان بیابیم که:

$$c_1 v_1 + c_2 v_2 + ... + c_n v_n = 0 \qquad (۱-۱)$$

فرض کنیم $v_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, v_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, ..., v_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$. پس معادله (۱-۱) به صورت

زیر در می آید :

$$
\begin{aligned}
a_{11}c_1 + a_{12}c_2 + \ldots + a_{1n}c_n &= 0 \\
a_{21}c_1 + a_{22}c_2 + \ldots + a_{2n}c_n &= 0 \\
\vdots \qquad \vdots \qquad \vdots \qquad & \vdots \\
a_{m1}c_1 + a_{m2}c_2 + \ldots + a_{mn}c_n &= 0
\end{aligned}
$$

ولی دستگاه بالا بنا بر قضیه ۱-۱-۵ به ازای $n > m$ بی نهایت جواب دارد.

لذا اسکالرهایی چون $c_1, c_2, \ldots, c_n$ که همه صفر نیستند وجود دارند که در دستگاه معادلات بالا صدق می‌کنند. بردارهای $v_1, v_2, \ldots, v_n$ وابسته خطی می‌باشند[۲]. $\square$

ملاحظه ۱-۱-۷. یک مجموعه از بردارهای مستقل خطی در $\mathbb{R}^n$ شامل حداکثر n بردار است.

این نتیجه را می‌توان به صورت زیر نیز بیان کرد: هر گاه nتا n-بردار مستقل خطی داشته باشیم،آنگاه با افزودن بردارهایی به این بردارها مجموعه وابسته خطی خواهد شد.

قضیه ۱-۱-۸. فرض کنید A یک ماتریس $n \times n$ باشد.در این صورت ، $\det(A) \neq 0$ اگر و فقط اگر ستونهای A مستقل خطی باشند.

برهان. [۲]. $\square$

تعریف ۱-۱-۹. ماتریس سطری - پلکانی. ماتریس A را یک ماتریس سطری پلکانی گوییم،هر گاه شرایط زیر برای آن برقرار باشد:

- تمام سطر های صفر در پایین ماتریس قرار داشته باشد.
- اولین درایه غیر صفر هر سطر از سمت چپ ۱باشدکه به آن یک پیشرو می گویند.
- یک پیشرو هر سطر در سمت راست یک پیشرو سطر بالای آن قرار داشته باشد.

تعریف ۱-۱-۱۰. ماتریس سطری - پلکانی تحویل یافته. ماتریس A را یک ماتریس سطری -پلکانی تحویل یافته گوییم،هر گاه ماتریس A سطری -پلکانی باشدو بعلاوه یک پیشرو تنها درایه غیر صفر ستون خود باشد [۲].

تعریف ۱-۱-۱۱. اسپن. فضای برداری V به روی میدان F را در نظر بگیرید ، این فضا توسط بردارهای $v_1, v_2, ...v_n$ اسپن شده است اگر کلیه این بردارها متعلق به V بوده و بتوان هر بردار $u \in V$ را به صورت ترکیب خطی از این بردارها نو شت .همچنین فضای اسپن شده توسط برداهای $v_1, v_2, ..., v_n$ را به صورت $sp\{v_1, v_2, ..., v_n\}$ نمایش می دهند.

تعریف ۱-۱-۱۲. پایه. مجموعه بردارهای $v_1, v_2, ..., v_n$ در فضای برداری V، برای آن فضا تشکیل یک پایه می دهند اگر:

$$V = sp\{v_1, v_2, ...v_n\}$$

و $v_1, v_2, ..., v_n$ مستقل خطی باشند.

تعریف ۱-۱-۱۳. بعد [1]. تعداد بردارهای پایه در یک فضای برداری مانند V را بعد آن فضا می نامند و با نماد $\dim(V)$نشان می دهند.به عبارتی بعد یک فضا برابر با حداکثر تعداد بردارهای مستقل خطی در آن فضا است.بنابراین در یک فضای n بعدی حداکثر بردارهای مستقل خطی n عدد می باشد.

تعریف ۱-۱-۱۴. کهاد. اگر A یک ماتریس مربعی باشد، دترمینان ماتریسی مربعی و کوچکتری که از حذف یک یا چند سطر و ستون A بدست می‌آید را کهاد ماتریس A می‌نامند. اگر فقط سطر iام و ستونj ام از ماتریس A حذف شود و دترمینان گرفته شود، آنگاه کهاد مرتبه اول iام و jام به دست می‌آید. اگر دو سطر و دو ستون حذف گردد و دترمینان گرفته شود یکی از کهادهای مرتبه دوم حاصل خواهد شد. منظور از کهاد معمولاً کهاد مرتبه اول است.

تعریف ۱-۱-۱۵. رتبه. رتبه ماتریس $A_{n\times n}$ برابر با ماکزیمم تعداد ستون های (سطر های)مستقل خطی در آن ماتریس است،که با نماد $\mathrm{Rank}(A)$ نشان داده می شود.

برای بدست آوردن ستون های مستقل خطی یک ماتریس می توان از فرم سطری پلکانی کاهش یافته آن کمک گرفت .

ازآنجائیکه رتبه یک ماتریس بصورت بزرگترین درجه کلیه کهاد های غیر صفر آن ماتریس تعریف می شود، می‌توان نتیجه گرفت که رتبه یک ماتریس مربعی مانند $A_{n\times n}$ حداکثر می تواند برابر n باشد و این زمانی است که تمامی ستون های (سطر های)ماتریس مستقل خطی باشند و در اینصورت $|A| \neq 0$ یعنی ماتریس $A_{n\times n}$ غیر منفرد است.در چنین حالتی ماتریس $A_{n\times n}$ را رتبه کامل [2] می نامند و اگر $|A| = 0$ باشد ماتریس منفرد بوده و تعدادی از ستون های آن وابستگی خطی دارند، چنین ماتریسی نقص رتبه [3] دارد.

[1] dimension [2] Full Rank [3] Rank Deficiency

برای ماتریس های $A_{n\times m}$ غیر مربعی ، $\mathrm{Rank}(A)\leqslant min(m,n)$ است ،که در صورت مساوی بودن می گوئیم ماتریس $A_{m\times m}$ رتبه کامل است و اگر کوچکتر باشد ماتریس $A_{m\times n}$ نقص رتبه دارد.

نکته:ضرب یک ماتریس غیر منفرد در ماتریس $A_{m\times n}$ رتبه آن را تغییر نمی دهد.

مثال ۱-۱-۱۶. رتبه ماتریس زیر را بدست آورید.

$$A=\begin{bmatrix} ۱ & ۳ & -۵ & ۱ & ۵ \\ ۱ & ۴ & -۷ & ۳ & -۲ \\ ۱ & ۵ & -۹ & ۵ & -۹ \\ ۰ & ۳ & -۶ & ۲ & -۱ \end{bmatrix}$$

ماتریس سطر ی پلکانی A را محاسبه کرده و با توجه به محل عناصر محوری درک می شود که ستون های اول ،دوم و چهارم مستقل خطی هستند.

$$A_{۴\times ۵}=\begin{bmatrix} ۱ & ۳ & -۵ & ۱ & ۵ \\ ۱ & ۴ & -۷ & ۳ & -۲ \\ ۱ & ۵ & -۹ & ۵ & -۹ \\ ۰ & ۳ & -۶ & ۲ & -۱ \end{bmatrix} \rightarrow \begin{bmatrix} ۱ & ۰ & ۱ & ۰ & ۱ \\ ۰ & ۱ & -۲ & ۰ & ۳ \\ ۰ & ۰ & ۰ & ۱ & -۵ \\ ۰ & ۰ & ۰ & ۰ & ۰ \end{bmatrix}$$

و چون ماتریس A سه ستون مستقل خطی دارد،لذا $\mathrm{Rank}(A)=۳$ است و ماتریس A نقص رتبه دارد.

تعریف ۱-۱-۱۷. فضای گستره [۱]. مجموعه ای است شامل عناصر b در فضای m بعدی $V_۲$ که برای آنها حداقل یک بردار مانند x در فضای n بعدی $V_۱$ وجود دارد،که رابطه $Ax=b$ را برآورده می سازد و آن را با نماد$R(A)$ نشان می دهند.

$$R(A)=\{b\in V_۲|\exists x\in V_۱\rightarrow Ax=b\} \qquad (۲-۱)$$

ملاحظه ۱-۱-۱۸. فضای گستره یک ماتریس همان فضای ستون های ماتریس است .

ملاحظه ۱-۱-۱۹. رتبه یک ماتریس معادل بعد فضای گستره آن ماتریس است.

$$\dim[R(A)]=\mathrm{Rank}(A)$$

[۱] Range space

مثال ۱-۱-۲۰. فضای گستره ماتریس زیر را بدست آورید.

$$A_{۴\times ۵} = \begin{bmatrix} ۱ & ۳ & -۵ & ۱ & ۵ \\ ۱ & ۴ & -۷ & ۳ & -۲ \\ ۱ & ۵ & -۹ & ۵ & -۹ \\ ۰ & ۳ & -۶ & ۲ & -۱ \end{bmatrix} \to \begin{bmatrix} ۱ & ۰ & ۱ & ۰ & ۱ \\ ۰ & ۱ & -۲ & ۰ & ۳ \\ ۰ & ۰ & ۰ & ۱ & -۵ \\ ۰ & ۰ & ۰ & ۰ & ۰ \end{bmatrix}$$

می دانیم فضای گستره ماتریس A کلیه ترکیبهای خطی ممکن ستون های A است.از آنجائیکه ستون های سوم و پنجم به ستون های اول ،دوم و چهارم وابسته می باشند،لذا $R(A)$ بصورت زیر تعریف می شود.

$$R(A) = sp\left\{\begin{bmatrix} ۱ \\ ۱ \\ ۱ \\ ۰ \end{bmatrix}, \begin{bmatrix} ۳ \\ ۴ \\ ۵ \\ ۳ \end{bmatrix}, \begin{bmatrix} ۱ \\ ۳ \\ ۵ \\ ۲ \end{bmatrix}\right\} \to dim[R(A)] = ۳$$

به عبارتی $R(A)$ برابر است با تمامی ترکیبهای خطی ستون های اول،دوم و چهارم ماتریس A . همچنین اگر فرم سطری پلکانی کاهش یافته ماتریس A را به دست آوریم ،با توجه به محل عناصر محوری می توان فهمید که ستون های اول ،دوم و چهارم مستقل خطی هستند.

۱-۱-۲ زیر فضا های اساسی ماتریس ها

برای یک ماتریس $A_{m\times n}$ با رتبه $r(r \leq min(m,n))$ چهار زیر فضای اساسی [۱] فضای پوچ، فضای ستون ها، فضای سطر ها و فضای پوچ چپ تعریف می شود. در ادامه به توضیح آنها پرداخته خواهد شد.

تعریف ۱-۱-۲۱. فضای پوچ. یک نگاشت خطی $A_{m\times n}$ مجموعه ای است شامل کلیه بردارهای $x_{n\times ۱}$ که رابطه $Ax = ۰$ را برآورده سازد. فضای پوچ با نماد $N(A)$ نشان داده می شود،

$$N(A) = \{x \in \mathbb{R}^n \to Ax = ۰\} \qquad (۳-۱)$$

[۱]Four fundamental subspace

بعد فضای پوچ را پوچی ماتریس A می نامند و با نماد $\text{nul}(.)$ یا $\nu(A)$ نشان می دهند.

$$\dim[N(A)] = \nu(A)$$

ملاحظه ۱-۱-۲۲. فضای پوچ $N(A)$ مجموعه تمامی پاسخ های معادله $Ax = 0$ است.

ملاحظه ۱-۱-۲۳. در صورتیکه تنها پاسخ معادله $Ax = 0$ همان پاسخ بدیهی(بردار صفر)باشد،بنابراین رتبه ماتریس A کامل است ،به عبارتی کلیه بردارهای ستونی (سطری)این ماتریس مستقل خطی هستند.

ملاحظه ۱-۱-۲۴. برای ماتریس $A_{m\times n}$ می توان نوشت ،

$$\text{Rank}(A) + \text{nul}(A) = n \tag{۱-۴}$$

پس بعد فضای پوچ برابر با $n - r$ است.این فضا مجموعه ای از $x \in \mathbb{R}^n$ است که عمود بر تمامی سطرهای ماتریس A (یا ستون های ماتریس A^T) هستند و از این جهت آن را مکمل متعامد فضای $R(A)$ یا کرنل [۱] ماتریس A نیز می نامند.

بعد فضای پوچ برابر تعداد ستون ها منهای رتبه ماتریس است.

$$\dim[N(A)] = \text{nul}(A) = n - \text{Rank}(A)$$

تعریف ۱-۱-۲۵. فضای ستون ها [۲].

در واقع همان فضای گستره ماتریس A یا$R(A)$ می باشد،که بعد آن برابر r است .این فضا مجموعه ای از ترکیبهای خطی ستون های ماتریس A است،به عبارتی توسط ستون های ماتریس A اسپن می شود.در فرم سطری پلکانی کاهشی ماتریس A ستونهایی که عناصر محوری در آن قرار دارند مطابق با بردارهای پایه فضای ستون ها خواهد بود .بعد فضای ستون ها برابر با رتبه ماتریس A می باشد.

$$R(A) = \{b \in \mathbb{R}^m | \exists x \in \mathbb{R}^n \to Ax = b\},$$

$$\dim[R(A)] = \text{Rank}(A)$$

[۱] kernel [۲] Column space

تعریف ۱-۱-۲۶. فضای سطرها [1]. فضای سطرها برای ماتریس A همان فضای گستره A^T است که با نماد $R(A^T)$ نشان داده می شود و بعد آن برابر r می باشد .فضای سطرها در واقع زیر فضایی است که توسط سطرهای ماتریس A اسپن شده یا به عبارتی شامل کلیه ترکیب های خطی سطرهای ماتریس A می باشد .در فرم سطری پلکانی کاهشی ماتریسA سطرهای غیر صفر معادل با بردارهای پایه برای فضای سطرها بوده که بعد فضای سطرها برابر با رتبه ماتریس A می باشد .

$$R(A^T) = \{b \in \mathbb{R}^n | \exists x \in \mathbb{R}^n \to A^T x = b\}, \tag{۱-۵}$$

$$\dim[R(A^T)] = \mathrm{Rank}(A)$$

تعریف ۱-۱-۲۷. فضای پوچ چپ [2]: در واقع همان فضای پوچ ماتریس A^T است که آن را با نماد $N(A^T)$ نشان می دهند و بعد آن برابر $m - r$ است .این فضا مجموعه ای از $x \in R^m$ است که عمود بر تمامی ستون های ماتریس A (سطر های ماتریس A^T) هستند و از این جهت آن را مکمل متعامد فضای $R(A)$ می نامند و با نماد $R(A)^\perp$نیز نشان داده می شود .بعد فضای پوچ چپ برابر با تعداد سطر ها منهای رتبه ماتریس است .

$$N(A^T) = \{x \in \mathbb{R}^m \to A^T x = ۰\}, \dim[N(A^T)] = m - \mathrm{Rank}(A) \tag{۱-۶}$$

حال بعد از بیان مقدمات اولیه به تعریف ماتریس خوش حالت خواهیم پرداخت. دستگاه معادلات خطی زیر را در نظر بگیرید:

$$A_{n\times n} x_{n\times ۱} = b_{n\times ۱}$$

با فرض اینکه A غیر منفرد باشد،می توان نوشت :

$$x = A^{-۱} b$$

حال اگر b شامل نویز یا خطای محاسباتی ناشی از گرد کردن مانند Δb باشد ،در این صورت این خطا بصورت زیر در پاسخ ظاهر خواهد شد،

$$x + \Delta x = A^{-۱}(b + \Delta b)$$

[1]Row space [2]Left Nullspace

لذا می توان نوشت ،

$$\Delta x = A^{-1}\Delta b$$

با توجه به خواص نرم ماتریس ها از این رابطه می توان نتیجه گرفت،

$$\|\Delta x\|_2 \leq \|A^{-1}\|_2\|\Delta b\|_2$$

از عبارت اخیر می توان تعبیر کرد که ،اگر $\|A^{-1}\|_2$ مقدار کوچکی داشته باشد ،برای تغییرات کم در b یعنی $\|\Delta b\|_2$ کوچک،مقدار $\|\Delta x\|$کم خواهدبود.ولی برای $\|A^{-1}\|_2$ های بزرگ،مقدار $\|\Delta x\|_2$ می تواند بزرگ باشد،حتی اگر $\|\Delta b\|$ مقدار کوچکی باشد .

لذا برای تشخیص خوش حالت بودن یک دستگاه معادلات ، پارامتری به نام عدد حالت [1] تعریف می گردد :[3]

$$\kappa = \|A\|_2\|A^{-1}\|_2 \qquad \kappa \geq 1$$

تعریف ۱-۱-۲۸. ماتریس خوش حالت. اگر عدد حالت کوچک باشد،بیان کننده آن است که ماتریس A و دستگاه معادلات حاصل خوش حالت(خوش وضع) [2] است و اگر عدد حالت مقدار خیلی بزرگی باشد،بیانگر آن است که ماتریس نزدیک به منفرد شدن است،لذا آن ماتریس را بد حالت (بد وضع) [3] می نامند و خطای محاسباتی در معکوس کردن ماتریس A زیاد است.

تعریف ۱-۱-۲۹. معین مثبت [4]. یک ماتریس متقارن حقیقی $A_{n\times n}$ مثبت معین نامیده می شود اگر برای هر $x \neq 0$ داشته باشیم:

$$x^T Ax > 0$$

تعریف ۱-۱-۳۰. یک دستگاه معادلات را سازگار می گوییم ،هر گاه حداقل یک جواب داشته باشدو ناسازگار گوییم ،هر گاه هیچ جوابی نداشته باشد. سازگار بودن سیستم را می توان به صورت زیر نیز بررسی کرد،که در اینجا b ماتریس افزوده A است.

$$\text{Rank}(A) = \text{Rank}(A|b)$$

[1] condition Number [2] Well conditioned [3] Ill conditioned [4] Positive Definite

۳-۱-۱ متعامد سازی

تعریف ۱-۱-۳۱. دو بردار

$$x = [x_1, x_2, ..., x_n]^T, y = [y_1, y_2, ..., y_n]^T$$

در $\mathbb{R}^n$ را در نظر بگیرید . ضرب داخلی x ، y به صورت زیر تعریف می شود:

$$< x, y >= \sum_{i=1}^{n} x_i y_i$$

ضرب داخلی بر روی یک ماتریس به صورت زیر بیان می شود،

$$< x, y >= y^T x = x^T y$$

اگر $x, x_1, x_2, y, y_1, v_2 \in \mathbb{R}^n$ و $a_1, a_2 \in \mathbb{R}$ باشد .آنگاه ویژگی های زیربرای ضرب داخلی برقرار است:

- $< x, y >=< y, x >$
- $< a_1 x_1 + a_2 x_2, y >= a_1 < x_1, y > +a_2 < x_2, y >$
- $< x, a_1 y_1 + a_2 y_2 >= a_1 < x, y_1 > +a_2 < x, y_2 >$
- $< x, x >= 0$ تساوی زمانی برقرار است که $x = 0$
- $||x||_2 = \sqrt{< x, x >}$

اگر x و y دو بردار غیر صفر باشند می توان زاویه بین دو بردار را از فرمول زیر بدست آورد:

$$cos(\theta) = \frac{< x, y >}{||x||_2 ||y||_2}$$

در شرایطی که $x = 0$ یا $y = 0$ باشد،آنگاه $\theta = \frac{\pi}{2}$ می شودکه در این صورت دو بردار x ، y را متعامدمی‌گویند.
واضح است که x ، y متعامد هستند اگر و تنها اگر $< x, y >= 0$ [۴].

تعریف ۱-۱-۳۲. ماتریس $n \times n$ ، A را متعامد می گوییم، هر گاه

$$A^T A = AA^T = I_n \qquad A^{-1} = A^T$$

تعریف ۱-۱-۳۳. اگر A یک ماتریس متعامد باشد، آنگاه $\det A = 1$.

تعریف ۱-۱-۳۴. بردار متعامد. بردار u را بر زیرفضای V_1 متعامد گویند ،اگر بردار u بر هر بردار در زیر فضای V_1 متعامد باشد .

دو زیر فضای V_1 و V_2 را متعامد گویند،اگر هر بردار در زیر فضای V_1 با هر بردار در زیر فضای V_2 متعامد باشد.

تعریف ۱-۱-۳۵. مکمل متعامد [1]. به مجموعه ای که شامل تمامی بردارهای متعامد بر زیر فضای V_1 باشد، مکمل متعامد زیر فضای V_1 گویند و آن را با نماد $V_1^{\perp}$ نشان می دهند .

فرض کنید V را فضای سه بعدی xyz در نظر بگیریم و w را زیر فضای xy، در این حالت محور z مکمل متعامد صفحه xy یعنی $w^{\perp}$ خواهد بود. همه بردار های موجود روی محور z بر هر کدام از بردارهای درون صفحه xy عمود است .

مثال ۱-۱-۳۶. ثابت کنید تمامی بردارهای یک مجموعه متعامد، مستقل خطی هستند.

یک مجموعه متعامد با بردارهای غیر صفر $v_1, v_2, ..., v_m$ را در نظر بگیرید. با استفاده از اسکالرهای $c_1, c_2, ..., c_m \in \mathbb{R}$ یک ترکیب خطی از این بردارها را بصورت زیر می توان نوشت.

$$c_1 v_1 + c_2 v_2 + ... + c_m v_m = 0$$

برای هر $i = 1, 2, ..., m$ داریم :

$$0 =< 0, v_i >=< c_1 v_1 + c_2 v_2 + ... + c_m v_m, v_i >$$

$$= c_1 < v_1, v_i > + c_2 < v_2, v_i > + ... + c_m < v_m, v_i >$$

$$= c_i < v_i, v_i >$$

از آنجائیکه بردارها متعامد هستند. بنابراین برای $i \neq j$ داریم $< v_i, v_j >= 0$ و از آنجائیکه $v_i \neq 0$ است . پس

[1] orthogonal complement

$0 \neq < v_i, v_i >$ لذا باید $c_i = 0$ باشد.بدین ترتیب نشان دادیم که ،

$$c_1 = c_2 = \ldots = c_m = 0$$

پس بردارهای $v_1, v_2, \ldots, v_m$ مستقل خطی هستند.

۱-۲ روش گرام اشمیت [۱]

جرج پترسون ارهام گرام [۲] دبیر ریاضیات بود و در دانمارک زندگی می کرد.مقالات مهم او عبارتند از: بسط سری با استفاده از روش حداقل مربعات و تحقیق در زمینه اعداد اول کمتر از یک .در فرایند گرام اشمیت مراحل یکامتعامد سازی بردارها بررسی می شود.

۱-۲-۱ روش یکامتعامد سازی گرام-اشمیت

تعریف ۱-۲-۱. فرض کنید بردارهای $v_1, v_2, \ldots, v_n$ بردارهای پایه فضای برداری n بعدی V_1 هستند. اگر بردارهای پایه $v_1, v_2, \ldots, v_n$ متعامد باشند به این مجموعه ، پایه های متعامد گویند و هر بردار مانند u متعلق به فضا ی برداری n بعدی V_1 را می توان بصورت ترکیب خطی زیر نمایش داد [۵].

$$u = \frac{\langle u, v_1\rangle}{||v_1||^2}v_1 + \frac{\langle u, v_2\rangle}{||v_2||^2}v_2 + \cdots + \frac{\langle u, v_n\rangle}{||v_n||^2}v_n \qquad (۷-۱)$$

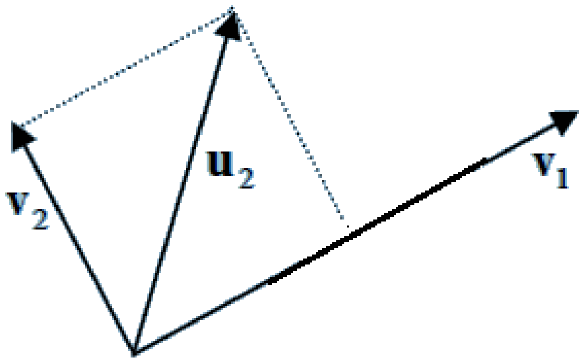
اگر بردارهای پایه $v_1, v_2, \ldots, v_n$ یکا متعامد باشند به آن پایه های یکامتعامد گویند.در اینصورت هر بردار مانند u متعلق به فضای برداری V_1 را می توان بصورت ترکیب خطی زیر نمایش داد.

$$u = \langle u, v_1\rangle v_1 + \langle u, v_2\rangle v_2 + \cdots + \langle u, v_n\rangle v_n \qquad (۸-۱)$$

بطور نمونه بردارهای پایه استاندارد $e_1, e_2, \ldots, e_n$ در فضای برداری $\mathbb{R}^n$ تشکیل یک مجموعه پایه های یکامتعامد را می دهند.

ایده اساسی بکار گرفته شده در فرآیند یکتا متعامد سازی گرام اشمیت را می توان بصورت زیر خلاصه کرد.

[۱] The Gram-Schmidt Process [۲] Gorgen pedersen Gram



شکل ۱-۱: تعبیر هندسی دو بردارعمود بر هم

دو بردار مخالف صفر v_1 و u_2 را در فضای برداری V_1 در نظر بگیرید که لزوما متعامد نیستند.هدف این است که با زدودن برخی از بردارهایی به شکل a_1v_1 از بردا ر u_2 آن را به یک بردار v_2 بصورت $v_2 = u_2 - a_1v_1$ تبدیل کنیم. به نحوی که بردارهای v_1 و v_2 متعامد باشند .تعبیر هندسی مسئله را می توانید در شکل۱-۱مشاهده کنید.به عبارتی ما به دنبال یافتن اعداد حقیقی مناسبی مانند a_1 هستیم بطوریکه شرط زیر برقرار گردد:

$$\langle v_1, v_2 \rangle = \langle v_1, u_2 - a_1v_1 \rangle = 0$$

از این رو می توان نوشت :

$$\langle v_1, u_2 \rangle - a_1\langle v_1, v_1 \rangle = 0 \rightarrow a_1 = \frac{\langle v_1, u_2 \rangle}{\langle v_1, v_1 \rangle} = \frac{\langle v_1, u_2 \rangle}{||v_1||^2}$$

بنابراین انتخابa_1 بصورت بالا مناسب خواهد بود و در نتیجه دوبردار v_1وv_2 متعامد خواهند بود :

$$v_2 = u_2 - \frac{\langle v_1, u_2 \rangle}{||v_1||^2}v_1$$

در حالت کلی بردارهای غیر صفر $v_1, v_2, \ldots, v_n$ و بردار غیر صفر u_{m+1} را در فضای برداری V_1 در نظر بگیرید، می خواهیم یک ترکیب خطی بصورت $a_1v_1 + a_2v_2 + \ldots + a_mv_m$ بیابیم به طوریکه بردارv_{m+1} که بصورت زیر تعریف می گردد بر هر یک از بردارهای $v_1, v_2, \ldots, v_m$ عمود باشد.

$$v_{m+1} = u_{m+1} - a_1v_1 - a_2v_2 - \cdots - a_mv_m \qquad (۱-۹)$$

به عبارتی در اینجا باید اسکالرهای حقیقی مناسب $a_1, a_2, ..., a_m$ را بیابیم به طوریکه شرط زیر برقرار باشد.

$$\langle v_i, v_{m+1}\rangle = \langle v_i, u_{m+1} - a_1 v_1 - a_2 v_2 - \cdots - a_m v_m\rangle = 0 \quad i = 1, 2, ..., m$$

بدین ترتیب رابطه بالا به شکل زیر قابل بیان است.

$$\langle v_i, u_{m+1}\rangle - a_1\langle v_i, v_1\rangle - a_2\langle v_i, v_2\rangle - \cdots - a_m\langle v_i, v_m\rangle = 0$$

بنابر این هر یک از a_i ها بصورت زیر تعریف می شوند.

$$a_i = \frac{\langle v_i, u_{m+1}\rangle}{\langle v_i, v_i\rangle} = \frac{\langle v_i, u_{m+1}\rangle}{||v_i||^2} \tag{۱۰-۱}$$

به این ترتیب بردار v_{m+1} با تعریف زیر بر هر یک از بردارهای $v_1, v_2, ...v_m$ عمود خواهد بود .

$$v_{m+1} = u_{m+1} - \frac{\langle v_1, u_{m+1}\rangle}{||v_1||^2}v_1 - \frac{\langle v_2, u_{m+1}\rangle}{||v_2||^2}v_2 - \cdots - \frac{\langle v_m, u_{m+1}\rangle}{||v_m||^2}v_m \tag{۱۱-۱}$$

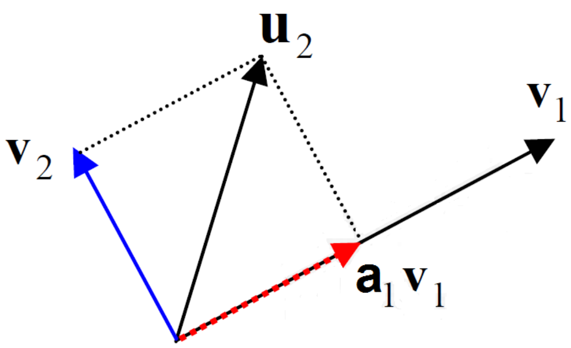
۱-۲-۲ تصاویر متعامد

فرض کنید V_2 یک زیر فضای برداری از فضای V_1 باشد.در اینصورت برای هر بردار مانند u در فضای برداری V_1 می توان عبارت زیر را نوشت:

$$u = proj_{V_2}u + proj_{V_2^\perp}u \tag{۱۲-۱}$$

که درآن $proj_{V_2}u$ یک بردار در زیر فضای V_2 است ،که به آن تصویر متعامد [1] بردار u بر روی V_2 گفته می شود و همینطور $proj_{V_2^\perp}u$ یک بردار در $V_2^\perp$ (مکمل متعامد V_2)است ،که به آن مولفه عمودی [2] بردار u عمود بر V_2 می گویند . یک تعبیر هندسی ساده از این گفته در شکل ۱-۲ نشان داده شده است، بردار u_2 را می توان بصورت

[1]orthogonal projection [2]orthogonal component



شکل ۱-۲: نمایش هندسی تصویر متعامد یک بردار

مجموع مولفه های عمودی و افقی آن نسبت به بردار $v_۱$ نوشت :

$$u_۲ = \alpha_۱ v_۱ + v_۲ = proj_{v_۱} u_۲ + proj_{v_۱^{\perp}} u_۲ \quad (۱۳-۱)$$

از طرفی با توجه به فرآیند گرام- اشمیت می توان نوشت:

$$proj_{v_۱} u_۲ = \frac{\langle v_۱, u_۲ \rangle}{||v_۱||^۲} v_۱ \quad (۱۴-۱)$$

حال اگر زیر فضای $V_۲$ مجموعه ای از بردارهای پایه متعامد بصورت $v_۱, v_۲, ..., v_n$ داشته باشد ،در اینصورت رابطه (۱-۱۴) را بصورت زیر می توان نوشت،

$$proj_{v_۲} u = \frac{\langle v_۱, u \rangle}{||v_۱||^۲} v_۱ + \frac{\langle v_۲, u \rangle}{||v_۲||^۲} v_۲ + \cdots + \frac{\langle v_n, u \rangle}{||v_n||^۲} v_n \quad (۱۵-۱)$$

در واقع هر یک از اجزای عبارت سمت راست رابطه۱-۱۵تصویر بردار u بر روی یک یک بردارهای پایه متعامد $v_۱, v_۲, ..., v_n$ می باشد و مجموع این تصاویر تصویر بردار u بر روی فضای$V_۲$ را نتیجه می دهد.حال اگر بردارهای پایه $v_۱, v_۲, ..., v_n$ یکامتعامد باشند داریم ،

$$proj_{V_۲} u = \langle v_۱, u \rangle v_۱ + \langle v_۲, u \rangle v_۲ + ... + \langle v_n, u \rangle v_n \quad (۱۶-۱)$$

یکی از مهمترین کاربرد های تصاویر متعامد در حل دستگاه معادلات خطی ناسازگار است .از آنجائیکه دستگاه معادلات خطی ناسازگار جواب ندارد، این سوال در ذهن ایجاد می شود که چرا ما به این موضوع می پردازیم؟برای روشن شدن مطلب به دو مثال زیر توجه کنید.

مثال ۱-۲-۲. معادله خطی را بیابید که از چهار نقطه زیر عبور کند،

$$(۱, -۱), (۴, ۱۱), (-۱, -۹), (-۲, -۱۳)$$

فرم کلی معادله خط را بصورت $y = mx+n$ در نظر می گیریم .برای اینکه خط مذکور از این نقاط عبور کند باید مختصات نقاط درآن صدق نماید.با قرار دادن هر یک از نقاط بالا در معادله خط ،معادلات زیر بدست می‌آیند،

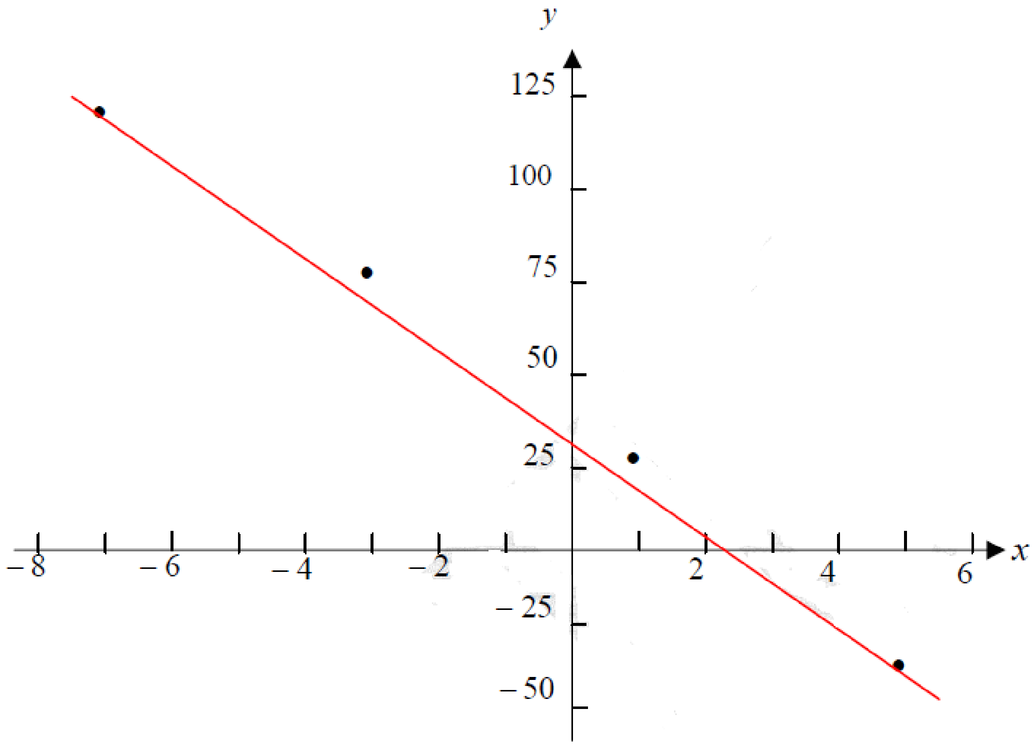
$$\begin{aligned} m+n &= -۱ \\ ۴m+n &= ۱۱ \\ -m+n &= -۹ \\ -۲m+n &= -۱۳ \end{aligned} \longrightarrow \begin{bmatrix} ۱ & ۱ \\ ۴ & ۱ \\ -۱ & ۱ \\ -۲ & ۱ \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} -۱ \\ ۱۱ \\ -۹ \\ -۱۳ \end{bmatrix}$$

با حل معادلات بالا جواب $m = ۴, n = -۵$ بدست می آید.بنابراین معادله خط مذکور بصورت $y = ۴x - ۵$ می باشد. لازم به ذکر است که این نمونه ای از یک دستگاه معادلات سازگار است .سازگار بودن سیستم را می‌توان بصورت زیر بررسی کرد،

$$Ax = b \rightarrow \mathrm{Rank}(A) = \mathrm{Rank}(A|b) = ۲$$

مثال ۱-۲-۳. معادله خطی را بیابید که از چهار نقطه زیر عبور کند،

$$(-۳, ۷۰), (۱, ۲۱), (-۷, ۱۱۰), (۵, -۳۵)$$



شکل ۱-۳: چهار نقطه از سیستم ناسازگار

همانند آنچه که در مثال قبل انجام شدبا قراردادن هر یک از نقاط در معادله خط،معادلات زیر بدست می آیند،

$$\begin{aligned} -۳m+n &= ۷۰ \\ m+n &= ۲۱ \\ -۷m+n &= ۱۱۰ \\ ۵m+n &= -۳۵ \end{aligned} \longrightarrow \begin{bmatrix} -۳ & ۱ \\ ۱ & ۱ \\ -۷ & ۱ \\ ۵ & ۱ \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} ۷۰ \\ ۲۱ \\ ۱۱۰ \\ -۳۵ \end{bmatrix}$$

از آنجائیکه این دستگاه معادلات ناسازگار است ،لذا پاسخی برای آن وجود ندارد.پس بر خلاف مثال قبل در این حالت نمی توان خطی را از این چهار نقطه عبور داد.ناسازگار بودن سیستم را می توان بررسی کرد،

$$Ax = b \rightarrow \mathrm{Rank}(A) = ۲, \mathrm{Rank}(A|b) = ۳$$

لذا $b \notin R(A)$ و سیستم ناسازگار است .برای اینکه این سیستم ناسازگار را بیشتر بررسی نماییم نمودار مختصات نقاط را رسم می نماییم، با توجه به شکل بالا چهار نقطه مذکور بر روی یک خط راست قرار ندارند و همانطور که گفته شد،دستگاه معادلات حاصل نیز ناسازگار می باشد. لیکن ممکن است این چهار نقطه از نتایج تجربی یک

سری آزمایشات بدست آمده و به خاطر برخی خطاهای فیزیکی و اندازه گیری از مقدار واقعی خود منحرف شده بر راستای یک خط راست قرار نگرفته باشند .در اینجا مسئله ای که مطرح می شود آن است که آیا می توان معادله خطی را بدست آورد که این چهار نقطه بطور تقریبی بر روی آن قرار گیرد؟برای یافتن جواب این سوال به بررسی مساله حداقل مربعات می پردازیم.

۱-۲-۳ مسئله حداقل مربعات [۱]

دستگاه معادلات خطی ناسازگار زیر را در نظر بگیرید،

$$Ax = b$$

چون سیستم ناسازگار است $b \notin R(A)$ و برای هیچ مقدار از x تساوی مذکور برقرار نیست.لذا داریم ،

$$\epsilon = b - Ax \tag{۱-۱۷}$$

که در آن ϵ بردار خطا [۲] می باشد،اندازه خطا با استفاده از نرم دو بصورت زیر تعریف می شود،

$$||\epsilon|| = ||b - Ax|| \tag{۱-۱۸}$$

برای یک سیستم ناسازگار $Ax = b$، هدف یافتن برداری مانند $\hat{x}$ است .بطوریکه خطای محاسبه شده بصورت $||\hat{\epsilon}|| = ||b - A\hat{x}||$ کوچکترین مقدار خطای ممکن باشد .در اینصورت بردار $\hat{x}$ را جواب حداقل مربعات می‌گویند.در واقع باید بردار $\hat{b} = A\hat{x}$ تا حد ممکن به بردار b شبیه باشد و مشخص است که باید $\hat{b} \in R(A)$ باشد.پس به دنبال بهترین تقریب برای بردار b در $R(A)$ هستیم.برای این منظور باید $\hat{b}$ رابصورت زیر انتخاب نماییم،

$$\hat{b} = proj_{R(A)} b \tag{۱-۱۹}$$

---
[۱] The least square problem [۲] Error vector

یعنی تصویر متعامد بردار b بر روی فضای گستره ماتریس A بهترین تقریب ،ممکن است.لذا برای حل مساله حداقل مربعات باید ابتدا $\hat{b} = proj_{R(A)}b$ را بیابیم و سپس معادله $\hat{b} = A\hat{x}$ را حل کنیم.پس با کمک گرفتن از مساله حداقل مربعات می توان برای یک سیستم ناسازگار با محاسبه مقدار خطای ممکن برای مساله جواب پیدا کنیم.

مساله حداقل مربعات خطا را می توان از روش های مختلفی حل نمود از جمله :

۱-با استفاده از تصاویر متعامد

۲-با استفاده از معادلات نرمال

۳- با استفاده از تجزیه QR

۴-با استفاده ار تجزیه چالسکی

در بخشهای بعد این روشها بیان خواهد شد.

۴-۲-۱ حل مساله کمترین مربعات با تصاویر متعامد

با ذکر یک مثال مراحل حل مساله حداقل مربعات با استفاده از تصاویر متعامد را بررسی می‌کنیم.

مثال ۴-۲-۱. برای سیستم ناسازگار زیر پاسخ حداقل مربعات را با استفاده از تصاویر متعامد بیابید.

$$Ax = b \Longrightarrow \begin{bmatrix} -۳ & ۱ \\ ۱ & ۱ \\ -۷ & ۱ \\ ۵ & ۱ \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} ۷۰ \\ ۲۱ \\ ۱۱۰ \\ -۳۵ \end{bmatrix} \Longrightarrow \begin{bmatrix} ۱ & ۰ \\ ۰ & ۱ \\ ۰ & ۰ \\ ۰ & ۰ \end{bmatrix} \begin{bmatrix} ۰ \\ ۰ \\ ۱ \\ ۰ \end{bmatrix}$$

ابتدا با استفاده از فرایند گرام -اشمیت پایه های یکامتعامدفضای گستره $R(A)$ را بدست می آوریم سپس تصویر متعامد بردار b را بر روی این پایه ها بررسی می کنیم.

$$R(A) = \left\{ \begin{bmatrix} -۳ \\ ۱ \\ -۷ \\ ۵ \end{bmatrix}, \begin{bmatrix} ۱ \\ ۱ \\ ۱ \\ ۱ \end{bmatrix} \right\}$$

$$v_۱ = a_۱ = [-۳, ۱, -۷, ۵] \Longrightarrow w_۱ = [\frac{-۳}{۲\sqrt{۲۱}}, \frac{۱}{۲\sqrt{۲۱}}, \frac{-۷}{۲\sqrt{۲۱}}, \frac{۵}{۲\sqrt{۲۱}}]$$

$$v_۲ = a_۲ - \frac{\langle v_۱, a_۲\rangle}{||v_۱||^۲} v_۱ = [۱, ۱, ۱, ۱] - \frac{-۴}{۸۴}[-۳, ۱, -۷, ۵] = [\frac{۱۸}{۲۱}, \frac{۲۲}{۲۱}, \frac{۱۴}{۲۱}, \frac{۲۶}{۲۱}]$$

$$w_۲ = [\frac{۱۸}{۴\sqrt{۱۰۵}}, \frac{۲۲}{۴\sqrt{۱۰۵}}, \frac{۱۴}{۴\sqrt{۱۰۵}}, \frac{۲۶}{۴\sqrt{۱۰۵}}]$$

حال تصویر بردار b را بر روی پایه های یکامتعامد $w_۱, w_۲$ بدست می آوریم،

$$\hat{b} = proj_{R(A)} b = \langle w_۱, b\rangle w_۱ + \langle w_۲, b\rangle w_۲$$

$$\hat{b} = (\frac{-۵۶۷}{\sqrt{۲۱}}) \begin{bmatrix} \frac{-۳}{۲\sqrt{۲۱}} \\ \frac{۱}{۲\sqrt{۲۱}} \\ \frac{-۷}{۲\sqrt{۲۱}} \\ \frac{۵}{۲\sqrt{۲۱}} \end{bmatrix} + (\frac{۵۸۸}{\sqrt{۱۰۵}}) \begin{bmatrix} \frac{۱۸}{۴\sqrt{۱۰۵}} \\ \frac{۲۲}{۴\sqrt{۱۰۵}} \\ \frac{۱۴}{۴\sqrt{۱۰۵}} \\ \frac{۲۶}{۴\sqrt{۱۰۵}} \end{bmatrix} = \begin{bmatrix} ۶۵٫۷ \\ ۱۷٫۳ \\ ۱۱۴٫۱ \\ -۳۱٫۱ \end{bmatrix}$$

حال با حل معادله $A\hat{x} = \hat{b}$ جواب مسئله حداقل مربعات را بدست می آوریم،

$$A\hat{x} = \hat{b} \Longrightarrow \begin{bmatrix} -۳ & ۱ \\ ۱ & ۱ \\ -۷ & ۱ \\ ۵ & ۱ \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} ۶۵٫۷ \\ ۱۷٫۳ \\ ۱۱۴٫۱ \\ -۳۱٫۱ \end{bmatrix} \Longrightarrow \begin{bmatrix} ۱ & ۰ \\ ۰ & ۱ \\ ۰ & ۰ \\ ۰ & ۰ \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} -۱۲٫۱ \\ ۲۹٫۴ \\ ۰ \\ ۰ \end{bmatrix}$$

بنابراین بهترین تقریب برای خطی که از چهار نقطه

$$(-۳, ۷۰), (۱, ۲۱), (-۷, ۱۱۰), (۵, -۳۵)$$

بگذردبصورت زیر می باشد،

$$y = mx + n = -۱۲٫۱x + ۲۹٫۴$$

۱-۲-۵ حل مساله کمترین مربعات بامعادلات نرمال

یکی از روش های حل مسئله حداقل مربعات با ماتریس رتبه کامل و خوش حالت A ،استفاده از معادلات نرمال است.

قضیه ۱-۲-۵. اگر بردار $\hat{x}$ جواب مسئله حداقل مربعات برای یک دستگاه معادلات خطی ناسازگار بصورت $Ax = b$ باشد ،می تواند جواب دستگاه معادلات خطی زیر نیز باشد،

$$A^T Ax = A^T b \tag{۱-۲۰}$$

که به آن معادلات نرمال [1] گفته می شود،از این رو اگر ماتریس A رتبه کامل و خوش حالت باشد،در اینصورت می توان یک پاسخ حداقل مربعات منحصر بفرد بصورت زیر بدست آورد،

$$\hat{x} = (A^T A)^{-۱} A^T b \tag{۱-۲۱}$$

این جوابی است که $||Ax - b||$ را مینیمم می نماید،

برهان. فرض کنید بردار $\hat{x}$ جواب مسئله حداقل مربعات باشد،لذا داریم :

$$A\hat{x} = proj_{R(A)} b$$

این رابطه را می توان بصورت زیر بازنویسی کرد

$$b - A\hat{x} = b - proj_{R(A)} b$$

می دانیم بردار $b - proj_{R(A)} b$ متعلق به زیر فضای $R(A)^{\perp}$ یا همان $N(A^T)$می باشد .

طبق تعریف $N(A^T)$ داریم ،

$$N(A^T) = \{x \in R^m \Longrightarrow A^T x = ۰\}$$

[1] Normal Equations

لذا بایدداشته باشیم

$$A^T(b - A\hat{x}) = A^T(b - proj_{R(A)}b) = ۰ \Longrightarrow A^Tb = A^TA\hat{x}$$

حال اگر ماتریس A رتبه کامل باشند ماتریس A^TA معکوس پذیر است .لذا جواب منحصر بفرد این معادله بصورت زیر بدست می آید،

$$\hat{x} = (A^TA)^{-۱}A^Tb$$

که همان جواب حداقل مربعات است .بنابراین می توان گفت که

$$\hat{x} = (A^TA)^{-۱}A^Tb$$

مقدار $||Ax - b||$ را مینیمم می کند. □

مثال ۱-۲-۶. حال دستگاه معادلات ناسازگار مثال قبل را با استفاده از معادلات نرمال حل می کنیم :

$$Ax = b \Longrightarrow \begin{bmatrix} -۳ & ۱ \\ ۱ & ۱ \\ -۷ & ۱ \\ ۵ & ۱ \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} ۷۰ \\ ۲۱ \\ ۱۱۰ \\ -۳۵ \end{bmatrix}$$

با توجه به معادلات نرمال داریم،

$$A^TA\hat{x} = A^Tb \Longrightarrow$$

$$\begin{bmatrix} -۳ & ۱ & -۷ & ۵ \\ ۱ & ۱ & ۱ & ۱ \end{bmatrix} \begin{bmatrix} -۳ & ۱ \\ ۱ & ۱ \\ -۷ & ۱ \\ ۵ & ۱ \end{bmatrix} \begin{bmatrix} \hat{m} \\ \hat{n} \end{bmatrix} = \begin{bmatrix} -۳ & ۱ & -۷ & ۵ \\ ۱ & ۱ & ۱ & ۱ \end{bmatrix} \begin{bmatrix} ۷۰ \\ ۲۱ \\ ۱۱۰ \\ -۳۵ \end{bmatrix}$$

$$\begin{bmatrix} ۸۴ & -۴ \\ -۴ & ۴ \end{bmatrix} \begin{bmatrix} \hat{m} \\ \hat{n} \end{bmatrix} = \begin{bmatrix} -۱۱۳۴ \\ ۱۶۶ \end{bmatrix}$$

$$\Longrightarrow \hat{m} = \frac{-۱۲۱}{۱۰} = -۱۲/۱, \hat{n} = \frac{۱۴۷}{۵} = ۲۹/۴$$

$$\hat{x} = \begin{bmatrix} \hat{m} \\ \hat{n} \end{bmatrix} = \begin{bmatrix} -۱۲/۱ \\ ۲۹/۴ \end{bmatrix}$$

بنابراین بهترین تقریب برای خطی که از چهار نقطه

$$(-۳, ۷۰), (۱, ۲۱), (-۷, ۱۱۰), (۵, -۳۵)$$

بگذرد بصورت زیر می باشد که همان نتیجه مثال قبل است،

$$y = -۱۲/۱x + ۲۹/۴$$

۱-۲-۶ حل مساله کمترین مربعات با تجزیه QR

در این روش ماتریس رتبه کامل $A_{m\times n}$ به حاصلضرب دو ماتریس $A = QR$ تجزیه می گردد،که در آن $Q_{m\times n}$ یک ماتریس متعامد و $R_{n\times n}$ یک ماتریس معکوس پذیر بالا مثلثی با عناصر قطری مثبت است ،با استفاده از این تجزیه معادلات نرمال پذیر را بصورت زیر می توان حل نمود،

$$A^T Ax = A^T b$$

$$(QR)^T QRx = (QR)^T b$$

$$R^T Q^T QRx = R^T Q^T b$$

از آنجائیکه $Q_{m\times n}$ یک ماتریس متعامد است $Q^T Q = I_n$ می باشد .لذا

$$R^T I_n Rx = R^T Q^T b \Longrightarrow R^T Rx = A^T b$$

از طرفی چون $R_{n\times n}$ یک ماتریس معکوس پذیر است ، R^T نیز معکوس پذیر می باشد،بنابراین داریم ،

$$Rx = Q^T b \qquad (۱-۲۲)$$

لذا برای بدست آوردن بردار x باید ابتدا تجزیه $A = QR$ را بدست آورده و سپس دستگاه معادلات زیر را حل نماییم،

$$\begin{cases} y = Q^T b \\ Rx = y \end{cases} \tag{۱-۲۳}$$

مثال ۱-۲-۷. برای دستگاه معادلات زیر مسئله حداقل مربعات را با استفاده از تجزیه QR حل نمایید.

$$۱/۵\begin{bmatrix} ۳ & -۶ & ۲۶ \\ ۴ & -۸ & -۷ \\ ۰ & ۴ & ۴ \\ ۰ & -۳ & -۳ \end{bmatrix}\begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} = \begin{bmatrix} ۱ \\ ۱ \\ ۱ \\ ۱ \end{bmatrix}$$

تجزیه $A = QR$ بفرم زیر می باشد،

$$A = QR \Longrightarrow \begin{bmatrix} \frac{۳}{۵} & \frac{-۶}{۵} & \frac{۲۶}{۵} \\ \frac{۴}{۵} & \frac{-۸}{۵} & \frac{-۷}{۵} \\ ۰ & \frac{۴}{۵} & \frac{۴}{۵} \\ ۰ & \frac{-۳}{۵} & \frac{-۳}{۵} \end{bmatrix} = \begin{bmatrix} \frac{۳}{۵} & ۰ & \frac{۴}{۵} \\ \frac{۴}{۵} & ۰ & \frac{-۳}{۵} \\ ۰ & \frac{۴}{۵} & ۰ \\ ۰ & \frac{-۳}{۵} & ۰ \end{bmatrix}\begin{bmatrix} ۱ & -۲ & ۲ \\ ۰ & ۱ & ۱ \\ ۰ & ۰ & ۵ \end{bmatrix}$$

بردار $y = Q^T b$ بصورت زیر می باشد

$$y = Q^T b = \begin{bmatrix} \frac{۳}{۵} & \frac{۴}{۵} & ۰ & ۰ \\ ۰ & ۰ & \frac{۴}{۵} & \frac{-۳}{۵} \\ \frac{۴}{۵} & \frac{-۳}{۵} & ۰ & ۰ \end{bmatrix}\begin{bmatrix} ۱ \\ ۱ \\ ۱ \\ ۱ \end{bmatrix} = \begin{bmatrix} \frac{۷}{۵} \\ \frac{۱}{۵} \\ \frac{۱}{۵} \end{bmatrix}$$

حال با حل معادله $Rx = y$ مقدار x را بدست می آوریم،

$$Rx = Q^T b \Longrightarrow \begin{bmatrix} ۱ & -۲ & ۲ \\ ۰ & ۱ & ۱ \\ ۰ & ۰ & ۵ \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} = \begin{bmatrix} \frac{۷}{۵} \\ \frac{۱}{۵} \\ \frac{۱}{۵} \end{bmatrix}$$

لذا جواب حداقل مربعات بصورت $x = [\frac{۴۱}{۲۵}, \frac{۴}{۲۵}, \frac{۱}{۲۵}]$ بدست می آید.

۱-۲-۷ حل مساله کمترین مربعات با تجزیه چالسکی

در این بخش ضمن بیان الگوریتم چالسکی به ذکر روش استفاده از آن در حل مسئله کمترین مربعات خواهیم پرداخت.

۱-۲-۷-۱ آشنایی با تجزیه چالسکی

آندره لوئس چالسکی [۱] (۱۵ اوت ۱۸۷۵ تا ۳۱ اگوست ۱۹۱۸) ریاضیدان فرانسوی بود که در زمینه نقشه برداری فعالیت می کرد.او توانست برای توسعه در زمینه نقشه برداری از یک روش تجزیه ماتریسی بهره بگیرد.که بعد از مرگش یکی از همراهان وی بنام فرمانده بنوا نسخه های او را منتشر کرد و این روش را چالسکی نامید.حال به توضیح تجزیه چالسکی می پردازیم.

تعریف ۱-۲-۸. تجزیه چالسکی. تجزیه چالسکی یک ماتریس مثبت معین A را بصورت حاصلضرب دو ماتریس به شکل $A = LL^T$ تجزیه می‌کند ،به طوریکه L یک ماتریس پایین مثلثی با عناصر قطری مثبت باشد.

فرض کنید

$$A = \begin{bmatrix} \alpha_{۱۱} & \alpha_{۱۲} & \dots & \alpha_{۱n} \\ \alpha_{۲۱} & \alpha_{۲۲} & \dots & \alpha_{۲n} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{n۱} & \alpha_{n۲} & \dots & \alpha_{nn} \end{bmatrix}$$

می خواهیم ماتریس را به صورت زیر تجزیه کنیم:

[۱] Cholesky

$$A=\begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & nn \end{bmatrix} \begin{bmatrix} l_{11} & l_{12} & \dots & l_{1n} \\ 0 & l_{22} & \dots & l_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & l_{nn} \end{bmatrix}$$

در ادامه الگوریتم چالسکی بیان خواهد شد.که در آن شرایط به دست آوردن یک ماتریس بالا مثلثی بیان شده است.

۱-۲-۷-۲ الگوریتم چالسکی

الگوریتم ۱-۱ الگوریتم چالسکی

Require: A
Ensure: LL^T
  for k=1 to n **do**
    $L_{kk} = (a_{kk} - \sum_{i=1}^{k-1} l_{kk}^2)^{1/2}$
    for j=k+1 to n **do**
      $l_{jk} = (a_{jk} - \sum_{i=1}^{k-1} l_{jt} l_{tk} / l_{kk})$
    end for
  end for

در جدول مثال زیرمراحل به دست آوردن ماتریس L بالا مثلثی با استفاده از تجزیه چالسکی شرح داده شده است.

مثال:فرض کنید ماتریس A به صورت زیر تعریف شده است، تجزیه چالسکی آن را بدست آورید.

$$A=\begin{bmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 5 \end{bmatrix}$$

Entry	General formula	Output
l_{11}	$\sqrt{a_{11}}$	**2**
l_{21}	a_{21}/l_{11}	**1**
l_{31}	a_{31}/l_{11}	**-1**
l_{22}	$\sqrt{a_{22}-l_{21}^2}$	**3**
l_{32}	$(a_{32}-l_{21}l_{31})/l_{22}$	**1**
l_{33}	$\sqrt{a_{33}-l_{31}^2-l_{32}^2}$	$\sqrt{3}$

در نتیجه

$$L=\begin{bmatrix} ۲ & ۰ & ۰ \\ ۱ & ۳ & ۰ \\ -۱ & ۱ & \sqrt{۳} \end{bmatrix}$$

یکی از کاربردهای تجزیه چالسکی استفاده از آن در حل مساله حداقل مربعات می باشد که به آن می پردازیم.

۱-۲-۷-۳ کمترین مربعات با تجزیه چالسکی

معادلات نرمال را در نظر بگیرید:

$$A^TAx=A^Tb$$

می دانیم زمانیکه ماتریس A رتبه کامل باشد،ماتریس A^TA یک ماتریس مثبت معین است.برای نشان دادن این موضوع می توان از تعریف مثبت معین استفاده نمود.

$$X^T(A^TA)X=(AX)^T(AX)=\begin{Bmatrix} \|Ax\|^۲>۰, X\neq ۰ \\ \|Ax\|^۲=۰, X=۰ \end{Bmatrix}$$

لذا می توان تجزیه چالسکی A^TA را بدست آورد و با استفاده از آن جواب مساله حداقل مربعات را محاسبه کرد.فرض کنید:

$$A^TAx=A^Tb \quad \xrightarrow{c=A^TA} Cx=A^Tb \quad \xrightarrow{d=Cx} \quad d=A^Tb$$

با بدست آوردن تجزیه چالسکی ماتریس C بصورت $C=LL^T$ و حل معادلات زیر می توان x را به راحتی محاسبه نمود.

$$\left\{\begin{array}{c} Lz = d \\ L^T x = z \end{array}\right\}$$

مثال :برای دستگاه معادلات زیر مساله حداقل مربعات را با استفاده از تجزیه چالسکی حل نمایید:

$$A = ۱/۵\begin{bmatrix} ۳ & -۶ & ۲۶ \\ ۴ & -۸ & -۷ \\ ۰ & ۴ & ۴ \\ ۰ & -۳ & -۳ \end{bmatrix} \quad b = \begin{bmatrix} ۱ \\ ۱ \\ ۱ \\ ۱ \end{bmatrix}$$

$$Ax = ۱/۵\begin{bmatrix} ۳ & -۶ & ۲۶ \\ ۴ & -۸ & -۷ \\ ۰ & ۴ & ۴ \\ ۰ & -۳ & -۳ \end{bmatrix}\begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} = \begin{bmatrix} ۱ \\ ۱ \\ ۱ \\ ۱ \end{bmatrix} = b$$

از آنجایی که $Rank(A) = ۳, Rank(A|b) = ۴$ است ،لذا سیستم ناسازگار است .یک دستگاه را سازگار می‌گویند اگر رابطه زیر برای آن برقرار باشد.

$$Rank(A) = Rank(A|b)$$

ابتدا مقدارهای $d = A^T b, C = A^T A$ را به دست می آوریم

$$C = \begin{bmatrix} ۱ & -۲ & ۲ \\ -۲ & ۵ & -۳ \\ ۲ & -۳ & ۳۰ \end{bmatrix}$$

$$d = \begin{bmatrix} \frac{۷}{۵} \\ \frac{-۱۳}{۵} \\ ۴ \end{bmatrix}$$

سپس معادله نرمال را حل می نماییم

$$Cx = d \rightarrow \begin{bmatrix} ۱ & -۲ & ۲ \\ -۲ & ۵ & -۳ \\ ۲ & -۳ & ۳۰ \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} = \begin{bmatrix} \frac{۷}{۵} \\ \frac{-۱۳}{۵} \\ ۴ \end{bmatrix}$$

تجزیه چالسکی ماتریس C بصورت زیر می باشد:

$$C = LL^T = \begin{bmatrix} ۱ & ۰ & ۰ \\ -۲ & ۱ & ۰ \\ ۲ & ۱ & ۵ \end{bmatrix} \begin{bmatrix} ۱ & -۲ & ۲ \\ ۰ & ۱ & ۱ \\ ۰ & ۰ & ۵ \end{bmatrix}$$

حال باید دستگاه معادلات زیر را حل نماییم: $Lz = d$

$$\begin{bmatrix} ۱ & ۰ & ۰ \\ -۲ & ۱ & ۰ \\ ۲ & ۱ & ۵ \end{bmatrix} \begin{bmatrix} z_۱ \\ z_۲ \\ z_۳ \end{bmatrix} = \begin{bmatrix} \frac{۷}{۵} \\ \frac{-۱۳}{۵} \\ ۴ \end{bmatrix}$$

از این معادلات مقدار $[z_۱, z_۲, z_۳] = [\frac{۷}{۵}, \frac{۱}{۵}, \frac{۱}{۵}]$ بدست می آید و سرانجام با حل دستگاه معادلات آخر پاسخ مسئله حداقل مربعات محاسبه می گردد. $L^T x = z$

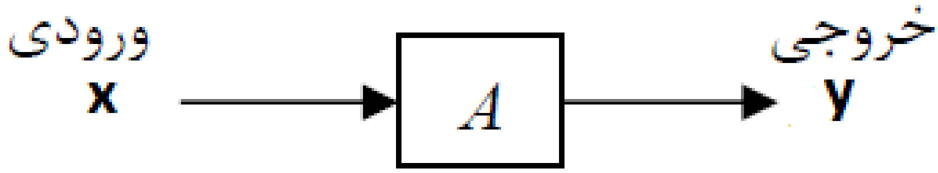
$$\begin{bmatrix} ۱ & -۲ & ۲ \\ ۰ & ۱ & ۱ \\ ۰ & ۰ & ۵ \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} = \begin{bmatrix} \frac{۷}{۵} \\ \frac{۱}{۵} \\ \frac{۱}{۵} \end{bmatrix}$$

و از اینجا $[x_۱, x_۲, x_۳] = [\frac{۴۱}{۲۵}, \frac{۴}{۲۵}, \frac{۱}{۲۵}]$ بدست می آید.

۳-۱ برازش داده با روش حداقل مربعات

شناسایی سیستم های فیزیکی از مباحث پر کاربرد در علوم تجربی و مهندسی است.در اینگونه مباحث طی آزمایشاتی داده هایی به عنوان ورودی و خروجی از سیستم مذکور بدست می آید و سعی می شود.بر اساس این

داده ها بهترین مدل ممکن برای سیستم تخمین زده شود و معمولا جهت سادگی در محاسبات مدل مذکور خطی در نظر گرفته می شود.سیستم زیر را در نظر بگیرید:



با فرض فوق رابطه بین داده های ورودی و خروجی را بصورت زیر می توان تعریف کرد،

$$Ax = y$$

هدف بدست آوردن مدلی برای سیستم است بطوریکه رابطه بالا برقرار گردد.در تخمین مدل سیستم انتخاب درجه مناسب برای تقریب در میزان دقت مدل تخمین زده شده تاثیر دارد.به مثال زیر توجه نمایید.

مثال ۱-۳-۱. پنج نقطه زیر را در نظر بگیرید:

$$(0,0)(5,8)(10,15)(15,19)(20,20)$$

با استفاده از روش حداقل مربعات

الف)معادله خطی به صورت $y = m_1x + m_2$ برای تقریب از نقاط بیابید .

ب) معادله منحنی مرتبه دومی به شکل $y = \alpha_2x^2 + \alpha_1x + \alpha_0$ را بیابید که از این پنج نقطه بگذرد .

ج) بردار خطا و نرم خطا رامحاسبه کنید و خطای برازش هر دو حالت را با هم مقایسه نمایید.

پاسخ: الف)ابتدا معادله خطی به فرم $y = m_1x + m_2$ را پیدا می کنیم که از این پنج نقطه بگذرد،

$$\begin{cases} 0m_1 + m_2 = 0 \\ 5m_1 + m_2 = 8 \\ 10m_1 + m_2 = 15 \\ 15m_1 + m_2 = 19 \\ 20m_1 + m_2 = 20 \end{cases} \Longrightarrow \begin{bmatrix} 0 & 1 \\ 5 & 1 \\ 10 & 1 \\ 15 & 1 \\ 20 & 1 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 15 \\ 19 \\ 20 \end{bmatrix} \Longrightarrow Am = y$$

با توجه به رابطه زیر

$$Rank(A) = 2, Rank(A|y) = 3$$

سیستم ناسازگار است .با استفاده از روش حداقل مربعات می توان معادله خط $y = m_1x + m_2$ را بدست آورد.

$$A^T A\hat{m} = A^T y \Longrightarrow \begin{bmatrix} 750 & 50 \\ 50 & 5 \end{bmatrix} \begin{bmatrix} \hat{m}_1 \\ \hat{m}_2 \end{bmatrix} = \begin{bmatrix} 875 \\ 62 \end{bmatrix} \Longrightarrow \hat{m}_1 = 1/02, \hat{m}_2 = 2/2$$

بنابراین بهترین تقریب برای خطی که ازاین چهار نقطه می گذرد بصورت $y = 1/02x + 2/2$ است. کد متلب مربوطه را در برنامه ۱-۱ ملاحظه می‌فرمایید.

برنامه ۱-۱: تقریب مناسب برای معادله خط

```
% define a continuous function
x=[0;5;10;15;20];
y=[0;8;15;19;20];
A=zeros(5,2);
for i=1:5
A(i,:)=[x(i) 1];
end
m=A\(A'\(A'*y))
% The output is:
m=1.0200
  2.2000
```

ب)حال معادله منحنی دوم به فرم $y = \alpha_2x^2 + \alpha_1x + \alpha_0$ را پیدا می کنیم که از این پنج نقطه بگذرد.

$$\begin{cases} \alpha_0 + 0\alpha_1 + 0\alpha_2 = 0 \\ \alpha_0 + 5\alpha_1 + 23\alpha_2 = 8 \\ \alpha_0 + 10\alpha_1 + 100\alpha_2 = 15 \\ \alpha_0 + 15\alpha_1 + 225\alpha_2 = 19 \\ \alpha_0 + 20\alpha_1 + 400\alpha_2 = 20 \end{cases} \Longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 1 & 5 & 25 \\ 1 & 10 & 100 \\ 1 & 15 & 225 \\ 1 & 20 & 400 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 15 \\ 19 \\ 20 \end{bmatrix} \Longrightarrow A\alpha = y$$

با توجه به رابطه

$$Rank(A) = 3, Rank(A|y) = 4$$

سیستم ناسازگار است.بنابراین با استفاده از روش حداقل مربعات بهترین تقریب برای منحنی مرتبه دوم

$y = \alpha_۲ x^۲ + \alpha_۱ x + \alpha_۰$ بصورت زیر است:

$$A^T A\hat{\alpha} = A^T y \Longrightarrow \begin{bmatrix} ۵ & ۵۰ & ۷۵۰ \\ ۵۰ & ۷۵۰ & ۱۲۵۰۰ \\ ۷۵۰ & ۱۲۵۰۰ & ۲۲۱۲۵۰ \end{bmatrix} \begin{bmatrix} \hat{\alpha_۰} \\ \hat{\alpha_۱} \\ \hat{\alpha_۲} \end{bmatrix} = \begin{bmatrix} ۶۲ \\ ۸۷۵ \\ ۱۳۹۷۵ \end{bmatrix}$$

$$\hat{\alpha_۰} = -۰/۲۲۸۶, \hat{\alpha_۱} = ۱/۹۹۱۴, \hat{\alpha_۲} = -۰/۰۴۸۶$$

$$y = -۰/۰۴۸۶x^۲ + ۱/۹۹۱۴x - ۰/۲۲۸۶$$

کد متلب مربوطه را در برنامه ۱-۲ ملاحظه می‌فرمایید.

برنامه ۱-۲: تقریب مناسب برای منحنی مرتبه دوم

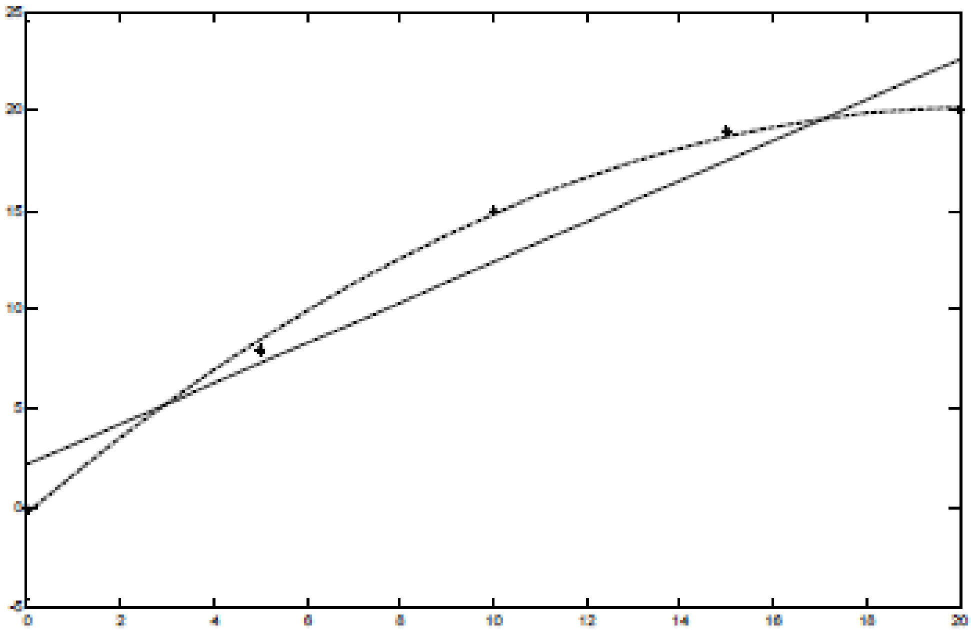
```
% define a continuous function
x=[0;5;10;15;20];
y=[0;8;15;19;20];
A=zeros(5,2);
for i=1:5
A(i,:)=[x(i)^2 ,x(i) , 1];
end
a=A\(A'\ (A'*y))
% The output is:
a=-0.0486
   1.9914
  -0.2286
```

نمودارهای خطی و منحنی بدست آمده به همراه نقاط مذکور در شکل ۱-۴ آورده شده است

ج)بردار خطا و نرم خطا نیز بصورت زیر بدست می آیند:

۱- محاسبه خطا برای تقریب خطی $y = ۱/۰۲x + ۲/۲$

$$\epsilon = b - Ax \Longrightarrow \begin{bmatrix} \epsilon_۱ \\ \epsilon_۲ \\ \epsilon_۳ \\ \epsilon_۴ \\ \epsilon_۵ \end{bmatrix} = \begin{bmatrix} ۰ \\ ۸ \\ ۱۵ \\ ۱۹ \\ ۲۰ \end{bmatrix} - \begin{bmatrix} ۰ & ۱ \\ ۵ & ۱ \\ ۱۰ & ۱ \\ ۱۵ & ۱ \\ ۲۰ & ۱ \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} ۰ - (۰m + n) \\ ۸ - (۵m + n) \\ ۱۵ - (۱۰m + n) \\ ۱۹ - (۲۰m + n) \end{bmatrix}$$



شکل ۱-۴: خط و منحنی بدست آمده از تخمین حداقل مربعات

$$\hat{x} = \begin{bmatrix} \hat{m} \\ \hat{n} \end{bmatrix} = \begin{bmatrix} ۱٫۰۲ \\ ۲٫۲ \end{bmatrix} \Longrightarrow \epsilon = \begin{bmatrix} \hat{\epsilon}_{۱} \\ \hat{\epsilon}_{۲} \\ \hat{\epsilon}_{۳} \\ \hat{\epsilon}_{۴} \\ \hat{\epsilon}_{۵} \end{bmatrix} = \begin{bmatrix} -۲٫۲ \\ ۰٫۷ \\ ۲٫۶ \\ ۱٫۵ \\ -۲٫۶ \end{bmatrix} \Longrightarrow \|\epsilon\| = ۴٫۵۹۳۵$$

کد متلب مربوطه را در برنامه ۱-۳ ملاحظه می‌فرمایید.

برنامه ۱-۳: تقریب نرم خطای خط

```
% define a continuous function
norm_e=norm(A*m-y)
norm_e=4.5935
```

۲- محاسبه خطا برای تقریب منحنی مرتبه دوم

$$y = -۰٫۰۴۸۶x^{۲} + ۱٫۹۹۱۴x - ۰٫۲۲۸۶$$

$$\epsilon = b - Ax$$

$$\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 15 \\ 19 \\ 20 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 1 & 5 & 25 \\ 1 & 10 & 100 \\ 1 & 15 & 225 \\ 1 & 20 & 400 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 - (\alpha_0 + 0\alpha_1 + 0\alpha_2) \\ 8 - (\alpha_0 + 5\alpha_1 + 23\alpha_2) \\ 15 - (\alpha_0 + 10\alpha_1 + 100\alpha_2) \\ 19 - (\alpha_0 + 15\alpha_1 + 225\alpha_2) \\ 20 - (\alpha_0 + 20\alpha_1 + 400\alpha_2) \end{bmatrix}$$

$$\hat{x} = \begin{bmatrix} \hat{\alpha}_0 \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{bmatrix} = \begin{bmatrix} -0/2286 \\ 1/9914 \\ -0/0486 \end{bmatrix} \Longrightarrow \epsilon = \begin{bmatrix} \hat{\epsilon}_1 \\ \hat{\epsilon}_2 \\ \hat{\epsilon}_3 \\ \hat{\epsilon}_4 \\ \hat{\epsilon}_5 \end{bmatrix} = \begin{bmatrix} 0/2286 \\ -0/5143 \\ 0/1714 \\ 0/2857 \\ -0/1714 \end{bmatrix} \Longrightarrow \|\epsilon\| = 0/6761$$

کد متلب مربوطه را در برنامه ۱-۴ ملاحظه می‌فرمایید.

برنامه ۱-۴: تقریب نرم خطای منحنی

```
% define a continuous function
% define a continuous function
norm_e=norm(A*a-y)
norm_e=0.6761
```

با توجه به اینکه نرم خطا در این حالت کمتر است لذا منحنی مرتبه دوم تقریب بهتری برای برازش این پنج نقطه است.

فصل ۲

تجزیه QR

در این بخش تجزیه QR یک ماتریس تعریف شده و روشهای مختلف محاسبه آن بیان خواهد شد.

تجزیه QR ماتریس $A_{m\times n}$ که $m \geq n$ است به صورت زیر می باشد.

$$A = QR$$

که در آن $Q \in \mathbb{R}^{m\times m}$ یک ماتریس متعامد و $R \in \mathbb{R}^{m\times n}$ یک ماتریس بالا مثلثی است.

تجزیه QR کاربرد های مختلفی دارد. اغلب روشهای عددی موجود برای یافتن پاسخ مساله حداقل مربعات سیستمهای ناسازگار مبتنی بر تجزیه متعامد ماتریس $A_{m\times n}$ بصورت $A = QR$ می باشد.

روش های مختلفی برای محاسبه تجزیه ی QR یک ماتریس وجود دارد .در این فصل سه روش مهم زیر را مورد مطالعه قرار می دهیم:

۱. تجزیه QR با استفاده از الگوریتم گرام -اشمیت

۲. تجزیه QR با استفاده ازماتریس های گیونز[1]

۳. تجزیه QR با استفاده از ماتریس های هوسهولدر [2]

[1]Givens [2]Householder

۱-۲ تجزیه QR با استفاده از الگوریتم گرام -اشمیت

برای به دست آوردن Q و R می توانیم از فرایند گرام اشمیت استفاده کنیم . ماتریس رتبه کامل

$$A_{m\times n} = [a_1|a_2|...|a_n]$$

را در نظر بگیرید ،از آنجائیکه رتبه ماتریس کامل است ،ستون های ماتریس $A_{m\times n}$ مستقل خطی می باشند ،لذا می توان آنها را به عنوان بردارهای پایه برای فضای گستره ماتریس $A_{m\times n}$ ،یعنی $R(A)$ در نظر گرفت . از طرفی با استفاده از فرایند گرام اشمیت می توان این بردارهای پایه را به بردارهای یکامتعامد تبدیل نمود.

$$v_1 = a_1$$

$$v_2 = a_2 - \frac{\langle v_1, a_2\rangle}{||v_1||^2}v_1$$

$$v_3 = a_3 - \frac{\langle v_1, a_3\rangle}{||v_1||^2}v_1 - \frac{\langle v_2, a_3\rangle}{||v_2||^2}v_2$$

$$\vdots$$

$$v_n = a_n - \frac{\langle v_1, a_n\rangle}{||v_1||^2}v_1 - \frac{\langle v_2, a_n\rangle}{||v_2||^2}v_2 - ... - \frac{\langle v_{n-1}, a_n\rangle}{||v_{n-1}||^2}v_{n-1}$$

با بازنویسی معادلات بالا داریم:

$$a_1 = v_1$$

$$a_2 = v_2 + \frac{\langle v_1, a_2\rangle}{||v_1||^2}v_1$$

$$a_3 = v_3 + \frac{\langle v_1, a_3\rangle}{||v_1||^2}v_1 + \frac{\langle v_2, a_3\rangle}{||v_2||^2}v_2$$

$$\vdots$$

$$a_n = v_n + \frac{\langle v_1, a_n\rangle}{||v_1||^2}v_1 + \frac{\langle v_2, a_n\rangle}{||v_2||^2}v_2 + ... + \frac{\langle v_{n-1}, a_n\rangle}{||v_{n-1}||^2}v_{n-1}$$

نمایش ماتریسی این معادلات به فرم زیر خواهد بود.

$$A=[v_1,v_2,...,v_n]\begin{bmatrix} 1 & \frac{\langle v_1,a_2\rangle}{||v_1||^2} & \frac{\langle v_1,a_3\rangle}{||v_1||^2} & \dots & \frac{\langle v_1,a_n\rangle}{||v_1||^2} \\ 0 & 1 & \frac{\langle v_2,a_3\rangle}{||v_2||^2} & \dots & \frac{\langle v_2,a_n\rangle}{||v_2||^2} \\ 0 & 0 & 1 & \dots & \frac{\langle v_3,a_n\rangle}{||v_3||^2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \frac{\langle v_{n-1},a_n\rangle}{||v_{n-1}||^2} \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

در این مرحله ماتریس $A_{m\times n}$ را بصورت حاصلضرب یک ماتریس با بردارهای ستونی متعامد در یک ماتریس بالا مثلثی با عناصر قطری یک نمایش دادیم .لذا برای بدست آوردن ماتریس $Q_{m\times n}$ کافی است که بردارهای ستونی متعامد را به بردارهای یکامتعامد تبدیل نماییم . به این ترتیب معادلات به فرم زیر در می آیند:

$$A=[q_1,q_2,...,q_n]\begin{bmatrix} ||v_1|| & \frac{\langle v_1,a_2\rangle}{||v_1||} & \frac{\langle v_1,a_3\rangle}{||v_1||} & \dots & \frac{\langle v_1,a_n\rangle}{||v_1||} \\ 0 & ||v_2|| & \frac{\langle v_2,a_3\rangle}{||v_2||} & \dots & \frac{\langle v_2,a_n\rangle}{||v_2||} \\ 0 & 0 & ||v_3|| & \dots & \frac{\langle v_3,a_n\rangle}{||v_3||} \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \dots & \frac{\langle v_{n-1},a_n\rangle}{||v_{n-1}||} \\ 0 & 0 & 0 & \dots & ||v_n|| \end{bmatrix}$$

که در آن $q_i=\frac{v_i}{||v_i||}$ می باشد. بنابراین توانستیم تجزیه $A=QR$ را بدست آوریم . که در آن ستون های ماتریس $Q_{m\times n}=[q_1|q_2|...|q_n]$ همان پایه های یکامتعامد شده می باشند و $R_{n\times n}$ یک ماتریس معکوس پذیر بالا مثلثی با عناصر قطری مثبت می باشد . حال می توان نشان داد که ماتریس بالا مثلثی $R_{n\times n}$ بصورت زیر قابل ساده سازی است

$$R_{n\times n}=\begin{bmatrix} ||v_1|| & \langle q_1,a_2\rangle & \langle q_1,a_3\rangle & \dots & \langle q_1,a_n\rangle \\ 0 & ||v_2|| & \langle q_2,a_3\rangle & \dots & \langle q_2,a_n\rangle \\ 0 & 0 & ||v_3|| & \dots & \langle q_3,a_n\rangle \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \dots & \langle q_{n-1},a_n\rangle \\ 0 & 0 & 0 & \dots & ||v_n|| \end{bmatrix}$$

مثال: برای ماتریس A معرفی شده در زیر تجزیه QR را با استفاده از فرایند گرام -اشمیت بدست آورید.

$$A = \begin{bmatrix} ۲ & ۱ & ۳ \\ -۱ & ۰ & ۷ \\ ۰ & -۱ & -۱ \end{bmatrix}$$

با توجه به اینکه $|A| \neq ۰$ است ، پس ستون های ماتریس A مستقل خطی هستند .اگر ماتریس A را بصورت $A = [a_۱, |a_۲|a_۳]$ در نظر بگیریم ، ستون های آن بصورت زیر بدست می آیند.

$$a_۱ = \begin{bmatrix} ۲ \\ -۱ \\ ۰ \end{bmatrix} a_۲ = \begin{bmatrix} ۱ \\ ۰ \\ -۱ \end{bmatrix} a_۳ = \begin{bmatrix} ۳ \\ ۷ \\ -۱ \end{bmatrix}$$

با اعمال فرآیند گرام- اشمیت بردارهای یکامتعامد زیر بدست می آیند .

$$v_۱ = \begin{bmatrix} ۲ \\ -۱ \\ ۰ \end{bmatrix} v_۲ = \begin{bmatrix} \frac{۱}{۵} \\ \frac{۲}{۵} \\ -۱ \end{bmatrix} v_۳ = \begin{bmatrix} \frac{۸}{۳} \\ \frac{۱۶}{۳} \\ \frac{۸}{۳} \end{bmatrix}$$

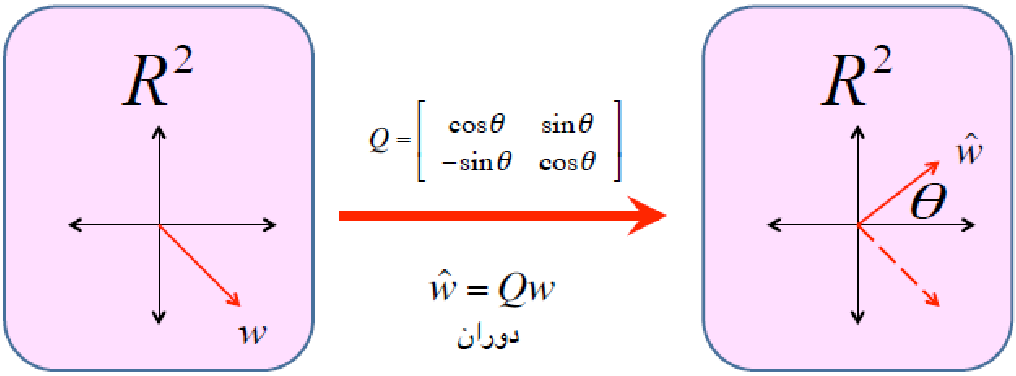
$$q_۱ = \begin{bmatrix} \frac{۲}{\sqrt{۵}} \\ \frac{-۱}{\sqrt{۵}} \\ ۰ \end{bmatrix} q_۲ = \begin{bmatrix} \frac{۱}{\sqrt{۳۰}} \\ \frac{۲}{\sqrt{۳۰}} \\ \frac{-۵}{\sqrt{۳۰}} \end{bmatrix} q_۳ = \begin{bmatrix} \frac{۱}{\sqrt{۶}} \\ \frac{۲}{\sqrt{۶}} \\ \frac{۱}{\sqrt{۶}} \end{bmatrix}$$

$$||v_۱|| = \sqrt{۵}, ||v_۲|| = \frac{\sqrt{۳۰}}{۵}, ||v_۳|| = \frac{۸\sqrt{۶}}{۳}$$

بنابراین،ماتریس Q بصورت$Q = [q_۱|q_۲|q_۳]$ بدست می آید .حال ماتریس R را بدست می آوریم.

$$R = \begin{bmatrix} ||v_۱|| & \langle q_۱, a_۲\rangle & \langle q_۱, a_۳\rangle \\ ۰ & ||v_۲|| & \langle q_۲, a_۳\rangle \\ ۰ & ۰ & ||v_۳|| \end{bmatrix} = \begin{bmatrix} \sqrt{۵} & \frac{۲}{\sqrt{۵}} & \frac{-۱}{\sqrt{۵}} \\ ۰ & \frac{۶}{\sqrt{۳۰}} & \frac{۲۲}{\sqrt{۳۰}} \\ ۰ & ۰ & \frac{۱۶}{\sqrt{۶}} \end{bmatrix}$$

بنابراین تجزیه QR ماتریس A بصورت زیر خواهد بود. $A = QR$



شکل ۲-۱: ماتریس دوران

$$A = \begin{bmatrix} \frac{۲}{\sqrt{۵}} & \frac{۱}{\sqrt{۳۰}} & \frac{۱}{\sqrt{۶}} \\ \frac{-۱}{\sqrt{۵}} & \frac{۲}{\sqrt{۳۰}} & \frac{۲}{\sqrt{۶}} \\ ۰ & \frac{-۵}{\sqrt{۳۰}} & \frac{۱}{\sqrt{۶}} \end{bmatrix} \begin{bmatrix} \sqrt{۵} & \frac{۲}{\sqrt{۵}} & \frac{-۱}{\sqrt{۵}} \\ ۰ & \frac{۶}{\sqrt{۳۰}} & \frac{۲۲}{\sqrt{۳۰}} \\ ۰ & ۰ & \frac{۱۶}{\sqrt{۶}} \end{bmatrix}$$

۲-۲ تجزیه QR با استفاده ازماتریس های گیونز

در ابتدا به توضیح مختصری درباره ماتریس های دوران که زیر بنای اصلی ماتریس های گیونز می باشند می‌پردازیم.

ماتریس دوران در جبر خطی، به ماتریسی گفته می شود که توسط آن میتوان اشکال را در فضای اقلیدسی دوران داد. برای مثال ماتریس

$$R = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

نقاط را در فضای xy دکارتی به میزان θ درجه حول نقطه مبدأ در جهت پاد ساعتگرد دوران می دهد. برای انجام عمل دوران با استفاده از یک ماتریس دوران (R) نیاز است تا مختصات نقطه به شکل برداری (v) مشخص باشد. برای بدست آوردن مختصات نقطه پس از عمل دوران کافی است که ضرب ماتریسی $R \times v$ انجام شود. از آنجایی که ضرب ماتریسی در بردار صفر (مختصات نقطه مبدا) تأثیری ندارد، از ماتریس های دوران تنها میتوان برای دوران نقاط حول نقطه مبدأ استفاده کرد. شکل ۳-۵ حاصل اعمال دوران را نشان می دهد.

تعریف ۲-۲-۱. ماتریس گیونز. ماتریس همانی $I_{n\times n}$ و اعداد حقیقی c و s را در نظر بگیرید به طوریکه

$$c^۲ + s^۲ = ۱$$

.همچنین فرض کنید $1 \leq i < j \leq n$.درایه های مکان $(i,i), (i,j), (j,i), (j,j)$ ازماتریس I را به ترتیب به $c, -s, s, c$ تبدیل می کنیم .ماتریس حاصل را با $J(i,j,c,s)$ نشان داده و یک ماتریس گیونز گویند.این گونه ماتریس ها به صورت زیرند

$$J(i,j,c,s) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & c & & s & & \\ & & & \ddots & & & \\ & & -s & & c & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \quad (1\text{-}2)$$

برای سادگی گاهی اوقات این ماتریس را با J_{ij} نشان می دهیم. ماتریس $J(i,j,c,s)$ متقارن نیست اما به سادگی می توان دید که متعامد است. بردار

$$x = (x_1, x_2, ..., x_i, ..., x_j, ..., x_n)^T$$

را در نظر بگیرید.مقادیر c و s را طوری می یابیم که

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix}^T \begin{pmatrix} x_i \\ x_j \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}$$

برای بر قراری رابطه اخیر کافی است داشته باشیم$-sx_i + cx_j = 0$.یا $c = \frac{sx_i}{x_j}$ با جایگذاری c در $c^2 + s^2 = 1$ خواهیم داشت

$$c = \frac{x_i}{\sqrt{x_i^2 + x_j^2}} \quad , \quad s = \frac{x_j}{\sqrt{x_i^2 + x_j^2}}$$

و

$$\alpha = cx_i + sx_j = \sqrt{x_i^2 + x_j^2}$$

حال ماتریس $J(i,j,c,s)$ را همانند رابطه ی ۲-۱ تعریف می کنیم.با محاسبه ی $J(i,j,c,s)x$ ملاحظه می‌کنیم

که درایه ی jام از این بردار برابر با صفر،درایه ی iام برابر است با $\sqrt{x_i^{۲} + x_j^{۲}}$ و سایر درایه های بردا ر x بدون تغییر باقی می مانند.

قضیه ۲-۲-۲. با مفروض بودن یک ماتریس $m \times n$ مانند A به تعداد $s = min(n, m - ۱)$ ماتریس متعامد $Q_۱, Q_۲, ..., Q_s$ تعریف شده توسط

$$Q_k = J_{km} J_{k,m-۱} ... J_{k,k+۱} \tag{۲-۲}$$

وجود دارد به طوری که اگر

$$Q = Q_۱^T Q_۲^T ... Q_s^T \tag{۳-۲}$$

آنگاه

$$A = QR \tag{۴-۲}$$

برهان. گام ۱:با استفاده از دوران ها گیونز ماتریس$Q_۱$ را چنان می سازیم که درایه های زیر $a_{۱۱}$ از ماتریس A صفر شود و قرار می دهیم :

$$A^{(۱)} = Q_۱ A$$

گام ۲:با استفاده از دوران های گیونز ماتریس $Q_۲$ را چنان می سازیم که درایه های زیر $a_{۲,۲}^{(۱)}$ از ماتریس $A^{(۱)}$ صفر شود و قرار می دهیم :

$$A^{(۲)} = Q_۲ A^{(۱)}$$

گام k: با استفاده از دوران های گیونز ماتریس Q_k را چنان می سازیم که درایه های زیر $a_{k,k}^{(k-۱)}$ از ماتریس $A^{k-۱}$ صفر شود و قرار می دهیم :

$$A^{(k)} = Q_k A^{(k-۱)}$$

$$k = ۱, ۲, ..., s = min(n, m - ۱)$$

بنابراین

$$R = A^{(s)} = Q_s A^{(s-۱)} = Q_s Q_{s-۱} A^{(s-۲)} = ... = Q_s Q_{s-۱} ... Q_۲ Q_۱ A.$$

قرار می دهیم $Q = Q_۱^T Q_۲^T ... Q_s^T$ در اینصورت $R = Q^T A$. در نتیجه $A = QR$.

□

مثال ۲-۲-۳. تجزیه ماتریس A را با استفاده از دوران های گیونز بدست آورید:

$$A = \begin{bmatrix} ۰ & ۱ & ۱ \\ ۱ & ۲ & ۳ \\ ۱ & ۱ & ۱ \end{bmatrix}$$

راهنمایی: $s = min(۳, ۲) = ۲$

گام ۱:

$$Q_۱ = J_{۱,۳} J_{۱,۲}$$

گام ۲:

$$Q_۲ = J_{۲,۳}$$

$$Q = Q_۱^T Q_۲^T \qquad R = Q^T A$$

$$A = QR$$

۲-۲-۱ ماتریس هوسهولدر

آلستون اسکات هوسهولدر [۱] ریاضیدان آمریکایی بودکه به طور تخصصی در زمینه زیست شناسی ریاضیات و آنالیز عددی فعالیت داشت و همچنین پدید آورنده تبدیلات هوسهولدر و روش هوسهولدر است. در گام اول به توضیح در زمینه ماتریس هوسهولدر می پردازیم و سپس روند به دست آوردن تجزیه QR با استفاده از ماتریس های هوسهولدر را بررسی می کنیم.

[۱] Alston Scott Householder

تعریف ۲-۲-۴. فرض کنید $v \in \mathbb{R}^n$ یک بردار غیر صفر باشد .در این صورت ماتریس

$$H = I - \frac{۲}{v^T v} v v^T$$

را یک ماتریس هوسهولدر و v را بردار هوسهولدر گویند [۶].

قبل از بیان قضیه یاد آوری می کنیم که اگر Z زیر فضایی از $\mathbb{R}^n$ باشد ،آنگاه فضای مکمل متعامد Z را $Z^{\perp}$ نشان می دهند و به صورت

$$Z^{\perp} = \{w \in R^n : z^T w = ۰, \forall z \in Z\}$$

تعریف می شودو بعلاوه جمع مستقیم Z و $Z^{\perp}$ برابر $\mathbb{R}^n$ است ،یعنی

$$R^n = Z \oplus Z^{\perp}$$

قضیه ۲-۲-۵. ماتریس هوسهولدر H متقارن و متعامد است و هر بردار $x \in \mathbb{R}^n$ را به بردار قرینه ی آن نسبت به فوق صفحه ی $span\{v\}^{\perp} = \{z \in R^n | z^{\perp} v = ۰\}$ می نگارد.

برهان:به سادگی می توان دید که $H^T = H$ از طرفی داریم

$$H^T H = H^۲ = (I - \frac{۲}{v^T v} v v^T)^۲ = I - \frac{۴}{v^T v} v v^T + \frac{۴}{(v^T v)^۲} v(v^T v) v^T = I.$$

بنابراین ماتریس H متعامد است .برای اثبات قسمت سوم قضیه فرض می کنیم $W = span\{v\}$. در این صورت داریم

$$R^n = W \oplus W^{\perp}$$

حال با توجه به اینکه $\dim(W) = ۱$ داریم $\dim(W^{\perp}) = n-۱$. فرض کنید مجموعه ی $\{\alpha_۱, \alpha_۲, ..., \alpha_{n-۱}\}$ یک پایه ی یکامتعامد برای $W^{\perp}$ باشد .به وضوح داریم $v^T \alpha_i = ۰$ ، $i = ۱, ۲, ..., n-۱$.بنابراین مقادیری حقیقی مثل $\beta, \beta_۱, ... \beta_{n-۱}$ موجودند به طوری که

$$x = \beta v + \beta_۱ \alpha_۱ + ... + \beta_{n-۱} \alpha_{n-۱}$$

از طرفی داریم $Hv=-v$، $H\alpha_i=\alpha_i$، $i=1,2,...,n-1$ بنابراین

$$Hx=\beta Hv+\beta_1\alpha_1+...+\beta_{n-1}\alpha_{n-1}=-\beta v+\beta_1\alpha_1+...+\beta_{n-1}\alpha_{n-1}.$$

این رابطه نشان می دهد که هر دو بردار x وHx دارای تصویر یکسان

$$P=\beta_1\alpha_1+...+\beta_{n-1}\alpha_{n-1}$$

روی $span\{v\}^{\perp}$ هستند و تصاویرشان روی بردارv در خلاف جهت یکدیگرند.

قضیه ۲-۲-۶. اگر $R^n \ni x \neq 0$، $v=x\pm\|x\|e_1$ و H ماتریس هوسهولدر متناظر با بردار v باشد،آنگاه $Hx=\mp\|x\|e_1$ که در آن e_1 ستون اول ماتریس همانی $n\times n$ است.

برهان :ما به دنبال یک ماتریس هوسهولدر مثل H هستیم به طوری که $Hx\in span\{e_1\}$.از طرفی ،داریم

$$Hx=(I-\frac{2}{v^Tv}vv^T)x=x-2\frac{v^Tx}{v^Tv}v.$$

$$-2\frac{v^Tx}{v^Tv}v=Hx-x$$

و در نتیجه $v\in span\{x,e_1\}$.بدون از دست دادن کلیت می توان فرض کرد $v=x+\alpha e_1$.حال داریم

$$v^Tx=x^Tx+\alpha e_1^Tx=x^Tx+\alpha\xi_1$$

$$v^Tv=(x^T+\alpha e_1^T)(x+\alpha e_1)=x^Tx+2\alpha\xi_1+\alpha^2$$

که درآن $\xi_1=e_1^Tx$ مولفه ی اول بردار x است. داریم

$$Hx=I-2\frac{v^Tx}{v^Tv}v$$

$$=x-2\frac{x^Tx+\alpha\xi_1}{x^Tx+2\alpha\xi_1+\alpha^2}(x+\alpha e_1)$$

$$=(1-2\frac{x^Tx+\alpha\xi_1}{x^Tx+2\alpha\xi_1+\alpha^2})x-2\alpha\frac{x^Tx+\alpha\xi_1}{x^Tx+2\alpha\xi_1+\alpha^2}e_1$$

حال برای اینکه Hx مضربی از $e_۱$ باشد کافی است داشته باشیم :

$$۱-۲\frac{x^Tx+\alpha\xi_۱}{x^Tx+۲\alpha\xi_۱+\alpha^۲}=۰$$

از این رابطه نتیجه می شود که $\alpha^۲=x^Tx$ یا $\alpha=\pm\|x\|_۲$. بنابراین $v=x\pm\|x\|_۲e_۱$ و در این صورت به سادگی می توان دید که

$$۲\alpha\frac{x^Tx+\alpha\xi_۱}{x^Tx+۲\alpha\xi_۱+\alpha^۲}=\pm\|x\|_۲$$

در نتیجه $Hx=\mp\|x\|_۲e_۱$.

در عمل برای جلوگیری از خطای ناشی از تفریق دو عدد نزدیک به هم اگر $\xi_۱=۰$ ،آنگاه‌قرارمی دهیم $v=x+\|x\|_۲e_۱$و در این صورت داریم $Hx=-\|x\|_۲e_۱$ در صورتی که$\xi\neq ۰$ قرار می دهیم: $v=x+sign(\xi_۱)\|x\|_۲e_۱$ که نتیجه می شود$Hx=-sign(\xi_۱)\|x\|_۲e_۱$.

مثال ۲-۲-۷. فرض کنید $x=(۲,۶,-۳)$ داریم $\xi_۱=۲$ و$\|x\|_۲=۷$. در نتیجه

$$v=x+sign(\xi_۱)\|x\|_۲e_۱=\begin{bmatrix}۲\\۶\\-۳\end{bmatrix}+۷\begin{bmatrix}۱\\۰\\۰\end{bmatrix}=\begin{bmatrix}۹\\۶\\-۳\end{bmatrix}$$

بنابراین

$$H=I-\frac{۲}{v^Tv}vv^T=\frac{۱}{۷}\begin{bmatrix}-۲&-۶&۳\\-۶&۳&۲\\۳&۲&۶\end{bmatrix}$$

در نتیجه

$$Hx=-sign(\xi_۱)\|x\|_۲e_۱=(-۷,۰,۰)^T$$

۲-۲-۲ تجزیه QRبا استفاده از تبدیل هوسهولدر

قضیه ۲-۲-۸. فرض کنید A یک ماتریس $n\times n$ دلخواه باشد. آنگاه ماتریس متعامد Q و ماتریس بالا مثلثی Rوجود دارند به طوری که $A=QR$

برهان. گام ۱:با استفاده از عناصر ستون اول ماتریس A،تبدیل هوسهولدر $H_۱$را طوری بسازید که

$$A^{(۱)} = H_۱ A = \begin{bmatrix} \alpha_۱ & \alpha_{۱,۲}^{(۱)} & \dots & \alpha_{۱,n}^{(۱)} \\ ۰ & \alpha_{۲,۲}^{(۱)} & \dots & \alpha_{۲,n}^{(۱)} \\ \vdots & \vdots & \vdots & \vdots \\ ۰ & \alpha_{n,۲}^{(۱)} & \dots & \alpha_{n,n}^{(۱)} \end{bmatrix}$$

گام ۲:با استفاده از عناصر زیر مولفه $\alpha_{۲,۲}^{(۱)}$ از ستون دوم ماتریس$A^{(۱)}$،تبدیل هوسهولدر $H_۲$ را طوری بسازید که

$$A^{(۲)} = H_۲ A^{(۱)} = \begin{bmatrix} \alpha_۱ & \alpha_{۱,۲}^{(۲)} & \alpha_{۱,۳}^{(۲)}\dots & \alpha_{۱,n}^{(۲)} \\ ۰ & \alpha^{(۲)} & \alpha_{۲,۳}^{(۲)}\dots & \alpha_{۲,n}^{(۲)} \\ \vdots & \vdots & \vdots & \vdots \\ ۰ & ۰ & \alpha_{n,۳}^{(۲)}\dots & \alpha_{n,n}^{(۲)} \end{bmatrix}$$

گام k ام: با استفاده از عناصر زیر مولفه $\alpha_{k,k}^{(k-۱)}$ از ستون k ماتریس $A^{(k-۱)}$ ،تبدیل هوسهولدر H_k را طوری بسازید که

$$A^{(k)} = H_k A^{(k-۱)} = \begin{bmatrix} \alpha_۱ & \alpha_{۱,۲}^{(k)} & \alpha_{۱,k}^{(k)}\dots & \dots & \alpha_{۱,n}^{(k)} \\ ۰ & \vdots & \vdots & \vdots & \vdots \\ ۰ & ۰ & \alpha^{(k)} & \alpha_{k,k+۱}^{(k)}\dots & \alpha_{k,n}^{(k)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ ۰ & ۰ & ۰ & \alpha_{n,k+۱}^{(k)}\dots & \alpha_{n,n}^{(k)} \end{bmatrix}$$

$$H_k A^{(k-۱)} = H_k H_{k-۱} A^{(k-۲)} = \dots = H_k \dots H_۲ H_۱ A$$

بعد از گام $n-۱$ ام یک ماتریس بالا مثلثی به شکل زیر نتیجه می شود.

$$H_{n-۱}...H_۲H_۱A = \begin{bmatrix} \alpha_۱ & \alpha_{۱,۲}^{(n-۱)} & \alpha_{۱,۳}^{(n-۱)} & \dots & \alpha_{۱,n-۱}^{(n-۱)} & \alpha_{۱,n}^{(n-۱)} \\ ۰ & \alpha_۲ & \alpha_{۲,۳}^{(n-۱)} & \vdots & \alpha_{۲,n-۱}^{(n-۱)} & \alpha_{۲,n}^{(n-۱)} \\ \vdots & \vdots & \vdots & \dots & \dots & \vdots \\ ۰ & ۰ & ۰ & \dots & \alpha_{n-۱} & \alpha_{n-۱,n}^{(n-۱)} \\ ۰ & ۰ & ۰ & \dots & ۰ & \alpha_{n,n}^{(n-۱)} \end{bmatrix}$$

قرار دهید

$$Q = H_۱H_۲...H_{n-۱}$$

$$R = H_{n-۱}...H_۲H_۱A$$

آنگاه از اینکه ماتریس های هوسهولدر متعامد و متقارنند ،نتیجه می شود که ماتریس Q نیز متعامد است و لذا

$$A = QR$$

الگوریتم ساخت ماتریس H_k در گام k ام $(k = ۱, ۲, ..., n-۱)$

۱-بردار $v \in \mathbb{R}^n$ را به صورت زیر در نظر بگیرید :

$$v = [\underbrace{۰, ۰, ..., ۰}_{\text{k-۱}}, \alpha_{k,k}^{(k-۱)}, \alpha_{k+۱,k}^{(k-۱)}, ..., \alpha_{n,k}^{(k-۱)}]$$

۲ - α_k را بصورت زیر تعریف کنید:

$$\alpha_k = -sign(\alpha_{k,k}^{(k-۱)})||v||_۲$$

۳-بردار w را به صورت زیر تعریف کنید:

$$w = v - \alpha_k e_k$$

۴-ماتریس H_k را به صورت زیر تعریف کنید:

$$H_k = I - ۲\frac{ww^t}{w^t w}$$

□

مثال ۲-۲-۹. تجزیه QR ماتریس A را به دست آورید:

$$A = \begin{bmatrix} ۴ & ۳ & ۰ \\ ۲ & ۱ & ۲ \\ ۴ & ۷ & ۰ \end{bmatrix}$$

حل: $n = ۳$ بنابر این $n - ۱ = ۲$، لذا کافیست مراحل الگوریتم فوق را برای $k = ۱, ۲$ انجام دهیم .

گام۱: $k = ۱$

$$A_۱ = \begin{bmatrix} ۴ \\ ۲ \\ ۴ \end{bmatrix} \Longrightarrow v_۱ = \begin{bmatrix} ۴ \\ ۲ \\ ۴ \end{bmatrix} \Longrightarrow \alpha_۱ = -۱\left\| \begin{matrix} ۴ \\ ۲ \\ ۴ \end{matrix} \right\|_۲ = -۶$$

$$w = \begin{bmatrix} ۴ \\ ۲ \\ ۴ \end{bmatrix} + ۶ \begin{bmatrix} ۱ \\ ۰ \\ ۰ \end{bmatrix} = \begin{bmatrix} ۱۰ \\ ۲ \\ ۴ \end{bmatrix}$$

$$H_۱ = I - ۲\frac{ww^t}{w^t w} = I - ۲\begin{bmatrix} \frac{۵}{۶} & \frac{۱}{۶} & \frac{۱}{۳} \\ \frac{۱}{۶} & \frac{۱}{۳۰} & \frac{۱}{۱۵} \\ ۱ & ۲ & ۳ \end{bmatrix} = \begin{bmatrix} -\frac{۲}{۳} & -\frac{۱}{۳} & -\frac{۲}{۳} \\ -\frac{۱}{۳} & \frac{۱۴}{۱۵} & -\frac{۲}{۱۵} \\ -\frac{۲}{۳} & -\frac{۲}{۱۵} & \frac{۱۱}{۱۵} \end{bmatrix}$$

آنگاه

$$A^{(۱)} = H_۱ A = \begin{bmatrix} -\frac{۲}{۳} & -\frac{۱}{۳} & -\frac{۲}{۳} \\ -\frac{۱}{۳} & \frac{۱۴}{۱۵} & -\frac{۲}{۱۵} \\ -\frac{۲}{۳} & -\frac{۲}{۱۵} & \frac{۱۱}{۱۵} \end{bmatrix} \begin{bmatrix} ۴ & ۳ & ۰ \\ ۲ & ۱ & ۲ \\ ۴ & ۷ & ۰ \end{bmatrix} = \begin{bmatrix} -۶ & -۷ & -\frac{۲}{۳} \\ ۰ & -۱ & \frac{۲۸}{۱۵} \\ ۰ & ۳ & -\frac{۴}{۱۵} \end{bmatrix}$$

گام۲: $k = ۲$

$$A_۲^{(۱)} = \begin{bmatrix} ۷ \\ -۱ \\ ۳ \end{bmatrix} \Longrightarrow v_۲ = \begin{bmatrix} ۰ \\ -۱ \\ ۳ \end{bmatrix} \Longrightarrow \alpha_۲ = ۱.\left\| \begin{matrix} ۰ \\ -۱ \\ ۳ \end{matrix} \right\|_۲ = \sqrt{۱۰}$$

$$w = \begin{bmatrix} ۰ \\ -۱ \\ ۳ \end{bmatrix} - \sqrt{۱۰} \begin{bmatrix} ۰ \\ ۱ \\ ۰ \end{bmatrix} = \begin{bmatrix} ۰ \\ -۱ - \sqrt{۱۰} \\ ۳ \end{bmatrix}$$

$$H_۲ = I - ۲\frac{ww^t}{w^tw} =$$

$$I - \frac{۲}{۲۶/۳۲۴۶} \begin{bmatrix} ۰ & ۰ & ۰ \\ ۰ & ۱۷/۳۲۴۶ & -۱۲/۴۸۶۸ \\ ۰ & -۱۲/۴۸۶۸ & ۹ \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & -۰/۳۱۶۲ & ۰/۹۴۸۷ \\ ۰ & ۰/۹۴۸۷ & ۰/۳۱۶۲ \end{bmatrix}$$

گام ۲:ادامه

$$A^{(k)} = H_kA^{(k-۱)} = \begin{bmatrix} \alpha_۱ & \alpha_{۱,۲}^{(k)} & \alpha_{۱,k}^{(k)}\cdots & \cdots & \alpha_{۱,n}^{(k)} \\ ۰ & \vdots & \vdots & \vdots & \vdots \\ ۰ & ۰ & \alpha^{(k)} & \alpha_{k,k+۱}^{(k)}\cdots & \alpha_{k,n}^{(k)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ ۰ & ۰ & ۰ & \alpha_{n,k+۱}^{(k)}\cdots & \alpha_{n,n}^{(k)} \end{bmatrix} =$$

$$\begin{bmatrix} -۶ & -۷ & -\frac{۲}{۳} \\ ۰ & ۳/۱۶۲۳ & -۰/۸۴۳۳ \\ ۰ & ۰ & ۱/۶۸۶۵ \end{bmatrix}$$

گام ۲:ادامه بنابراین

$$H_۲A^{(۱)} = H_۲H_۱A = R$$

در نتیجه

$$A = H_۱^{-۱} H_۲^{-۱} R = H_۱ H_۲ R = QR$$

که

$$Q = H_۱ H_۲$$

$$Q = H_۱ H_۲ = \begin{bmatrix} -\frac{۲}{۳} & -\frac{۱}{۳} & -\frac{۲}{۳} \\ -\frac{۱}{۳} & \frac{۱۴}{۱۵} & -\frac{۲}{۱۵} \\ -\frac{۲}{۳} & -\frac{۲}{۱۵} & \frac{۱۱}{۱۵} \end{bmatrix} \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & -۰٫۳۱۶۲ & ۰٫۹۴۸۷ \\ ۰ & ۰٫۹۴۸۷ & ۰٫۳۱۶۲ \end{bmatrix} =$$

$$\begin{bmatrix} -۰٫۶۶۶۷ & -۰٫۵۲۷۱ & -۰٫۵۲۷۰ \\ -۰٫۳۳۳۳ & -۰٫۴۲۱۶ & ۰٫۸۴۳۳ \\ -۰٫۶۶۶۷ & ۰٫۷۳۷۹ & ۰٫۱۰۵۴ \end{bmatrix}$$

فصل ۳

بروزرسانی تجزیه QRو کاربردهای آن

برنامه های کاربردی متعددی وجود دارد که نیاز است ماتریس داده ها در آن ها بارها و بارها بروزرسانی شود. به عنوان مثال در یک برنامه پردازش سیگنال ،هر سطر از ماتریس مربوط به اندازه گیری داده ها در یک زمان معین است ،هر اندازه گیری جدید باعث اضافه شدن یک سطر جدید به ماتریس می شود. یا فرض کنید مساله حداقل مربعات برای یک سری داده با استفاده از تجزیه QR حل شده باشد .حال اگریک سری داده جدید به مساله داده شود یا باید از ابتدا دوباره مساله را حل کرد و تجزیه QR آن را بدست آورد یا اینکه از بروزرسانی تجزیه QR با استفاده از روش هائی که در آینده بیان خواهد شد استفاده کرد که این امر به تبع باعث کاهش زمان محاسباتی می‌شود. پس از ذکر مفاهیم و روش‌های مورد نیاز، دو مثال کاربردی از بروزرسانی تجزیه QR بیان خواهد شد.

۳-۱ بروز رسانی تجزیه QR

با اضافه کردن یک سری داده جدید به مساله می توان از بروزرسانی تجزیه QR آن برای بدست آوردن Q وR جدید استفاده کرد.در ادامه وقتی داده ها به عنوان یک سطر یا ستون جدید اضافه شوند را بررسی خواهیم کرد.

۳-۱-۱ اضافه کردن یک ستون

فرض کنید تجزیه QR ماتریس A را بدست آورده ایم.اگر یک ستون به ماتریس A اضافه کنیم می توانیم تجزیه QR ماتریس جدید را با بروزرسانی تجزیه QR ماتریس A بدست آوریم و از آن برای بروزرسانی مساله کمترین مربعات بهره بگیریم [۷، ۸، ۴].

سیستم خطی $Ax = b$ را در نظر بگیرید که در آن: $A \in \mathbb{R}^{m\times n}$، $x \in \mathbb{R}^n$و $b \in \mathbb{R}^m$. مساله کمترین مربعات که در آن A یک ماتریس رتبه کامل است به صورت زیر تعریف می شود.

$$\min_x ||Ax - b||_۲$$

فرض کنید $A = QR$ و $d = Q^T b$ باشد.در این صورت

$$||Ax - b||_۲^۲ = ||Q^T Ax - Q^T b||_۲^۲ = ||Rx - d||_۲^۲$$

برقرار خواهد بود. اگر

$$\tilde{A} = \tilde{Q}\tilde{R}$$

آنگاه $\tilde{A}$ همان ماتریس A بروزرسانی شده می باشد.برای محاسبه $\tilde{d}$ نیاز داریم که

$$||\tilde{A}x - \tilde{b}||_۲ = \|\tilde{R}x - \tilde{d}\|_۲$$

که $\tilde{b}$ بروزرسانی شده b است و

$$\tilde{d} = \tilde{Q}^T\tilde{b}$$

اکنون مراحل بروزرسانی تجزیه QR ماتریس A را بعد از اضافه شدن یک ستون بررسی می کنیم.
فرض کنید $A = QR$ که $Q \in \mathbb{R}^{n\times n}$ متعامد است و $R \in \mathbb{R}^{n\times m}$ یک ماتریس بالا مثلثی است . اگر

$$A = [A_۱ \;\; A_۲]$$

آنگاه

$$\tilde{A} = [A_۱ \;\; u \;\; A_۲]$$

$u \in \mathbb{R}^m$ یک بردار ستونی است که درموقعیت kام ماتریس $A_{m\times n}$ اضافه شده است ،در این صورت خواهیم داشت:

$$\tilde{A} = [A(۱ : m, ۱ : k - ۱) \;\; u \;\; A(۱ : m, k : n)]$$

پس

$$Q^T\tilde{A} = [R(۱:m, ۱:k-۱) \;\; v \;\; R(۱:m, k:n)] \tag{۳-۱}$$

که $v = Q^T u$.

با اضافه شدن این ستون به ماتریس R این ماتریس از حالت مثلثی خارج خواهد شد که باید دوباره آن را به حالت مثلثی در آورد. فرض کنید v ستونی است که به ماتریس R اضافه شده است ،برای تبدیل آن به فرم مثلثی باید $m-k$ درایه از ستون v را از پایین صفر کرد. با توجه به تعریف ماتریس گیونزکه در فصل ۲ بیان شد.

$$J(i,j,c,s) = \begin{pmatrix} ۱ & & & & & & \\ & \ddots & & & & & \\ & & c & & s & & \\ & & & \ddots & & & \\ & & -s & & c & & \\ & & & & & \ddots & \\ & & & & & & ۱ \end{pmatrix} \tag{۳-۲}$$

بایدمقادیر c و s را طوری بیابیم که

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix}^T \begin{pmatrix} x_i \\ x_j \end{pmatrix} = \begin{pmatrix} \alpha \\ ۰ \end{pmatrix}$$

برقرار باشد. با فرض ثابت بودن زاویه دوران، برای سادگی G را به صورت زیر تعریف می کنیم:

$$G_m = J(m-۱, m, c, s)$$

ماتریس G_m را یک ماتریس دوران بر روی سطر های $m, m-۱$ در نظر میگیریم که $G_m v$ موقعیت mام را صفر می کند.همچنین اگر $G_{m-۱}$ بر روی سطرهای $m-۱, m-۲$ اعمال شود،آنگاه $G_{m-۱}G_m v$ موقعیت $m-۱$ام را صفر می کند. صفرهایی که قبلا وجود داشتند بدون تغییر باقی می مانند. در ادامه این روش ماتریس های دوران $G_m, ..., G_{k+۱}$ را می توان معرفی کردکه $G_m, ..., G_{k+۱}v$ موقعیت های $k+۱, ..., m$ را صفر

می‌کند. به سادگی نتیجه می شود که $\tilde{R}$ یک ماتریس بالا مثلثی است.

ماتریس های دوران بر روی R_1 ها تغییر ایجاد نمی کنند،ولی تا حدی که بتوان با پر کردن R به یک ماتریس بالا مثلثی رسید بر روی R_2 تاثیر می گذارند.

همچنین

$$(G_{k+1}^T ... G_m^T Q^T)\tilde{A} = \tilde{Q}^T \tilde{A} = \tilde{R}$$

، که $\tilde{R} \in \mathbb{R}^{m\times(n+1)}$ بالا مثلثی است و $\tilde{Q} \in \mathbb{R}^{m\times m}$ پس

$$\tilde{A} = \tilde{Q}\tilde{R}$$

فرض کنید ماتریس ۶ × ۸ وجود داردکه در موقعیت $k = 4$ آن یک ستون اضافه شده است. با بررسی تغییرات بر روی ماتریس R بحث را ادامه می دهیم.همانطور که در ماتریس زیر مشاهده می شودیک ستون در موقعیت $k = 4$ ماتریس R اضافه شده است که ماتریس را از فرم بالا مثلثی خارج کرده است با توجه به توضیحاتی که داده شد ابتدا باید با کمک ماتریس های گیونز درایه های که با $\ominus$ نشان داده شده است را صفر کرد. پس به $m - k$ ماتریس گیونز $G(i,j) \in R^{m\times m}$ نیاز داریم تا بتواند $v(k+1 : m)$ را صفر کند. درایه های که بعدازاضافه کردن ستون غیر صفر باقی می مانند را با $+$ وعناصری صفری که می توانند مقدار بگیرند با $\oplus$ نمایش داده شده‌است.

$$\begin{bmatrix} + & + & + & + & + & + & + \\ \circ & + & + & + & + & + & + \\ \circ & \circ & + & + & + & + & + \\ \circ & \circ & \circ & + & + & + & + \\ \circ & \circ & \circ & \ominus & \oplus & + & + \\ \circ & \circ & \circ & \ominus & \circ & \oplus & + \\ \circ & \circ & \circ & \ominus & \circ & \circ & \oplus \\ \circ & \circ & \circ & \ominus & \circ & \circ & \circ \end{bmatrix}$$

حال می توان بروزرسانی b را نیز محاسبه کرد.

$$G(k,k+1)^T ... G(m-1,m)^T Q^T b = \tilde{d}$$

درالگوریتم (۳-۱) مراحل بروزرسانی $\tilde{R}$ بیان شده است.

همچنین $\tilde{d}$ برای محاسبه کمترین مربعات با استفاده از تجزیه QR به صورت زیر محاسبه می شود.

$$||\tilde{A}x - b||_۲ = ||\tilde{R}x - \tilde{d}||_۲$$

الگوریتم ۳-۱ بروزرسانی ماتریس بالا مثلثی R بعد از اضافه شدن یک ستون به ماتریس Aدر تجزیه QR

$u = Q^T u$
$\tilde{d} = Q^T b$
for i= m:-1:k+1 **do**
  $[c(i), s(i)] = givens(u(i-1), u(i))$
  $u(i-1) = c(i)u(i-1) - s(i)u(i)$
  update R if there is a nonzero row
  if i <= n+1 **then**
    $R(i-1:i, i-1:n) = \begin{bmatrix} c(i) & s(i) \\ -s(i) & c(i) \end{bmatrix}^T R(i-1:i, i-1:n)$
  end if
end for
Update R
$\tilde{d}(i-1:i) = \begin{bmatrix} c(i) & s(i) \\ -s(i) & c(i) \end{bmatrix}^T \tilde{d}(i-1:i)$
if k=1 **then**
  $\tilde{R} = upper\ triangular\ part\ of[u\ R]$
else
  if k=n+1 **then**
    $\tilde{R} = upper\ triangular\ part\ of[R\ u]$
  else
    $\tilde{R} = upper\ triangular\ part\ of[R(1:m, 1:k-1)\ u\ R(1:m, k:n)]$
  end if
end if
Compute the residual
$resid = ||\tilde{d}(n+1:m)||_2$

$\tilde{Q}$ را می توان از الگوریتم (۳-۲) بدست آورد.

در الگوریتم (۳-۲) بردارهای c ، s از الگوریتم (۳-۱) بدست می آیند و $\tilde{Q} \in R^{m\times m}$ یک ماتریس متعامد است که $\tilde{A} = \tilde{Q}\tilde{R}$.

الگوریتم ۳-۲ بروزرسانی ماتریس متعامد Q بعد از اضافه شدن یک ستون به ماتریس A در تجزیه QR

for i= m:-1:k+1 **do**

$$Q(1:m, i-1:i) = Q(1:m, i-1:i)\begin{bmatrix} c(j) & s(j) \\ -s(j) & c(j) \end{bmatrix}$$

end for

$\tilde{Q} = Q$

۳-۱-۲ اضافه کردن یک سطر

فرض کنید $A \in \mathbb{R}^{n\times m}$ ، $n \geq m$ ، $Rank(A) = m$. تجزیه $A = QR$ داده شده است که در آن $Q \in \mathbb{R}^{n\times n}$ متعامد و $R \in \mathbb{R}^{n\times m}$ ماتریسی بالا مثلثی است و $m \times m$ بلوک های بالای R بالا مثلثی هستند و باقیمانده R صفر است [۴]. فرض کنید :

$$\tilde{A} = \begin{bmatrix} A_۱ \\ z^T \\ A_۲ \end{bmatrix} \in \mathbb{R}^{n+۱\times m}$$

که

$$A = \begin{bmatrix} A_۱ \\ A_۲ \end{bmatrix}$$

z^T سطر جدیدی خواهد بود که به سطرهای ماتریس A اضافه شده است.

حال مراحل بدست آوردن تجزیه $\tilde{A} = \tilde{Q}\tilde{R}$ که همان A بروزرسانی شده بعد از اضافه شدن یک سطر می‌باشد با استفاده از تجزیه $A = QR$ بررسی می‌شود.

R،Q را به صورت زیر تقسیم بندی می‌کنیم

$$\begin{bmatrix} A_۱ \\ A_۲ \end{bmatrix} = \begin{bmatrix} Q_۱ \\ Q_۲ \end{bmatrix} R$$

مستقیما می توان رابطه زیر را که به تجزیه QR نزدیک است نتیجه گرفت:

$$\tilde{A} = \begin{bmatrix} 0 & Q_1 \\ 1 & 0^T \\ 0 & Q_2 \end{bmatrix} \begin{bmatrix} z^T \\ R \end{bmatrix}$$

مشکلی که در اینجا مطرح است ماتریس سمت راست رابطه فوق می باشد که بالا مثلثی نیست،باید آن را به فرم بالا مثلثی تبدیل کرد. در ادامه نحوه کار بیان می شود. از آنجا که

$$\begin{bmatrix} z^T \\ R \end{bmatrix} = \begin{bmatrix} z_1 & z_2 & \cdots \\ r_{11} & r_{12} & \cdots \\ 0 & r_{22} & \cdots \\ 0 & 0 & \ddots \\ \vdots & \vdots & \end{bmatrix}$$

اگر $G(2,1)^T$ یک ماتریس دوران روی سطر های ۱و۲فرض شود آنگاه

$$G(2,1)^T \begin{bmatrix} z^T \\ R \end{bmatrix}$$

درایه r_{11} را صفر می کند. واضح است که این تبدیلات صفری که قبلا در ماتریس به وجود آمده است را از بین نمی برد. با استفاده از ماتریس دوران $G(2,1)^T$ در یک مرحله r_{11} صفر می شود. می توان با استفاده از $G(3,2)^T$،r_{22} را صفر کرد که روی سطرهای ۲و۳ عمل می کند.صفرهایی که در موقعیت r_{11} تولید شده است همچنان صفر باقی می مانند زیرا $r_{21} = 0$ است. در گام بعدی دوران $G(4,3)^T$ را خواهیم داشت که روی سطرهای ۳و۴عمل می‌کند و r_{33} را صفر می کند.

اگر این روند را تا m مرحله ادامه پیدا کند در انتها یک ماتریس بالا مثلثی خواهیم داشت که

$$\tilde{R} = G_m^T \dots G_2^T G_1^T \begin{bmatrix} z^T \\ R \end{bmatrix} \in \mathbb{R}^{(n+1)\times m}$$

همچنین

$$\tilde{Q} = \begin{bmatrix} 0 & Q_1 \\ 1 & 0^T \\ 0 & Q_2 \end{bmatrix} G_1 G_2 ... G_m,$$

در نتیجه $\tilde{A} = \tilde{Q}\tilde{R}$ که $\tilde{Q} \in \mathbb{R}^{(n+1)\times(n+1)}$ متعامد و $\tilde{R} \in \mathbb{R}^{(n+1)\times m}$ بالا مثلثی است.

۳-۱-۳ محاسبه QR با ماتریس های گیونز [1]

از دوران های گیونز می توان برای محاسبه تجزیه QR استفاده کرد.که جزئیات کار را با یک مثال بر روی یک ماتریس ۴×۳ توضیح داده می شود.

$$\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} \xrightarrow{(3,4)} \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ 0 & \times & \times \end{bmatrix} \xrightarrow{(2,3)} \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix} \xrightarrow{(1,2)}$$

$$\begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix} \xrightarrow{(3,4)} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{bmatrix} \xrightarrow{(2,3)} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \\ 0 & 0 & \times \end{bmatrix} \xrightarrow{(3,4)} R$$

در دوران های گیونز دو بردار مطرح می شود که زیر بنای اصلی تبدیلات گیونز بر روی این بردارها است.روشن است که اگرماتریس های گیونزبا G_j نشان داده شود،آنگاه با تاثیر دادن این ماتریس های دوران و رابطه $Q^T A = R$ ماتریس R را می توان محاسبه کردکه یک ماتریس بالا مثلثی است و $Q = G_1 ... G_t$ و t اندیس همه ماتریس های دوران است.

در الگوریتم (۳-۳) $A \in \mathbb{R}^{m\times n}$ و $m \geq n$ فرض شده است. همچنین می توان با توجه به تجزیه $A = QR$ رابطه ی $Q^T A = R$ را نتیجه گرفت که در آن R بالا مثلثی و Q متعامد است.

[1] Givens matrixs

الگوریتم ۳-۳ محاسبه تجزیه QR با استفاده از تبدیلات گیونز

for j=1:n **do**
  for i= m:-1:j+1 **do**
    $[c,s] = givens(A(i-1,j), A(i,j))$
    $A(i-1:i, j:n) = \begin{bmatrix} c(i) & s(i) \\ -s(i) & c(i) \end{bmatrix}^T A(i-1:i, j:n)$
  end for
end for

۳-۱-۴ تجزیه QR ماتریس هسنبرگی بااستفاده از تبدیلات گیونز [۱]

در این قسمت بررسی می شود که چگونه می توان با استفاده از دوران های گیونز تجزیه QR یک ماتریس بالا هسنبرگی را به دست آورد. با یک مثال کوچک هدف اصلی شرح داده می شود. ماتریس بالا هسنبرگی زیر را در نظر بگیرید .

$$\begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \circ & \times & \times & \times & \times & \times \\ \circ & \circ & \times & \times & \times & \times \\ \circ & \circ & \circ & \times & \times & \times \\ \circ & \circ & \circ & \circ & \times & \times \end{bmatrix}$$

حال می خواهیم با استفاده از تبدیلات گیونز تجزیه QR آن را به دست آوریم. بعد از دو گام که ماتریس های دوران گیونز بر روی آن اعمال می شود خواهیم داشت:

$$G(۲,۳)^T G(۱,۲)^T A \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \circ & \times & \times & \times & \times & \times \\ \circ & \circ & \times & \times & \times & \times \\ \circ & \circ & \times & \times & \times & \times \\ \circ & \circ & \circ & \times & \times & \times \\ \circ & \circ & \circ & \circ & \times & \times \end{bmatrix}$$

می توان درایه $(۴,۳)$ را با استفاده از ماتریس دوران $G(۳,۴)$ به صفر تبدیل کرد در زیر مشاهده می شود به همین ترتیب می توان از دوران های گیونز برای صفر کردن درایه های $(۵,۴)$ و $(۶,۵)$ نیز استفاده کرد که سرانجام یک

[۱] Hessenberg matrix

ماتریس بالا مثلثی خواهیم داشت.

$$G(۳,۴)^TG(۲,۳)^TG(۱,۲)^TA = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \circ & \times & \times & \times & \times & \times \\ \circ & \circ & \times & \times & \times & \times \\ \circ & \circ & \times & \times & \times & \times \\ \circ & \circ & \circ & \times & \times & \times \\ \circ & \circ & \circ & \circ & \times & \times \end{bmatrix}$$

الگوریتم به دست آوردن تجزیه QR یک ماتریس بالا هسنبرگی را در الگوریتم (۳-۴) مشاهده می کنید که $A \in \mathbb{R}^{n\times n}$ یک ماتریس بالا هسنبرگی است و در اینجا نیزبه آسانی با توجه به اینکه $A = QR$ می‌باشد، می‌توانیم رابطه $Q^TA = R$ را نتیجه بگیریم که Q متعامد است و R بالا مثلثی است.

همچنین $Q = G_۱...G_{n-۱}$ که G_j همان ماتریس های دوران هستند به طوری که $G_j = G(j, j+۱)$

الگوریتم ۳-۴ محاسبه تجزیه QR ماتریس بالا هسنبرگی با استفاده از تبدیلات گیونز

for j=1:n-1 **do**

$[c, s] = givens(A(j,j), A(j+1,j))$

$A(j:j+1, j:n) = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}^T A(j:j+1, j:n)$

end for

۳-۱-۵ بروزرسانی رتبه یک ماتریس [۱]

فرض کنید تجزیه $QR = B \in \mathbb{R}^{m\times m}$ موجود است.تجزیه QR رابطه (۳-۳)که $u, v \in \mathbb{R}^n$ باشد، به صورت زیر بررسی می شود[۸].

$$B + uv^T = Q_۱R_۱ \tag{۳-۳}$$

می توان نوشت:

$$B + uv^T = Q(R + wv^T) \tag{۴-۳}$$

[۱] Rank -one update

که $w = Q^T u$. فرض کنید دوران های $J_{n-1}, ..., J_2, J_1$ محاسبه شده است . به طوریکه

$$J_1^T ... J_{n-1}^T w = \pm ||w||_2 e_1.$$

برقرار باشد. در اینجا هر J_k یک دوران در فضای k،$k+1$ است.اگر دوران های گیونز که در بالا معرفی شدند بر روی R اعمال شود آن را می توان به صورت رابطه (۳-۵) نشان داد.

$$H = J_1^T ... J_{n-1}^T R \tag{۵-۳}$$

که یک ماتریس بالا هسنبرگی است. برای مثال فرض کنید $n = 4$ است و R ، w را به صورت زیر تعریف شده‌است.

$$R = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \end{bmatrix} \qquad w = \begin{bmatrix} \times \\ \times \\ \times \\ \times \end{bmatrix}$$

می توان w،R را به صورت زیر بروزرسانی کرد.

$$R = J_3^T R = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix} \qquad w = J_3^T w = \begin{bmatrix} \times \\ \times \\ \times \\ 0 \end{bmatrix}$$

$$R = J_2^T R = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix} \qquad w = J_2^T w = \begin{bmatrix} \times \\ \times \\ 0 \\ 0 \end{bmatrix}$$

$$R = J_1^T R = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix} \qquad w = J_1^T w = \begin{bmatrix} \times \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

در نتیجه

$$(J_1^T ... J_{n-1}^T)(R + wv^T) = H \pm ||w||_2 e_1 v^T = H_1 \qquad (۳-۶)$$

که ماتریس H_1 بدست آمده بالا هسنبرگی است. در الگوریتم (۳-۴) بیان شد که چگونه می توان تجزیه QR یک ماتریس بالا هسنبرگی را با استفاده از تبدیلات گیونز بدست آورد. پس می توان ماتریس های دوران G_k را که $k = 1 : n - 1$ را پیدا کرد که با تاثیر بر روی H_1 آن را به یک ماتریس بالا مثلثی تبدیل کند. بنابراین داریم:

$$G_{n-1}^T ... G_1^T H_1 = R_1 \qquad (۳-۷)$$

با ترکیب معادلات (۳-۷)،(۳-۴) می توان بروزرسانی رتبه -یک ماتریس B را به صورت زیر نوشت.

$$B + uv^T = Q_1 R_1$$

که

$$Q_1 = QJ_{n-1} ... J_1 G_1 ... G_{n-1}$$

در ادامه دو کاربرد از بروزرسانی تجزیه QR در یادگیری ماشین بیان خواهد شد.اولین کاربرد، استفاده از آن در شبکه عصبی و دومین کاربرد آن در تحلیل تفکیک خطی می‌باشد.

۳-۲ تحلیل تفکیک خطی [۱]

تحلیل تفکیک خطی (آنالیز افتراقی خطی یا (LDA)یا تشخیص خطی فیشر)روش‌های آماری هستند که از جمله در یادگیری ماشین و شناسایی الگو برای پیدا کردن ترکیب خطی ویژگیهایی که به بهترین صورت دو یا چند کلاس از اشیا را از هم جدا می‌کند، استفاده می‌شوند. یک کاربرد عمدهٔ این روش، کاستن تعداد بعدهای داده است. LDA ارتباط نزدیکی با تحلیل واریانس و تحلیل رگرسیون دارد که سعی دارند یک متغیر مستقل را به عنوان ترکیبی خطی از ویژگی های دیگر بیان کنند. این متغیر مستقل در LDA به شکل برچسب یک کلاس است. همچنین LDA ارتباطی تنگاتنگ با تحلیل مولفه های اصلی PCA دارد. چرا که هر دو متد به دنبال ترکیبی خطی از متغیرهایی هستند که به بهترین نحو داده ها را توصیف می کنند.

کاربرد اصلی LDA در مسایل طبقه بندی و دسته بندی است [۹]. نکات زیر را در مسایل دسته بندی داریم:

- در مسایل دسته بندی یک بردار ورودی X به یکی از K کلاس مجزای C_k اختصاص داده میشود.
- برای این کار فضای ورودی به نواحی تصمیم گیری تقسیم بندی میشود که مرزهای آنرا سطوح تصمیم گیری می نامند.
- سطوح تصمیم گیری مدل LDA از توابع خطی تشکیل شده است. برای جدا سازی فضای ورودی D بعدی از ابرصفحه های $D-۱$ بعدی استفاده میشود.

در مسایل دو کلاسی تابع هدف بصورت زیر است:

$$t \in \{۰, ۱\}$$

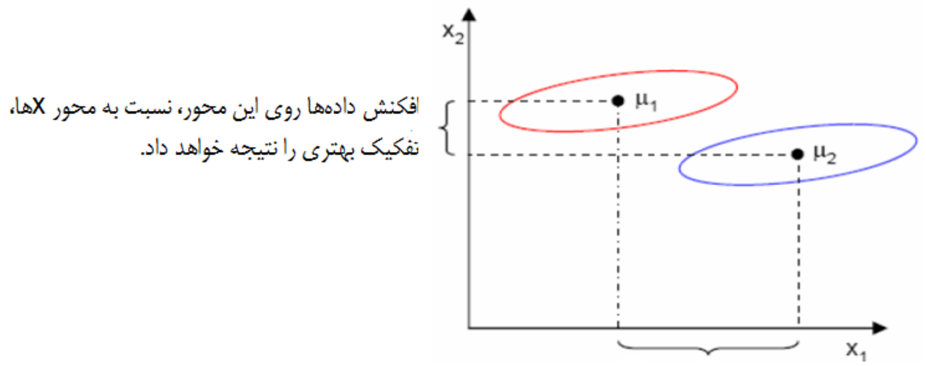
که مقدار $t = ۰$ برای کلاس $C_۱$ و $t = ۱$ برای کلاس $C_۲$ استفاده میشود.

حال مدل تحلیل تفکیک خطی (LDA) را بررسی می کنیم .

فرض کنید که بردار x که یک داده D بعدی است را با استفاده از رابطه $y = W^T x$ به یک بعد تصویر کنیم. اینکار باعث میشود تا داده هایی که در D بعد بخوبی جدا پذیر بودند در یک بعد در هم فرو رفته و جدا نشوند. اما میتوان با انتخاب صحیح W تصویرسازی [۲](افکنش) داده را طوری انجام داد که حداکثر جدا سازی بدست آید.

در یک مسئله دو کلاسه فرض کنید که تعداد $N_۱$ داده به کلاس $C_۱$ و تعداد $N_۲$ داده به کلاس $C_۲$ تعلق داشته

---
[۱] Linear Discriminant Analysis
[۲] projection



شکل ۳-۱: تفکیک کلاس ها

باشد در این صورت میانگین هر کلاس برابر است با:

$$\mu_1 = \frac{1}{N_1}\sum_{n\in C_1} x_n \qquad \mu_2 = \frac{1}{N_2}\sum_{n\in C_2} x_n$$

میانگین هر کلاس برای داده های تصویر شده به صورت زیر محاسبه می شود.

$$\tilde{\mu}_1 = \frac{1}{N_1}\sum_{y\in C_i} y = \frac{1}{N_1}\sum_{x\in C_i} W^T x = W^T \frac{1}{N_1}\sum_{x\in C_i} x = W^T \mu_1$$

$$\tilde{\mu}_2 = \frac{1}{N_2}\sum_{y\in C_i} y = \frac{1}{N_2}\sum_{x\in C_i} w^T x = W^T \frac{1}{N_2}\sum_{x\in C_i} x = W^T \mu_2$$

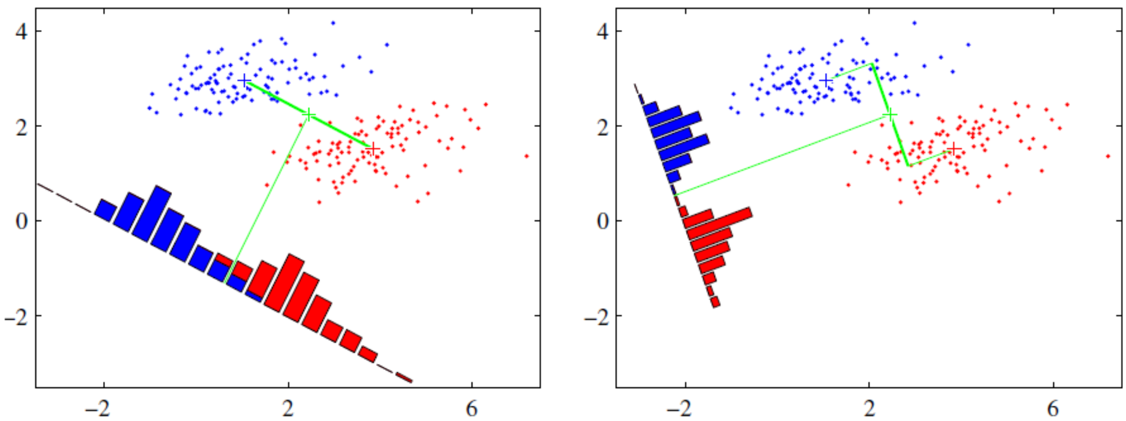
یک راه ساده برای جداسازی داده های تصویر شده این است که فاصله بین میانگین داده های تصویر شده افزایش یابد. در این حالت میتوان W را طوری انتخاب کرد که فاصله زیر افزایش یابد:

$$\tilde{\mu}_2 - \tilde{\mu}_1 = W^T(\mu_2 - \mu_1)$$

تابع هدف به صورت زیر تعریف می شود.

$$J(W) = |\tilde{\mu}_2 - \tilde{\mu}_1| = |W^T\mu_1 - W^T\mu_2| = |W^T(\mu_2 - \mu_1)|$$

فاصله بین میانگین دو کلاس لزوما معیار خوبی برای جداپذیر ساختن دو کلاس نیست.به شکل ۳-۱ توجه کنید. این مطلب اهمیت بررسی واریانس داخل کلاسی را روشن می کند.



شکل ۳-۲: افکنش کلاس ها

داده های شکل ۳-۲ در فضای اصلی جدا پذیر هستند اما وقتی به خطی که میانگین های آنها را به هم وصل میکند تصویر میشوند همپوشانی زیادی دارند.

ایده فیشر: تابعی را ماکزیمم کنیم که جدایی بین کلاسهای تصویر شده را ماکزیمم کند در حالیکه در داخل هر کلاس واریانس را کم نماید تا بدینوسیله هم پوشانی کاهش یابد. واریانس داخل کلاسی داده های تصویر شده متعلق به کلاس C_i به صورت

$$\tilde{s}_i^۲ = \sum_{y \in C_i} (y - \tilde{\mu}_i)^۲$$

تعریف می شود. ماتریس s را ماتریس پراکندگی [۱] می نامند. $\tilde{s}_۱^۲ + \tilde{s}_۲^۲$ مجموع واریانس درون کلاسی رامشخص می کند. قاعده فیشر بصورت نسبت واریانس بین کلاسها به واریانس درون کلاسی تعریف میشود:

$$J(w) = \frac{|\tilde{\mu}_۲ - \tilde{\mu}_۱|}{|\tilde{s}_۱^۲ + \tilde{s}_۲^۲|} \qquad (۳-۸)$$

کواریانس داده های x در کلاس c_i را با S_i و کواریانس بین کلاسی با S_W نشان داده می شود که به صورت زیر تعریف می گردد.

$$S_i = \sum_{x \in C_i} (x - \mu_i)(x - \mu_i)^T$$

$$S_W = S_۱ + S_۲$$

[۱] scatter matrix

اکنون واریانس داده های تصویر شده y را می توان با کمک ماتریس پراکندگی داده های x نشان داد

$$\tilde{s}_i^{۲} = \sum_{y \in C_i} (y - \tilde{\mu}_i)^{۲} = \sum_{x \in c_i} (W^T x - W^T \mu_i)^{۲}$$

$$= \sum_{x \in c_i} W^T (x - \mu_i)(x - \mu_i)^T W$$

$$= W^T S_i W$$

$$\tilde{s}_{۱}^{۲} + \tilde{s}_{۲}^{۲} = W^T S_{۱} W + W^T S_{۲} W = W^T (S_{۱} + S_{۲}) W = W^T S_W W = \tilde{S}_W$$

که $\tilde{S}_W$ ماتریس پراکندگی بین کلاسی داده های تصویر شده y است. به طور مشابهه می توان اختلاف میانگین بین کلاسی داده های تصویر شده را با رابطه زیر نشان داد.

$$(\tilde{\mu}_{۲} - \tilde{\mu}_{۱})^{۲} = (W^T \mu_{۱} - W^T \mu_{۲})^{۲}$$

$$= W^T (\mu_{۱} - \mu_{۲})(\mu_{۱} - \mu_{۲})^T W$$

$$= W^T S_B W = \tilde{S}_B$$

که S_B پراکندگی بین کلاسی برای داده های اولیه x و $\tilde{S}_B$ پراکندگی بین کلاسی برای داده های تصویر شده y است.

حال می توان رابطه ۳-۸ را بصورت زیر مرتب نمود:

$$J(W) = \frac{|\tilde{\mu}_{۲} - \tilde{\mu}_{۱}|}{|\tilde{s}_{۱}^{۲} + \tilde{s}_{۲}^{۲}|} = \frac{W^T S_B W}{W^T S_W W}$$

برای پیدا کردن بیشترین مقدار تابع $J(w)$مشتق آن را محاسبه کرده و مقدار مشتق ،برابر صفر قرار داده می شود.

$$\frac{d}{dW} J(W) = \frac{d}{dW} \left(\frac{W^T S_B W}{W^T S_W W} \right) = ۰$$

$$\Rightarrow (W^T S_W W) \frac{d}{dW} (W^T S_B W) - (W^T S_B W) \frac{d}{dW} (W^T S_W W) = ۰$$

$$\Rightarrow (W^T S_W W) ۲ S_B W - (W^T S_B W) ۲ S_W = ۰$$

طرفین رابطه بالا را بر $۲W^T S_W W$ تقسیم می کنیم:

$$\Rightarrow \left(\frac{W^T S_W W}{W^T S_W W}\right) S_B W - \left(\frac{W^T S_B W}{W^T S_W W}\right) S_W W = ۰$$

$$\Rightarrow S_B W - J(W) S_W W = ۰$$

$$\Rightarrow S_W^{-۱} S_B W - J(W) W = ۰$$

$$S_W^{-۱} S_B W = \lambda W$$

که

$$\lambda = J(W) = scalar$$

می توان مسئله را با استفاده از مقادیر ویژه حل کنیم اما در اینجا بدون استفاده از مقادیر ویژه مسئله بالا را با استفاده از روش زیر حل می کنیم:

$$J(W) = \frac{W^T S_B W}{W^T S_W W} = \frac{(W^T(\mu_۱ - \mu_۲)(\mu_۱ - \mu_۲)^T W)}{W^T S_W W} = \frac{(W^T(\mu_۱ - \mu_۲))^۲}{W^T S_W W}$$

$$\frac{\partial J}{\partial W} = \frac{۲(\mu_۱ - \mu_۲)(W^T(\mu_۱ - \mu_۲)W^T S_W W - ۲ S_W W (W^T(\mu_۱ - \mu_۲))^۲}{(W^T S_W W)^۲}$$

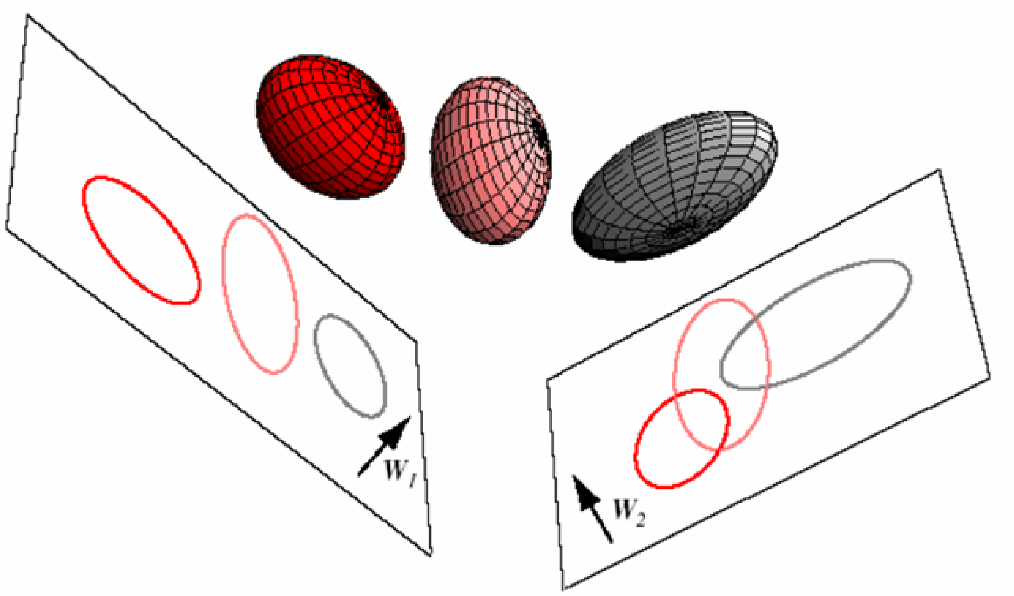
$$= \frac{W^T(\mu_۱ - \mu_۲)[۲(\mu_۱ - \mu_۲)W^T S_W W - W^T(\mu_۱ - \mu_۲) S_W W]}{(W^T S_W W)^۲}$$

$$= \frac{W^T(\mu_۱ - \mu_۲)}{W^T S_W W}[۲(\mu_۱ - \mu_۲) - \frac{W^T(\mu_۱ - \mu_۲) S_W W}{W^T S_W W}]$$

اگر $\frac{W^T(\mu_۱ - \mu_۲)}{W^T S_W W}$ را مقدار c در نظر بگیریم:

$$= (\mu_۱ - \mu_۲) - c S_W W = ۰ \rightarrow W = ۲ c S_W^{-۱}(\mu_۱ - \mu_۲)$$

$$\rightarrow W \propto S_W^{-۱}(\mu_۱ - \mu_۲) \quad (۹\text{-}۳)$$



شکل ۳-۳: کاهش بعد کلاسها

رابطه ۳-۹ تشخیص خطی فیشر نامیده می شود اگر چه یک تابع جداساز نیست اما جهتی را مشخص می کند که با نگاشت داده ها در آن جهت به فضای یک بعدی خواهیم رسید که داده ها به راحتی با یک حد آستانه جدا می‌شوند.

الگوریتم فیشر را میتوان به ابعاد بالاتر نیز گسترش داد. در شکل ۳-۳ داده های ۳ بعدی به زیر فضا های ۲بعدی تصویر شده اند (از فضای D بعدی به $D-۱$ بعدی) که توسط بردارهای $W_۱$، $W_۲$ توصیف میشوند. در اینحالت زیرفضا های بهینه به نحوی پیدا میشوند که حداکثر جداسازی را برای یک مقدار ثابت از واریانس داخل کلاسی داشته باشند. در این شکل $W_۱$ بخوبی از عهده اینکار بر می آید.

۳-۲-۱ کاهش بعد مرحله ای از طریق تجزیه QR (IDR/QR)[۱]

در ابتدای بحث برای راحتی کار نمادهای مهم مورد استفاده در این بخش را در جدول ۳-۱ معرفی می کنیم. فرض کنید $A \in \mathbb{R}^{d\times n}$ ماتریسی از داده ها باشد. با استفاده از معادله $y_j = W^T a_j$ می توان ستون های ماتریس A یعنی a_j ها را از فضای d- بعدی به فضای ℓ - بعدی تصویر کرد.که این امر توسط تبدیل خطی $W \in \mathbb{R}^{d\times\ell}$ انجام می گردد.

همانطور که بیان شد هدف LDA پیدا کردن تبدیلات خطی و کاهش فضا های با ابعاد زیاد با حفظ ساختار آن‌ها است ،با توجه به این امر برای دستیابی به مقصود که همان یافتن ماتریس تبدیلات خطی W است ازLDA استفاده شده است.

[۱] An Incremental Dimension Reduction Algorithm via QR Decomposition

جدول ۳-۱: نمادهای مهم مورد استفاده در بخشIDR/QR

نمادها	توضیحات
n	تعداد نقاط داده های آموزشی
n_i	تعداد داده ها در کلاس iام
d	بعد داده های آموزشی
c	تعداد کلاس ها
W	ماتریس تبدیلات
A	ماتریس داده ها
a_j	ستون های ماتریس داده ها
A_i	ماتریس داده ها در کلاس iام
S_B	ماتریس پراکندگی بین کلاسی
S_W	ماتریس پراکندگی درون کلاسی
C	ماتریس مرکزی
m_i	میانگین داده ها در کلاس iام
m	میانگین کل دادها

فرض کنید ماتریس داده های A به c کلاس تقسیم بندی شده است .بنابراین

$$A = [A_۱, ..., A_c]$$

که $A_i \in \mathbb{R}^{d \times n_i}$ و

$$\sum_{i=۱}^{c} n_i = n$$

I_i ها را مجموعه ای از اندیس ستون هایی که متعلق به کلاس i ام هستند در نظر میگیریم:

به طور کلی اگر هر کلاس یک گروه منسجم باشد به عبارتی واریانس درون گروهی هر چه کمتر (خوشه ها متراکم تر)داده ها بهتر دسته بندی می شوند و همچنین فاصله میانگین دسته ها هر چه بیشتر باشد به خوبی کلاس ها از یکدیگر جدا می شوند.

برای بررسی این ویژگی کلاس ها ،ماتریس کواریانس درون کلاسی (S_W) و ماتریس کواریانس بین کلاسی (S_B) به صورت زیرنمایش داده می شود.

$$S_W = \sum_{i}^{c} \sum_{j \in I_i} (a_j - m_i)(a_j - m_i)^T$$

$$S_B = \sum_{i}^{c} \sum_{j \in I_i} (m_i - m)(m_i - m)^T$$

$$= \sum_{i}^{c} n_i(m_i - m)(m_i - m)^T$$

m_i میانگین iامین کلاس و m میانگین کل داده ها می باشد.

ماتریس های H_W, H_B را به صورت زیر تعریف می کنیم.

$$H_W = [A_1 - m_1.e_1^T, ...A_c - m_c.e_c^T] \in \mathbb{R}^{d\times n} \quad (3\text{-}10)$$

$$H_B = [\sqrt{n_1}(m_1 - m), ..., \sqrt{n_k}(m_k - m)] \in \mathbb{R}^{d\times n} \quad (3\text{-}11)$$

که

$$e_i = (1, ..., 1)^T \in \mathbb{R}^{n_i}$$

می توان S_W و S_B را به صورت زیر نیز تعریف کرد.

$$S_W = H_W H_W^T$$

$$S_B = H_B H_B^T$$

تعریف ۳-۲-۱. در جبر خطی تریس [1] یک ماتریس مربعی $n \times n$ برابر است با حاصلجمع درایه های قطر اصلی آن یا به عبارت دیگر

$$Tr(A) = a_{11} + a_{22} + ... + a_{nn} = \sum_{i=1}^{n} a_{ii}$$

که a_{ii} درایه های واقع در سطر iام و ستون iام ماتریس A است.

همچنین تریس یک ماتریس برابر مجموع مقادیر ویژه آن نیز می باشد. قابل ذکر است که تریس فقط برای ماتریس های مربعی تعریف می شود:

تریس ماتریس های پراکندگی را به صورت زیر محاسبه می شود:

$$trace(S_W) = \sum_{i=1}^{c} \sum_{j\in I_i} ||a_j - m_i||_2^2$$

[1] trace

$$trace(S_B) = \sum_{i=1}^{c} \sum_{j \in I_i} ||m_i - m||_2^2$$

ماتریس های پراکندگی تصویر شده (افکنش شده) با استفاده از تبدیل خطی W به صورت زیر می باشند.

$$\tilde{S}_W = (W^T H_W)(W^T H_W)^T = W^T S_W W$$

$$\tilde{S_B} = (W^T H_B)(W^T H_B)^T = W^T S_B W$$

تبدیل خطی بهینه W، $trace(\tilde{S}_B)$ را بیشینه و $trace(\tilde{S}_W)$ را کمینه خواهد کرد. یک بهینه سازی متداول در روش LDA پیدا کردن W مناسب با استفاده از معادله زیر است .

$$W = \arg \max_{w_i^T S_W w_j = 0, \forall i \neq j} \text{trace}((W^T S_W W)^{-1}(W^T S_B W)) \qquad (۱۲-۳)$$

که این معادله همان تجزیه خطی فیشر است.

با توضیحات فوق به بیان الگوریتم IDR/QR [۱۰] می‌پردازیم.

هدف الگوریتم IDR/QR حل مسئله بهینه سازی (۳-۱۳) است.

$$W = \arg \max_{W^T W = I} \text{trace}(W^T S_B W) \qquad (۱۳-۳)$$

توجه داشته باشید که این بهینه سازی تنها به حداکثر رساندن فاصله بین کلاس ها مربوط می شود. یک راه حل برای معادله (۳-۱۳) استفاده از تجزیه QR روی ماتریس مرکزی C است که به اصطلاح روش متعامد مرکزی OCM[۱] [۱۱] نامیده می شود که در ادامه به آن خواهیم پرداخت.

ماتریس مرکزی C به صورت زیر تعریف می شود:

$$C = [m_1, m_2, ..., m_c] \qquad (۱۴-۳)$$

با توجه به تجزیه QR ماتریس C که در آن $Q \in R^{n \times c}$ دارای ستون های یکامتعامد است و $R \in R^{c \times c}$ یک

[۱] Orthogonal Centroid Method

ماتریس بالا مثلثی است.در روش متعامد مرکزی داریم:

$$W = QM \tag{۱۵-۳}$$

اکنون می توان مسئله بهینه سازی معادله (۳-۱۳) را با کمک گرفتن از ماتریس متعامد M حل کرد.انتخاب ماتریس M اختیاری است. رابطه

$$trace(W^T S_B W) = trace(M^T W^T S_B W M)$$

برای هر ماتریس متعامد M برقرار است.

در روش متعامد مرکزی OCM [۱۱] برای سادگی کار M یک ماتریس همانی در نظر گرفته می شود.

مرحله دوم الگوریتم IDR/QR با درنظر گرفتن فاصله بین کلاسی مرحله اول را بهبود می بخشد. در این مرحله ما به دنبال یافتن ماتریس تبدیلات W به نحوی هستیم که $W = QM$ باشد .دقت داشته باشید در اینجا لازم نیست M متعامد باشد.

$$(W^T S_B W) = (M^T Q^T S_B Q M),$$

$$(W^T S_W W) = (M^T Q^T S_W Q M),$$

مسئله بهینه سازی اصلی پیدا کردن مقدار W بهینه است که معادل است با پیدا کردن M ، با فرض اینکه

$$B = Q^T S_B Q \ \ , \ \ F = Q^T S_W Q$$

نکته اینکه اندازه B از ماتریس پراکنده S_B کوچکتر است.(به طور مشابهه برای F نیز برقرار است). هنگامی که B، F داری اندازه های کوچکتری از $c \times c$ باشد میتوان M را از روش های مختلف موجود در LDA بدست آورد. می‌توان از روش LDA منظم برای محاسبه M استفاده کرد.در اینجا M را با حل کردن مسئله مقادیر ویژه کوچک روی $(F + \mu I_c)^{-۱} B$ برای مقدار ثابت μ محاسبه می شود.شبه کد در الگوریتم (۳-۵) آورده شده است.

الگوریتم ۳-۵ Batch IDR/QR

Require: data matrix A;

Ensure: : optimal transformation matrix W;

/*Stage 1:*/

1: construct centorid matrix C;

2: Compute QR Decomposition of C as C = QR;
where $Q \in \mathbb{R}^{d\times c}, R \in \mathbb{R}^{c\times c}$;

/* Stage II: */

3: $Z \leftarrow H_W^T Q$;

4: $Y \leftarrow H_B^T Q$;

5: $F \leftarrow Z^T Z$;/*Reduced within-class scatter matrix*/

6: $B \leftarrow Y^T Y$;/*Reduced between-class scatter matrix*/

7: Compute the c eigenvectors $\phi_i$of $(F + \mu I_c)^{-1} B$,
with decreasing eigenvalues;

8: $W \leftarrow QM$ where $M = [\phi_1, ..., \phi_c]$.

۳-۲-۲ IDR/QRافزایشی [۱]

در این بخش جزئیات الگوریتم IDR/QR افزایشی بررسی می شود. قراردادهای زیررا اتخاذ می کنیم.

هر متغییر X پس از درج یک نمونه جدید و بروز رسانی شدن مدل با $\tilde{X}$ نشان داده می شود.

برای مثال: تعداد n_i از اعضای کلاس iام با $\tilde{n}_i$ و همچنین m_i مرکزی با $\tilde{m}_i$ نشان داده می شود.

با درج یک نمونه جدید در ماتریس مرکزی C ، H_W, H_B با F ، B نمایش داده می شوند.

بروزرسانی الگوریتم افزایشی در سه مرحله انجام می شود:

مرحله(۱) بروزرسانی تجزیهQR ماتریس مرکزی $C = [m_1, ..., m_k]$} خط ۲الگوریتم(۳-۵)}

مرحله (۲) بروزرسانی ماتریس پراکندگی کاهشی درون کلاسی (F){خط ۵لگوریتم(۳-۵)}

مرحله (۳) بروزرسانی ماتریس پراکندگی کاهشی بین کلاسی (B){خط۶ الگوریتم(۳-۵)}

فرض کنیدx یک نمونه جدید درج شده است، که متعلق به کلاسiام است. بدون از دست دادن کلیت مسئله فرض می کنیم قبل از وارد کردن متغییر x داده ها از کلاس ۱تا k امین کلاس پشت سر هم قرار داده شده است .به طور کلی می توان با تعویض برچسب بین طبقات مختلف کلاسی این کار را انجام داد.

می توان با توجه به اضافه کردن یک متغییر جدید دو حالت مختلف برای بروزرسانی IDR/QR افزایشی در نظر گرفت.

حالت اول:متغییر x متعلق به یک کلاس موجود باشد: $i \leq k$

[۱] Incremental IDR/QR

حالت دوم :متغییر x متعلق به یک کلاس جدید باشد $i > k$.

تکنیک بررسی این دو حالت کاملا متفاوت است .در این بخش به بررسی حالت اول می پردازیم. توجه دارید زمانی که میخواهیم یک نمونه جدید به یک کلاسی که وجود دارد $(i \leq k)$ اضافه کنیم از آنجا که کلاس اول تا k ام پشت سر هم قرار دارندو x جدید متعلق به کلاس i ام است و $۱ \leq i \leq k$ است پس درج x یک کلاس جدید ایجاد نمی کند.

همان طور که بیان شد بروزرسانی الگوریتم افزایشی در سه مرحله انجام می شودکه حال با توجه به اضافه کردن یک نمونه جدید مراحل آن بررسی می شود.

مرحله اول:به روز رسانی تجزیه QR ماتریس مرکزی C

هنگامی که x درج شده متعلق به کلاس i ام باشد .در این حالت

$$\tilde{C} = [m_۱, ..., m_i + f, ..., m_k]$$

$$f = \frac{x - m_i}{\tilde{n_i}} \quad \tilde{n_i} = n_i + ۱$$

بنابراین $\tilde{C}$ می تواند به صورت (۳-۱۶) بازنویسی شود

$$\tilde{C} = C + f \cdot g^T \quad g = (۰, ..., ۱, ..., ۰)^T \tag{۳-۱۶}$$

۱ در موقعیت i ام قرار دارد. مسئله بروزرسانی تجزیه QR ماتریس مرکزی C به صورت زیر فرمول بندی می شود.

$$C = QR, Q \in \mathbb{R}^{d \times k}, R \in \mathbb{R}^{k \times k}$$

که تجزیه QR، $\tilde{C}$ را می توان با استفاده از تجزیه QR ماتریس C محاسبه کرد.

هنگامی که

$$\tilde{C} = C + f.g^T$$

بروزرسانی تجزیه QR ماتریس مرکزی C رامی توان با استفاده از بروزرسانی رتبه-یک QR فرمول بندی کرد. برای این کار نمی توانیم از روش ارائه شده در کتاب گلوب [۸] استفاده کنیم چون در آنجا از تجزیه QR کامل استفاده می شود به عبارت دیگر ماتریس Q باید مربعی باشد.در حالی که در مسئله مورد بحث در اینجا Q یک ماتریس مستطیلی است.به این منظور با تغییر مختصری در الگوریتم ارائه شده در [۱۲] تجزیه QR را در دو مرحله

می توان محاسبه کرد.

مرحله اول: بروزرسانی رتبه–یک بر روی یک ماتریس کوچک .

مرحله دوم:بروزرسانی QR با اضافه کردن یک سطر جدید. جزئیات در ادامه آورده شده است.

f به دو قسمت تقسیم بندی می شود:افکنش آن روی پایه های متعامد Q و مکمل متعامد آن.

f را به صورت زیر تعریف می شود :

$$f = QQ^T + (I - QQ^T)f$$

باتوجه به تعریف ماتریس متعامد خواهیم داشت

$$(Q^T - Q^TQQ^T)f = (Q^T - Q^T)f = ۰ \times f = ۰$$

در نتیجه

$$Q^T(I - QQ^T)f = ۰$$

پس می توان بیان کرد که $(I - QQ^T)f$ بر روی فضای تولید شده توسط ستون های Q متعامد است.

در نتیجه

$$\tilde{C} = C + f.g^T$$

$$= QR + QQ^Tf.g^T + (I - QQ^T)f.g^T$$

$$= Q(R + f_۱.g^T) + f_۲.g^T.$$

که

$$f_۱ = Q^Tf, \quad f_۲ = (I - QQ^T)f$$

در ادامه ، تجزیه QR ، $\tilde{C}$ را بدست می آوریم.

تجزیه QR ماتریس $Q(R + f_۱.g^T)$ بروزرسانی می شود که متناظر با یک بروزرسانی رتبه –یک است.

نتیجه تجزیه QR بروزرسانی شده به صورت

$$Q(R + f_۱.g^T) = Q_۱R_۱$$

می باشد که در آن $Q_1 = QP_1$ و $P_1 \in R^{k\times k}$ متعامد است.

فرض کنید $||f_2|| \neq 0$. قرار می دهیم $q = \frac{f_2}{||f_2||}$. از آنجا که q بر زیر فضای ساخته شده بوسیله ستون های Q متعامد است لذا Q بر زیر فضای ساخته شده از ستون های $Q_1 = QP_1$ هم متعامد است. .به عبارت دیگرستون های $[Q_1, q]$ متعامد هستند.

بعد بروز رسانی تجزیه QR ماتریس زیر محاسبه می شود.

$$\tilde{C} = [Q_1, q] \begin{pmatrix} R_1 \\ ||f_2||g^T \end{pmatrix}$$

که متناظر با حالتی است که $||f_2||g^T$ به عنوان یک سطر جدید درج می شود.

تجزیه QR بروزرسانی شده به صورت زیر است .

$$[Q_1, q] \begin{pmatrix} R_1 \\ ||f_2||g^T \end{pmatrix} = [\tilde{Q}, \tilde{q}] \begin{pmatrix} \tilde{R} \\ 0 \end{pmatrix} = \tilde{Q}\tilde{R},$$

که

$$[\tilde{Q}, \tilde{q}] = [Q_1, q]P_2$$

برای برخی از ماتریس های متعامد p_2 برقرار است. سرانجام ،تجزیه QR بروزرسانی شده ماتریس $\tilde{C}$ را به صورت زیر خواهیم داشت(با فرض $||f_2|| \neq 0$.)

$$\tilde{C} = Q_1R_1 + ||f_2||q.g^T = [Q_1, q] \begin{pmatrix} \tilde{R} \\ ||f_2||g^T \end{pmatrix} = \tilde{Q}\tilde{R}$$

اگر $||f_2|| = 0$ پس $\tilde{C} = Q_1R_1$ که تجزیه QR بروزرسانی شده $\tilde{C}$ است.

گام ۲:بروزرسانی F

در این مرحله بروزرسانی ماتریس پراکندگی درون کلاسی کاهش بعد یافته (F)بررسی می شود.

$$F = Q^T H_W H_W^T Q$$

فرض کنید $\tilde{F}$ را فرم بروزرسانی شده F در نظر بگیرید:

$$\tilde{F} = \tilde{Q}^T \tilde{H}_W \tilde{H}_W^T \tilde{Q}$$

است. توجه دارید که

$$H_W = [A_۱ - m_۱.e_۱^T, ..., A_k - m_k.e_k^T] \in \mathbb{R}^{d \times n}$$

که بروزرسانی شده آن را با $\tilde{H}_w$ نشان می دهند .تفاوت $\tilde{H}_w$ و H_W بر روی i ام بلوک آن است. فرض کنید i امین بلوک H_W به صورت $H_i = A_i - m_i e_i^T$نشان داده شود .

بعد از بروزرسانی $\tilde{H}_W$، i امین بلوک آن را به صورت زیر خواهیم داشت.

$$\tilde{H}_i = \tilde{A}_i - \tilde{m}_i.\tilde{e}_i^T = [A_i, x] - \tilde{m}_i.\tilde{e}_i^T$$

$$= [A_i - m_i.e_i^T, x - m_i] - (\tilde{m}_i - m_i).\tilde{e}_i^T$$

$$= [H_i, u] - v.\tilde{e}_i^T, \tag{۱۷-۳}$$

که در آن

$$u = x - m_i, v = \tilde{m}_i - m_i, \tilde{e}_i = \begin{pmatrix} e_i \\ ۱ \end{pmatrix} \in \mathbb{R}^{n_i+۱}.$$

حال می توان $\tilde{H}_i \tilde{H}_i^T$ را به صورت زیر محاسبه کرد .

$$\tilde{H}_i \tilde{H}_i^T = ([H_i, u] - v.\tilde{e}_i^T)([H_i, u] - v.\tilde{e}_i^T)^T$$

$$= [H_i, u] \begin{pmatrix} H_i^T \\ u^T \end{pmatrix} - v.\tilde{e}_i^T \begin{pmatrix} H_i^T \\ u^T \end{pmatrix}$$

$$-[H_i, u]\tilde{e}_i.v^T + (v.\tilde{e}_i^T)(\tilde{e}_i.v^T)$$

$$= H_i H_i^T + u.u^T - v.u^T - u.v^T + (n_i + ۱)v.v^T$$

$$= H_iH_i^T + (u-v).(u-v)^T + n_iv.v^T, \quad (۱۸\text{-}۳)$$

که از روابط زیر در استنتاج فوق استفاده شده است.

$$[H_i, u]\tilde{e_i} = \sum_{j\in I}(a_j - m_i) + u = u$$

$$(v.\tilde{e_i}^T)(\tilde{e_i}.v^T) = vv^T(\tilde{e_i}^T.\tilde{e_i}) = (n_i + ۱)vv^T$$

.

از آنجا که

$$H_WH_W^T = \sum_{j=۱}^{k} H_jH_j^T$$

خواهیم داشت

$$\tilde{H}_W\tilde{H}_W^T = \sum_{j=۱}^{k} \tilde{H}_j\tilde{H}_j^T$$

$$= \sum_{۱\le j\le k, j\ne i} \tilde{H}_j\tilde{H}_j^T + \tilde{H}_i\tilde{H}_i^T$$

$$= \sum_{j=۱}^{k} H_jH_j^T + (u-v).(u-v)^T + n_iv.v^T.$$

در نتیجه خواهیم داشت:

$$\tilde{F} = \tilde{Q}^T\tilde{H}_W\tilde{H}_W^T\tilde{Q}$$

$$= \tilde{Q}^T\tilde{H}_W\tilde{H}_W^T\tilde{Q} + \tilde{Q}^T(u-v).(u-v)^T\tilde{Q} + n_i\tilde{Q}^Tv.v^T\tilde{Q}$$

$$= \tilde{Q}^T\tilde{H}_W\tilde{H}_W^T\tilde{Q} + (\tilde{u}-\tilde{v}).(\tilde{u}-\tilde{v})^T + n_i\tilde{v}.\tilde{v}^T$$

$$\approx Q^TH_W{H_W}^TQ + (\tilde{u}-\tilde{v}).(\tilde{u}-\tilde{v})^T + n_i\tilde{v}.\tilde{v}^T$$

$$= F + (\tilde{u}-\tilde{v}).(\tilde{u}-\tilde{v})^T + n_i\tilde{v}.\tilde{v}^T, \quad (۱۹\text{-}۳)$$

که $\tilde{u} = \tilde{Q}^Tu$ ، $\tilde{v} = \tilde{Q}^Tv$. فرض بر این است در تقریب (۳-۱۹) $\tilde{Q}$ همان Q بروزرسانی شده می باشد که

نمونه جدیدی در آن درج شده است .

مرحله۳:بروزرسانی B

در گام پایانی، بروزرسانی ماتریس پراکندگی بین کلاسی کاهش بعد یافته مورد بررسی قرار می گیرد. همان گونه که بیان شد

$$B = Q^T H_B {H_B}^T Q$$

که نسخه بروزرسانی شده آن به صورت

$$\tilde{B} = \tilde{Q}^T \tilde{H}_B \tilde{H}_B^T \tilde{Q}$$

است. نکته کلیدی در بروزرسانی B آن است که $\tilde{H}_B$ به صورت معادله ی (۳-۲۰) تعریف شود

$$\tilde{H}_B = [\sqrt{\tilde{n_1}}(\tilde{m_1} - \tilde{m}), ..., \sqrt{\tilde{n_k}}(\tilde{m_k} - \tilde{m})] \quad (۳-۲۰)$$

سپس می توان معادله (۳-۲۰)را به صورت زیر بازنویسی کرد:

$$\tilde{H}_B = [\tilde{m_1}, \tilde{m_2}, ..., \tilde{m_k}, \tilde{m}]F = [\tilde{C}, \tilde{m}]F,$$

که

$$F = \begin{pmatrix} D \\ -h^T \end{pmatrix} \quad , D = diag(\sqrt{n_1}, ..., \sqrt{n_k} \quad , h = [\sqrt{\tilde{n_1}}, ..., \sqrt{\tilde{n_k}}]^T$$

با استفاده از تجزیه QR بروزرسانی شده $\tilde{C} = \tilde{Q}\tilde{R}$، خواهیم داشت:

$$\tilde{Q}^T \tilde{H}_B = [\tilde{Q}^T \tilde{C}, \tilde{Q}^T \tilde{m}]F = [\tilde{R}, \tilde{Q}^T \tilde{m}]F = \tilde{R}D - \tilde{Q}^T \tilde{m}.\tilde{h}^T.$$

به آسانی بررسی می شود که

$$\tilde{m} = \frac{1}{\tilde{n}} \tilde{C}.r$$

که

$$r = (\tilde{n}_1, ..., \tilde{n}_k)^T.$$

الگوریتم ۳-۶ بروزرسانی کلاس های موجود

Require: centroid matrix $C = [m_1, m_2, ...m_k]$ ،its QR Decomposition $C = QR$،the matrix W،the size n_j of thej_{th} class for each j،and a new point x from the i-the class،$i \leq k$

Ensure: updated matrix$\tilde{F}$,updated centroid matrix $\tilde{C}$,its QR Decomposition $\tilde{C} = \tilde{Q}\tilde{R}$ and updated matrix $\tilde{B}$.

$\tilde{n}_j \leftarrow n_j$،for $j \neq i$;$\tilde{n}_i \leftarrow n_i + 1$;$f \leftarrow \frac{x-m_i}{\tilde{n}_i}$;

$\tilde{m}_i \leftarrow m_i + f$;$\tilde{m}_j \leftarrow m_j$,for each $j \neq i$;

$\tilde{C} \leftarrow [\tilde{m}_1, ..., \tilde{m}_i, ..., \tilde{m}_k]$

$f_1 \leftarrow Q^T f$;$f_2 \leftarrow (I - QQ^T)f$;

do rank-one QR-updating of $Q(R + f_1.g^T)$as $Q(R + f_1.g^T) = Q_1 R_1$

if $\|f_2\| = 0$ **then**

  $\tilde{Q} \leftarrow Q_1$;$\tilde{R} \leftarrow R_1$;

else

  $q \leftarrow \frac{(I-QQ^T)f}{||(I-QQ^T)||}$; $g \leftarrow (0, ..., 1, ..., 0)^T$

  do QR -updating $of [Q_1, q] \begin{pmatrix} R_1 \\ ||f_2||g^T \end{pmatrix}$ as $[Q_1, q] \begin{pmatrix} R_1 \\ ||f_2||g^T \end{pmatrix} = Q_2 R_2$;

  $\tilde{Q} \leftarrow Q_2$; $\tilde{R} \leftarrow R_2$

end if

$u \leftarrow x - m_i$; $v \leftarrow \tilde{m}_i - m_i$;

$\tilde{u} \leftarrow Q^T u$, $\tilde{v} \leftarrow \tilde{Q}^T v$

$\tilde{F} \leftarrow F + (\tilde{u} - \tilde{v}).(\tilde{u} - \tilde{v})^T + n_i \tilde{v}.\tilde{v}^T$

$D \leftarrow diag(\sqrt{n_1}, ..., \sqrt{n_k}$, $h \leftarrow [\sqrt{\tilde{n}_1}, ..., \sqrt{\tilde{n}_k}]^T$;

$r \leftarrow (\tilde{n}_1, ..., \tilde{n}_k)^T$; $\frac{1}{\tilde{n}}\tilde{R}.r$;

$\tilde{B} \leftarrow (\tilde{R}D - \tilde{r}.h^T)(\tilde{R}D - \tilde{r}.h^T)^T$;

بنابراین

$$\tilde{Q}\tilde{m} = \tilde{Q}^T \frac{۱}{\tilde{n}} \tilde{C}.r = \frac{۱}{\tilde{n}} \tilde{R}.r.$$

که از آن نتیجه می شود

$$\tilde{B} = \tilde{Q}^T \tilde{H}_B \tilde{H}_B^T \tilde{Q} = (\tilde{R}D - \tilde{Q}^T \tilde{m}.h^T)(\tilde{R}D - \tilde{Q}^T \tilde{m}.h^T)^T$$

$$(\tilde{R}D - (\frac{۱}{\tilde{n}}\tilde{R}.r).h^T)(\tilde{R}D - (\frac{۱}{\tilde{n}}\tilde{R}.r).h^T)^T.$$

می توان مراحل کار را در الگوریتم ۳-۶ مشاهده کرد. با به دست آوردن $\tilde{B}$ می توان W را مطابق الگوریتم ۳-۵ محاسبه کرد.

۳-۳ ELM افزایشی بر اساس تجزیه QR

شبکه عصبی ELM [۱] [۱۳] یکی از انواع جدید شبکه عصبی می باشد که در صورت افزایش ورودی ها، وزنهای شبکه عصبی یا باید دوباره محاسبه شوند و یا وزنهای قبلی بروزرسانی شوند. در این بخش نحوه بروزرسانی وزنهای شبکه عصبی بر اساس وزنهای قبلی با استفاده از تجزیه QR بیان می شود[۱۴]. همان گونه که در فصل ۲ بیان کردیم روش های مختلفی برای تجزی QR وجود دارد.با دقت و آزمایشاتی که انجام گرفته است روش گرام اشمیت برای توسعه الگوریتم افزایشی مناسب تر است.

روش گرام اشمیت ،روشی است برای متعامد سازی مجموعه ای از بردارها در یک فضای داخلی.نکته قابل توجه در روش گرام اشمیت آن است که ماتریس باید رتبه کامل باشد یا به عبارتی ستون ها باید مستقل خطی باشند. فرض کنید در یک شبکه عصبیELM لایه پنهان با ماتریس H نمایش داده شود. می خواهیم تجزیه QR ماتریس H را محاسبه کنیم.

$$H = QR$$

Q متعامد و Rیک ماتریس بالا مثلثی است. فرض کنید

$$H = [h_۱, h_۲, ..., h_k], Q = [q_۱, q_۲, ..., q_k]$$

و

$$R = \begin{bmatrix} r_{۱۱} & r_{۱۲} & \dots & r_{۱k} \\ & r_{۲۲} & \dots & r_{۲k} \\ & & \ddots & \vdots \\ & & & r_{kk} \end{bmatrix}$$

با توجه به اینکه تجزیه QR ماتریس H را می توان به صورت $H = QR$ نوشت خواهیم داشت:

$$[h_۱, h_۲, ..., h_k] = [q_۱, q_۲, ..., q_k]. \begin{bmatrix} r_{۱۱} & r_{۱۲} & \dots & r_{۱k} \\ & r_{۲۲} & \dots & r_{۲k} \\ & & \ddots & \vdots \\ & & & r_{kk} \end{bmatrix} \qquad (۳-۲۱)$$

[۱] Extreme Learning Machine

با استفاده از رابطه بالا می توان ستون های k...۱ ماتریس لایه پنهان H را به صورت زیر محاسبه کرد.

$$h_1 = q_1.r_{11} \tag{۳-۲۲}$$

$$h_2 = q_1.r_{12} + q_2.r_{22} \tag{۳-۲۳}$$

$$\vdots$$

$$h_k = q_1.r_{1k} + ... + q_k.r_{kk} \tag{۳-۲۴}$$

از آنجا که Q یک ماتریس متعامد است خواهیم داشت.

$$Q^TQ = I \rightarrow q_i^T q_j = \begin{cases} 0 \ , i \neq j \\ 1 \ , i = j \end{cases}$$

حال می توان درایه های ماتریس بالا مثلثی R را به صورت زیر بدست آوریم.

$$q_1^T q_1 = 1 = \frac{h_1^T}{r_{11}}.\frac{h_1}{r_{11}} \rightarrow r_{11} = \sqrt{h_1^T h_1} \tag{۳-۲۵}$$

$$q_2^T q_1 = 0$$

$$q_1^T h_2 = q_1^T q_1 r_{12} + q_1^T q_2 r_{22}$$

$$q_1^T h_2 = q_1^T q_1 r_{12} + 0$$

با توجه به اینکه $q_1^T q_1 = 1$، داریم $r_{12} = q_1^T h_2$ به طور کلی اگر $1 \leq i < k$ خواهیم داشت:

$$r_{ik} = q_i^T h_k \ \ i \neq k \tag{۳-۲۶}$$

حال قرار می دهیم $\delta h = q_k.r_{kk} = h_k - (q_۱.r_{۱k} + ... + q_{k-۱}.r_{k-۱})$ بنابراین حالت کلی به صورت زیر است.

$$r_{kk} = \sqrt{\delta h_k^T \delta h_k} \quad , \quad q_k = \delta h_k / r_{kk}$$

فرض کنید می خواهیم یک نود به k نود موجود اضافه کنیم. در واقع می خواهیم ستون $h_{k+۱}$ را به ماتریس H_k اضافه کنیم پس خواهیم داشت:

$$H_{k+۱} = [H_k | h_{k+۱}]$$

تجزیه QR، $H_{k+۱}$ بصورت

$$H_{k+۱} = Q_{k+۱} R_{k+۱}$$

می باشد که

$$Q_{k+۱} = [Q_k | q_{k+۱}]$$

و

$$R_{k+۱} = \left[\begin{array}{c|c} R_k & \delta r_{k+۱} \\ ۰ & r_{k+۱,k+۱} \end{array}\right] = \left[\begin{array}{cccc|c} r_{۱۱} & r_{۱۲} & \dots & r_{۱k} & r_{۱,k+۱} \\ & r_{۲۲} & \dots & r_{۲k} & r_{۲,k+۱} \\ & & \ddots & \vdots & \vdots \\ & & & r_{kk} & r_{k,k+۱} \\ & & ۰ & & r_{k+۱,k+۱} \end{array}\right]$$

با توجه به معادله ۲۶ − ۳ داریم:

$$\delta r_{k+۱} = Q_k^T h_{k+۱} \qquad (۲۷-۳)$$

$$\delta_{k+۱} = q_{k+۱}.r_{k+۱,k+۱} = h_{k+۱} - (q_۱.r_{۱,k+۱} + ... + q_k.r_{k+۱,k+۱}) = h_{k+۱} - Q_k \delta r_{k+۱}$$

$$= h_{k+۱} - Q_k \delta r_{k+۱}$$

$$r_{k+۱,k+۱} = \sqrt{\delta h_{k+۱}{}^T \delta h_{k+۱}}$$

$$q_{k+۱} = \delta h_{k+۱} / r_{k+۱,k+۱}$$

تعریف ۳-۳-۱. برای هر ماتریس متناهیA با درایه های مختلط یک ماتریس B وجود دارد که در چهار معادله زیر صدق می کند:

- $ABA = A$
- $BAB = B$
- $(AB)^* = AB$
- $(BA)^* = BA$

چنین ماتریسی ،معکوس مور -پنروز A نامیده می شود و با $A^\dagger$ نمایش داده می شود.

با توجه به تعریف می توانیم معکوس H را بصورت زیر نشان داد.

$$H^\dagger = R^{-1}Q^T \tag{۳-۲۸}$$

اساس این روش محاسبه R_{k+1}^{-1} از روی R_k^{-1} است که تعداد گره های پنهان k را افزایش می دهد که در زیر روش کار بیان شده است . $\tilde{\beta}_{k+1}$ از $\tilde{\beta}_k$ و R_k^{-1} هستیم.

$$R_{k+1}^{-1} = \left[\begin{array}{c|c} R_k & \delta r_{k+1} \\ 0 & r_{k+1,k+1} \end{array}\right]^{-1} = \left[\begin{array}{c|c} R_k^{-1} & -R_k^{-1}\delta r_{k+1} r_{k+1,k+1}^{-1} \\ 0 & r_{k+1,k+1}^{-1} \end{array}\right] \tag{۳-۲۹}$$

همچنین به دنبال گرفتن خروجی وزن $\tilde{\beta}_{k+1}$ از $\tilde{\beta}_k$ و R_k^{-1} هستیم . در یک شبکه عصبی ELM خروجی وزن به صورت

$$\tilde{\beta}_k = H^\dagger.T$$

محاسبه می شود.پس می توان وزن خروجی $\tilde{\beta}_{k+1}$ را به صورت زیر محاسبه کنیم که باعث کاهش تعداد عملیات پیچیدگی محاسباتی می شود.

$$\tilde{\beta}_{k+1} = \left[\begin{array}{c|c} R_k^{-1} & -R_k^{-1}\delta r_{k+1} r_{k+1,k+1}^{-1} \\ 0 & r_{k+1,k+1}^{-1} \end{array}\right] \left[\begin{array}{c} Q_k^T \\ q_{k+1}^T \end{array}\right].T =$$

$$\begin{bmatrix} \tilde{\beta}_k - R_k^{-1}\delta r_{k+1}\beta_{k+1} \\ \beta_{k+1} \end{bmatrix}$$

سرانجام

$$\tilde{\beta}_{k+1} = H_{k+1}^{\dagger}T = R_{k+1}^{-1}Q_{k+1}^{T}T$$

شیوه فوق در بروزرسانی شبکه عصبی ELM با نام QRI-ELM مورد استفاده قرار گرفته است .

فصل ۴

افزایش سرعت شناسایی زمینه ویدیو با محاسبه مرحله‌ای تجزیه QR

۴-۱ مقدمه

شناسایی نواحی متحرک در ویدئو یا پیش زمینه (Foreground) از قسمت ثابت صحنه یا زمینه (Background) یکی از مراحل اساسی در بسیاری از مسایل حوزه پردازش تصویر و بینایی ماشین می‌باشد که تحقیقات متعددی را به خود معطوف نموده است. تفاضل هر فریم ویدیو از زمینه تصویر یک راهکار معمول برای شناسایی اشیاء متحرک است. در بسیاری موارد زمینه ثابت فیلم از قبل مشخص نیست و باید تخمین زده شود. به این زمینه ثابت، تصویر زمینه یا مدل زمینه گفته می‌شود. تخمین دقیق‌تر این مدل منجر به شناسایی بهتر اشیاء خواهد شد که روشهای متعددی در این زمینه پیشنهاد شده است [۱۵].

یکی از روشهای مواجهه با مسئله استفاده از تجزیه QR است که در [۱۶] به آن پرداخته شده است. ایده اصلی در شیوه فوق‌الذکر تبدیل فریمهای ویدیو به یک ماتریس و شناسایی فریمهای زمینه از طریق میزان وابستگی خطی ستونهای ماتریس بوده است. حتی در وضعیتی که دوربین ثابت باشد و هیچ شیء متحرکی در صحنه وجود نداشته باشد، فریمهای این ویدیو کاملا با هم برابر نبوده و مقادیر پیکسلها تا حدودی با هم متفاوت هستند. اگر هر فریم مثلا با ابعاد ۳۲۰×۲۴۰ به صورت یک بردار ستونی با $۷۶۸۰۰ = ۳۲۰ * ۲۴۰$ عنصر دربیاید، هر فریم را می‌توان یک بردار در فضای ۷۶۸۰۰ تصور کرد. نزدیک بودن بردارها در فضا می‌تواند به این معنی باشد که هر دو بردار (فریم) محتوای نسبتا یکسانی دارند. در این فصل فرض بر این است که دوربین ثابت بوده و اشیاء متحرک در صحنه توقف طولانی ندارند؛ مانند ویدیوهای اخذ شده از دوربین‌های نظارتی اتوبانها که شکل ۴-۱

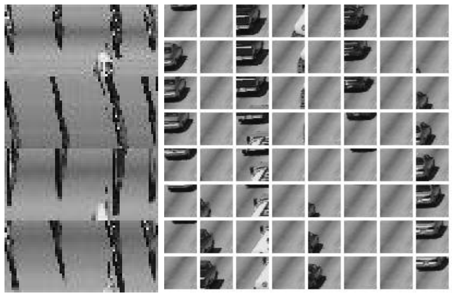
شکل ۴-۱: یک فریم از یک ویدیوی نمونه و زمینه استخراج شده بر اساس روش مورد استفاده و نتیجه شناسایی اشیاء متحرک مبتنی بر زمینه بدست آمده.

یک نمونه را نشان می‌دهد. فرض کنیم بردارهایی که در واقع متعلق به زمینه بوده‌اند به صورت پراکنده در ماتریس A قرار گرفته‌اند. با این فرض که شیء متحرک همرنگ با زمینه نیست، یکی از بردارهای متعلق به زمینه به عنوان یک پایه متعامد در ماتریس Q از تجزیه QR ماتریس A ظاهر خواهد شد و سایر ستونهای متعامد یکه ماتریس Q که مرتبط با این بردارها هستند مقدار R (روی قطر اصلی) بسیار کوچکی خواهند داشت. اگر فرض عدم توقف اشیاء متحرک را داشته باشیم، تمام بردارهای متعامد یکه ماتریس Q که مقدار R کوچکی دارند متعلق به زمینه خواهند بود. به این ترتیب می‌توان این فریمها را شناسایی و مدل زمینه را بر اساس آنها ایجاد کرد.

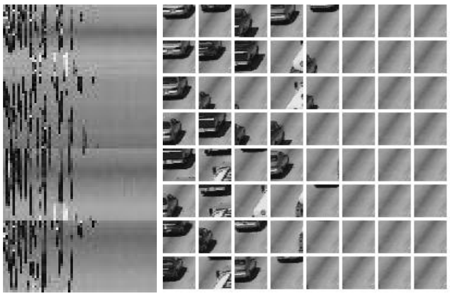
در بسیاری از مواقع همانند شکل ۴-۱ هیچ وقت کل صحنه خالی از اشیاء متحرک نیست، لذا به جای اعمال روش فوق روی کل فریم آن را روی بلاکهای کوچکتری از تصویر اعمال نموده و نتایج را در کنار هم قرار می‌دهیم. شکل ۴-۲ یک بلاک از ۶۴ فریم متوالی ویدیوی شکل ۴-۱ به همراه ماتریس ایجاد شده از کنار هم قرار دادن ستونی‌شده‌ی این فریمها را نشان می‌دهد. در این مثال اندازه هر بلوک ۳۰ در ۳۰ پیکسل بوده است و لذا A یک ماتریس ۶۴ $\times$ ۹۰۰ خواهد بود. برای نمایش بهتر فقط ۱۲۰ سطر ماتریس A (متناظر با عناصر روی قطر اصلی، قطر فرعی و سطر و ستون وسط هر بلاک) نمایش داده شده است.

ترتیب جدید این بلاکها بر اساس مقادیر R تجزیه QR ماتریس مربوط به آنها در شکل ۴-۳ نشان داده شده است. همان‌گونه که مشاهده می‌شود در ترتیب جدید بلاکهای متناظر با زمینه تصویر به انتهای لیست منتقل شده‌اند. نمایش ستونی این بلاکها در تصویر سمت چپ شکل ۴-۳ مؤید همین مطلب است که ستونهای مرتبط با زمینه تصویر به دلیل وابستگی زیادشان به هم در انتها قرار گرفته‌اند.

عمل شناسایی فریمهای زمینه می‌تواند روی ماتریس حاصل از کل ویدیو و یا روی یک برش از ویدیو انجام شود. عموما در کاربردهای برخط به جای اعمال روش روی کل ویدیو، یک پنجره لغزان با طول ثابت روی ویدیو درنظر گرفته می‌شود و مدل زمینه بر اساس فریمهای موجود در این پنجره برآورد می‌گردد. سپس فریمهای موجود در پنجره بروزرسانی شده و مجددا عملیات تکرار می‌شود. به عنوان مثال اگر طول پنجره N فریم باشد، عملیات برای فریمهای ۱ تا N انجام شده (ماتریس A و مدل زمینه ساخته می‌شود) سپس با دریافت فریمهای جدید، ماتریس بر اساس فریمهای ۲ تا $N + ۱$ ساخته شده و عملیات ادامه پیدا می‌کند.



شکل ۴-۲: توالی یک بلاک تصویر در ۶۴ فریم متوالی (از چپ به راست و از بالا به پایین) و بخشی از ماتریس A حاصل از آن (اولین ستون متناظر با اولین بلاک، دومین ستون متناظر با دومین بلاک و تا آخر).



شکل ۴-۳: بازترتیب بلاکهای شکل ۴-۲ بر اساس مقادیر R از تجزیه QR به همراه ماتریس A که ستونهای آن بر اساس مقادیر روی قطر ماتریس R از تجزیه QR ماتریس A در شکل ۴-۲ مرتب شده‌اند.

۴-۲ دستاوردهای پژوهش

مشکلی که در روش فوق‌الذکر وجود دارد نیاز به محاسبه مجدد تجزیه QR در هر پنجره می‌باشد. اگر ماتریس حاصل از فریمهای یک تا N را A و ماتریس حاصل از فریمهای دو تا $N+1$ را A_1 بنامیم، تنها اختلاف آنها ستونهای اول و آخر آنهاست. در واقع $N-1$ ستون اول ماتریس A_1، همان $N-1$ ستون آخر A هستند و آخرین ستون A_1 فریم جدیدی است که وارد شده است.

اگر بتوان از خروجی تجزیه A تجزیه A_1 را با عملیات محاسباتی کمتری بدست آورد، عمل تخمین زمینه هم با سرعت بیشتری انجام خواهد گرفت. الگوریتم‌های متعددی برای روش گرام اشمیت در منابع مختلف ذکر شده است [۸، ۴]؛ از این میان الگوریتم ارائه شده در [۴] برای منظور موردنظر ما مناسب تشخیص داده شد که در الگوریتم ۱ نشان داده شده است. برای اصلاح آن به نحوی که بتوان با داشتن تجزیه QR ماتریس A، تجزیه ماتریس A_1 را بدست آورد، فقط کافیست مقداردهی اولیه ماتریس v را به صورت مناسب انجام داد و در اولین حلقه الگوریتم، به جای شروع از اولین ستون، از ستون iام عملیات را ادامه دهیم. الگوریتم ۲ وضعیت جدید را نشان می‌دهد. به این ترتیب ورودی الگوریتم۲ به جز ماتریس ورودی، ماتریسهای Q و R بدست آمده از تجزیه ماتریس فریمهای قبلی (A) و i، اندیس آخرین فریم اضافه شده می‌باشد.

اگر فریم جدید در ستون آخر قرار گیرد و ستون اول حذف شود، مقدار i در الگوریتم ۲ همیشه باید یک باشد و لذا دو الگوریتم تفاوتی با هم نخواهند داشت، به منظور رفع این مشکل و حفظ سازگاری الگوریتم ۲ با الگوریتم ۱، ماتریس A از آخر به اول پر می‌شود، به این صورت که آخرین ستون آن، اولین فریم ویدیو و اولین ستون آن، آخرین فریم ویدیوی ما در پنجره جاری خواهد بود. فریم جدید ورودی در یک صف اولین ورودی اولین خروجی جایگزین اولین فریم قبلی (آخرین ستون ماتریس) خواهد شد. اگر از الگوریتم ۱ برای محاسبه تجزیه QR ماتریس A_1 استفاده کنیم، با توجه به اینکه پیدا کردن بردارهای متعامد یکه Q و ماتریس R در الگوریتم ۱

الگوریتم ۱ الگوریتم گرام اشمیت

Input: $A_{m\times n}$
Output: Q, R
$v \leftarrow Q$
for k=1,...,n **do**
  for j=1,...,k-1 **do**
    $r_{jk} \leftarrow \langle v_k, v_j \rangle$ (v_j همان q_j هست)
    $v_k \leftarrow v_k - v_j r_{jk}$
  end for
  $r_{kk} \leftarrow \|v_k\|_2$
  $v_k \leftarrow v_k / r_{kk}$
end for

الگوریتم ۲ الگوریتم گرام اشمیت مرحله‌ای

Input: A_1, Q, R, i
Output: Q, R
$v \leftarrow [Q(1:m, 1:i-1)|A(1:m, i:n)]$
for k=i,...,n **do**
  for j=1,...,k-1 **do**
    $r_{jk} \leftarrow \langle v_k, v_j \rangle$
    $v_k \leftarrow v_k - v_j r_{jk}$
  end for
  $r_{kk} \leftarrow \|v_k\|_2$
  $v_k \leftarrow v_k / r_{kk}$
end for

متاثر از اولین ستون داده‌های ما هستند و اولین ستون آن در A_2 حذف شده است، تمام محاسبات الگوریتم ۱ باید از ابتدا انجام شوند. به منظور رفع این مشکل دو کار انجام شده است: الف) ماتریس A از آخر به اول پر می‌شود، به این صورت که آخرین ستون آن، اولین فریم ویدیو و اولین ستون آن، آخرین فریم ویدیو ما خواهد بود. فریم جدید ورودی در یک صف جایگزین اولین فریم قبلی (آخرین ستون ماتریس) خواهد شد و لذا محاسبه مجدد ماتریس‌های Q و R می‌تواند به راحتی و بدون تاثیر از فریم جدید انجام پذیرد که کار دوم مورد نیاز می‌باشد. ب) الگوریتم ۱ به نحوی اصلاح شده است که به جای محاسبه مجدد کل ماتریسهای Q و R، فقط عناصر مرتبط با فریمهای جدید ورودی محاسبه شوند.

در یک فرآیند دوری، هر زمان که فریم جدیدی وارد می‌شود، در سمت چپ ستون قبلی در ماتریس قرار می‌گیرد و پس از رسیدن به ابتدای ماتریس، فریم جدید در آخرین ستون قرار می‌گیرد. اگر مرتبه زمانی محاسبه هر بردار پایه متعامد ماتریس Q را $O(1)$ فرض کنیم، مرتبه زمانی الگوریتم ۱ برای محاسبه N ستون ماتریس Q، $O(N)$ و لذا مرتبه زمانی محاسبه ماتریس Q برای تمام N فریم یک پنجره، $O(N^2)$ می‌باشد. حال اگر محاسبات موردنیاز بر طبق الگوریتم اصلاح شده ۲ انجام شود، اگر فریم وارد شده iامین فرم ورودی باشد (البته پس از یک بار پر شدن پنجره)، مرتبه زمانی الگوریتم $O(i)$ خواهد بود. چون i در هر دور پر شدن پنجره مقادیر یک تا N را بخود می‌گیرد، پس مرتبه زمانی انجام محاسبات برای N فرم یک پنجره با الگوریتم دوم، $\sum_{i=1}^{N} i = N(N+1)/2 \approx N^2/2$ خواهد بود که تقریباً معادل نصف زمان لازم برای انجام محاسبات با الگوریتم ۱ است. نتایج برنامه‌های نوشته شده روی چند نمونه ویدیو نشان داد که در عمل به دلیل بالاسری مربوط به سایر دستورات برنامه شناسایی زمینه، روش استخراج مدل زمینه با الگوریتم دوم به صورت میانگین ۳۵ درصد نسبت به الگوریتم ۱ کاهش پیدا کرده است. خروجی شناسایی زمینه و شناسایی اشیاء متحرک برای یک فریم از یک ویدیوی نمونه در شکل ۴-۱ آمده است. هنوز در الگوریتم ۲ محاسبات صرفا بر اساس فریم جدید و تجزیه ماتریس قبلی انجام نمی‌شود و به قسمتی از ماتریس ورودی هم وابسته است، به عنوان کارهای آتی می‌توان به حذف وابستگی محاسبات به ماتریس ورودی اشاره نمود.

فهرست منابع

[۱] کنت هافمن، ری کنزی. جبرخطی. ترجمه‌ی فرشیدی، جمشید. مرکز نشر دانشگاهی، ۱۳۷۰.

[۲] آی.گراسمان، استانلی. جبر خطی مقدماتی. ترجمه‌ی عالم‌زاده، علی اکبر. پژوهش، ۱۳۷۵.

[۳] دیوید کینکید، وارد چنی. جبرخطی. ترجمه‌ی منصوره صایمی، فایزه توتونیان. انتشارات دانشگاه امام رضا(ع)، ۱۳۸۱.

[4] Watkins, David S. *Fundamentals of matrix computations*. Pure and applied mathematics. Wiley-Interscience, New York, 2010.

[5] Meyer, Carl D., ed. . *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.

[۶] روش های عددی درجبرخطی. ویراستار سالکویه، داوود خجسته. آموزشهای بنیادی، ۱۳۹۰.

[7] Hammarling, Sven and Lucas, Craig. Updating the QR factorization and the least squares problem, november 2008.

[8] Golub, G.H. and Van Loan, C.F. *Matrix Computations*. Johns Hopkins University Press, Baltimore, 2nd ed. , 1989.

[9] Bishop, Christopher M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, 1 ed. , 2007.

[10] Ye, Jieping, Li, Qi, Xiong, Hui, Park, Haesun, Janardan, Ravi, and Kumar, Vipin. IDR/QR: an incremental dimension reduction algorithm via qr decomposition. in Kim, Won, Kohavi, Ron, Gehrke, Johannes, and DuMouchel, William, eds. , *KDD*, pp. 364–373. ACM, 2004.

[11] Howland, Peg and Park, Haesun. Generalizing discriminant analysis using the generalized singular value decomposition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(8):995–1006, 2004.

[12] Daniel, J. W., Gragg, W. B., Kaufman, L., and Stewart, G. W. Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization. *Mathematics of Computation*, 30(136):772–795, 1976.

[13] Huang, Gao, Huang, Guang-Bin, Song, Shiji, and You, Keyou. Trends in extreme learning machines: A review. *Neural Networks*, 61:32 – 48, 2015.

[14] Yibin Ye, Yang Qin. QR factorization based incremental extreme learning machine with growth of hidden nodes. *Pattern Recognition Letters*, 65:177–183, 2015.

[15] Sobral, Andrews and Vacavant, Antoine. A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos. *Computer Vision and Image Understanding*, 122:4–21, 2014.

[16] Amintoosi, M., Farbiz, F., and Fathy, M. A QR Decomposition based mixture model algorithm for background modeling. in *ICICS2007, Sixth International Conference on Information, Communication and Signal Processing*, pp. 1–5, Singapore, December 2007.

پیوست آ

نمونه برنامه‌های نوشته شده

برنامه آ-۱: برنامه تجزیه چالسکی مرتبط با الگوریتم ۱-۱، صفحه ۲۹.

```
% define a continuous function
%function L=cholesky(A)
% clc
 %A=input('A=')
%
 A = [1 -2 2;-2 5 -3;2 -3 30]
n=rank(A);
l=zeros(n,n);
L(1,1)=sqrt(A(1,1));
for i=2:n

L(i,1)=A(i,1)/L(1,1);
end
for k=2:n
for j=1:k-1
s(j)=L(k,j)^2;
end

S=sum(s);
L(k,k)=sqrt(A(k,k)-S);
for j=1:k-1
b1(j)=L(i,j)*L(k,j);
end
B=sum(b1);
L(i,k)=(A(i,k)-B)/L(k,k);
end
L
L'
A=L*L'
```

برنامه آ-۲: روش گرام اشمیت (فصل ۲)

```
% M. Amintoosi, m.amintoosi@gmail.com
% S. PourSediq
% Page 230 Watkins,
function [Q,R,rIdx] = qr_MGS_watkins%(A)
% QR with GramS
A=[2,1,3;-1,0,7;0,-1,-1]
[m,n]=size(A);
v=A;

for k=1:n
    for j=1:k-1
        R(j,k)=v(:,k)'*v(:,j);
        v(:,k)=v(:,k)-v(:,j)*R(j,k);
    end
    R(k,k)=norm(v(:,k));
    v(:,k)=v(:,k)/R(k,k);
end
Q = v
R
R=Q'*A
[~,rIdx] = sort(abs(diag(R)),'descend');
```

کد

برنامه آ-۳: ماتریس گیونز(فصل ۲)

```
function [ c,s] = givens( a,b )
%if b=0
 %    c=1
 %    s=0
%else
    if abs(b)>=abs(a)
        t=-a/b
        s=1/sqrt(1+t^2)
        c=s*t
    else
        t=-b/a
        c=1/sqrt(1+t^2)
        s=c*t

end
G=[c,s;-s,c]
x=G'*[a,b]'
end
```

برنامه آ-۴: بروزرسانی تجزیه QRبعد از اضافه شدن یک ستون(فصل ۳)

```
% define a continuous function
x=[0;5;10;15;20];% Make room and insert x before j-th column.
%Updating the QR Factorization and the Least Squares Problem page 34
A = magic(5);
u = A(:,4)
A(:,4:end) = [];
```

```
[Q,R]=qr(A)
[m,n]=size(R);
k=4;
A2 = A;
A2(:,k) = u;
A2(:,k+1:n+1) = A(:,k:n);
u = Q'*u;
for i = m:-1:k+1
    [c(i),s(i)]=givens(u(i-1),u(i));
    ui = s(i)*u(i-1)+c(i)*u(i);
    u(i-1)=c(i)*u(i-1)-s(i)*u(i);
    u(i) = ui;
    %u(i)=s(i)*u(i-1)+c(i)*u(i)
    % Update R if there is a nonzero row
    if i<=n+1
        R(i-1:i,i-1:n) =  [c(i),s(i);-s(i),c(i)]'* R(i-1:i,i-1:n);
    end
end
if k==1
    Rmad = [ u R];
else if k == n + 1
        Rmad = [R u ];
    else
        Rmad = [R(:,1:k-1) u R(:,k:n)];
    end
end
Rmad
[Q2,R2] = qr(A2)
```

برنامه آ-۵: روش IDR/QR(فصل ۳)

```
% define a continuous function
X1 = [4,2 ; 2,4; 2,3 ; 3,6; 4 ,4];
X2 = [9,10;6,8;9,5;8,7];
A1 = X1';
A2 = X2';
A = [A1 , A2]
mu1 = mean(A1')'
mu2 = mean(A2')'
%% Paper Method
n1 = size(A1,2);
e1 = ones(n1,1);
m1 = mu1*e1'
A1_m1 = A1-m1

n2 = size(A2,2);
e2 = ones(n2,1);
m2 = mu2*e2';
A2_m2 = A2-m2

Hw = [A1_m1  A2_m2]

m = mean(A')';
Hb = [sqrt(n1)*(mu1-m) , sqrt(n2)*(mu2-m) ]

% Page 2
```

```
Sw = Hw*Hw'
Sb = Hb*Hb'

% Page 4, Algorithm 1
C = [mu1 mu2]
[Q,R] = qr(C);
Z = Hw' * Q;
Y = Hb' * Q;
W = Z'*Z;
B = Y'*Y;
mu = 10;%.001;
[phi,D] = eig((W+mu*eye(2))'*B);
M = phi;
G = Q*M;

y = G(:,1)'*A

plot(A1(1,:),A1(2,:),'rs',A2(1,:),A2(2,:),'bo')
hold on
quiver(G(1,1),G(2,1))
quiver(G(1,2),G(2,2),'r')
hold off
axis square
% axis equal
```

برنامه آ-۶: بروزرسانی IDR/QR(فصل ۳)

```
% Algorithm 2 of IDR
k = 9;
C = magic(k);
[Q,R] = qr(C);

x = randperm(k)';
i = 3;
D = C;

f = (x-C(:,i))/2
% D(:,i) = mean([D(:,i) x],2)
D(:,i) = D(:,i)+f
[QD,RD] = qr(D)

f1 = Q'*f;
f2 = (eye(k) - Q*Q')*f
% f22 = f-Q*(Q'*f)

g = zeros(k,1);
g(i) = 1;

Q*(R+f1*g')+f2*g'

[Q1,R1] = qrupdate(Q,R,f,g);

if norm(f2) <= 1e5
    Qmad = Q1; Rmad = R1;
else
```

```
    q = f2/norm(f2);
    [Q2,R2] = qrupdate([Q1 q],[R1;norm(f2)*g'],f,g);
    Qmad = Q2; Rmad = R2;
end

Qmad*Rmad
```

واژه‌نامه انگلیسی به فارسی

Background . . . . پس‌زمینه
Cholesky Decomposition . . . . تجزیه چالسکی
Foreground . . . . پیش‌زمینه
Givens . . . . گیونز
Householder . . . . هوسهولدر
Ill-posed . . . . بدطرح
Ill-conditioned . . . . بدوضع
Invertible . . . . معکوس‌پذیر
Image editing . . . . ویرایش تصویر
Linear Discriminant Analysis . . . . تحلیل تفکیک خطی
Linear combination . . . . ترکیب خطی
Linear model . . . . مدل خطی
Normal equation . . . . معادلات نرمال
orthogonal complement . . . . مکمل متعامد
Positive Definite . . . . مثبت معین
Pseudo-inverse . . . . شبه‌معکوس
Range space . . . . فضای گستره
Rank -one update . . . . بروزرسانی رتبه -یک
scatter matrix . . . . ماتریس پراکندگی
The Gram-Schmidt Process . . . . روش گرام -اشمیت
The error Least Squares Problem . . . . مساله کمترین مربعات خطا

Hakim Sabzevari University

An Outline of MSc. Thesis

Surname:Pourseddigh	Name:Somayeh	Student No.:9213133020
Supervisor: Dr. Mahmood Amintoosi		
Advisor: Dr. Amin Rafiei		
Faculty of Mathematics and Computer Science	Applied Mathematics	Operational Research
Title of thesis: Background's Frames Identification using QR Decomposition		
Keywords:Bakground Modeling, QR Decomposition, Matrix Factorization, Gram-Schmit Process		

Abstract: Background modeling is one of the most fundamental tasks in computer vision and video surveillance. In this thesis a QR decomposition based algorithm for background modeling is investigated. The key idea of this method is identification of background frames based on R values of QR decomposition. A new faster version of this algorithm by incremental QR updating based on Gram - Schmidt method is presented and its good performance is demonstrated on a real video. Also, two applications of incremental QR updating in machine learning area are considered here.

Hakim Sabzevari University
Faculty of Mathematics and Computer Science

A Thesis Submitted in Partial Fulfilment of the Requirement for the Degree of Master of Science in Applied Mathematics

Background's Frames Identification using QR Decomposition

Supervisor:
Dr. Mahmood Amintoosi

Advisor:
Dr. Amin Rafiei

By:
Somayeh Pourseddigh

January 2016