بسم الله الرحمن الرحیم

دانشگاه قم
دانشکده علوم پایه
پایان‌نامه دوره کارشناسی ارشد رشته ریاضی کاربردی

عنوان:

کاربرد جبر خطی در بازیابی اطلاعات

استاد راهنما:

دکترعفت گلپر رابوکی

نگارنده:

منصوره مهرپویا

زمستان ۱۳۹۳

تقدیم به

پدر ومادرم، که همواره دعای خیرشان بدرقه‌ی راهم است.

همسر مهربانم، که بی‌یاری او پیمودن این راه ممکن نبود.

حمد و سپاس خدای را که توفیق کسب دانش و معرفت را به ما عطا فرمود.

در اینجا وظیفه خود می‌دانم از استاد گرامی و بزرگوار سرکار خانم دکتر عفت گلپر رابوکی که بنده را در انجام این تحقیق، پژوهش و نگارش این پایان نامه راهنمایي نموده‌اند، صمیمانه تشکر و قدردانی کنم.

از جناب آقای دکتر مهدی احمدی‌نیا که داوری اینجانب را برعهده داشتند، کمال تشکر را دارم.

چکیده

موضوع این پایان‌نامه بررسی مساله‌ی بازیابی اطلاعات[1] بااستفاده از کاربرد جبرخطی[2] می‌باشد. در مورد رتبه‌بندی صفحات وب، کاربرد مقادیرویژه و بردار ویژه برای یافتن صفحات که مورد تقاضای استفاده کننده می‌باشد مفصلا بحث می‌شود. ورودی اول مساله، تعدادی سندمتنی[3] که ازموضوعات مختلف بدست آمده است می‌باشد. ورودی دوم نیز یک پروسه‌ی سوال[4]است. هدف ما این است که سوال را در اسناد جستجو نموده و سند مرتبط با آن را پیدا کنیم. مشکلی که در اینجا وجود دارد این است که سند متنی بازگشت داده شده، همواره دارای درصد خطایی در شناسایی می‌باشد که موجب می‌شود معادل با سوال کاربر در نظر گرفته نشوند و در نتیجه سند، مرتبط تشخیص داده نشود. در این پایان‌نامه با استفاده از جبرخطی، بازیابی اطلاعات را به گونه‌ای انجام می‌دهیم که درصد خطا به کمترین مقدار خود برسد. در این راستا ابتدا تجزیهQR، SVD و سپس SDD را بیان می‌کنیم. روش QR یک روش تجزیه ماتریس است که در آن ماتریس A به صورت $A = QR$ تجزیه می‌شود که در آن Q یک ماتریس متعامد و R یک ماتریس بالا مثلثی می‌باشد. روش SVD یک روش تجزیه و فشرده‌سازی ماتریس است که یک تقریب رتبه-پایین بهینه را ارائه می‌دهد. در این روش می‌توان یک ماتریس حقیقی $m \times n$ مانند A را به صورت $A = U\Sigma V^T$ تجزیه کرد که در آن U و V ماتریس‌های متعامد و Σ یک ماتریس قطری از مقادیرتکین A به صورت نزولی می‌باشد. در روش تجزیه نیمه گسسته SDD، ماتریس A بصورت $A = XDY^T$ تجزیه می‌شود. در این روش درایه‌های ماتریس‌های X و Y محدود به مجموعه $\{-1, 0, 1\}$ می‌شوند و ماتریس قطری D شامل مقادیرحقیقی مثبت است. به طور خلاصه می‌توان گفت پردازشی ارائه شده است که انجام آن إعمال الگوریتم بازیابی اطلاعات، موجب کارایی مناسب‌ترآن در حوزه‌ی بازیابی اسناد مرتبط می‌شود.

کلمات کلیدی: فضاهای برداری - موتور جستجو - رتبه‌بندی - تجزیه‌ی QR - تجزیه‌ی مقادیر تکین - تجزیه‌ی نیمه گسسته - بازیابی اطلاعات.

[1] Information retrieval

[2] Linear algebra

[3] Document

[4] Query

فهرست مطالب

فهرست جداول

فهرست تصاویر

فهرست الگوریتم‌ها

مقدمه

تجزیه‌ی ماتریس‌ها کاربردهای فراوانی در زمینه‌های مختلف مانند مهندسی، زیست‌شناسی، پزشکی، داده‌کاوی، پردازش تصویر، حذف نویز، بازیابی اطلاعات و... دارد. یکی از کاربردهای مهم تجزیه‌ی ماتریس‌ها، تجزیه و تحلیل ماتریس‌های با ابعاد زیاد شامل استخراج داده‌های مرتبط، کشف اطلاعات معنی‌دار در داده‌ها، خوشه بندی[۱] و پیدا کردن نماینده‌ای کارآمد برای داده‌ها و طبقه‌بندی آن‌هاست. با توجه به این‌که در دنیای امروز حجم اطلاعات و تنوع ویژگی‌های آن‌ها بسیار زیاد شده است، حجم یا بعد ماتریس داده‌ها هم بسیار بزرگ بوده و به الگوریتم‌هایی نیاز داریم که در عین سریع بودن، بتوانند از دقت بالایی در تجزیه و تحلیل داده‌ها برخوردار باشند و در عین حال بدون از دست دادن ویژگی‌های اصلی داده‌ها، حجم کم‌تری از فضای حافظه را اشغال کنند. از مهم‌ترین روش‌های تجزیه‌ی ماتریسی می توان به روش LU، SVD، QR، SDD[۲]، چولسکی، NMF[۳]، BMF[۴] و... اشاره کرد.

یکی از روش‌های متداول و مفید در تجزیه و فشرده‌سازی ماتریس‌ها روش SVD[۵] است. اما این روش معمولاً در داده‌های شلوغ و با ابعاد زیاد به آسانی قابل تفسیر نیست بعلاوه زمان اجرای آن در این ماتریس‌ها به صورت غیر قابل قبولی بزرگ است و نمی‌تواند به سرعت و وضوح ویژگی‌های اصلی داده‌ها را نمایش دهد[۲۱] .

در این پایان‌نامه روش‌های تجزیه‌ی QR، SVD و سپس SDD بیان شده و روش فشرده‌سازی داده‌ها در این الگوریتم‌ها مورد بررسی قرار می‌گیرند. از مهم‌ترین خصوصیات این روش‌های تجزیه این است که می توان تقریبی از ماتریس‌های حاصل از تجزیه را به گونه‌ای انتخاب کرد که علاوه بر اینکه حجم حافظه‌ی کمی اشغال می‌کنند ماتریس‌های حاصل مینیمم اختلاف را با ماتریس داده‌ی اصلی داشته باشند.

[۱] Clustering

[۲] Semidiscrete Decomposition

[۳] Nonnegative Matrix Factorization

[۴] Binary Matrix Factorization

[۵] Singular Value Decomposition

این پژوهش مشتمل بر ۴ فصل است:

فصل اول شامل مفاهیم و تعاریف مقدماتی است که خواننده برای فهم بهتر مطالب فصول بعدی به دانستن آن نیاز دارد.

در فصل دوم در مورد الگوریتم رتبه‌بندی گوگل و معایب محاسبه‌ی این روش می‌پردازیم.

در فصل سوم به بررسی روش‌های بازیابی اطلاعات با استفاده از تجزیه‌ی QR می‌پردازیم.

در فصل چهارم به بررسی روش‌های بازیابی اطلاعات با استفاده از تجزیه‌های SVD، SDD و LSI یا اندیس گذاری معنایی پنهان را شرح می‌دهیم. سپس سه روش تجزیه‌ی QR، SVD و SDD را با هم مقایسه می‌کنیم [۳۴، ۳۲].

برای بررسی اعتبار الگوریتم، در مورد چهار معیار نسبت فشردگی ماتریس‌ها، میانگین خطای سطر، دقت و یادآوری این روش‌ها بحث کرده و آن‌ها را با هم مقایسه خواهیم کرد. بررسی‌ها نشان می‌دهند که این روش‌ها در ماتریس‌های بزرگ به آسانی همگرا هستند و دارای سرعت قابل ملاحظه‌ای در تجزیه‌ی ماتریس‌ها می‌باشند.
نتایج به‌دست آمده نشان می‌دهند که با استفاده از این الگوریتم‌ها می توان ویژگی‌های اصلی در ماتریس داده‌ی اصلی را به خوبی تشخیص داد و آن‌ها را تجزیه و تحلیل کرد.

فصل ۱

مفاهیم اولیه

۱.۱ مقدمه

در این فصل به معرفی مفاهیم و تعاریفی می‌پردازیم که در فصول دیگر به آن‌ها نیاز داریم. این مفاهیم شامل معرفی انواع ماتریس‌ها، خواص و قضایای مربوط به آن‌ها است.

۲.۱ تعاریف

تعریف ۱.۲.۱. ***زنجیره‌ی مارکوف*** [۱]: *زنجیره‌ی مارکوف یک فرآیند تصادفی بدون حافظه است بدین معنی که توزیع احتمال شرطی حالت بعد تنها به حالت فعلی بستگی دارد و به وقایع قبل از آن وابسته نیست. زنجیره‌ی مارکوف در مدل‌سازی دنیای واقعی کاربردهای زیادی دارد. در حقیقت یک زنجیره‌ی مارکوف دنباله‌ای از متغیرهای تصادفی* $X_۱, X_۲, ...$ *است که دارای خاصیت مارکوف هستند، یعنی:*

$$Pr(X_{n+۱} = x|X_۱ = x_۱, X_۲ = x_۲, ..., X_n = x_n) = Pr(X_{n+۱} = x|X_n = x_n)$$

مقادیر ممکن برای x_i *مجموعه‌ی قابل شمارشی را می‌سازند که فضای حالت* [۲] *نام دارند.*

تعریف ۲.۲.۱. ***ماتریس تصادفی*** [۳]: *ماتریس مربعی* $P \in R^{n\times n}$ *یک ماتریس تصادفی نامیده می‌شود اگر همه‌ی درایه‌هایش نامنفی بوده و جمع درایه‌ها در هر سطر یا ستون برابر با یک باشد.*
اگر $\sum_{j=۱}^{n} p_{ij} = ۱$ *،* $i = ۱, ..., n$ *آن را یک ماتریس تصادفی سطری و اگر* $\sum_{i=۱}^{n} p_{ij} = ۱$ *،* $j = ۱, ..., n$ *آن را یک ماتریس تصادفی ستونی نامند.*
یک ماتریس تصادفی با زنجیره‌های مارکوف X_t *در فضای حالت* S *تعریف می‌شود. در اینصورت* p_{ij} *احتمال حرکت از مرحله‌ی* i *به مرحله‌ی* j *در یک بازه‌ی زمانی است.*

تعریف ۳.۲.۱. ***چند جمله‌ای مشخصه:*** *چند جمله‌ای مشخصه‌ی ماتریس* $A \in R^{n\times n}$ *که با* $P_A(\lambda)$ *نشان*

[۱] Markov Chain
[۲] State
[۳] Stochastic Matrix

می‌دهند به صورت زیر تعریف می‌کنند:

$$P_A(\lambda) = det(\lambda I - A) = \lambda^n + c_1\lambda^{n-1} + ... + c_n = 0 \qquad (1.1)$$

که در آن

$$c_i = (-1)^i \sum_{j_1=1}^{n} ... \sum_{j_i=1}^{n} \lambda_{j_1}...\lambda_{j_i}$$

ریشه‌های این معادله مقادیر ویژه‌ی ماتریس A هستند. از بسط رابطه‌ی فوق بر حسب ضرایب ماتریس A داریم:

$$c_1 = -(a_{11} + a_{22} + ... + a_{nn}) = -tr(A) = -\sum_{i=1}^{n} a_{ii}, c_n = (-1)^n\lambda_1...\lambda_n$$

چند جمله‌ای مشخصه از درجه‌ی n است و دقیقا n ریشه دارد که ممکن است برخی تکراری باشند.

تعریف ۴.۲.۱. ***مقدار ویژه***[1] ***و بردار ویژه***[2]***:*** *فرض کنید $A \in R^{n\times n}$ باشد. اسکالر $\lambda \in \mathbb{C}$ را یک مقدار ویژه ماتریس A گویند هر گاه بردار مخالف صفر $x \in \mathbb{C}^n$ موجود باشد به طوری که $Ax=\lambda x$, یعنی λ جواب معادله*[3] *چند جمله‌ای مشخصه ماتریس A باشد. در این صورت x را بردار ویژه‌ی نظیر مقدار ویژه‌ی λ گویند.*

تعریف ۵.۲.۱. ***مرتبه‌ی جبری و هندسی:*** *مرتبه‌ی جبری یک مقدار ویژه برابر است با توان ریشه‌ی نظیر در چند جمله‌ای مشخصه‌ی آن و مرتبه‌ی هندسی عبارت است از تعداد بردارهای ویژه‌ی مستقل خطی نظیر آن مقدار ویژه.*

قضیه ۱.۲.۱. *هر ماتریس تصادفی یک مقدار ویژه‌ی یک دارد.*

اثبات. فرض کنید P یک ماتریس تصادفی ستونی باشد. بردار $\mathbf{1} = [1...1]^T$ را در نظر بگیرید. داریم:

$$\mathbf{1}^T P = [\sum_{i=1}^{n} p_{i1}, ..., \sum_{i=1}^{n} p_{in}] = [1...1]$$

بنابراین $\lambda = 1$ یک مقدار ویژه‌ی ماتریس P و بردار $[1...1]^T$ بردار ویژه‌ی نظیر آن است. از آنجا که مقادیر ویژه‌ی P و P^T برابرند پس هر ماتریس تصادفی سطری نیز دارای یک مقدار ویژه‌ی ۱

[1] Eigenvalue

[2] Eigenvector

[3] Equation

می‌باشد. □

تعریف ۶.۲.۱. بردار ایستا[۱]: *بردار ویژه‌ی نظیر مقدار ویژه‌ی ۱ در ماتریس تصادفی را بردار ایستا گوییم.*

تعریف ۷.۲.۱. شعاع طیفی: *مجموعه‌ی تمام مقادیر ویژه‌ی ماتریس $A \in R^{n\times n}$ را طیف A می‌گویند و با $\sigma(A)$ نشان می‌دهند. شعاع طیفی ماتریس A را با $\rho(A)$ نشان می‌دهند و به صورت زیر تعریف می‌کنند:*

$$\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$$

تعریف ۸.۲.۱. الگوریتم روش توانی[۲]: *روش توانی روشی کارا و یکی از معروف‌ترین روش‌های تکراری است در روش توانی ما به دنبال یافتن بزرگترین مقدار ویژه از نظر قدر مطلق و بردار ویژه‌ی نظیر آن هستیم. این روش زمانی کاربرد دارد که بزرگترین مقدار ویژه از نظر قدرمطلق تکراری نباشد [۱۱].*

تعریف ۹.۲.۱. فضای سطری[۱] **وستونی**[۲] **ماتریس:** *فرض کنید $A \in R^{m\times n}$. مجموعه‌ی تمام ترکیبات خطی سطرهای ماتریس A را فضای سطری و مجموعه‌ی تمام ترکیبات خطی ستون‌های ماتریس A را فضای ستونی A گویند.*

تعریف ۱۰.۲.۱. ماتریس متعامد [۳]: *ماتریس $A \in R^{n\times n}$ یک ماتریس متعامد نامیده می‌شود اگروفقط اگر $A^T = A^{-1}$، یعنی*

$$AA^T = A^T A = I$$

یا به عبارت دیگر:

$$a_i^T a_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad , i, j = 1, 2, ..., n \tag{۲.۱}$$

که a_i ستون i ام ماتریس A می‌باشد.

تعریف ۱۱.۲.۱. ماتریس قطری شدنی [۴]: *ماتریس A را قطری شدنی گویند هرگاه ماتریس وارون*

[۱] Stationary vector
[۲] Power method
[۱] Row space
[۲] Column space
[۳] Orthogonal Matrix
[۴] Diagonalizing matrix

پذیر P و ماتریس قطری D وجود داشته باشند که:

$$P^{-1}AP = D$$

تعریف ۱.۲.۱۲. ماتریس قابل کاهش[۵] : ماتریس $A \in R^{n\times n}$,$n \geq 2$ را قابل کاهش گویند هرگاه ماتریس جایگشت $P \in R^{n\times n}$ و عدد صحیح r که $1 \leq r \leq n-1$ موجود باشد بطوری که:

$$P^T AP = \begin{bmatrix} B & C \\ O & D \end{bmatrix}$$

که در آن $B \in R^{r\times r}$, $C \in R^{r\times n-r}$, $O \in R^{n-r\times r}$, $D \in R^{n-r\times n-r}$.

یک ماتریس را غیر قابل کاهش[۱] گویند هرگاه قابل کاهش نباشد.

قضیه ۱.۲.۲. *(پرون-فروبنیوس[۲] در ماتریس‌های مثبت): فرض کنید $A = (a_{ij})$ یک ماتریس مثبت $n \times n$ باشد یعنی $a_{ij} > 0$، $1 \leq i, j \leq n$. در این صورت گزاره‌های زیر برقرار است:*

۱. یک عدد حقیقی مثبت r که ریشه‌ی پرون یا مقدار ویژه‌ی پرون فروبنیوس نامیده می‌شود وجود دارد به طوری که r یک مقدار ویژه‌ی A است و هر مقدار ویژه‌ی دیگر λ (مختلط یا حقیقی) اکیدا از نظر قدرمطلق از r کوچکتر است، $|\lambda| < r$. بنابراین شعاع طیفی $\rho(A)$ برابر r است.

۲. مقدار ویژه‌ی پرون فروبنیوس ساده است، یعنی r یک ریشه‌ی ساده‌ی چند جمله‌ای مشخصه‌ی A است. در نتیجه فضای ویژه‌ی نظیر r یک بعدی است.

۳. یک بردار ویژه مثلا x از A نظیر r وجود دارد که همه‌ی عناصر آن مثبت است.

اثبات. برای اثبات به [۲۶] مراجعه کنید.

این قضیه کاربردهای مهمی در نظریه‌ی آمار، سیستم‌های دینامیکی، اقتصاد، موتورهای جستجو و حتی رتبه‌بندی تیم‌های ورزشی دارد. □

قضیه ۱.۲.۳. *فرض کنید $A \in R^{n\times n}$ یک ماتریس نامنفی است یعنی تمام درایه‌های آن نامنفی می‌باشد. آنگاه $\rho(A)$ یک مقدار ویژه‌ی A است. بعلاوه یک بردار ویژه‌ی نامنفی وجود دارد که $Ax = \rho(A)x$.*

[۵] Reducible

[۱] Irreducible

[۲] Perron Frobenius

اثبات. برای هر $\varepsilon > ۰$ ماتریس $A(\varepsilon) = [a_{ij} + \varepsilon] > ۰$ را تعریف می‌کنیم.
فرض کنید $x(\varepsilon)$ بردار پرون[۲۶] ماتریس $A(\varepsilon)$ باشد به طوری که $x(\varepsilon) > ۰$ و $\sum_{i=۱}^{n} x(\varepsilon)_i = ۱$. از آنجا که مجموعه‌ی بردارهای $x(\varepsilon)$ برای $\varepsilon > ۰$ در گوی فشرده‌ی $\{x \in R^n | \quad \|x\|_۱ \leq ۱\}$ قرار دارد پس یک دنباله‌ی نزولی $\varepsilon_۱, \varepsilon_۲, ...$ وجود دارد به طوری که $\lim_{k\to\infty} \varepsilon_k = ۰$ و $\lim_{k\to\infty} x(\varepsilon_k) = x$. با توجه به اینکه $x(\varepsilon_k) > ۰$، برای $k = ۱, ۲, ...$، بنابراین:

$$x = \lim_{k\to\infty} x(\varepsilon_k) \geq ۰$$

بعلاوه $x \neq ۰$ زیرا:

$$\sum_{i=۱}^{n} x_i = \lim_{k\to\infty} \sum_{i=۱}^{n} x(\varepsilon_k)_i = ۱$$

و داریم:

$$\rho(A(\varepsilon_k)) \geq \rho(A(\varepsilon_{k+۱})) \geq ... \geq \rho(A) \qquad , \qquad k = ۱, ۲, ...$$

بنابراین دنباله‌ی اعداد حقیقی $\rho(A(\varepsilon_k))$ یک دنباله‌ی نزولی است. پس $\rho = \lim_{k\to\infty} \rho(A(\varepsilon_k))$ وجود دارد و $\rho \geq \rho(A)$. اما داریم:

$$\begin{aligned} Ax &= \lim_{k\to\infty} A(\varepsilon_k)x(\varepsilon_k) = \lim_{k\to\infty} \rho(A(\varepsilon_k))x(\varepsilon_k) \\ &= \lim_{k\to\infty} \rho(A(\varepsilon_k)) \lim_{k\to\infty} x(\varepsilon_k) \\ &= \rho x \end{aligned}$$

با توجه به اینکه $x \neq ۰$ پس ρ یک مقدار ویژه‌ی A است اما $\rho \geq \rho(A)$ بنابراین $\rho = \rho(A)$. □

ماتریس‌هایی که با روش بازیابی اطلاعات تولید می‌شوند نامنفی هستند. پس برای این ماتریس‌ها قضییه‌ی ۳.۲.۱ تضمین می‌کند که $\rho(A)$ یک مقدار ویژه‌ی A است.
برای ماتریس‌های نامنفی غیر قابل‌کاهش قضییه‌ی زیر برقرار می‌باشد.

قضیه ۴.۲.۱. *اگر $A \in R^{n\times n}$ یک ماتریس نامنفی غیرقابل‌کاهش باشد آنگاه :*

الف. $\rho(A) \geq ۰$.

ب. $\rho(A)$ *یک مقدار ویژه‌ی* A *است.*

ج. یک بردار ویژه‌ی مثبت x وجود دارد به طوری که: $Ax = \rho(A)x$.

د. $\rho(A)$ یک مقدار ویژه‌ی جبری ساده یا به عبارت دیگر یک مقدار ویژه‌ی ساده‌ی هندسی است.

اثبات. برای اثبات به [۲۶] مراجعه کنید. □

با توجه به اینکه ماتریس‌های تصادفی ماتریس‌های نامنفی هستند، اگر غیرقابل‌کاهش باشند قضیه‌ی (۴.۲.۱) برای آن‌ها برقرار است.

تذکر ۱. *در عمل برای ماتریس‌های رتبه‌بندی گوگل که بعدا بحث خواهد شد، شرط غیرقابل‌کاهش بودن و قضییه‌ی پرون-فروبنیوس برقرار است.*

تعریف ۱۳.۲.۱. عدد حالت ماتریس [۱]: *فرض کنید ماتریس $A \in \mathbb{R}^{n\times n}$ وارون‌پذیر باشد. عدد حالت ماتریس A به صورت زیر تعریف می‌شود:*

$$C(A) = \|A\|\|A^{-1}\| \tag{۳.۱}$$

عدد حالت یک ماتریس منحصربه‌فرد نیست زیرا می‌توان با نرم‌های مختلف آن را محاسبه کرد، ولی طبق هم‌ارزی نرم‌ها این اعداد باهم رابطه دارند.
به‌طور کلی $C_p(A) = \|A\|_p\|A^{-1}\|_p$ و $C(A) \geq 1$. اگر عدد حالت ماتریس A بزرگ باشد A را بدوضع [۱] و در غیر این‌صورت خوش‌وضع [۲] می‌نامند.

تعریف ۱۴.۲.۱. ضرب داخلی دو بردار [۳]: *فرض کنید $x, y \in \mathbb{R}^n$. ضرب داخلی دو بردار که ضرب نقطه‌ای [۴] یا ضرب اسکالر [۵] نیز نامیده می‌شود، به صورت زیر تعریف می‌شود:*

$$< x, y >= x^T y = \sum_{i=1}^{n} x_i y_i \tag{۴.۱}$$

تعریف ۱۵.۲.۱. نرم اقلیدسی [۶]: *فرض کنید $x = (x_1, x_2, ..., x_n)^T \in \mathbb{R}^n$. نرم‌اقلیدسی، طول‌اقلیدسی*

---

[۱] Condition Number
[۱] Ill condition
[۲] Well condition
[۳] Inner Product
[۴] Dot Produc
[۵] Scalar Product
[۶] Euclidean norm

[1] یا اندازه‌ی بردار x به صورت زیر تعریف می‌شود [۹]:

$$\|x\|_۲ = (\sum_{i=۱}^{n} |x_i|^۲)^{۱/۲} = \sqrt{< x, x >} \qquad (۵.۱)$$

تعریف ۱.۲.۱۶. نرم فروبنیوس [2]: فرض کنید A یک ماتریس $m \times n$ باشد. نرم فروبینیوس A به صورت زیر تعریف می‌شود [۹]:

$$\|A\|_F = (\sum_{i=۱}^{m}\sum_{j=۱}^{n} |a_{ij}|^۲)^{۱/۲} \qquad (۶.۱)$$

نرم فروبنیوس برابر با مجموع عناصر قطری ماتریس $A^T A$ است، در نتیجه رابطه‌ی زیر برقرار می‌باشد:

$$\|A\|_F = \sqrt{Trace(A^T A)} = \sqrt{Trace(AA^T)} \qquad (۷.۱)$$

بعلاوه فرض کنید P یک ماتریس متعامد $m \times m$ باشد آنگاه داریم:

$$\|PA\|_F = \sqrt{Trace((PA)^T(PA))} = \sqrt{Trace(A^T P^T PA)} = \sqrt{Trace(A^T A)} = \|A\|_F$$

تعریف ۱.۲.۱۷. فاصله‌ی اقلیدسی [3]: فرض کنید $x = (x_۱, x_۲, ..., x_n)^T$ و $y = (y_۱, y_۲, ..., y_n)^T$ دو بردار از فضای $\mathbb{R}^n$ باشند. فاصله‌ی اقلیدسی دوبردار به صورت زیر تعریف می‌شود [۹]:

$$\begin{aligned} d(x,y) &= \|x-y\|_۲ = \sqrt{< (x-y), (x-y) >} \\ &= \sqrt{\|x\|^۲ + \|y\|^۲ - ۲ < x, y >} = (\sum_{i=۱}^{n}(x_i - y_i)^۲)^{۱/۲} \end{aligned} \qquad (۸.۱)$$

تعریف ۱.۲.۱۸. گراف: یک گراف G زوج مرتبی مانند (V, E) است که در آن V مجموعه‌ای متناهی و ناتهی و E زیرمجموعه‌ای از مجموعه‌ی تمام زیرمجموعه‌های دو عضوی V می‌باشد. اعضای V را راس‌ها یا گره‌های G و اعضای E را یال‌های G می‌نامند.

تعریف ۱.۲.۱۹. گراف جهتدار: در گراف $G = (V, E)$ اگر ترتیب قرار گرفتن راس‌ها در مجموعه‌ی E

[1] Euclidean length

[2] Frobenius norm

[3] Euclidean distance