#خوانش_گروهی_با_فیلاگر

# مطالعه کتاب و مقالات هوش‌مصنوعی
# به صورت گروهی در فیلاگر

**رقیه فرجی**
کارشناسی ارشد فیزیک اتمی

filoger

فیلاگر|جامعه هوش‌مصنوعی ایران

# فصل دوم

## آشنایی با جبر خطی

- ✍️ معرفی ماتریس‌های خاص
- ✍️ تجزیه ماتریس‌ها
- ✍️ پیدا کردن معکوس ماتریس‌ها
- ✍️ عملگر Trace
- ✍️ دترمینان
- ✍️ روش PCA

# ماتریس‌ها و بردارهای خاص

ماتریس قطری:

$D_{i,j} = 0$ for all unequal $i$ , $j$

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

نمایش

$$\mathbf{V} = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix} \longrightarrow \mathrm{diag}(\mathbf{v}) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

41

کاربردها

diag(v)x = v.x

$$\begin{bmatrix} -3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 7 \end{bmatrix}^{-1} = \begin{bmatrix} -1/3 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1/7 \end{bmatrix}$$

ماتریس همانی

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

# ماتریس‌ها و بردارهای خاص

ماتریس متقارن:

$A^T = A$

$$\begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 5 \end{bmatrix}$$

مثال

اگر بخواهیم فاصله نقاط را بصورت ماتریس بیان کنیم چون فاصله دو نقطه مستقل از این است که ما فاصله را از نقطه اول تا دوم حساب کنیم یا برعکس حاصل ماتریسی متقارن می‌باشد

41

بردار واحد:

$$||\boldsymbol{x}||_2 = 1.$$

بردارهای متعامد:

$$\boldsymbol{x}^\top \boldsymbol{y} = 0.$$

$$A^\top A = AA^\top = I.$$

$$\Downarrow$$

$$A^{-1} = A^\top$$

ماتریسی که سطر های آن بر هم عمود باشند یا همچنین
ستون‌های آن بر هم عمود باشند را ماتریس متعامد گوییم

ویژگی حائز اهمیت این ماتریس در آن است که معکوس آن
براحتی قابل محاسبه است

41

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{pmatrix}\begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{I}$$

**A** is orthogonal.

# تجزیه ماتریس‌هابه روش Eigendecomposition

در این روش به‌دنبال خصوصیاتی از ماتریس‌ها هستیم که universal باشد

$$\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}.$$

روش decomposition اطلاعاتی از ماتریس‌ها را بدست می‌دهد که با مشاهده ماتریس قابل تفسیر نیست

یکی از روش‌های Matrix decomposition استفاده از روش Eign decompositiom می‌باشد

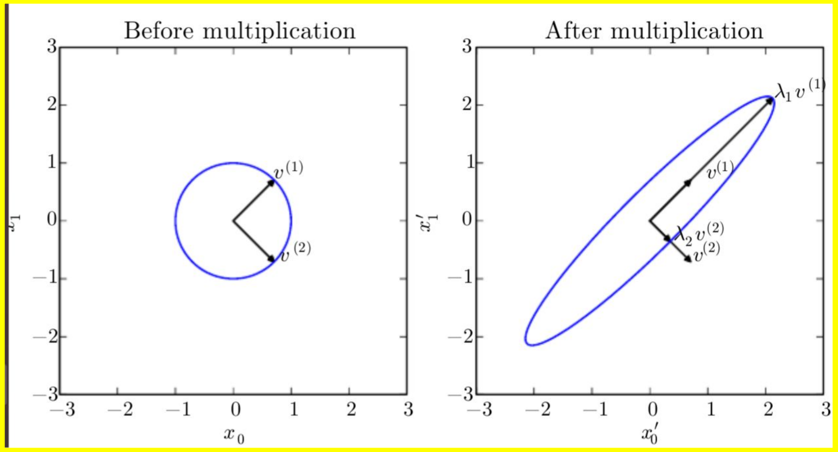
42
43

$$V = [\, v^{(1)}\, ,\, \dots\, ,\, v^{(n)} ]$$

$$\lambda = [\, \lambda_1\, ,\, \dots\, ,\, \lambda_n ]$$

$$\boxed{\boldsymbol{A} = \boldsymbol{V} \operatorname{diag}(\boldsymbol{\lambda}) \boldsymbol{V}^{-1}.}$$

# تجزیه ماتریس‌هابه روش Eigendecomposition




در شکل روبرو تاثیر بردارهای ویژه و مقادیر ویژه را مشاهده می‌کنید.

در شکل سمت چپ قبل از اعمال ضرب ماتریس A می‌باشد

در شکل سمت راست بعد از ضرب در ماتریس A مشاهده می‌شود که در همان راستاهای بردار ویژه‌ها باقی مانده منتها با مقدارهای ویژه

43

# تجزیه ماتریس‌هابه روش Eigendecomposition

لزوما همه‌ی ماتریس‌ها دارای بردارویژه و مقدار ویژه نیستند

**Q**ها ماتریس‌های متعامد هستند

تجزیه ماتریس متقارن

$$A = Q \Lambda Q^{\top}$$

هر ماتریس متقارنی حتما دارای مقادیر ویژه است اما لزوما میتواند تجزیه واحدی نباشد

هرگاه برای دو بردار ویژه مقدار ویژه یکسان داشته باشیم می‌توان از span ها برای ساختن ماتریس متعامد استفاده کرد

44

# تجزیه ماتریس‌هابه روش Eigendecomposition

$$f(x) = x^\top A x$$

هرگاه تابع مقابل را بخواهیم بهینه کنیم در شرایطی که ماتریس متقارن باشد و x بردار ویژه آن باشد می‌توان مقدار بیشینه یا کمینه آن را از روی مقادیر ویژه محاسبه کرد

positive definite : به ماتریس‌هایی گفته می‌شود که مقادیر ویژه‌ها مثبت است

positive semidefinite: به ماتریس‌هایی گفته می‌شود که مقادیر ویژه‌ها مثبت و یا صفر است

negative definite : به ماتریس‌هایی گفته می‌شود که مقادیر ویژه‌ها منفی است

negative semidefinite: به ماتریس‌هایی گفته می‌شود که مقادیر ویژه‌ها منفی و یا صفر است

44

# Singular Value Decomposition

روش singular value decomposition راه دیگری برای فاکتوربندی ماتریس‌ها ارائه می‌دهد
در این روش همانند قبل ویژگی‌هایی از ماتریس استخراج می‌شود که singular vectors و singular values نامیده‌می‌شوند

بطور کلی این روش برا همه‌ی ماتریس‌ها جواب دارد

از روش قبلی داشتیم

$$A = V \operatorname{diag}(\boldsymbol{\lambda}) V^{-1}.$$

A ماتریسی m در n می‌باشد

در روش singular داریم

$$A = UDV^{\top}.$$

U:m × m

ماتریس متعامدی می‌باشد که ستون‌های آن left-singular vectors نامیده می‌شوند

D:m × n

ماتریس قطری بوده و عناصر آن singular values نامیده می‌شوند

V:n × n

ماتریس متعامدی بوده که ستون‌های آن right-singular vectors نامیده می‌شوند

45

# Singular Value Decomposition

بردار ویژه‌های $A^\top A$ ستون‌های **V** را می‌سازند

بردار ویژه‌های $AA^\top$ ستون‌های **U** را می‌سازند

ریشه دوم مقدار ویژه‌های $AA^\top$ عناصر قطری ماتریس **D** را می‌سازند

# محاسبه معکوس Moore-Penrose

$$A^{+} = VD^{+}U^{\top},$$

**D** ،**U** و **V** ماتریس‌هایی مربوط به singular value decomposition هستند

$D^{+}$ ماتریس D است که عناصر قطری آن معکوس شده‌اند

$U^{\top}$ ترانهاده ماتریس U می‌باشد

# عملگر Trace

این عملگر جمع عناصر روی قطر اصلی را محاسبه می‌کند

$$\mathrm{Tr}(\boldsymbol{A}) = \sum_i \boldsymbol{A}_{i,i}.$$

خیلی از محاسبات ریاضیاتی روی ماتریس‌ها با استفاده از این عملگر تسهیل می‌شوند

مثال

می‌توان نرم فروبنیویس را بسادگی با استفاده از عمگر Trace محاسبه کرد

$$||A||_F = \sqrt{\mathrm{Tr}(\boldsymbol{A}\boldsymbol{A}^\top)}.$$

خواص عملگر

$$\mathrm{Tr}(\boldsymbol{ABC}) = \mathrm{Tr}(\boldsymbol{CAB}) = \mathrm{Tr}(\boldsymbol{BCA})$$

$$\mathrm{Tr}(\boldsymbol{A}) = \mathrm{Tr}(\boldsymbol{A}^\top).$$

47

# دترمینان

دترمینان ماتریس مربعی را با علامت det ( A ) نشان می‌دهیم.

دترمینان تابعی است که ماتریس را به فضای اسکالر نگاشت می‌کند.

دترمینان معادل با ضرب مقادیر ویژه ماتریس می‌باشد

دترمینان ماتریس معیاری است از اینکه بعد از ضرب ماتریسی فضا فشرده یا منبسط می‌شود

47

# Principal Components Analysis or PCA

فرض کنید m نقطه بصورت مقابل داشته باشیم و بخواهیم برای این نقاط lossy compression حساب کنیم

$$\{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(m)}\}$$

$$\boldsymbol{x}^{(i)} \in \mathbb{R}^n \longrightarrow \boldsymbol{c}^{(i)} \in \mathbb{R}^l.$$

48

فرض کنید تابعی وجود دارد مثل $f(x) = c$ که این کار را انجام می‌دهد در این صورت تابعی مانند $x \approx g(f(x))$ وجود دارد که مقادیر رو از فضای فشرده شده به فضای اصلی برمی‌گرداند

# Principal Components Analysis or PCA

1

$$g(\boldsymbol{c}) = \boldsymbol{D}\boldsymbol{c}$$

2

$$\boldsymbol{c}^* = \arg\min_{\boldsymbol{c}} ||\boldsymbol{x} - g(\boldsymbol{c})||_2.$$

3

$$\boldsymbol{c}^* = \arg\min_{\boldsymbol{c}} ||\boldsymbol{x} - g(\boldsymbol{c})||_2^2.$$

$$= \arg\min_{\boldsymbol{c}} -2\boldsymbol{x}^\top \boldsymbol{D}\boldsymbol{c} + \boldsymbol{c}^\top \boldsymbol{c}$$

49

4

$$\nabla_{\boldsymbol{c}}(-2\boldsymbol{x}^\top \boldsymbol{D}\boldsymbol{c} + \boldsymbol{c}^\top \boldsymbol{c}) = \boldsymbol{0}$$

5

$$\boldsymbol{c} = \boldsymbol{D}^\top \boldsymbol{x}.$$

6

$$f(\boldsymbol{x}) = \boldsymbol{D}^\top \boldsymbol{x}.$$

7

$$r(\boldsymbol{x}) = g(f(\boldsymbol{x})) = \boldsymbol{D}\boldsymbol{D}^\top \boldsymbol{x}.$$

# Principal Components Analysis or PCA

$$\boldsymbol{D}^* = \arg\min_{\boldsymbol{D}} \sqrt{\sum_{i,j} \left( x_j^{(i)} - r(\boldsymbol{x}^{(i)})_j \right)^2} \text{ subject to } \boldsymbol{D}^\top \boldsymbol{D} = \boldsymbol{I}_l$$

$$\boldsymbol{d}^* = \arg\min_{\boldsymbol{d}} \sum_i ||\boldsymbol{x}^{(i)} - \boldsymbol{d}\boldsymbol{d}^\top \boldsymbol{x}^{(i)}||_2^2 \text{ subject to } ||\boldsymbol{d}||_2 = 1.$$

$$\boldsymbol{d}^* = \arg\min_{\boldsymbol{d}} \sum_i ||\boldsymbol{x}^{(i)} - \boldsymbol{x}^{(i)\top} \boldsymbol{d}\boldsymbol{d}||_2^2 \text{ subject to } ||\boldsymbol{d}||_2 = 1.$$

# Principal Components Analysis or PCA

$$d^* = \arg\min_{d} ||X - Xdd^\top||_F^2 \text{ subject to } d^\top d = 1.$$

$$= \arg\min_{d} -\mathrm{Tr}(X^\top X d d^\top) \text{ subject to } d^\top d = 1$$

$$\boxed{= \arg\max_{d} \mathrm{Tr}(d^\top X^\top X d) \text{ subject to } d^\top d = 1}$$

در فرمول نهایی پیداست که مقدار نهایی می‌تواند وابسته به مقادیر ویژه ماتریس ورودی‌هاست

51
52

# سپاس از همراهی شما

filoger

فیلاگر | جامعه هوش‌مصنوعی ایران