2022.02.22

Singular Value Decomposition

$n\times m$ matrix 来表示 X (n个 sample, m个 features)

Goal: 1. 用个更小的 matrix 来 approximate X，但同时存留了 similar info.

2. Dimensinality Reduction / Feature Extraction.

3. Anomaly Detection & Denoising

- n个 n-vector linearly independent iff 由这些 vectors 组成的 matrix 的 determinant 为 non-zero.

- 若A为 $n\times m$，则需存 mn 个 values. 若 $\text{rank}(A)=k$，则可将 $A=UV$.
  U为 $n\times k$, V为 $k\times m$. 则 $k(m+n)$ 将小于 mn (占更少 memory)

- Frobenius Distance

$$d_F(A,B)=\|A-B\|_F=\sqrt{\sum_{i,j}(a_{ij}-b_{ij})^2}$$

pairwise.

Goal: 找 B ($\text{rank}(B)=k \ll m/n$) 使 $d_F(A,B)$ 最小.

Def: When $k<\text{rank}(A)$, the rank k approximation of A 为

$$A^{(k)}=\underset{\{B\,|\,\text{rank}(B)=k\}}{\arg\min}\ d_F(A,B)$$

B和A有同样 dimension，但B的 rank 更小

Def: Singular Value Decomposition

of a rank-r matrix A has the form

$A = U \Sigma V^T$

$U_{n\times r}$, $V_{m\times r}$ > column orthonormal $\quad \Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 & \cdots \\ 0 & \sigma_2 & & \\ & & \ddots & \\ 0 & \cdots & & \sigma_r \end{pmatrix}$ $r\times r$

$A = (U_1 \ U_2) \begin{pmatrix} \Sigma_1 & \\ & \Sigma_2 \end{pmatrix} (V_1 \ V_2)$

$A^{(k)} = U_1 \Sigma_1 V_1^T$

· Property: $d_F(A, A^{(k)})^2 = \sum_{i=k+1}^{r} \sigma_i^2$

· The i^{th} singular vector represents the direction of the i^{th} most variance 指的是 U中的 vector.

· 找k值

1. 在singular value plot 中找 elbow point
2. 看选不同k后的residual error

Latent Semantics Analysis

文章 → words count

	brain	cell	data
CS1	0	0	4
CS2	0	0	3
Medicine 2		1	0

→ 对此 matrix 用 Singular Value Decomposition

TfiDf.

tf · idf → $\log\left(\dfrac{\text{number of documents}}{\text{number of} \downarrow \text{ that contains that term}}\right)$

term frequency in the documents

- **Determinant**

$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \det(A) = ad - bc$

$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \quad \det(A) = a \cdot \det\begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \cdot \det\begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \cdot \det\begin{pmatrix} d & e \\ g & h \end{pmatrix}$

- $\mathrm{rank}(A) = \mathrm{rank}(A^T)$

- Singular Value Decomposition

$A = \cancel{U\Sigma I}\ U\Sigma V^T \qquad A_{n\times m} \qquad U_{n\times r} \qquad V_{r\times m}$

orthonormal

$\underset{k\times k}{\Sigma} = \begin{pmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_r \end{pmatrix}$

$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0$

σ_i is the square root of the eigenvalues of A^TA. 称 singular values.

- $A = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} \Sigma_1 & \\ & \Sigma_2 \end{pmatrix} \begin{pmatrix} V_1 & V_2 \end{pmatrix}$

$A^{(k)} = U_1 \Sigma_1 V_1^T$

$d_F(A, A^{(k)})^2 = \sum_{i=k+1}^{r} \sigma_i^2$