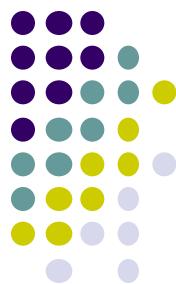


Solution: 奇异值分解

Singular Value Decomposition



Let \hat{Y} be the **centered** $d \times N$ data matrix (assume $N > d$).

$$\sum_i \mathbf{y}_i = \mathbf{0} \quad \mathbf{Y} = \begin{pmatrix} y_{1,1} & y_{2,1} & y_{3,1} & \cdots & y_{N,1} \\ y_{1,2} & y_{2,2} & y_{3,2} & \cdots & y_{N,2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{1,d} & y_{2,d} & y_{3,d} & \cdots & y_{N,d} \end{pmatrix}_{d \times N}$$

Singular values ↓
 $= \mathbf{U} \mathbf{S} \mathbf{V}^T$
↑
夹矩×自己

is the SVD of \hat{Y} , where

正交即 $\mathbf{A}^T \mathbf{A} = \mathbf{I}$

\mathbf{y}_2

Unitary Matrices

酉矩阵

- \mathbf{U} is $d \times d$ orthogonal, the left singular vectors.
- \mathbf{V} is $N \times N$ orthogonal, the right singular vectors.
- \mathbf{S} is $d \times N$ diagonal, with $s_1 \geq s_2 \geq \dots \geq s_d \geq 0$, the singular values.

对角阵

- ✓ The SVD always exists, and is unique up to signs.
- ✓ The columns of \mathbf{V} are the principal components

SVD 奇异值分解

相当于将特征值分解(仅针对满秩矩阵)

扩展到任意 $M \times N$ 矩阵

* 在特征分解中, 某一 $M \times N$ 矩阵 A 在作用于了映射空间.

现有一组 n 维空间已知基, 但 A 矩阵对应另一个空间 K ,

空间 K 的维数与 A 的秩相同, 即 $K = \text{rank}(A)$,

现利用 A 将 n 维空间已知基 映射(降维)到 K 维空间

现有：

- n 维空间正交基 $\{v_1, \dots, v_n\}$

↓ 任矩阵 $A_{(m \times n)}$ 映射，此时仅可保保 k 维， $k = \text{Rank}(CA)$

$$\{Av_1, Av_2, \dots, Av_n\}$$

若映射后仍为正交基，则有

$$Av_i \cdot Av_j = (Av_i)^T Av_j = v_i^T A^T A v_j = 0 \quad \forall i, j \in \{1, \dots, n\}$$

现将上式如下：

由了序n维空间已知必有

$$v_i^T v_j = v_i \cdot v_j = 0$$

$$\therefore v_i^T A^T A v_j = v_i^T \lambda_j v_j = \lambda_j v_i^T v_j = \lambda_j v_i \cdot v_j = 0$$

则已变化新in 已知基：

$$Av_i \cdot Av_i = \lambda_i v_i \cdot v_i = \lambda_i \quad \Rightarrow \quad u_i = \frac{Av_i}{\|Av_i\|} = \frac{1}{\sqrt{\lambda_i}} Av_i$$

$$\therefore |Av_i|^2 = \lambda_i \geq 0$$

$$\begin{cases} Av_i = \sigma_i u_i & 0 < i < \text{Rank}(A) \\ \sigma_i = \frac{1}{\sqrt{\lambda_i}} \quad (\text{奇异值}) & \end{cases}$$

由上可得：

已知基 u_1, \dots, u_k (若 $k < m$, 则将其扩充到 m 个)

则有 m 维空间 \mathbb{R}^m 中已知基 $\{u_1, \dots, u_m\}$

现在有：

$\text{rank}(A)=k$: 仅有 k 维无关

$$A \left[\begin{matrix} v_1, v_2, \dots, v_k \\ \hline v_{k+1}, \dots, v_n \end{matrix} \right] = \left[\begin{matrix} u_1, u_2, \dots, u_k \\ \hline u_{k+1}, \dots, u_m \end{matrix} \right]$$

由矩阵 A 的列向量构成的基

由矩阵 A 的行向量构成的基

即从 m 维空间 \mathbb{R}^m 到 n 维空间 \mathbb{R}^n 的线性映射

设将 v_i 变换为 u_i

设 v_i 为奇数

设 v_i 为偶数

$\left[\begin{array}{c|c} \sigma_1 & \cdots & \sigma_k & | & 0 \\ \hline 0 & & 0 & | & 0 \end{array} \right]$

σ_i 为奇数

由此可得

A矩阵的奇异值分解为

$$A = U \Sigma V^T$$

$m \times m$ 正交阵

其中 U 为 AA' 的特征向量

称为 A 的左奇异向量



$m \times n$ 对角阵

$n \times n$ 对角阵

其中 Σ 为 $A'A$ 的特征向量

即右奇异向量

矩阵 A 的高斯分解

$$A = U \Sigma V^T$$

$$A = [u_1 \cdots u_k | u_{k+1} \cdots u_m]$$

$$\begin{array}{c} \\ \downarrow \\ \hline \end{array} \quad \begin{array}{c} \left[\begin{array}{cc|c} \sigma_1 & \dots & \sigma_k & 0 \\ 0 & \dots & 0 & 0 \end{array} \right] \\ \hline \end{array} \quad \begin{array}{c} \left[\begin{array}{c} v_1^T \\ \vdots \\ v_k^T \\ \hline v_{k+1}^T \\ \vdots \\ v_n^T \end{array} \right] \\ \hline \end{array}$$

分块乘法 ↓

$$A = [u_1 \cdots u_k] \left[\begin{array}{cc} \sigma_1 & \\ & \ddots \\ & & \sigma_k \end{array} \right] \left[\begin{array}{c} v_1^T \\ \vdots \\ v_k^T \end{array} \right]$$

今

$$\left\{ \begin{array}{l} X = [u_1 \cdots u_k] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix} = [\sigma_1 u_1 \cdots \sigma_k u_k] \\ Y = \begin{bmatrix} v_1^T \\ \vdots \\ v_k^T \end{bmatrix} \end{array} \right.$$

則 $A = XY$ 即為 A 在高斯分解

Solution: Singular Value Decomposition



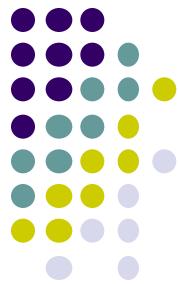
Let \hat{Y} be the **centered** $d \times N$ data matrix (assume $N > d$).

$$\sum_i \mathbf{y}_i = \mathbf{0} \quad \hat{\mathbf{Y}} = \begin{pmatrix} y_{1,1} & \boxed{y_{2,1}} & y_{3,1} & \cdots & y_{N,1} \\ y_{1,2} & y_{2,2} & y_{3,2} & \cdots & y_{N,2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{1,d} & y_{2,d} & y_{3,d} & \cdots & y_{N,d} \end{pmatrix}_{d \times N} \quad \begin{matrix} \text{Singular values} \\ \downarrow \\ = \mathbf{U} \mathbf{S} \mathbf{V}^T \\ \uparrow \\ \text{Unitary Matrices} \end{matrix}$$

$$\arg \min_{\mathbf{X}, \mathbf{W}} \|\mathbf{Y} - \mathbf{X}^T \mathbf{W}\|$$

$$\mathbf{X} = \mathbf{U}^T, \mathbf{W} = \mathbf{S} \mathbf{V}^T$$

Compute eigen~ (vectors and values)



- Eigen problem $\underset{\Delta}{\mathbf{z}\mathbf{v}} = \lambda \mathbf{v}$
- Characteristic polynomial

λ为特征值
v为特征向量

$$\det(\mathbf{Z} - \lambda \mathbf{I}) = 0$$

- Iterative method (when matrix is very **huge**)
 - Simplest method: $\mathbf{v}^{(n+1)} = \mathbf{Z} \mathbf{v}^{(n)}$ (迭代五次收敛示例)
 - Mostly used method: **Lanczos method**
 - http://en.wikipedia.org/wiki/Lanczos_algorithm

EVD 特征值分解

- 特征值 λ_i
- 單位特征向量 x_i
- 正交对称阵 A
($m \times m$)

$$\begin{aligned} Ax_1 &= \lambda_1 x_1 \\ Ax_2 &= \lambda_2 x_2 \\ &\vdots \\ Ax_m &= \lambda_m x_m \\ &\Downarrow \\ A = U \Lambda & \end{aligned}$$

由3对称矩阵特征向量而得

$$\therefore U \text{为正交阵 } (U^T = U^{-1})$$

$$A = U \Lambda U^{-1} = U \Lambda U^T$$

正交性
转置=逆

$$\left\{ \begin{array}{l} U = [x_1 \ x_2 \ \cdots \ x_m] \\ \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots & \lambda_m \end{bmatrix} \end{array} \right.$$

也将该向量 X 分解到 A 的子空间中：

$$A X = U \Lambda U^T X \quad (\text{即映射到 } A \text{ 的子空间向量})$$

换成坐标系中)

$$X = a_1 X_1 + a_2 X_2 + \dots + a_m X_m$$

$$\underbrace{U \Lambda U^T X}_{\text{分解到-部分}} = U \Lambda \begin{bmatrix} X_1^T \\ X_2^T \\ \vdots \\ X_m^T \end{bmatrix} (a_1 X_1 + a_2 X_2 + \dots + a_m X_m) = U \Lambda \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$$

用 X_i 表示的形式

△

A 的子空间向量]

$$U \wedge \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = U \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = U \begin{bmatrix} \lambda_1 a_1 \\ \lambda_2 a_2 \\ \vdots \\ \lambda_m a_m \end{bmatrix}$$

[△]
打开 \wedge

* 此时若 A 不满秩
则有 $\lambda_k a_k = 0$
即存在程度退化