证明: $A_{m\times n} = U_{m\times m} \times \Sigma_{m\times n} \times V^T_{n\times n}$ (SVD)

要证明SVD首先需要有以下证明:

① $r(A^TA) = r(A)$

证明: $\boxed{\begin{array}{c} Ax=0 \\ \Downarrow \\ A^T(Ax)=0 \end{array}}$ ~~($A^TA)x=$~~

$A^TAx = 0$

考虑: $(Ax)^T \cdot Ax = x^T \underline{A^T \cdot Ax} = 0$

又因为 $(Ax)^T \cdot Ax = \|Ax\|^2 = 0$

$\therefore \quad Ax = 0$

$A_{m\times n}$

根据秩零和定理

$rank(A) + nullity(A) = n$

因为 Ax 与 A^TAx 的零空间互为子集，所以 Ax 与 A^TAx 的零空间相等，即 $nullity(A) = nullity(A^TA)$

又因为

$rank(A^TA) + nullity(A^TA) = n$

所以 $rank(A^TA) = rank(A)$

② A^TA 为对称矩阵

$A^TA = (A^TA)^T = A^TA$

$\because$ A^TA 为 $n\times n$ 矩阵 λ值为关于λ^n的方程，故共有 n 个λ值，

非零特征值为 $rank(A^TA) = r$

$v_1 \quad v_2 \cdots v_r \cdots\cdots v_n$

$\underbrace{\lambda_1 \quad \lambda_2 \cdots \lambda_r}_{\lambda > 0} \quad \underbrace{\cdots \quad \lambda_n}_{\substack{r+1 \to n \\ \lambda = 0}}$

③ 证明 A^TA 的特征值为非负数

$A^TA \cdot x = \lambda x$

$x^T A^TA \cdot x = x^T \lambda x$

$(Ax)^T \cdot Ax = \lambda x^T x \geq 0$

$\lambda \|x\|^2 \geq 0$

$\lambda \geq 0$

有了以上证明，现开始证明 $A_{m\times n} = U_{m\times m} \cdot \Sigma_{m\times n} V_{n\times n}$

设 V_i 为 A^TA 的特征向量 V_i为n维向量

构造以下形式

$$\overset{1\times n}{V_i^T} \overset{n\times n}{(A^TA)} \overset{n\times 1}{V_j} = V_i^T \times \lambda_j V_j = \lambda_j \cdot V_i^T \cdot V_j$$

$$V_i^T (A^T A) \cdot V_j = \begin{cases} \lambda_j & i=j \leq r \\ 0 & i=j > r \\ 0 & i \neq j \end{cases}$$

$$V_i^T (A^T \cdot A) \cdot V_j = (AV_i)^T \cdot AV_j$$

考虑两种情况：① $i=j \leq r$ ② $i=j > r$

$$\|AV_i\|^2 = \begin{cases} \lambda_i & i=j \leq r \\ 0 & i=j > r \end{cases}$$

$$\therefore \|AV_i\| = \begin{cases} \sqrt{\lambda_i} & i=j \leq r \\ 0 & i=j > r \end{cases}$$

设 $u_i = \dfrac{AV_i}{\|AV_i\|}$ （u_i为 m维向量）

则 $$AV_i = \begin{cases} \sqrt{\lambda_i} \cdot u_i & i=j \leq r \\ 0 & i=j > r \end{cases}$$

$$\overset{m\times n}{A}\overset{n\times n}{V} = A(V_1\ V_2\ V_3 \cdots V_n)$$

$$= (AV_1\ AV_2 \cdots AV_n)$$

$$= (\sqrt{\lambda_1} u_1\ \sqrt{\lambda_2} u_2 \cdots \sqrt{\lambda_n} u_n)$$

现考虑 $u_i^T \cdot u_j$ （$i \neq j$ 且 i和$j \leq r$）

$$u_i^T \cdot u_j = \frac{V_i^T A^T \cdot A^T V_j^T}{\|AV_i\| \cdot \|AV_j\|} = \frac{V_i^T \lambda_j V_j}{\|AV_i\| \cdot \|AV_j\|} = \frac{\lambda_j V_i^T \cdot V_j}{\|AV_i\| \cdot \|AV_j\|}$$

$\because V_i$ 与 V_j 为正交特征向量 $\therefore V_i^T \cdot V_j = 0$ $\therefore u_i \cdots u_r$ 互为正交关系

$$AV = A(V_1\ V_2\ V_3 \cdots V_n)$$

$$= (\sqrt{\lambda_1}\cdot u_1, \sqrt{\lambda_2}\,u_2 \cdots \sqrt{\lambda_n}\,V_n)$$

$\because$ λ_i 在 $i>r$ 时 λ_i 为零

$$\therefore\ AV = (\sqrt{\lambda_1}\,u_1\ \sqrt{\lambda_2}\,u_2 \cdots \sqrt{\lambda_r}\cdot u_r \cdots 0 \cdots)$$

因为u为 m维向量，在m维空间中，每组标准正交基中最多有m个向量，所以 $r \leqslant m$，可以使用施密特正交化构造出 m个正交的u向量

$$\therefore\ AV = (\sqrt{\lambda_1}\,u_1\ \sqrt{\lambda_2}\,u_2 \cdots \sqrt{\lambda_r}\ u_r \cdots 0\cdot u_m)$$

$$= (u_1\ u_2\ u_3 \cdots u_r \cdots u_m)\cdot \begin{matrix} \sqrt{\lambda_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & 0 & \cdots & 0 \\ 0 & 0 & \sqrt{\lambda_3} & 0 \cdots & 0 \\ \vdots & \vdots & 0 & \sqrt{\lambda_r} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{matrix}$$

$$\therefore\ AV = U\Sigma$$

$\because$ V为正交矩阵

$$A = U\Sigma\cdot V^{-1} = U\Sigma V^T$$

最终 $A_{m\times n} = \underset{m\times m}{U}\cdot\underset{m\times n}{\Sigma}\cdot\underset{n\times n}{V^T}$

由于本人经常搞不清楚什么列观点、行观点、

故而把 SVD 写成下以下行式

$$\underset{m\times n}{A}=\underset{m\times m}{U}\ \underset{m\times n}{\Sigma}\ \underset{n\times n}{V^T}$$

$$A=\begin{pmatrix}u_1 & u_2 & u_3 & \cdots & u_m\end{pmatrix}\times\begin{pmatrix}\sqrt{\lambda_1} & 0 & 0 & \cdots & 0 & \cdots & 0\\ 0 & \sqrt{\lambda_2} & 0 & \cdots & 0 & \cdots & 0\\ 0 & 0 & \sqrt{\lambda_3} & \cdots & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & & & \vdots\\ 0 & 0 & 0 & \cdots & \sqrt{\lambda_r} & \cdots & 0\\ \vdots & \vdots & \vdots & & & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0\end{pmatrix}\begin{pmatrix}v_1^T\\ v_2^T\\ v_3^T\\ \vdots\\ v_m^T\end{pmatrix}$$

从左到右计算

$$A=\begin{pmatrix}\sqrt{\lambda_1}u_1 & \sqrt{\lambda_2}u_2 & \sqrt{\lambda_3}u_3 & \cdots & \sqrt{\lambda_r}u_r & \cdots & 0\cdot u_m\end{pmatrix}\begin{pmatrix}v_1^T\\ v_2^T\\ v_3^T\\ \vdots\\ v_r^T\\ \vdots\\ v_m^T\end{pmatrix}$$

$$=\sqrt{\lambda_1}u_1v_1^T+\sqrt{\lambda_2}u_2v_2^T+\sqrt{\lambda_3}u_3v_3^T+\cdots+\sqrt{\lambda_r}u_rv_r^T+\cdots+0\cdot u_m\cdot v_m^T$$

值得注意的是 u_i 为 $m\times 1$ 的列向量，v_i 为 $1\times n$ 的行向量

u_iv_i 为一个 $m\times n$ 的矩阵

所以 A 等于 m 个 $m\times n$ 的矩阵相加的结果

最后我简要说明一下具体SVD是怎样计算的。

已知 $\underset{m\times n}{A}=\underset{m\times m}{U}\ \underset{m\times n}{\Sigma}\ \underset{n\times n}{V^T}$

① $A^TA=(U\Sigma V^T)^T(U\Sigma V^T)$

$=V\Sigma^TU^TU\Sigma V^T$

$=V(\Sigma)^2V^T$

② $AA^T=(U\Sigma V^T)\cdot(U\Sigma V^T)^T$

$=U\Sigma V^T\cdot V\Sigma^TU^T$

$=U(\Sigma)^2U^T$

③ 校正：（因为 $-U$ 和 $-V$ 也满足）

$AV=U\Sigma V^TV$

$AV=U\Sigma$

$$A(V_1\ V_2)=(u_1\ u_2)\begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix}$$

$$(AV_1\ AV_2)=(u_1s_1\ \ u_2s_2)$$

$$u_1=\frac{AV_1}{s_1}$$

$$u_2=\frac{AV_2}{s_2}$$

为了说明矩阵乘法，现考虑以下矩阵

$$A=\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} \qquad Q=\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

$$A\cdot Q=\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix}\cdot\begin{pmatrix} A & B \\ C & D \end{pmatrix}=\begin{pmatrix} Aa_1+Cb_1 & Ba_1+Db_1 \\ Aa_2+Cb_2 & Ba_2+Db_2 \\ Aa_3+Cb_3 & Ba_3+Db_3 \end{pmatrix}$$

$$=\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}\cdot A+\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} C\ ,\ \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}\cdot B+\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\cdot D\right)=\left(A\begin{pmatrix} A \\ C \end{pmatrix}, A\begin{pmatrix} B \\ D \end{pmatrix}\right)$$

如果 Q为对角矩角，可以改写成以下形式

Q为对角矩阵即 $C=B=0$

所以 $A=(P_1A\ ,\ P_2D) \qquad P_1=\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \qquad P_2=\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$

再考虑 $A\cdot Q^T$

$$=\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix}\cdot\begin{pmatrix} A & C \\ B & D \end{pmatrix}$$

$$=(P_1, P_2)\cdot\begin{pmatrix} m_1^T \\ m_2^T \end{pmatrix}$$ 其中 $m_1=\begin{pmatrix} A \\ C \end{pmatrix}$ $m_2=\begin{pmatrix} B \\ D \end{pmatrix}$，$m_1/m_2$ 为Q的列向量

$$=P_1\cdot m_1^T+P_2\cdot m_2^T$$

可以展开检查一下：

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}(A\ \ C)+\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}(B\ \ D)$$

$$=\begin{pmatrix} Aa_1 & Ca_1 \\ Aa_2 & Ca_2 \\ Aa_3 & Ca_3 \end{pmatrix}+\begin{pmatrix} Bb_1 & Db_1 \\ Bb_2 & Db_2 \\ Bb_3 & Db_3 \end{pmatrix}=\begin{pmatrix} Aa_1+Bb_1 & Ca_1+Db_1 \\ Aa_2+Bb_2 & Ca_2+Db_2 \\ Aa_3+Bb_3 & Ca_3+Db_3 \end{pmatrix}$$