Lecture 3: singular value decomposition (SVD)

1. 回顾特征值和特征向量

首先回顾下特征值和特征向量的定义如下：

$$Ax = \lambda x$$

其中 A 是一个 $n \times n$ 矩阵，x 是一个 n 维向量，则 λ 是矩阵 A 的一个特征值，而 x 是矩阵 A 的特征值 λ 所对应的特征向量。

求出特征值和特征向量有什么好处呢？ 就是我们可以将矩阵A特征分解。如果我们求出了矩阵A的n个特征值 $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ ，以及这 n 个特征值所对应的特征向量 $w_1, w_2, \ldots, w_n$ ，

那么矩阵A就可以用下式的特征分解表示：

$$A = W\Sigma W^{-1}$$

其中W是这n个特征向量所张成的n×n维矩阵，而Σ为这n个特征值为主对角线的n×n维矩阵。

一般我们会把W的这n个特征向量标准化，即满足 $||w_i||_2 = 1$ ，或者 $w_i^T w_i = 1$ ，此时W的

n个特征向量为标准正交基，满足 $W^T W = I$ ，即 $W^T = W^{-1}$ ，也就是说W为酉矩阵。

这样我们的特征分解表达式可以写成

$$A = W\Sigma W^T$$

注意到要进行特征分解，矩阵A必须为方阵。

那么如果A不是方阵，即行和列不相同时，我们还可以对矩阵进行分解吗？答案是可以，此时我们的SVD登场了。

Singular value decomposition

- **Linear algebra (线性代数)** plays a central role in almost every application area of mathematics in the physical, engineering and biological sciences. It is perhaps the most important theoretical framework to be familiar with as a student of mathematics. Thus it is no surprise that it also plays a key role in data analysis and computation.

- In what follows, emphasis will be placed squarely on the **singular value decomposition (SVD) (奇异值分解)**. It forms one of the most powerful techniques for **analyzing a myriad (大量的) of application areas**.

In even the earliest experience of linear algebra, the concept of a **matrix transforming** a vector via multiplication was defined(矩阵变换是通过向量的乘法来定义的).

For instance, the vector x when multiplied by a matrix A produces a new vector y that is now aligned (对齐的), generically, in a new direction with a new length. To be more specific, the following example illustrates a particular transformation:

$$\mathbf{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \rightarrow \mathbf{y} = \mathbf{Ax} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

Figure 15.1(a) shows the vector x and its transformed version, y, after application of the matrix A. Thus generically, **matrix multiplication** will rotate (旋转) and stretch (compress) (拉伸) a given vector as prescribed by the matrix $\boldsymbol{A}$ (see Fig. 15.1(a)).

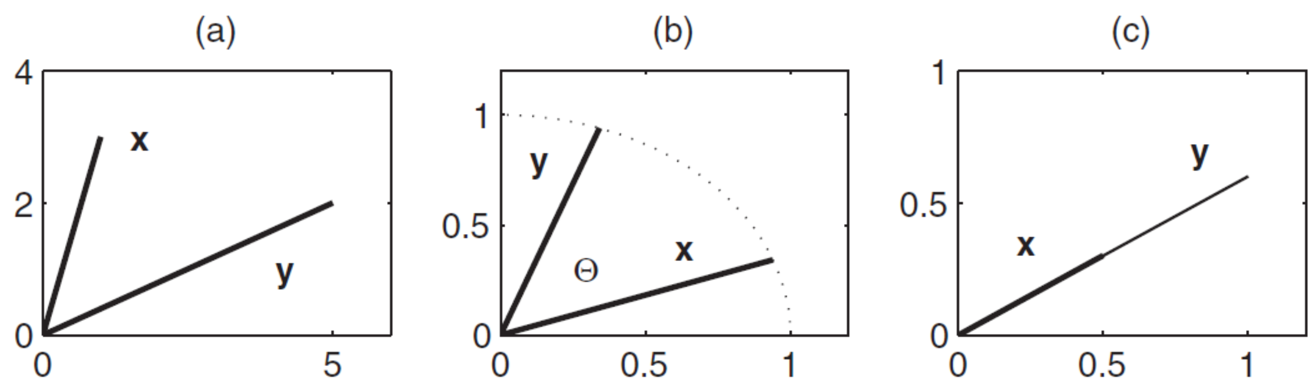


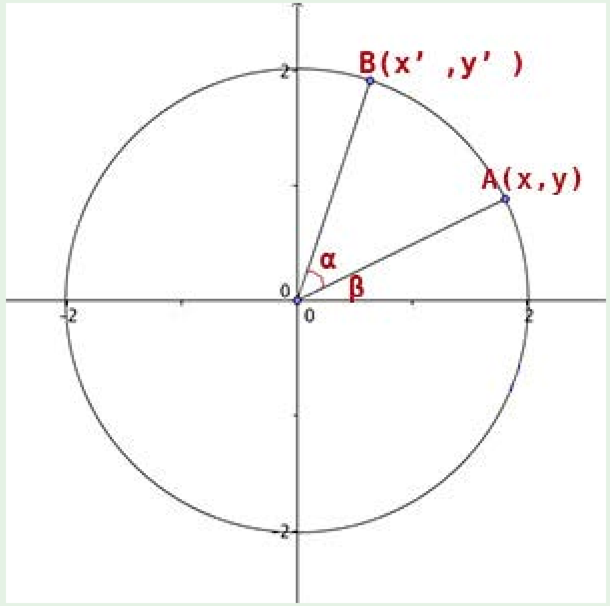
Figure 15.1: Transformation of a vector $\mathbf{x}$ under the action of multiplication by the matrix $\mathbf{A}$, i.e. $\mathbf{y} = \mathbf{A}\mathbf{x}$. (a) Generic rotation and stretching of the vector as given by Eq. (15.1.1). (b) Rotation by 50° of a unit vector by the rotation matrix (15.1.2). (c) Stretching of a vector to double its length using (15.1.3) with $\alpha = 2$.

The rotation (旋转) and stretching (拉伸) of a transformation can be precisely (精确地) controlled by proper construction of the matrix *A*. In particular, it is well known that in a two-dimensional space, the **rotation matrix**

$$\mathbf{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

takes a given vector *x* and rotates it by an angle θ to produce the vector *y*.

The transformation produced by *A* is known as a **unitary transformation (酉变换)** since the matrix inverse is $A^{-1} = \bar{A}^T$ where the bar denotes **complex conjugation (复共轭)**. Thus rotation can be directly specified without the vector being scaled.



两角和公式

sin(A+B) = sinAcosB+cosAsinB

sin(A-B) = sinAcosB-cosAsinB

cos(A+B) = cosAcosB-sinAsinB

cos(A-B) = cosAcosB+sinAsinB

tan(A+B) = (tanA+tanB)/(1-tanAtanB)

tan(A-B) = (tanA-tanB)/(1+tanAtanB)

cot(A+B) = (cotAcotB-1)/(cotB+cotA)

cot(A-B) = (cotAcotB+1)/(cotB-cotA)

```
x' = r*cos(α+β)
   = r*( cos(α)*cos(β)-sin(α)*sin(β) )
   = r*cos(β)*cos(α)-r*sin(β)*sin(α)
   = x*cos(α)-y*sin(α)

y' = r*sin(α+β)
   = r*( sin(α)*cos(β)+cos(α)sin(β) )
   = r*cos(β)*sin(α)+r*sin(β)*cos(α)
   = x*sin(α)+y*cos(α)
```

```
|x'|   |cos(α)  -sin(α)|   |x|
|  | = |               | * | |
|y'|   |sin(α)   cos(α)|   |y|
```

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} c_x \\ c_y \end{bmatrix} \right) + \begin{bmatrix} c_x \\ c_y \end{bmatrix}$$

https://zhuanlan.zhihu.com/p/433389563

To **scale (缩放) the vector in length**, the matrix

$$\mathbf{A} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}$$

can be applied to the vector x. This multiplies the length of the vector x by α. If $\alpha = 2\ (0.5)$, then the vector is twice (half) its original length.

The combination of the above **two matrices gives arbitrary control of rotation and scaling** in a two-dimentional vector space. Figure 15.1 demonstrates some of the various operations associated with the above matrix transformations.

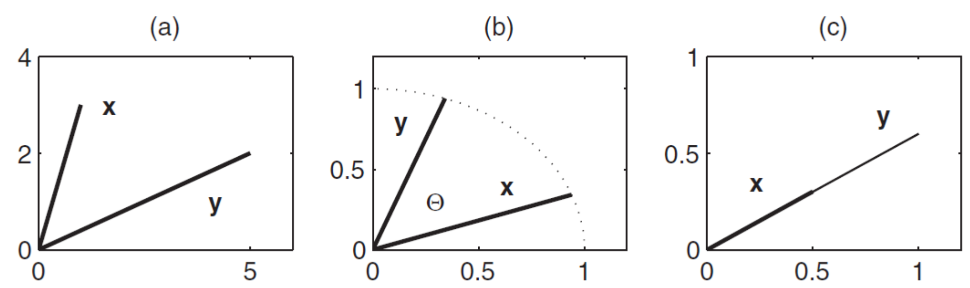


Figure 15.1: Transformation of a vector $\mathbf{x}$ under the action of multiplication by the matrix $\mathbf{A}$, i.e. $\mathbf{y} = \mathbf{A}\mathbf{x}$. (a) Generic rotation and stretching of the vector as given by Eq. (15.1.1). (b) Rotation by 50° of a unit vector by the rotation matrix (15.1.2). (c) Stretching of a vector to double its length using (15.1.3) with $\alpha = 2$.

Singular value decomposition (SVD)

A **singular value decomposition (SVD)** is a factorization (分解) of a matrix into a number of constitutive (组成) components all of which have a specific meaning in applications. The SVD, much as illustrated in the preceding paragraph, is essentially a transformation that **stretches/compresses and rotates a given set of vectors**.

In particular, the following **geometric principle** (几何原理) will guide our forthcoming(后续) discussion:

the image of a unit sphere (单位圆) under any $m \times n$ matrix is a hyper-ellipse (超椭圆). A hyper-ellipse in $\mathbb{R}^m$ is defined by the surface obtained upon stretching a unit sphere in $\mathbb{R}^m$ by some factors $\sigma_1, \sigma_2, \cdots, \sigma_m$ in the orthogonal(正交) directions $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m \in \mathbb{R}^m$. The stretchings σ_i can possibly be zero. For convenience, consider the u_j to be unit vectors so that $\|\mathbf{u}_j\|_2 = 1$. The quantities $\sigma_j \mathbf{u}_j$ are then the principal semi-axes of the hyper-ellipse with the length σ_j.

Figure 15.2 demonstrates a particular hyper-ellipse (超椭圆) created under the matrix transformation $\boldsymbol{A}$.

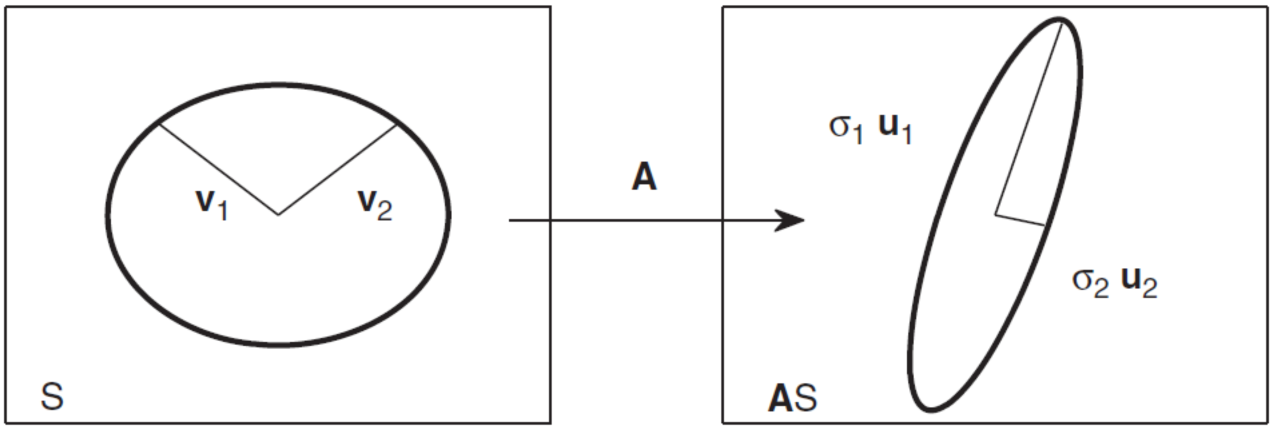
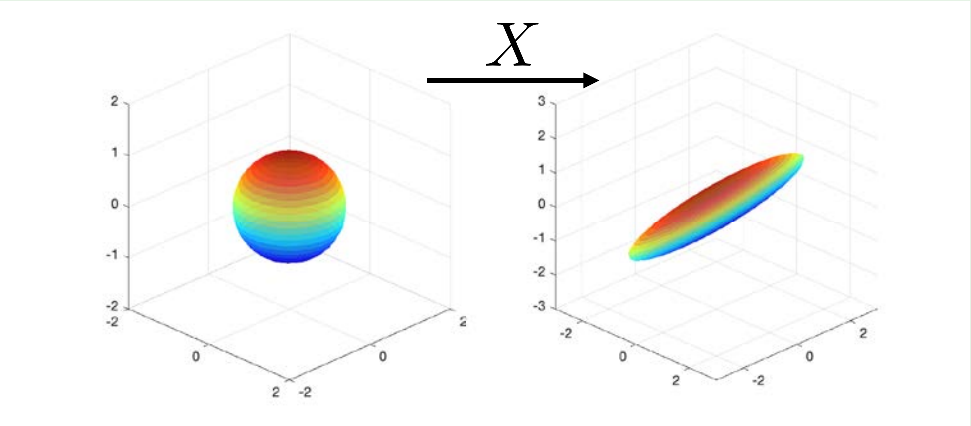


Figure 15.2: Image of a unit sphere S transformed into a hyper-ellipse $\mathbf{A}S$ in $\mathbb{R}^2$. The values of σ_1 and σ_2 are the singular values of the matrix $\mathbf{A}$ and represent the lengths of the semi-axes of the ellipse.

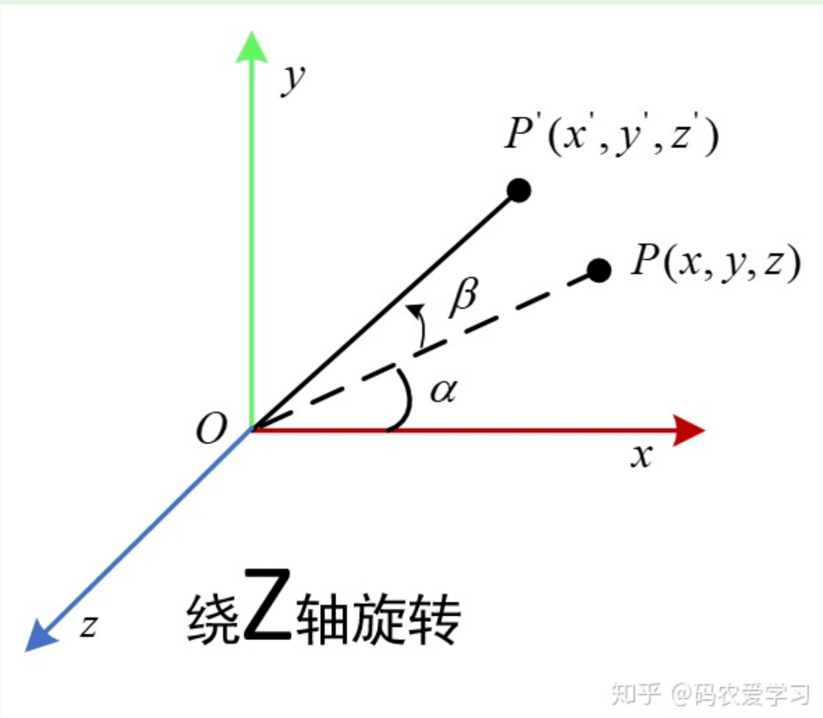
Geometric interpretation

SVD can be interpreted geometrically based on how a hyper sphere, given by $S^{n-1} \triangleq \{x \,|\, \|x\|_2 = 1\} \subset \mathbb{R}^n$ maps into an ellipsoid, $\{y \,|\, y = Xx \text{ for } x \in S^n - 1\} \subset \mathbb{R}^m$, through X.



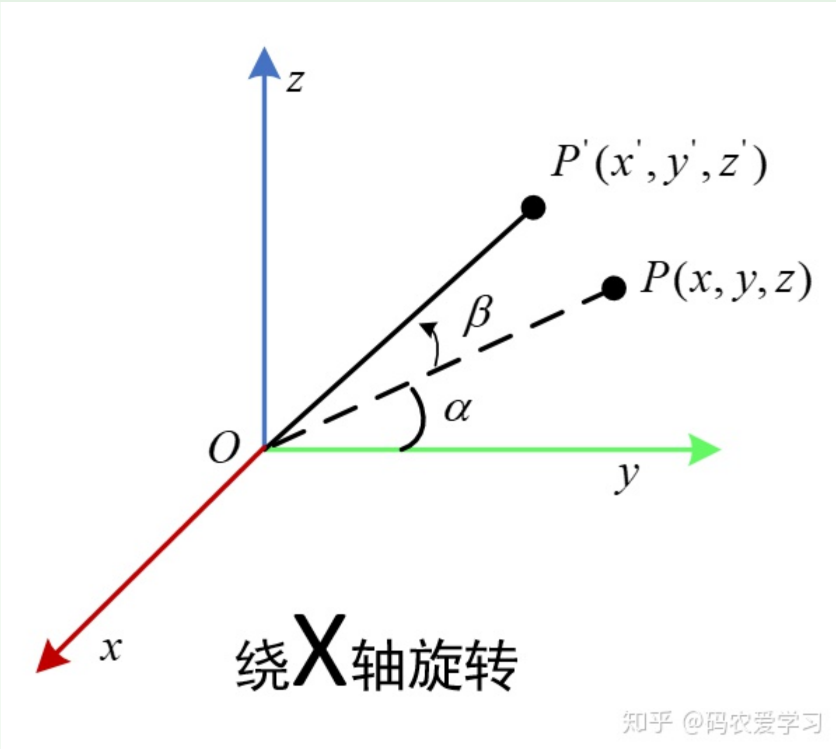
$$\mathbf{X} = \underbrace{\begin{bmatrix} \cos(\theta_3) & -\sin(\theta_3) & 0 \\ \sin(\theta_3) & \cos(\theta_3) & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{R}_z} \underbrace{\begin{bmatrix} \cos(\theta_2) & 0 & \sin(\theta_2) \\ 0 & 1 & 0 \\ -\sin(\theta_2) & 0 & \cos(\theta_2) \end{bmatrix}}_{\mathbf{R}_y} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_1) & -\sin(\theta_1) \\ 0 & \sin(\theta_1) & \cos(\theta_1) \end{bmatrix}}_{\mathbf{R}_x} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix}$$

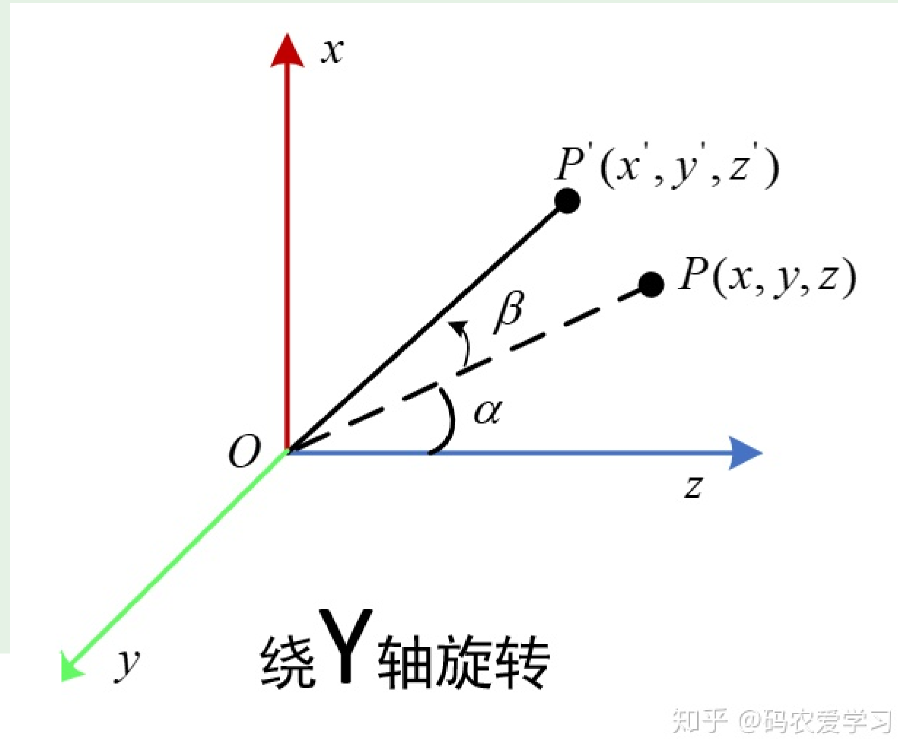
https://zhuanlan.zhihu.com/p/183973440



$$\begin{cases} x' = x \cdot \cos\beta - y \cdot \sin\beta \\ y' = x \cdot \sin\beta + y \cdot \cos\beta \\ z' = z \end{cases}$$

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$





$$\begin{cases} y' = y\cdot\cos\beta - z\cdot\sin\beta \\ z' = y\cdot\sin\beta + z\cdot\cos\beta \\ x' = x \end{cases}$$

$$\begin{bmatrix} y' \\ z' \\ x' \end{bmatrix} = \begin{bmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} y \\ z \\ x \end{bmatrix} \Rightarrow \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \boxed{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{bmatrix}} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

 础

$$X\quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{bmatrix} \leftarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \leftarrow \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}$$

$$Y\quad \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \leftarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \leftarrow \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}$$

$$Z\quad \begin{bmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \leftarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \leftarrow \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}$$

A few things are worth noting at this point.

First, if A has rank (秩) r, exactly r of the lengths σ_j will be nonzero. And if the matrix A is an $m \times n$ matrix where $m > n$, at most n of the σ_j will be nonzero. Consider for the moment a full rank (满秩) matrix A. Then the n singular values of A are the lengths of the principal semi-axes (主半轴) AS as shown in Fig. 15.2.

Convention assumes(约定) that the singular values are ordered with the largest first and then in descending order: $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0$.

On a more formal level, **the transformation from the unit sphere to the hyper-ellipse** can be more succinctly (简洁地) stated as follows:

$$\mathbf{A}\mathbf{v}_j = \sigma_j \mathbf{u}_j \quad 1 \leq j \leq n$$

https://zhuanlan.zhihu.com/p/108093909

https://zhuanlan.zhihu.com/p/108097834

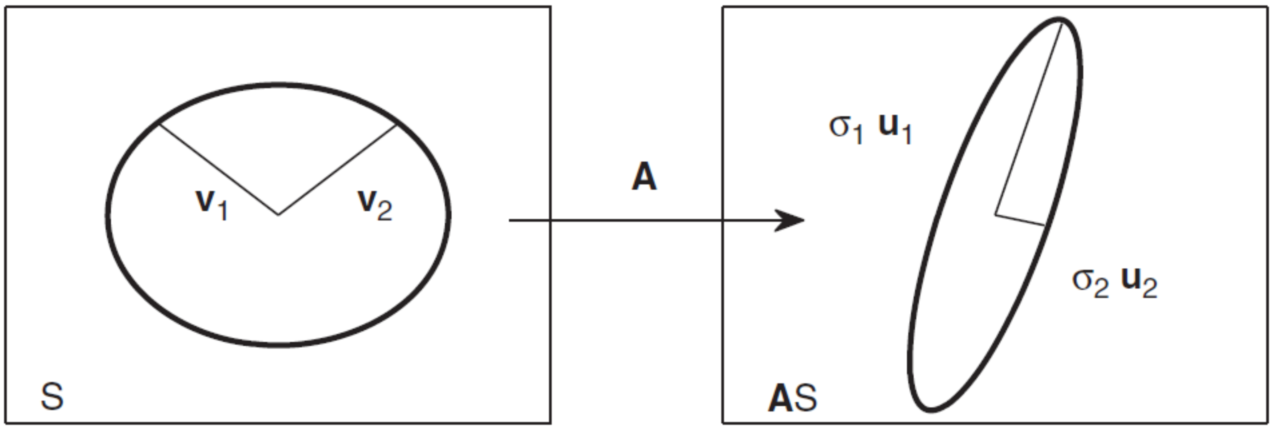


Figure 15.2: Image of a unit sphere S transformed into a hyper-ellipse $\mathbf{A}S$ in $\mathbb{R}^2$. The values of σ_1 and σ_2 are the singular values of the matrix $\mathbf{A}$ and represent the lengths of the semi-axes of the ellipse.

Thus in total, there are n vectors that are transformed under A. A more compact way to write all of these equations simultaneously is with the representation

$$\begin{bmatrix} \\ \\ & \mathbf{A} & \\ \\ \\ \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \\ \\ \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{bmatrix}$$

so that in compact matrix notation this becomes

$$\mathbf{AV} = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}$$

The matrix $\hat{\mathbf{\Sigma}}$ is an $n \times n$ diagonal matrix with positive entries (正数) provided the matrix A is of full rank. The matrix $\hat{\mathrm{U}}$ is an $m \times n$ matrix with orthonormal columns (正交列向量), and the matrix V is an $n \times n$ unitary matrix (酉矩阵).

Since V is unitary (酉矩阵), the above equation can be solved for A by multiplying on the right with V^* (共轭转置) so that

$$\mathrm{A} = \hat{\mathrm{U}}\hat{\Sigma}\mathrm{V}^*$$

This factorization (分解) is known as the reduced singular value decomposition, or **reduced SVD**, of the matrix A.

Graphically, the factorization is represented in Fig. 15.3.

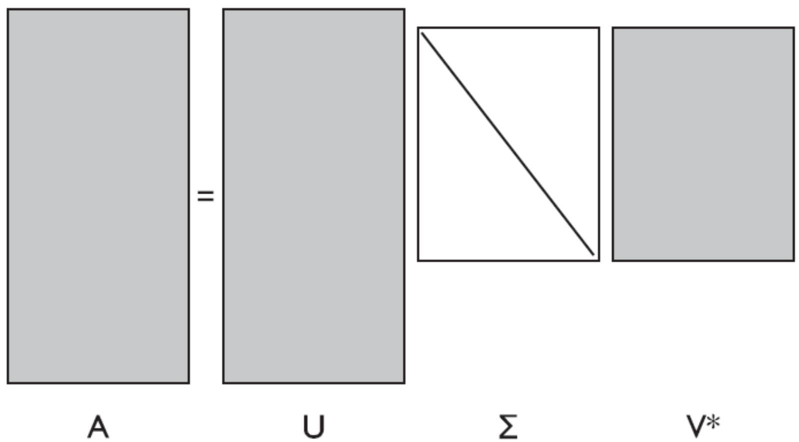


Figure 15.3: Graphical description of the reduced SVD decomposition.

酉矩阵

- 特殊的正规矩阵 $U^*U = UU^* = I_n$
- U, U^*都是酉矩阵
- 酉矩阵的特征值都是模为1的复数，即分布在复平面的单位圆上，因此酉矩阵行列式的值也为1
- 酉矩阵 与对角阵关系 $U = V\Sigma V^*$ V 是酉矩阵，Σ 是主对角线上元素绝对值为1的对角阵
- 例子

The reduced SVD is not the standard definition of the SVD used in the literature. What is typically done to augment the treatment above is to construct a matrix U from $\hat{\mathbf{U}}$ by adding an additional $m - n$ columns that are orthonormal (正交) to the already existing set in $\hat{\mathbf{U}}$. Thus **the matrix U becomes an $m \times m$ unitary matrix (酉矩阵)**.

In order to make this procedure work, an additional $m - n$ rows of zeros is also added to the $\hat{\boldsymbol{\Sigma}}$ matrix. These “silent” (无作用) columns of U and rows of $\boldsymbol{\Sigma}$ are shown graphically in Fig. 15.4.

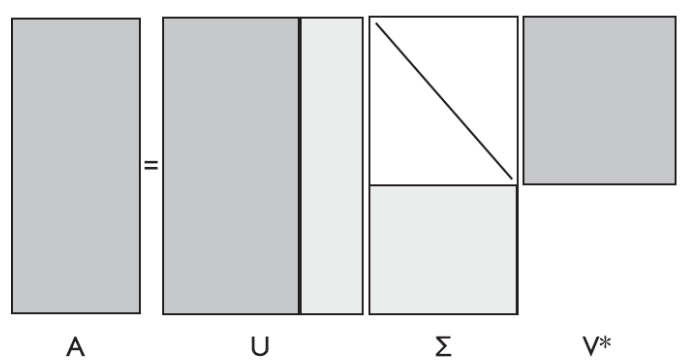


Figure 15.4: Graphical description of the full SVD decomposition where both **U** and **V** are unitary matrices. The light shaded regions of **U** and **Σ** are the silent rows and columns that are extended from the reduced SVD.

In performing this procedure, it becomes evident that rank deficient matrices (秩缩减矩阵) can easily be handled by the SVD decomposition.

In particular, instead of $m - n$ silent rows and matrices, there are now simply $m - r$ silent rows and columns added to the decomposition. Thus the matrix $\mathbf{\Sigma}$ will have r positive diagonal entries, with the remaining $n - r$ being equal to zero.

The **full SVD decomposition(完全SVD分解)** thus take the form

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^*$$

with the following three matrices

$$\mathbf{U} \in \mathbb{C}^{m\times m} \text{ is unitary}$$
$$\mathbf{V} \in \mathbb{C}^{n\times n} \text{ is unitary}$$
$$\mathbf{\Sigma} \in \mathbb{R}^{m\times n} \text{ is diagonal.}$$

Additionally, it is assumed that the diagonal entries of $\sum$ are nonnegative and ordered from largest to smallest so that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$ where $p = min(m, n)$.

The SVD decomposition of the matrix A thus shows that the matrix **first applies a unitary transformation (酉变换) preserving (保留) the unit sphere (单位圆) via V^* (共轭转置).**

This is followed by a **stretching(伸长) operation that creates an ellipse (椭圆形) with principal semi-axes (半主轴) given by the matrix** $\sum$.

Finally, the generated **hyper-ellipse is rotated(旋转) by the unitary transformation U**.

Thus the statement: the image of a unit sphere under any $m \times n$ matrix is a hyper-ellipse, is shown to be true. The following is the primary theorem concerning the SVD.

Theorem *Every matrix* $\mathbf{A} \in \mathbb{C}^{m\times n}$ *has a singular value decomposition (15.1.8). Furthermore, the singular values* $\{\sigma_j\}$ *are uniquely determined, and, if* $\mathbf{A}$ *is square and the* σ_j *distinct, the singular vectors* $\{\mathbf{u}_j\}$ *and* $\{\mathbf{v}_j\}$ *are uniquely determined up to complex signs (complex scalar factors of absolute value 1).*

Computing the SVD

The above theorem guarantees the existence of the SVD, but in practice, it still remains to be computed. This is a fairly straightforward process if one considers the following matrix products:

$$\begin{aligned}\mathbf{A}^T\mathbf{A} &= \left(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*\right)^T\left(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*\right)\\ &= \mathbf{V}\boldsymbol{\Sigma}\mathbf{U}^*\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*\\ &= \mathbf{V}\boldsymbol{\Sigma}^2\mathbf{V}^*\end{aligned}$$

and

$$\begin{aligned}\mathbf{A}\mathbf{A}^T &= \left(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*\right)\left(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*\right)^T\\ &= \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*\mathbf{V}\boldsymbol{\Sigma}\mathbf{U}^*\\ &= \mathbf{U}\boldsymbol{\Sigma}^2\mathbf{U}^*\end{aligned}$$

$$A^H = \overline{\left(A^T\right)}$$

共轭转置

$$(A+B)^H = A^H + B^H$$

$$(kA)^H = \bar{k}A^H$$

$$(AB)^H = B^H A^H$$

$$\left(A^H\right)^H = A$$

如果 A 可逆，则 $\left(A^H\right)^{-1} = \left(A^{-1}\right)^H$

正规矩阵 normal matrix

正规矩阵（英语：normal matrix）A 是与自己的共轭转置满足交换律的复系数方块矩阵，也就是说，A 满足

$$\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*$$

A^* 是 A 的共轭转置。

如果 A是实系数矩阵，则 $A^* = A^T$ ，从而条件简化为 $AA^T = A^T A$.

正规矩阵的概念十分重要，因为它们正是能使谱定理成立的对象：矩阵 A 正规当且仅当它可以被成 $A = U\Lambda U^*$ 的形式。其中的 $\Lambda = diag(\lambda_1, \lambda_2, \ldots)$ 为对角矩阵，U 为酉矩阵。

总而言之，就是 正规矩阵 一定可以 特征分解/频谱分解/谱定理。

例 若 $A=\begin{pmatrix}1 & -1\\ 0 & 1\\ 2 & 3\end{pmatrix}$, 则 $A^T=\begin{pmatrix}1 & 0 & 2\\ -1 & 1 & 3\end{pmatrix}$.

矩阵的转置满足下述运算规律

1. $(A^T)^T=A$
2. $(A+B)^T=A^T+B^T$
3. $(\lambda A)^T=\lambda A^T$
4. $(AB)^T=B^TA^T$

$(ABC)^T=C^TB^TA^T$

1.酉矩阵（unitary matrix）若n阶复矩阵A满足

$$A^HA=AA^H=E$$

则称A为酉矩阵，记之为 $A\in U^{N\times N}$。其中， A^H是A的共轭转置。

2.性质

如果A是酉矩阵

(1) $A^{-1}=A^H$;

(2) A^{-1}也是酉矩阵；

(3) $det(A)=1$;

(4)充分条件是它的n个列向量是两两正交的单位向量。

对于多个矩阵相乘，有 $(A_1A_2\cdots A_t)^T=A_t^T\cdots A_2^TA_1^T$

Multiplying (15.1.10) and (15.1.11) on the right by V and U, respectively, gives the two self-consistent (自一致性) eigenvalue problems

$$\mathbf{A}^T\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{\Sigma}^2$$
$$\mathbf{A}\mathbf{A}^T\mathbf{U} = \mathbf{U}\mathbf{\Sigma}^2$$

Thus if the **normalized eigenvectors(归一化特征向量)** are found for these two equations, then the orthonormal basis vectors (标准正交基向量) are produced for U and V. Likewise, **the square root of the eigenvalues of these equations produces the singular values σ_j. (特征值的平方根为奇异值)**

刚才提到，线性变换的其中一种方式即为缩放，但如果一个矩阵只是对向量有缩放的影响，就比较有趣了，这一定是一个方阵：

$$Au = \lambda u$$

也就是对向量 u 做方阵 A 的线性变换，其结果就是其做 λ 倍的缩放，这里的 λ 可以大于1，也可以小于1，可以大于0，也可以小于0，小于0就相当于做了镜像变成了原来相反的方向。对于给定的一个方阵 A ，具有这样有趣性质的向量 u 和数值 λ 的个数是有限的，而且方阵 A 既然表示线性变换，那么这种维持线性不变性的向量是否就表示了这个变换的某种本质特征呢，起个名字吧，就叫本征向量或者**特征向量**，数值 λ 作为缩放的方向和比例就叫本征值或者**特征值**，特征值越大说明方阵在这个方向上的能量越大，比如PCA中选择特征值大的部分特征向量组成的空间中做投影就是希望能够尽量保持足够大的方差或者信息。既然说到了特征值和特征向量，这都是对方阵而言的，那么不是方阵的话有没有这种能够衡量矩阵属性的概念呢？也是有的，那就是**奇异值**，不止方阵，任何矩阵都可以析构傅成这样一种形式：

$$A = \mu \Sigma \delta^T$$

这就是奇异值分解，其中 μ 和 δ 为正交矩阵， Σ 为只含对角线元素的矩阵，其中对角线的值即为奇异值，它表示矩阵 A 的作用是将一个向量 δ 这组基构成的空间线性变换到 μ 这组基构成的空间中去，一个矩阵的奇异值是固定的，二左右奇异矩阵则是不定的，可以找到多种形式，这里的变换同样可以包含旋转、缩放和投影。这个 Σ 矩阵所代表的奇异值很有意思，其中奇异值也就是类似特征值在对应奇异向量上的缩放因子，不为0的奇异值个数为该矩阵的秩。因为矩阵的行列式的绝对值为所有奇异值的积，所以如果有一个对角线元素为0，则称该矩阵为**奇异矩阵**，反之为非奇异矩阵，其中矩阵的谱范数就是最大奇异值。事实上，奇异值刻画了矩阵的奇异性，也就是其行或列的线性相关性，相关性越强，则非奇异矩阵越接近奇异。

这里我们谈到特征值和奇异值，那么它们之间有什么异同呢？首先从形式上来说，奇异值适合于任何矩阵，而特征值只是对方阵而言的，事实上，即使对于同一个方阵来说，特征值与奇异值之间也没有什么内在的关联，但是矩阵 A_{m*n} 的非零奇异值是 A^TA 或 AA^T 非零特征值的正平方根。

Example

Consider the SVD decomposition of

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

The following quantities are computed:

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}$$

$$\mathbf{A}\mathbf{A}^T = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}$$

The eigenvalues are clearly $\lambda = \{9, 4\}$, giving singular values (奇异值) $\sigma_1 = 3$ and $\sigma_2 = 2$.

The eigenvectors can similarly be constructed and the matrices *U* and *V* are given by

$$\begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix}$$

where there is an indeterminate sign (不确定的标志) in the eigenvectors.

However, a self-consistent choice of signs must be made. One possible choice gives

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

The SVD can be easily computed in MATLAB with the following command:

```
1 [U,S,V] = SVD(A)
```

where the *[U, S,V]* correspond to *U*, Ε and *V*, respectively. This decomposition (分解) is a critical tool for analyzing many data driven phenomena. But before proceeding to such applications, the connection of the SVD with standard and well-known techniques is elucidated (阐明).