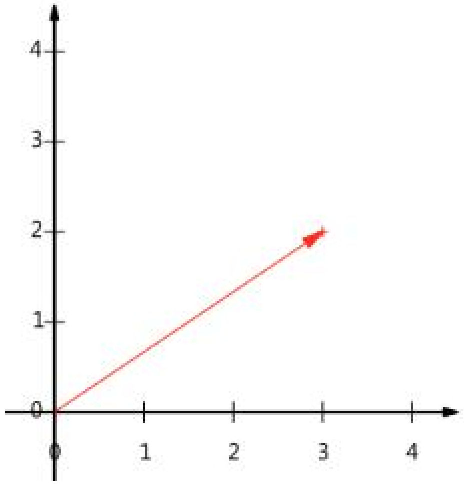
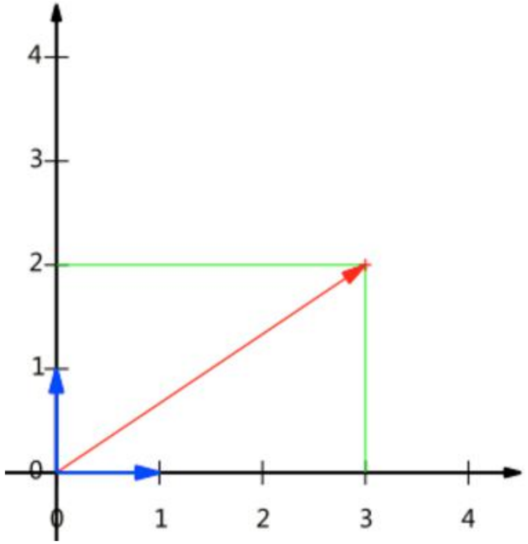
SVD矩阵分解

向量的表示及基变换

- 向量可以表示为(3,2)
  实际上表示线性组合：$x(1,0)^\mathsf{T} + y(0,1)^\mathsf{T}$



- 基：(1,0)和(0,1)叫做二维空间中的一组基

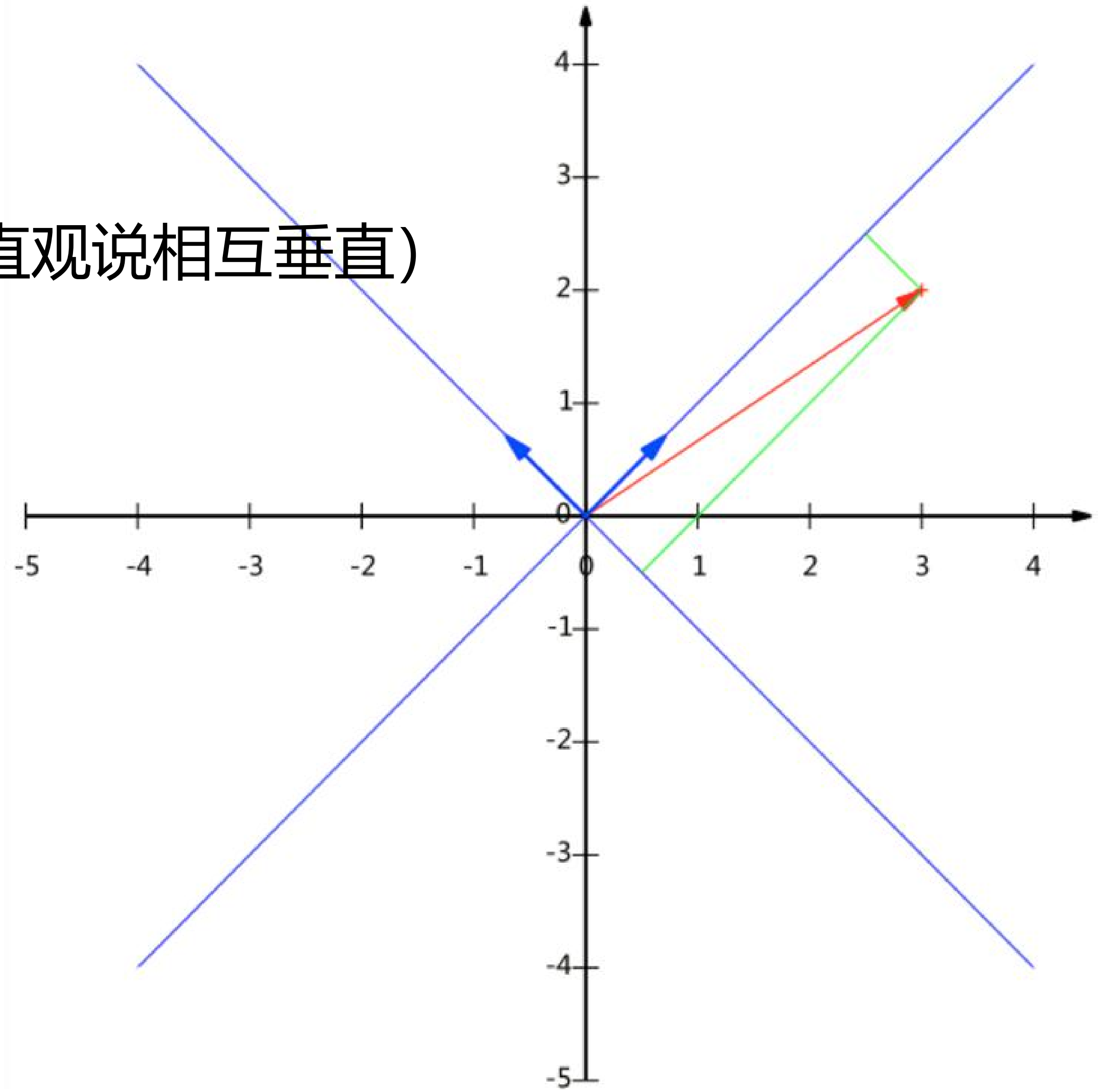


SVD矩阵分解

- 基变换

  - 基是正交的（即内积为0，或直观说相互垂直）

  - 要求：线性无关



SVD矩阵分解

基变换

变换：数据与一个基做内积运算，结果作为第一个新的坐标分量，然后与第二个基做内积运算，结果作为第二个新坐标的分量

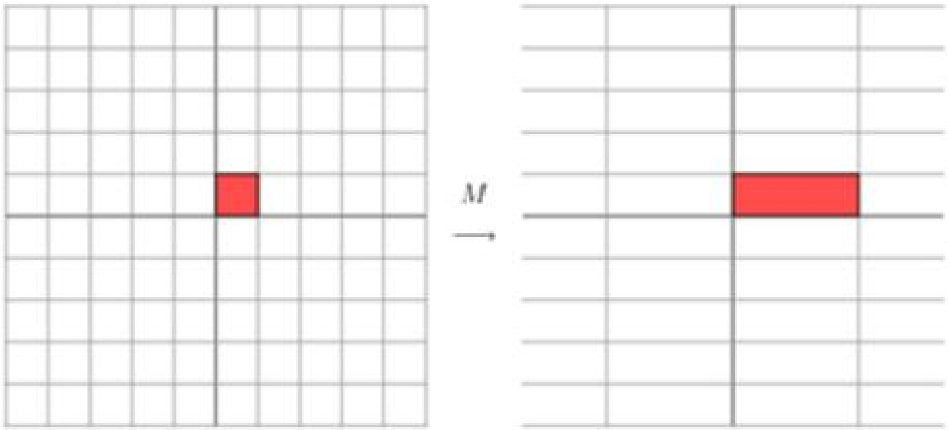
数据（3，2）映射到基中坐标：$$\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 5/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

SVD矩阵分解

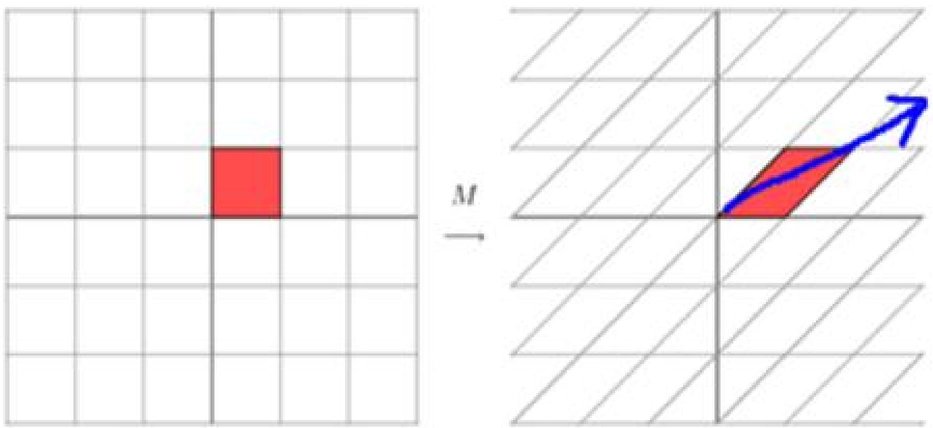
- 矩阵乘以一个向量

- 结果仍是一个向量

$$\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x \\ y \end{bmatrix}$$



$$M = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$



SVD矩阵分解

- 特征值分解

- 矩阵里面的信息有很多呀？来分一分吧！ $A = U\Lambda U^{-1}$ U：特征向量 ^：特征值

- 当矩阵是N*N的方阵且有N个线性无关的特征向量时就可以来玩啦！

- 这时候我们就可以在对角阵当中找比较大的啦，他们就是代表了！

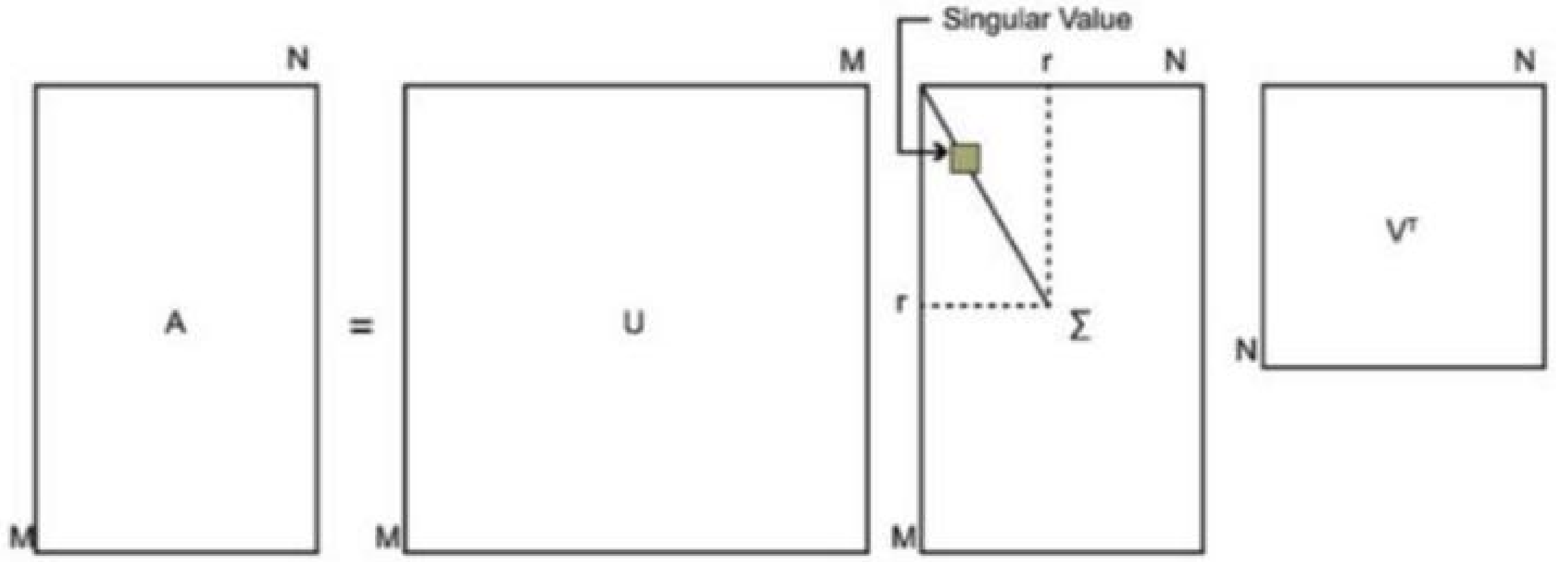
  用于矩阵分解，矩阵近似，矩阵降维

SVD矩阵分解

- SVD

特征值分解不挺好的嘛，但是它被限制住了，如果我的矩阵形状变了呢?

但是问题也来了，如果M和N都很大呢?



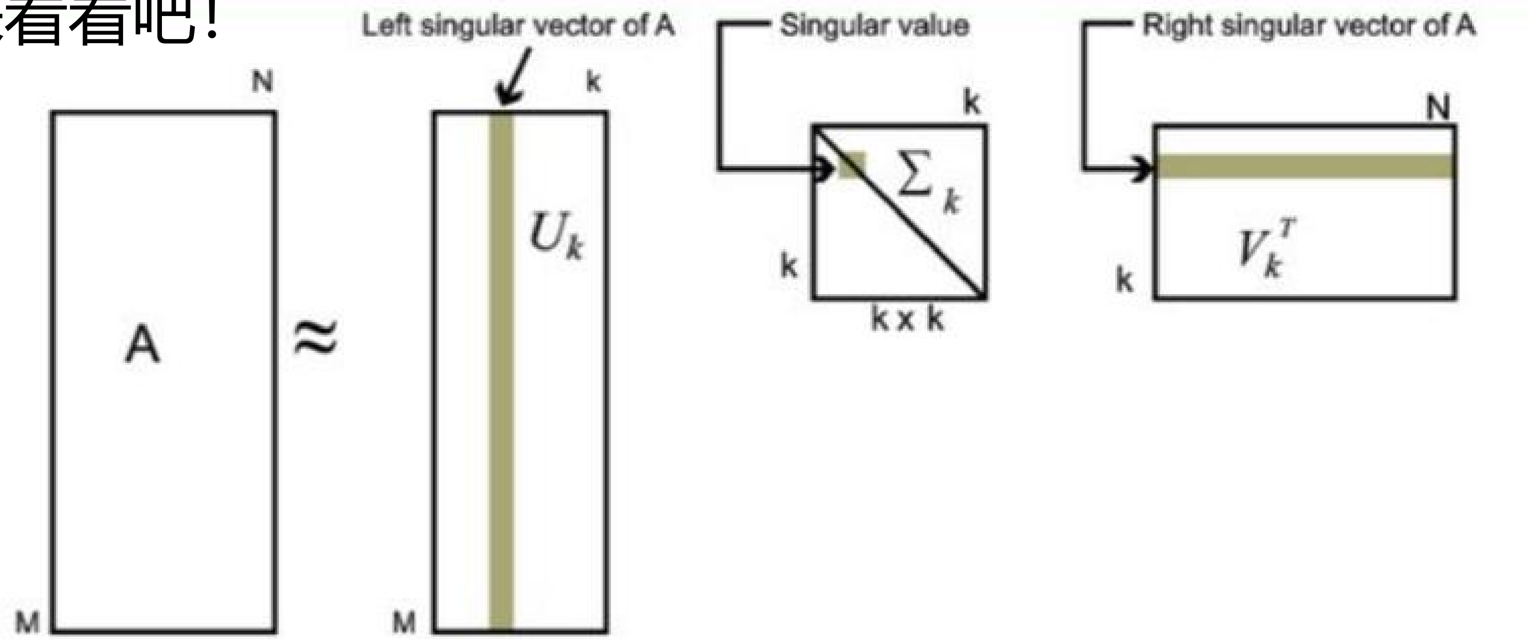
SVD矩阵分解

✓ SVD

SVD用于特征压缩，特征降维；
推荐系统基于SVD提取主要信息

- 照样按照特征值的大小来进行筛选，一般前10%的特征值（甚至更少）的和就占到了总体的99%了。

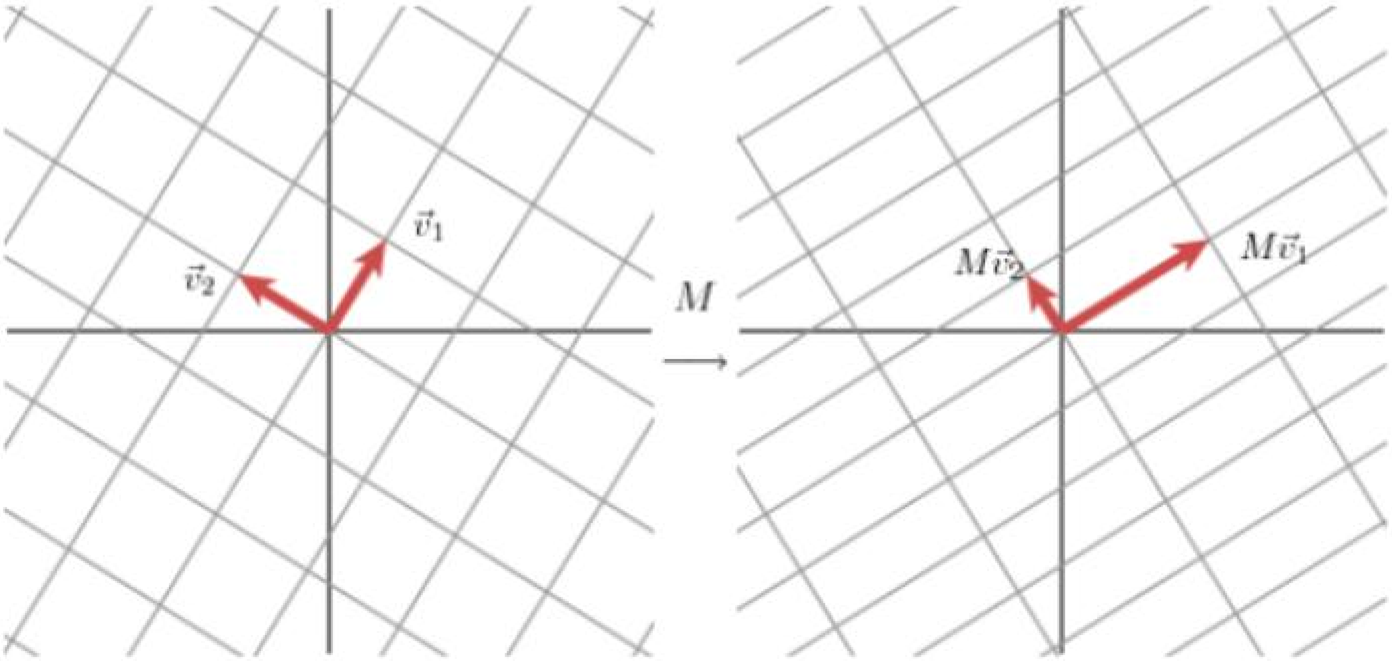
- 取前K个来看看吧！



SVD矩阵分解

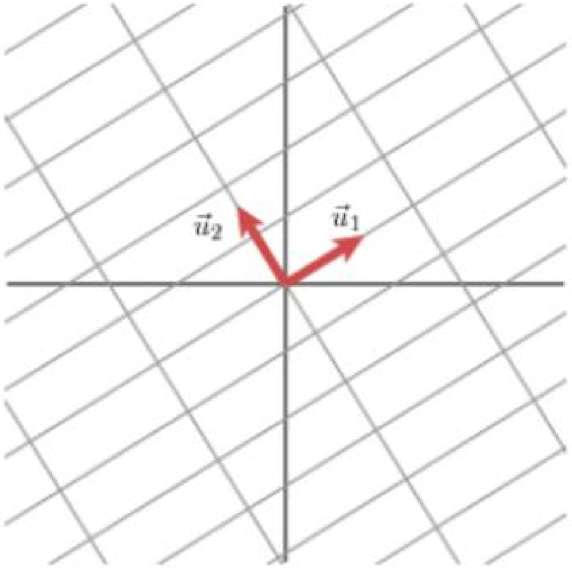
SVD推导

前提：对于一个二维矩阵M可以找到一组标准正交基v1和v2使得Mv1和Mv2是正交的。



SVD矩阵分解

SVD推导



- 使用另一组正交基u1和u2来表示Mv1和Mv2的方向
- 其长度分别为：$\|MV_1\| = \sigma_1$ ，$\|MV_2\| = \sigma_2$ 可得：$$MV_1 = \sigma_1 u_1$$ $$MV_2 = \sigma_2 u_2$$
- 对于向量X在这组基中的表示：$\mathbf{x} = (\mathbf{v}_1 \cdot \mathbf{x})\mathbf{v}_1 + (\mathbf{v}_2 \cdot \mathbf{x})\mathbf{v}_2$
  （点积表示投影的长度，可转换成行向量乘列向量 $\mathbf{v} \cdot \mathbf{x} = \mathbf{v}^T \mathbf{x}$ ）
- 可得：$$M\mathbf{x} = (\mathbf{v}_1 \cdot \mathbf{x})M\mathbf{v}_1 + (\mathbf{v}_2 \cdot \mathbf{x})M\mathbf{v}_2$$ $$M\mathbf{x} = (\mathbf{v}_1 \cdot \mathbf{x})\sigma_1 \mathbf{u}_1 + (\mathbf{v}_2 \cdot \mathbf{x})\sigma_2 \mathbf{u}_2$$ 从而：$$M\mathbf{x} = \mathbf{u}_1 \sigma_1 \mathbf{v}_1^T \mathbf{x} + \mathbf{u}_2 \sigma_2 \mathbf{v}_2^T \mathbf{x}$$ $$M = \mathbf{u}_1 \sigma_1 \mathbf{v}_1^T + \mathbf{u}_2 \sigma_2 \mathbf{v}_2^T$$ 化简得：$M = U \Sigma V^T$