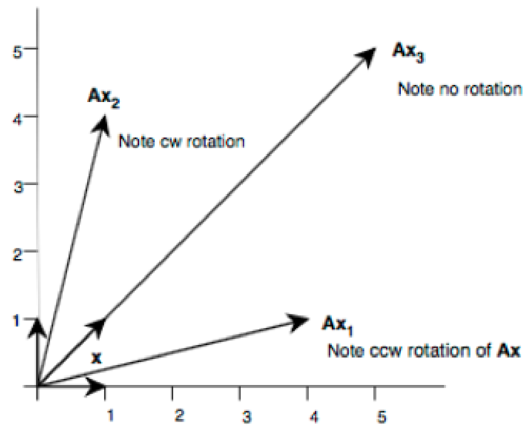
三、矩阵分析与应用

1、特征值、特征向量的意义（专门指的是方阵）

方阵的特征值(Eigenvalues)与特征向量(Eigenvectors)

$\mathbf{Ax}=\lambda\mathbf{x}$几何意义，并思考如何计算$\mathbf{A}^{1000}$

给定一个矩阵$\mathbf{A}=\begin{bmatrix}4 & 1\\ 1 & 4\end{bmatrix}$，对于$\mathbf{x}_1=\begin{bmatrix}1\\0\end{bmatrix}$,则有$\mathbf{Ax}_1=\begin{bmatrix}4\\1\end{bmatrix}$；对于$\mathbf{x}_3=\begin{bmatrix}1\\1\end{bmatrix}$，则有$\mathbf{Ax}_3=5\begin{bmatrix}1\\1\end{bmatrix}$



特征值的意义是对向量乘以A，共线

对称矩阵的特征值与特征向量

- 如果一个对称矩阵的特征值不同，则其相应的所有的特征向量正交($\mathbf{U}\mathbf{U}^T=\mathbf{U}^T\mathbf{U}=\mathbf{I}$)

$$\mathbf{A}=\mathbf{U}\Lambda\mathbf{U}^T \tag{7}$$

$$=[\mathbf{u}_1,\ldots,\mathbf{u}_n]\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}\begin{bmatrix}\mathbf{u}_1^T\\ \vdots\\ \mathbf{u}_n^T\end{bmatrix} \tag{8}$$

$$=\sum_{i=1}^{n}\lambda_i\mathbf{u}_i\mathbf{u}_i^T \tag{9}$$

2、PCA解释

PCA目的在于降维，降维的意义就是乘以一个矩阵之后，新的向量协方差矩阵，方差最大（包含的信息最多），协方差为0（正交）

- 问题：假设变换矩阵为$\mathbf{Y} = \mathbf{QX}$，并先假设$\mathbf{Q}$是方阵(先不降维)，则有

$$\mathbf{C}_Y = \frac{1}{n}\mathbf{Y}\mathbf{Y}^T = \mathbf{Q}\mathbf{C}_X\mathbf{Q}^T$$

如何使得$\mathbf{C}_Y$是一个对角矩阵？回忆$\mathbf{C}_X = \mathbf{U}\Lambda\mathbf{U}^T \Rightarrow \Lambda = \mathbf{U}^T\mathbf{C}_X\mathbf{U}$. 如果$\mathbf{Q} = \mathbf{U}^T$？

因此选取特征值较大的特征向量出来计算PCA

3、SVD

特征值分解针对方阵或者对称矩阵，对于一般矩阵来说，就是SVD

任何秩为r的矩阵$\mathbf{A} \in \mathbb{R}^{m\times n}$，可以被分解为

$$\begin{aligned}
\mathbf{A} &= \underbrace{[\mathbf{U}_1\mathbf{U}_2]}_{\mathbf{U}}\underbrace{\begin{bmatrix} \Sigma_1 & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{bmatrix}}_{\mathbf{\Sigma}}\underbrace{\begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix}}_{\mathbf{V}^T} && (12) \\
&= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T && (13) \\
&= \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T && (14) \\
&= \sum_{i=1}^{r}\sigma_i\mathbf{u}_i\mathbf{v}_i^T && (15)
\end{aligned}$$

其中$\mathbf{U} \in \mathbb{R}^{m\times m}$和$\mathbf{V} \in \mathbb{R}^{n\times n}$是正交矩阵。$\sum_1 = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$且有$\sigma_r > 0$。

(13)便于分析，但并不计算有效；(14)计算有效，但有时不方便分析；(15)方便展开，用于低秩矩阵计算。