# Lecture 3 Singular Value Decomposition(SVD)（奇异值分解）

## matrix decomposition

- LU,QR,Cholsky...
- SVD is good for "big data"
- extract most important features

> 考虑n*m的矩阵$X=(...;x_1,x_2,...x_m;...)$，每一列都是一个高维向量，可以理解为有m组”样本“。
> 若$n\gg m$，理解为测量量数远大于样本数；
> 或者$m\gg n$，理解为大规模调研（survey），股票预测（stack market）等；
> 这些矩阵的形状是“细长”型的，因此需要新的方法来作为分解工具。

## Start With an Example

### motivation

在前面我们讨论过，$Ax$实际上是对向量作两种变换：

- 旋转变换（rotation）：$A=(cos\theta,-sin\theta;sin\theta,cos\theta)$
- 伸缩变换（stretching）：$A=(\alpha,0;0,\alpha)$
  在二维情况下，假设两个标准正交基，作变换以后仍正交，此时由圆变为（椭）圆。
  扩展到n维情况是：有n个标准正交基，经过变换之后仍正交，变为超椭圆（hyoer-ellpse）。

## 推导过程

这样，我们就能写出变换式：$Av=\sigma u$，其中**u,v可以不同**
写成向量形式：
即$AV=\hat{U}\hat{\Sigma}$

> 其中$V$为正交矩阵（unitary transformation），具有性质：$V^TV=VV^T=I$
> $\hat{\Sigma}$为对角阵，负责对向量作伸缩变换，且$\sigma_1\geq\sigma_2\geq...\geq\sigma_n\geq 0$

> $\hat{U}$是结果向量构成的矩阵，向量之间正交，具有性质：$\hat{U}^T\hat{U}=I_{n*n}$
> 因此，当$m>n$时，得到$\hat{U}$的公式称为Reduced SVD。

由于我们对$\hat{U}$的形式不太满意，于是将$V$放到右边，有
$A=\hat{U}\hat{\Sigma}V^T$。
运用施密特正交化原理得到完整正交基，将$\hat{U}$变为m*m的方阵，对$\hat{\Sigma}$下方全部补0，得正式的SVD
$A=U\Sigma V^T$

- **注意当$m<n$也成立**

# SVD

## Definition

Every matrix $A\in C^{m*n}$ has a singular value decomposition.

> 对于矩阵没有要求，适用范围更广。

- singular values $\sigma_j$ are uniquely determined;
- if A is square and the $\sigma_j$ distinct,the singular vectors $u_j$ and $v_j$ are uniquely determined up to complex signs.

> 在上面的叙述中认为$u,v$扩大多少倍都认为是同一向量。该叙述确定了唯一条件。

- $\Sigma$：$\sigma_1\geq\sigma_2\geq...\geq\sigma_{maxm,n}\geq 0$
- $U^TU=UU^T=I_{m*m},V^TV=VV^T=I_{n*n}$

> 在matlab中，可以用svd函数实现奇异值分解：
>
> ```
> svd(A);
> [U,S,V]=svd(A);
> ```

# SVD与eigen-decomposition的关系

$A^TA=V{\Sigma_v}^2V^T,AA^T=U{\Sigma_u}^2U^T$，其中$\Sigma_v=\Sigma^T\Sigma,\Sigma_u=\Sigma\Sigma^T$

等号右边的形式就是对角化分解。这就是说，我们可以通过对$AA^T$或$A^TA$作对角分解得出奇异值分解的结果。其中中间的对角矩阵的元素就是奇异值分解中对角矩阵元素的平方。

# 性质

- 如果A的秩是r，那么有且只有r个非零的singular values

> proof:
> $rank(A) = rank(U\Sigma V^T) = rank(\Sigma) = r$

## Optional

- $|A| = \Pi_{j=1}^{min m,n} \sigma_j$

> proof:
> 行列式展开成奇异值分解形式即可。