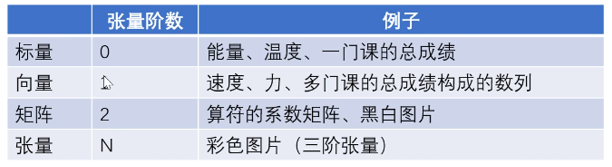
**一、张量及Python基础**

**1.1 什么是张量**

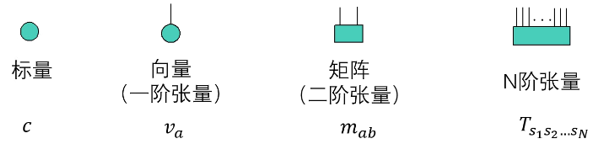


* N阶张量又称**N维数列**
* 张量阶数即为张量指标的个数
* 向量、矩阵也称为1阶张量、2阶张量
* 每个指标可取的值的个数，被称为指标的维数
* 张量的“直观”定义：多个指标标记下的一堆数

比如张量，其指标的个数为，称其为阶张量，若指标可取的值的个数为，称指标的维数为。

**1.2 张量的图形表示**

张量用连接着个腿的圆圈或方块表示，为张量的阶数，如下图所示。



每个腿对应指标，比如一阶张量的腿对应指标，二阶张量的腿对应指标和。

**1.3 张量的基本运算**

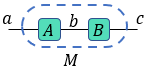
1) 向量内积：



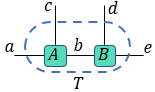
2) 向量乘矩阵：



3) 矩阵乘矩阵：



4) 张量缩并：



**二、线性代数基础**

**2.1 本征值分解与最大本征值问题**

1) 本征值分解

给定一个的矩阵，其本征值分解定义为：



其中，称为变换矩阵，满足正交归一性，其每一列均为矩阵M的本征向量。为对角矩阵，称为**本征谱**。

满足的矩阵称为厄米矩阵，其一定存在本征值分解，且本征谱为实数。

2) 最大本征值问题：假设矩阵的本征值为实数，求解给定矩阵的最大本征值及其本征态。

最大本征值问题对应于如下优化问题：给定矩阵，求解归一化向量，使函数的值极大化。

该最优化问题的解为的(绝对值)的最大本征态，对应值为最大本征值。

3) 最大本征值问题的幂级数求解法：

考虑实对称矩阵，设和为其绝对值最大的唯一本征值和本征向量，则有，为的第0列。

证明：设的本征值分解为，有：



其中，，可推得：



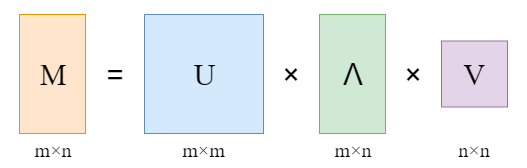
**2.2 奇异值分解(SVD)**

给定矩阵，有：



其中，与的每一列分别被称为的左、右奇异向量，且满足正交归一性。为**非负定实对角矩阵**，称为奇异谱，其元素称为**奇异值**，且**从大到小排列**。

* 任何矩阵都存在奇异值分解，且极大化第一个元素实部时，分解唯一。
* **矩阵的秩定义为矩阵的非零奇异值的个数**。



**2.3 矩阵的低秩近似问题**

给定的矩阵，设其秩为，求解秩为的矩阵，有，且极小化二矩阵间的范数

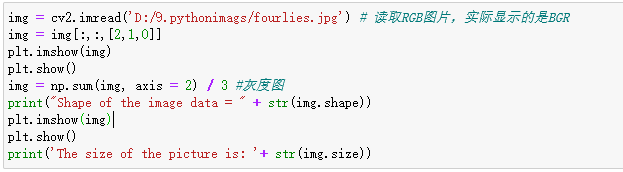


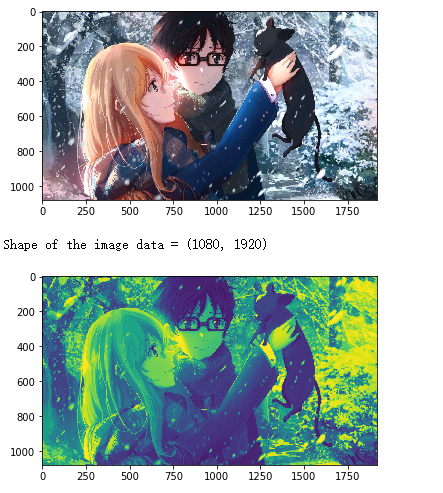
低秩近似问题的最优解为：



**2.4 基于奇异值分解低秩近似的图形压缩**

1) 将图片导入并做灰度处理。



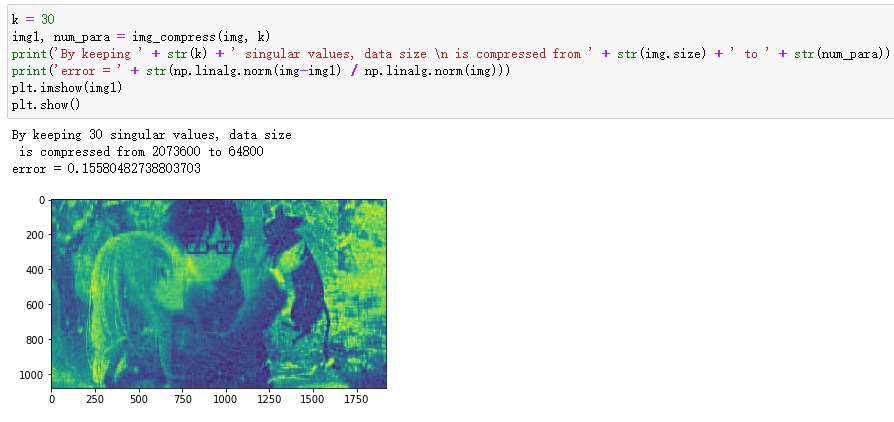


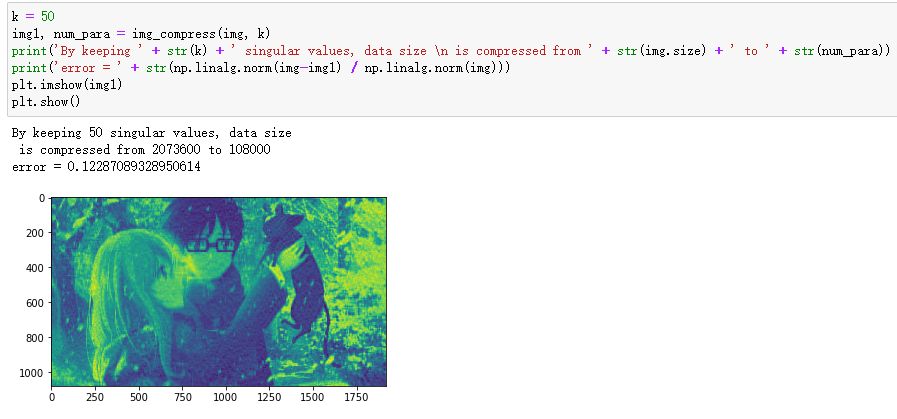
2) 将图片进行奇异值分解后，取前10个最大的奇异值对图片进行重组。

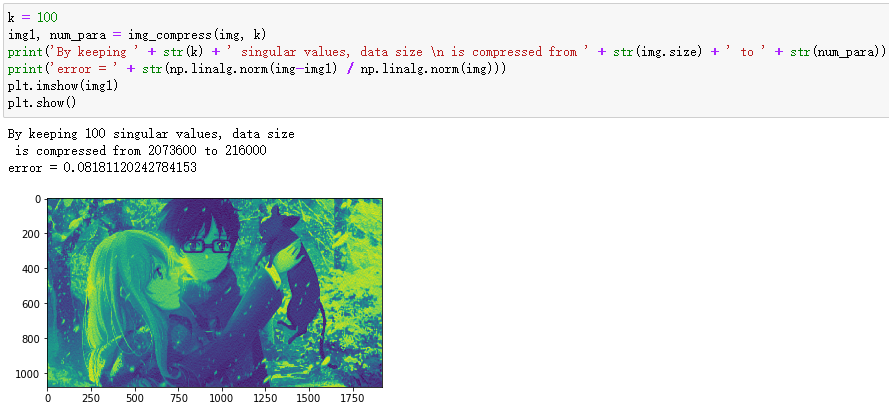


这时候可以看到，图片很模糊，此时的图片与原图片之间的误差约为，而大小仅为原图片的左右。

3) 取多个要保留的奇异值进行上述操作，这里分别取，和。





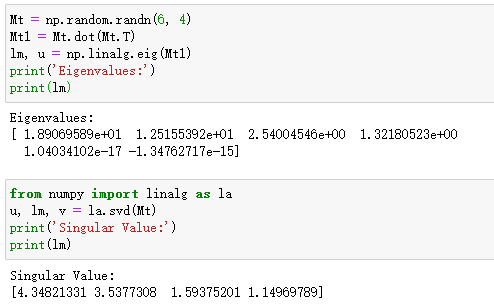


不难看出，随着保留的奇异值个数的增加，图片越来越清晰，当要保留的奇异值为时，可以看出，此时的图片与原图片之间的误差小于，而大小仅为原图片的左右。

SVD的优点：简化数据，去除噪声点，提高算法的结果。适用的数据类型为数值型。

通过SVD对数据的处理，可以**使用小得多的数据集来表示原始数据集**，实际上去除了噪声和冗余信息，以此达到优化数据、提高算法结果的目的。比如上面对图片的处理：图片实际上是数字矩阵，通过SVD将该矩阵降维，只使用其中的重要特征来表示该图片，达到图片压缩的目的。

奇异值是矩阵的特征值的平方根，验证如下：



从上面的结果中可以看到，矩阵的奇异值(4个)是矩阵的特征值(前4个)的平方根。