Assignment: Dimensionality Reduction

Qianlang Chen (u1172983)

CS 5140 Spring 2021

```
import numpy

A = numpy.loadtxt('./data/A.csv', delimiter=',')
print(A.shape)
```

```
(3500, 20)
```

Problem 1

Part A

```
import numpy
from numpy import linalg

U, S, Vt = linalg.svd(A)
S = numpy.identity(len(S)) * S

def Ak(k):
    Uk = U[:,:k]
    Sk = S[:k,:k]
    Vtk = Vt[:k,:]
    return Uk @ Sk @ Vtk
```

```
K = tuple(range(1, 11))
for k in K:
    print(f'k = {k}: {linalg.norm(A - Ak(k), 2)}')
```

```
k = 1: 100.00246446709691
k = 2: 92.11139537014573
k = 3: 87.43842799887248
k = 4: 70.84270143876726
k = 5: 58.89263665883072
k = 6: 57.77458152758134
k = 7: 25.86135552441823
k = 8: 24.871609868369216
k = 9: 24.395000417106157
k = 10: 22.997273663616927
```

Part B

```
norm_A = linalg.norm(A, 2)

K = tuple(range(1, 21))
for k in K:
    if linalg.norm(A - Ak(k), 2) < .2 * norm_A:
        print(k)
        break
```

```
10
```

Part C

```
```