1. MATH347 HOMEWORK 2

Topic:	Math@UNC environment
Post date:	May 23, 2023
Due date:	May 24, 2023

1.1. Background

This homework investigates the matrix fundamental spaces, and the concept of linear dependence and independence.

1.2. Theoretical questions (1 course point each)

1. Is the zero vector a linear combination of any non-empty set of vectors?
2. If $S \subseteq V$, with $(V, \mathbb{R}, +, \cdot)$ a vector space then span(S) equals the intersection of all subspaces of V that contain S.
3. Can a vector space have more than one basis? Give an example.
4. Prove that the dimension of the vector set $\mathscr{Z} = (\{\mathbf{0}\}, \mathbb{R}, +, \cdot)$ is equal to zero.
5. Describe the geometry of $C(\boldsymbol{P})$, $N(\boldsymbol{P}^T)$ the column and left null space of the projection matrix $\boldsymbol{P} = \boldsymbol{q}\boldsymbol{q}^T$, with $\boldsymbol{q} \in \mathbb{R}^m$, $\|\boldsymbol{q}\| = 1$.
6. Describe the geometry of $C(\boldsymbol{P})$, $N(\boldsymbol{P}^T)$ the column and left null space of the projection matrix $\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{Q}^T$, with $\boldsymbol{Q} \in \mathbb{R}^{m\times n}$, $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}_n$.

1.3. Fundamental spaces and linear dependence in image processing

1.3.1. Data input

Images are often processed using techniques from linear algebra. In this assignment a database with $p = 1000$ facial images will be used to investigate the fundamental vector subspaces associated with a matrix (linear mapping) and concepts of linear dependence. Each facial image has been centered, downsampled to $p_x \times p_y$ pixels, $p_x = p_y = 128 = 2^7$, scaled to account for different face size, and background illumination has been equalized to enable comparison. A single image is represented as a vector of size $m = p_x p_y = 16384 = 2^{14}$.

The question that arises is whether a facial image with $m = 2^{14}$ components can be represented more economically as the scaling coefficients in a linear combinations of only $n = 99$ vectors. An average of all p images has been computed and is stored in vector $\boldsymbol{a}$. Subsequently, a matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ has been constructed with orthonormal columns from which a face image can be obtained as

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{a}, \tag{1}$$

with $\boldsymbol{x}$ a vector of scaling coefficients. The columns of $\boldsymbol{A}$ represent deviations of facial features with respect to the average face $\boldsymbol{a}$. Since there are only $n = 99$ column vectors in $\boldsymbol{A}$ it is not expected to exactly capture any particular face $\boldsymbol{f}$. There will be some error that can be computed through a norm

$$\varepsilon = \|\boldsymbol{f} - \boldsymbol{b}\|.$$

The question investigated here is the efficiency and accuracy of face representation through (1).

```
∴ using MAT
```

```
∴ DataFileName = homedir()*"/courses/MATH347DS/data/faces/faces.mat";
```

```
∴ DataFile = matopen(DataFileName,"r");
```

```
∴ A=read(DataFile,"A"); a=read(DataFile,"a");
```

```
∴ m, n = size(A); px=Int(sqrt(m)); py=Int(m/px); [m n px py]
```

$$[\ 16384\ \ 99\ \ 128\ \ 128\ ] \tag{2}$$

```
∴
```

After loading the database, a directory is created for this assignment and set as the current directory

```
∴ mkpath(homedir()*"/courses/MATH347DS/homework/hw02");
```

```
∴ cd(homedir()*"/courses/MATH347DS/homework/hw02");
```

```
∴
```

1.3.2. Utility functions

Define functions to display a facial image given as a vector, save an image to a file, and load a face from the collection of p facial images.

```
∴ function shwface(a,px,py)
    im=reshape(a,px,py)'
    clf(); imshow(im,cmap="gray")
  end;
```

```
∴ shwface(a,px,py);
```

```
∴ function savface(a,px,py,fname)
    im=reshape(a,px,py)'
    imshow(im,cmap="gray")
    savefig(fname*".png")
  end;
```

```
∴ savface(a,px,py,"averageface");
```

```
∴
```

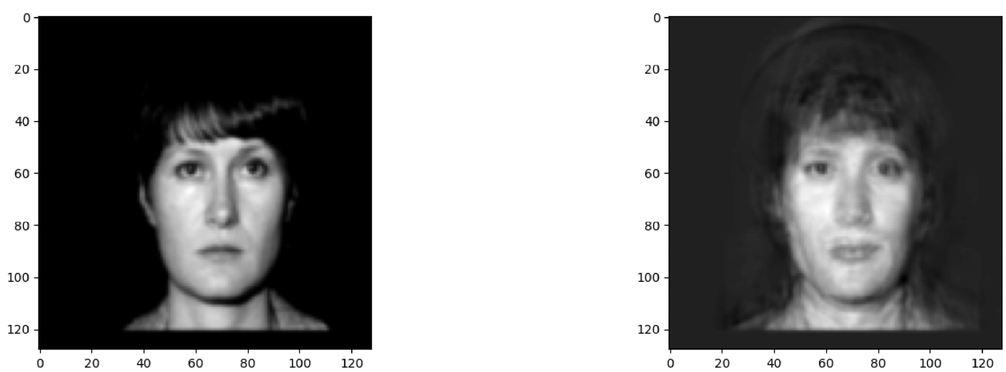

Figure 1. Sample reconstructed faces

```
∴ fFile=open(homedir()*"/courses/MATH347DS/data/faces/testfaces/3002");
```

```
∴ f=read(fFile);
```

```
∴ size(f)
```

$$[\ 16384\ ] \tag{3}$$

```
∴ shwface(f,px,py);
```

```
∴ savface(f,px,py,"3002");
```

```
∴ cf=f-a;
```

```
∴ xf=A\cf;
```

```
∴ figure(2); shwface(A*xf+a,px,py);
```

```
∴ savface(A*xf+a,px,py,"x3002");
```

```
∴ norm(A'*A-I)
```

$3.710352739902468\, e-5$

```
∴
```

1.3.3. Questions to investigate (2 course points each)

The calculations within the questions require no more that 5 matrix operations, sometimes only 1 or 2. Carefully think about the appropriate matrix operations. Specify those operations as mathematical equations before writing and executing the appropriate Julia code. Example:

The norm of $\boldsymbol{a}$ is computed as $\|\boldsymbol{a}\| = (\boldsymbol{a}^T \boldsymbol{a})$.

```
∴ nrma = sqrt(a'*a)
```

$$[\ 7063.574438491783\ ] \tag{4}$$

```
∴
```

The above approach is one of the principal benefits of the TeXmacs/Julia notebook format: the mathematical formulas are presented and considered first, with implementation done afterwards. Additional Julia instructions are required for visualization or for data input and output.

1. Verify that you are working with a well-chosen data set. Is $\boldsymbol{A}$ orthonormal?
2. Consider some subset of the available test faces. For each face $\boldsymbol{f}$, the FTLA states that

   $$\boldsymbol{f} = \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{u} \in C(\boldsymbol{A}), \boldsymbol{v} \in N(\boldsymbol{A}^T).$$

   Compute $\|\boldsymbol{v}\|/\|\boldsymbol{u}\|$ for the chosen faces through concise matrix operations. Discuss the significance of these ratios.
3. For the subset chosen above, find the angles between $\boldsymbol{u}$ and $\boldsymbol{v}$. Discuss their significance.
4. Assess the quality of the $\boldsymbol{A}$ data set, by determining the maximum error encountered in representing deviations of face with respect to the mean face as a linear combination of columns of $\boldsymbol{A}$. Do so for all faces within your chosen data set.
5. Assess the quality of the specified mean face $\boldsymbol{a}$ by constructing a different mean face $\boldsymbol{c}$ for your chosen subset. Determine how close $\boldsymbol{c}$ is to $\boldsymbol{a}$.
6. Determine the face from your chosen subset best represented by $\boldsymbol{A}\boldsymbol{x} + \boldsymbol{a}$, and display the original and reconstructed faces. Do the same for the face from your chosen subset worst represented by $\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$.