1. MATH547 HOMEWORK 3

Topic: Math@UNC environment
Post date: May 20, 2020
Due date: May 21, 2020

1.1. Background

This homework investigates consequences of the fundamental theorem of algebra and application of the singular value decomposition.

1.2. Theoretical questions

Consider a linear mapping $f: U \to V$, from vector space $\mathcal{U} = (U, \mathbb{R}, +, \cdot)$ with basis $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_n\}$, to $\mathcal{V} = (V, \mathbb{R}, +, \cdot)$, with basis $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$.

1. Is $\{f(\boldsymbol{u}_1), \dots, f(\boldsymbol{u}_n)\}$ a basis for $\mathcal{V}$?
   SOLUTION. Not necessarily, for example if $f(\boldsymbol{u}_j) = \boldsymbol{0}$.

2. If $\mathrm{nullity}(f) = 0$ must $m = n$?
   SOLUTION. No. Let $A \in \mathbb{R}^{m \times n}$ be matrix associated with f. If $\mathrm{nullity}(f) = 0$ then $N(A) = \{\boldsymbol{0}\}$, and the only solution to $f(\boldsymbol{x}) = \boldsymbol{0}$ is $\boldsymbol{x} = \boldsymbol{0}$. Consider $f: \mathbb{R} \to \mathbb{R}^2$, $f(x) = [\; x \;\; 2x \;]^T = A x$, with $A = [\; 1 \;\; 2 \;]^T$, and $m = 1$, $n = 2$, $m \neq n$.

3. If $m = n$ and $\boldsymbol{u}_i = \boldsymbol{v}_i$ for $i = 1, \dots, m$, what is the matrix A representing f?
   SOLUTION. The identity matrix $A = I$. Apply the linear mapping $\boldsymbol{v} = f(\boldsymbol{u}) = A\boldsymbol{u}$ to the basis vectors $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_m\}$, to obtain

$$V = [\; \boldsymbol{v}_1 \;\; \boldsymbol{v}_2 \;\; \dots \;\; \boldsymbol{v}_m \;] = AU = [\; A\boldsymbol{u}_1 \;\; A\boldsymbol{u}_2 \;\; \dots \;\; A\boldsymbol{u}_m \;].$$

Since U is a basis, the column vectors are linearly independent and the matrix is invertible,

$$A = VU^{-1}.$$

The statement $\boldsymbol{u}_i = \boldsymbol{v}_i$ can be interpreted in one of two ways (these descriptions are said to be duals of one another, and distinguish between a vector as a geometrical entity and its coordinates in some given basis):

   i. $\boldsymbol{u}_i$ is a vector that contains the coordinates of the ith basis vector of U in the I basis, $\boldsymbol{u}_i = I\boldsymbol{u}_i$. Similarly, $\boldsymbol{v}_i$ is a vector that contains the coordinates of the ith basis vector of V in the I basis, $\boldsymbol{v}_i = I\boldsymbol{v}_i$. In this case $\boldsymbol{u}_i = \boldsymbol{v}_i$ implies $U = V$, and $A = I$. This interpretation was emphasized in this course, and is appropriate for $U = V = \mathbb{R}^m$.

   ii. $\boldsymbol{u}_i$ is the ith basis vector of U, whose coordinates in that basis are $\boldsymbol{b}_i$ given by $\boldsymbol{u}_i = U\boldsymbol{b}_i$. Similarly $\boldsymbol{v}_i = V\boldsymbol{c}_i$. The statement $\boldsymbol{u}_i = \boldsymbol{v}_i$ now implies $U\boldsymbol{b}_i = V\boldsymbol{c}_i$, such that

$$\boldsymbol{c}_i = V^{-1}U\boldsymbol{b}_i$$

and the matrix associated with the mapping would be $A = V^{-1}U$. This interpretation is appropriate when mapping between different vector spaces of the same dimension, $U \neq V$, $\dim(U) = \dim(V)$.

4. Determine the singular value decomposition and pseudo-inverse of a matrix $A \in \mathbb{R}^{1 \times n}$ (i.e., a row vector).
   SOLUTION. The SVD is $A = U\Sigma V^T$, $U \in \mathbb{R}^{m \times m}$ orthogonal, $V \in \mathbb{R}^{n \times n}$ orthogonal, $\Sigma \in \mathbb{R}_+^{m \times n}$ diagonal. For A a row vector, $m = 1$, and

$$\Sigma = [\; \sigma_1 \;\; 0 \;\; \dots \;\; 0 \;]$$

with $\sigma_1 = \|A\|_2$. By definition of the matrix 2-norm

$$\sigma_1 = \sup_{\|\boldsymbol{x}\| = 1} \|A\boldsymbol{x}\|.$$

Since $A = a^T$,

$$A x = a^T x = \sum_{i=1}^{n} a_i x_i, \quad \sum_{i=1}^{n} x_i^2 = 1.$$

The largest possible value of $\|A x\|$ when $\|x\| = 1$, is $\sigma_1 = \|A\| = \max_{1 \leqslant i \leqslant n} |a_i|$. The SVD is

$$A = [1][\; \sigma_1 \;\; 0 \;\; \ldots \;\; 0 \;] \begin{bmatrix} \dfrac{a^T}{\sigma_1} \\ V_{n-1}^T \end{bmatrix}$$

with $V_{n-1}^T a = 0$, $V_{n-1}^T V_{n-1} = I_{n-1}$ (orthogonal).

1.3. Ordered bases for the fundamental spaces and painting motifs

The fudamental theorem of linear algebra partitions the domain and codomain of a linear mapping. The singular value decomposition provides orthogonal bases for each of the subspaces arising in the partition. The bases are ordered according to the amplification behavior of the linear mapping, expressed through the norm of successive restrictions of the mapping. This approach is closely aligned with typical problems in data science, and can be used in a variety of scenarios. In this homework linear algebra methods will first be used in a field far removed from the physical sciences: extracting the quirks of painter style from the overall composition of a painting, and applying one artist's style to another artist's composition. This is often-encountered data science problem: distinguishing between small and large scale features of data.

First steps in solving the homework questions will be carried out in class. Each of the following subsections is a homework question, with 1 grade point awarded for a correct solution. Here are some initial Octave instructions to define a directory to save images into and define a function to read an image.

```
octave] cd homework; mkdir hw03; cd hw03
```

```
octave] im=imread("/home/student/courses/MATH547ML/data/paintings/
        Andy_Warhol_127.jpg");
```

```
octave] imshow(im)
```

```
octave] function im=pread(name)
           im=imread(strcat("/home/student/courses/MATH547ML/data/paintings/",name));
           im=rgb2gray(im);
        end
```

```
octave] im=pread("Andy_Warhol_127.jpg");
```

```
octave] imshow(im); size(im)
```

```
ans =

   360   357
```

```
octave]
```

1.3.1. Data input mappings

Define a linear mapping that rescales data within an image file to some specified size $p_x \times p_y$. Determine whether the mapping is data-preserving, and if not, quantify the amount of data loss.

SOLUTION. One such linear mapping is to simply take a portion of the image. If the initial image was $q_x \times q_y$ the data loss is $q_x \times q_y - p_x \times p_y$.

```
octave] function im1=psize(im0,px,py)
          [mx,my]=size(im0);
          cx=floor(mx/2); cy=floor(my/2);
          px2=floor(px/2); py2=floor(py/2);
          im1 = im0(cx-px2:cx+px2,cy-py2:cy+py2);
        end
octave] imshow(psize(im,256,256))
octave] imshow(psize(pread("Vincent_van_Gogh_167.jpg"),256,256))
octave]
```

1.3.2. Images as data: change of basis

Take the largest possible portion of a painting of size $p \times p$ with $p = 2^q$. Interpret the resulting image as a vector $\boldsymbol{b} \in \mathbb{R}^m$ with $m = p^2 = 2^{2q}$. The image is thus specified as a linear combination of the columns of the identity matrix $\boldsymbol{I} \in \mathbb{R}^{m \times m}$,

$$\boldsymbol{b} = \boldsymbol{I}\boldsymbol{b} = b_1\boldsymbol{e}_1 + b_2\boldsymbol{e}_2 + \cdots + b_m\boldsymbol{e}_m,$$

and describes illumination on a pixel-by-pixel basis. A column vector of $\boldsymbol{I}$ can be interpreted as the binary base representation of a natural number from the set $P = \{1, 2, 3, \ldots, 2^m\}$, namely $2^{j-1} \overset{g}{\to} \boldsymbol{e}_j$.

$$\boldsymbol{e}_j^T = [\, 0 \;\; \ldots \;\; 0 \;\; 1 \;\; 0 \;\; \ldots \;\; 0 \,] \in \mathbb{R}^m$$

$$n_j = 00\ldots010\ldots0_2 = 2^{j-1}$$

Note that is one choice among the possible combinations of m objects chosen from a set with 2^m elements, $C(2^m, m)$. Even for small m, the number of choices is enormous

$$C(2^m, m) = \frac{2^m!}{m!(2^m - m)!}, m = 16 \Rightarrow C(2^{16}, 16) = \frac{2^{16}!}{16!\,(2^{16} - 16)!} \cong 5.5 \times 10^{63}.$$

Construct a new basis by random choice of numbers from P and the g mapping. Check if the basis is orthogonal.

SOLUTION. Let M be the largest representable integer on a computer. If $M < m = p^2$ then the binary number representation would not construct a basis since the vector $\boldsymbol{u}$ associated with binary number $B_{\boldsymbol{u}} > M$ is not within $\text{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ associated with binary numbers $\{B_1, \ldots, B_m\}$, all less than M, $B_i \leqslant M$. Choose $m < M$. Orthogonality of vectors associated with binary numbers B_i, B_j corresponds to a zero result from the bitwise and operation.

```
octave] m=4*4;
octave] B=randi(flintmax()-1,m,1);
octave] C=zeros(m,m);
        for i=1:m
          for j=1:m
            C(i,j)=bitand(B(i),B(j));
          end;
        end
octave] max(max(C))
ans =     8.9756e+15
octave]
```

Linear dependence is the same as equality in this binary representation.

```
octave] D=zeros(m,m);
        for i=1:m
          for j=i+1:m
            D(i,j) = B(i)==B(j);
          end;
        end
octave] max(max(D))
```

ans = 0

```
octave]
```

From above, a non-orthogonal basis is obtained.

1.3.3. Images as data: positive checkerboard basis

Construct another basis that corresponds to successive halving of image regions by factors $(2^k, 2^l)$ along the horizontal and vertical dimensions. Are basis vectors orthogonal?

SOLUTION. *There are many possible approaches to this problem. In the following a full solution is presented highlighting that application of linear algebra concepts to practical data science problems usually involves non-trivial preprocessing of the data. Since the solution does involve considerable technique from outside of linear algebra full credit is awarded to all attempts of this problem.*

First consider the simplest possible case. A 1×1 image cannot be halved, hence start with 2×2 images for which

$$
\begin{matrix} 1\ 0 \\ 1\ 0 \end{matrix}, \begin{matrix} 0\ 1 \\ 0\ 1 \end{matrix}, \begin{matrix} 1\ 1 \\ 0\ 0 \end{matrix}, \begin{matrix} 0\ 0 \\ 1\ 1 \end{matrix}, \begin{matrix} 1\ 0 \\ 0\ 1 \end{matrix}, \begin{matrix} 0\ 1 \\ 1\ 0 \end{matrix}
$$

give 6 possible patterns encoded by row vectors

$$
\boldsymbol{B}^T = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.
$$

Only four basis vectors are required and can be chosen as

$$
\boldsymbol{C} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.
$$

```
octave] B=[1 0 1 0; 0 1 0 1; 1 1 0 0; 0 0 1 1; 1 0 0 1; 0 1 1 0];
        C=[1 0 1 0; 1 1 0 0; 0 1 1 0; 0 0 1 1];
octave] [rank(B) rank(C)]
```

ans =

   4   4

```
octave]
```

The pattern suggested by above is alternating groups of n 1's and n 0's, that are then shifted right by $1, \ldots, n$ places. Define a function to generate alternating 1,0 bit groupings for images of size $m = p \times p$, $p = 2^q$. There are 2^{2q-1} such numbers.

```
octave] q=2; p=2^q; m=p^2; disp([q p m])
```

   2   4   16

```
octave] g=zeros(2^(2*q-1),1); o=2^m-1;
         for i=0:q
           k=2^i; l=2^k; disp([k l]);
           g(i+1) = bitshift(l-1,2^i);
           dec2bin(g(i+1))
         end
```

```
   1   2
ans = 10
   2   4
ans = 1100
   4  16
ans = 11110000
```

```
octave]
```

1.3.4. Images as data: mixed-sign checkerboard basis

Map the binary digits $\{0,1\}$ from the positive checkerboard basis to integers $\{-1,1\}$. Check if the newly obtained basis is orthogonal. Display approximations of the image that result from the first 2^l basis vectors, $l = 2q, 2q-2, 2q-4, 2q-6, 2q-8$.

SOLUTION. As above, s.ince the solution does involve considerable technique from outside of linear algebra full credit is awarded to all attempts of this problem.

1.3.5. Images as mappings

Alternatively, an image can be interpreted as a matrix A, hence a mapping. From

$$A = AI = A\,[\ e_1\ \ e_2\ \ \ldots\ \ e_m\ ] = [\,A e_1\ \ A e_2\ \ \ldots\ \ A e_m\,],$$

the image can be interpreted as the transformation of the image encoded by I. Denote by A_k the kth-rank approximation of A from the singular value decomposition $A = U\Sigma V^T$

$$A_k = \sum_{l=1}^{k} \sigma_l u_l v_l^T.$$

Display the images that correspond to $k = m, m/2, m/4, m/8$.

SOLUTION. Construct and represent each A_k.

```
octave] function im=svdim(S,U,V,p)
           im=S(1,1)*U(:,1)*V(:,1)';
           for k=2:p
             im=im+S(k,k)*U(:,k)*V(:,k)';
           end;
         end
```

```
octave] figure(3); clf; imshow(im)
```

```
octave] [U,S,V]=svd(im,1);
```

```
octave] figure(2); plot(log10(diag(S)),'o');
```

```
octave] figure(1); clf; imagesc(svdim(S,U,V,10)); colormap(gray); print -deps
         HW03Fig01.eps
```

```
octave] figure(1); clf; imagesc(svdim(S,U,V,20)); colormap(gray); print -deps
         HW03Fig02.eps
```

```
octave] figure(1); clf; imagesc(svdim(S,U,V,40)); colormap(gray); print -deps
         HW03Fig03.eps
```

```
octave]
```

Figure 1. Successive SVD approximations of an image with $k = 10, 20, 40$ rank-one updates.

1.3.6. Extracting and applying motifs

Consider images from two different artists, A, B and their singular value decompositions

$$A = S \Lambda T^T, B = U \Sigma V^T.$$

Let $q = \operatorname{rank}(A)$, $r = \operatorname{rank}(B)$. Construct and display images that take the large scale features from A combined with small scale features from B,

$$C = \sum_{l=1}^{\min(k,q)} \lambda_l s_l t_l^T + \sum_{l=\max(1,r-k)}^{r} \sigma_l u_l v_l^T,$$

for $k = m/2, m/4, m/8, m/16$ where $r = \operatorname{rank}(B)$.

SOLUTION. Choose first 40 modes from the composition image, and mmodes 50 to 60 from the style image.

```
octave] ims=pread("Vincent_van_Gogh_50.jpg");
octave] figure(4); imshow(ims)
octave] imC=psize(im,256,256); imS=psize(ims,256,256);
octave] figure(1); clf; imshow(imC); print -deps Hw03Fig04.eps;
octave] figure(2); clf; imshow(imS); print -deps HW03Fig05.eps;
octave] [Uc,Sc,Vc]=svd(imC,1); [Us,Ss,Vs]=svd(imS,1);
octave] function im=svdimpq(S,U,V,p,q)
            im=S(p,p)*U(:,p)*V(:,p)';
            for k=p+1:q
              im=im+S(k,k)*U(:,k)*V(:,k)';
            end;
         end
octave] imCS = svdim(Sc,Uc,Vc,40) + svdimpq(Ss,Us,Vs,50,60);
octave] figure(5); clf; imagesc(imCS); colormap(gray); print -deps Hw03Fig06.eps;
octave]
```
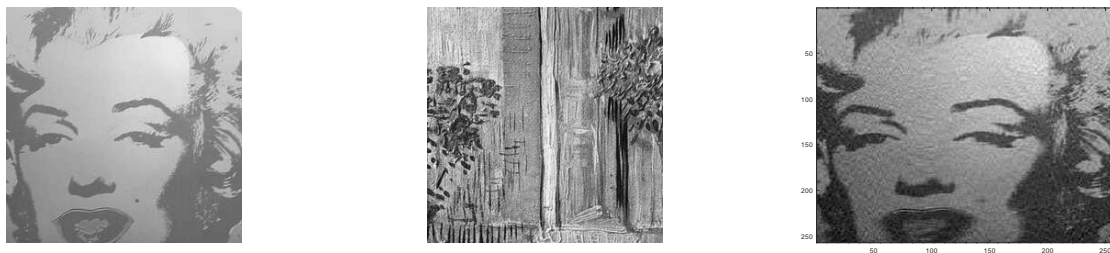

Figure 2. Applying the style of Van Gogh to a work by Warhol.