Teaching with Technology in MATH547 Linear Algebra for Applications in Data Science

Background and motivation

Imparting data literacy is a topic of timely interest for UNC. A key foundation for data science is linear algebra, the study of scaling and combining multidimensional objects. These objects can be images, medical records, sounds, social networks. They all tend to be represented by groupings of numbers (vectors, $\boldsymbol{x}, \boldsymbol{b}$), and combining them to form new objects is encoded in yet other groupings of numbers (matrices $\boldsymbol{A}, \boldsymbol{B}$). Traditional teaching of the associated mathematical concepts typically use hand exercises for small vectors, i.e.,

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}, \boldsymbol{x} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \boldsymbol{b} = \boldsymbol{A} \cdot \boldsymbol{x} = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ -4 \end{pmatrix}.$$

Though correct, such presentations do not explore the rich variety of applications of linear algebra to the large-scale applications driving the current interest in data science.

Teaching technology to be introduced

This course builds upon previous experience to provide live course material that combines theoretical concepts with embedded computation and visualization. The new technologies to be used are:

- An editing platform for live documents (TeXmacs, http://texmacs.org/tmweb/home/welcome.en.html)
- Embedded Octave/Matlab computation in TeXmacs (through interfaces I have personally developed for Windows, OS X)
- Data sets gathered from a variety of sources to provide realistic applications (Kaggle, MIT Faces)

Instructional benefits

- Elimination of drudgery. After a few basic examples have been mastered, live documents allow effortless exploration of "what if" type questions, e.g., what if the vectors are colinear?, what if the vectors are orthogonal?, what if the vectors are nearly colinear?
- Appreciation of application range. Live documents allow realistic applications, for instance the face recognition example below, but also examples from the humanities (see syllabus).

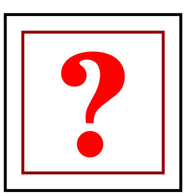
- Familiarity with data input/output. Students gain an appreciation of the steps needed to transform real data into mathematical constructs, e.g., from electroencephalograms to vectors.

- Job-market skills to complement theory. Basic coding, computer architecture and operations are learned along with the mathematics.

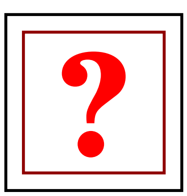
Example:

In this example, the MIT data base of face images is processed to produce various representations of an "average individual". All operations are modifiable directly from within the document allowing student exploration of the topic of facial recognition.

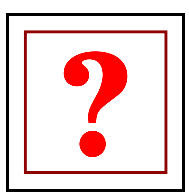
```
octave: function cimg=cmprim(img,p)
          [h,w]=size(img); mat=zeros(h,w); mat=double(img)/255.;
          [U S V]=svd(img); s=diag(S); cmat=zeros(h,w);
          for i=1:p
            cmat = cmat + s(i)*U(:,i)*V(:,i)';
          end;
          cmn=min(min(cmat)); cmat=cmat+cmn*ones(h,w);
          cmx=max(max(cmat)); sc=255./cmx; cmat=sc*cmat; cimg=uint8(cmat);
          fname=strcat("/home/student/courses/MATH547ML/cimg",num2str(p),
        ".png");
          imwrite(cimg,fname);
        endfunction;
octave: load /home/student/courses/MATH547/lessons/mitfaces/faces.mat;
octave: face=reshape(a,128,128)';
        facep01=cmprim(face,1); facep02=cmprim(face,2); facep04=cmprim(face,
        4);
        facep08=cmprim(face,8); facep16=cmprim(face,16); facep32=cmprim(face,
        32);
        facep64=cmprim(face,64); facep128=cmprim(face,128);
octave:
```



$p=1$



$p=2$



$p=4$

$p=8$

$p=16$

$p=32$