Homework 10

Thursday, November 19, 2015 9:30 AM

1) In class, we saw that the nonzero eigenvalues of AA^T are the same as the nonzero eigenvalues of A^TA —— just the squares of the nonzero singular values of A. In this problem, you'll relate the singular values of A to the eigenvalues of $\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}$.

ⓐ Diagonalize the matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

ⓑ Diagonalize the matrix

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

ⓒ Let A be an arbitrary $m \times n$ real matrix, with $m \le n$. Let

$$B = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \left(\begin{array}{c|c} 0 & A \\ \hline A^T & 0 \end{array}\right) \end{array}.$$

B is an $(m+n) \times (m+n)$ symmetric matrix. (Therefore it is unitarily diagonalizable.)

In terms of the singular values and left- and right-singular vectors of A, specify the eigenvalues and eigenvectors of B.

Note: B has $m+n$ eigenvalues, so don't forget any! Parts ⓐ and ⓑ should be helpful special cases, but feel free to experiment more with Matlab until you see the pattern. Commands like these might be helpful:

```
m = 2;
n = 3;
A = randn(m, n);
[U, S, V] = svd(A)  % <-- returns left singular vectors, singular values, right singular vectors

B = [zeros(m,m) A; A' zeros(n,n)];
[W, D] = eig(B) % <-- returns eigenvectors, eigenvalues
```

2) A real symmetric matrix is "positive definite" if its eigenvalues

are all >0, positive semi-definite if its eigenvalues are all ≥ 0, and "indefinite" otherwise.

ⓐ Classify each of the following matrices as positive definite, positive semidefinite, or indefinite. Try to do it by hand.

$$A = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 4 \\ 3 & 4 & 9 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 6 & -2 & 0 \\ 0 & -2 & 5 & -2 \\ 0 & 0 & -2 & 3 \end{pmatrix}$$

$$D = B^{-1}$$

$$E = C^{-1}$$

$$F = \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}$$

ⓑ - If A is positive semidefinite and $\alpha \geq 0$, prove that αA is positive semi-definite.

- If A and B are positive semi-definite, prove that $A+B$ is positive semidefinite.
(<u>Hint</u>: You'll probably want to use a theorem from class...)
- Conclude that if A and B are positive semi-definite, then so are the matrices $pA + (1-p)B$, for all $p \in [0,1]$.

(Thus the set of positive semi-definite matrices of a given dimension is **convex**. This is extremely important in optimization theory: namely, in semi-definite programming.)

ⓒ The definition of a positive semi-definite matrix can be used to define a partial order on symmetric matrices.

Definition: For symmetric matrices A and B of the same dimensions, define
"$A \succeq B$"

if $A-B$ is positive semi-definite.

- Give an example of two symmetric matrices A and B such that neither $A \succcurlyeq B$ nor $B \succcurlyeq A$. These matrices are incomparable; that's why it is called a partial order.

- Give an example of symmetric A and B so that

$$A \succcurlyeq B \text{ is true,}$$
$$\text{but } A^2 \succcurlyeq B^2 \text{ is false!}$$

(In other words $A-B$ is positive semidefinite, but A^2-B^2 is not.)

(Hint: You can use 2×2 matrices. Play around until you find an example, and then try to simplify it to understand how it works.)

3) In 1929, Hubble famously showed that the universe is expanding. Specifically, he showed a roughly linear relationship between the distances of other galaxies and their velocities away from us.

Here is the data he used:

NGC #	Distance (x10^6 parsecs)	Radial velocity (km/sec)	Right ascension	Declination	Adjusted radial velocity
NA	0.032	170	NA	NA	170
NA	0.034	290	NA	NA	290
6822	0.214	-130	{19, 44, 57.8}	{-14, 48, 11}	60
598	0.263	-70	{1, 33, 51.}	{30, 39, 37}	15
221	0.275	-185	{0, 42, 41.9}	{40, 51, 57}	-30
224	0.275	-220	{0, 42, 44.3}	{41, 16, 9}	-65
5457	0.45	200	{14, 3, 12.5}	{54, 20, 53}	395
4736	0.5	290	{12, 50, 52.6}	{41, 7, 9}	405
5194	0.5	270	{13, 29, 52.4}	{47, 11, 41}	430
4449	0.63	200	{12, 28, 11.}	{44, 5, 33.4}	305
4214	0.8	300	{12, 15, 39.2}	{36, 19, 41}	370
3031	0.9	-30	{9, 55, 33.2}	{69, 3, 55}	90
3627	0.9	650	{11, 20, 15.1}	{12, 59, 22}	580
4826	0.9	150	{12, 56, 43.9}	{21, 41, 0}	205
5236	0.9	500	{13, 37, 0.8}	{-29, 51, 59}	425
1068	1	920	{2, 42, 40.8}	{0, 0, 48}	830
5055	1.1	450	{13, 15, 49.3}	{42, 1, 47}	585
7331	1.1	500	{22, 37, 4.3}	{34, 24, 59}	740
4258	1.4	500	{12, 18, 57.5}	{47, 18, 14}	610
4151	1.7	960	{12, 10, 32.7}	{39, 24, 20}	1035
4382	2	500	{12, 25, 24.2}	{18, 11, 27}	515
4472	2	850	NA	NA	850
4486	2	800	{12, 30, 49.4}	{12, 23, 28}	800
4649	2	1090	{12, 43, 40.2}	{11, 33, 9}	1100

The second column gives the distance to each galaxy, and the last column gives the velocity away from us. (Hubble actually started with the data in column 3, but I have adjusted these velocities for the motion of our Sun.)

(a) Run linear regression of distance versus velocity to find the best-fitting line. Make sure your line goes through (0,0)!!

(b) Now run linear regression of velocity versus distance. Why does this give a different answer than part (a)?

(c) Now use PCA to find the best-fitting line. Plot all three lines, and the data, on one labeled graph. Why is PCA more appropriate for analyzing this data than either linear regression?

(d) Some of these data points are more precise than others. For example, they may have been collected by different telescopes.

In class, we saw how to get the k-dimensional subspace S that minimizes

$$\sum_{j=1}^{m} \| \vec{x}_j - P_S \vec{x}_j \|^2 ,$$

the sum of the squared distances from the data points to their projections on S. (The answer was to set $S = \text{Span}\{k \text{ largest e-value e-vectors of } \sum_j x_j x_j^T\}$.)

Extend this analysis to show how to get the k-dim'l subspace S that minimizes

$$2 \cdot \| \vec{x}_1 - P_S \vec{x}_1 \|^2 + \sum_{j=2}^{m} \| \vec{x}_j - P_S \vec{x}_j \|^2 .$$

(This situation would arise if data point $\vec{x}_1$ was more precise than the others.)