Eddie Price **A few uses of eigenvalues** **Spring 2018**

In this class, we don't see any applications of eigenvalues and eigenvectors, so I thought I would show you a couple.

Multivariable Optimization. To generalize the second derivative test to a function $f(x_1, \ldots, x_n)$ of n variables, we look at the Hessian matrix

$$\mathcal{H} = \begin{bmatrix} f_{x_1x_1} & f_{x_1x_2} & \cdots & f_{x_1x_n} \\ f_{x_2x_1} & f_{x_2x_2} & \cdots & f_{x_2x_n} \\ \vdots & \vdots & & \vdots \\ f_{x_nx_1} & f_{x_nx_2} & \cdots & f_{x_nx_n} \end{bmatrix}$$

where the (i, j)-entry of $\mathcal{H}$ is the 2nd order partial derivative of your function f with respect to x_i and then x_j.
For a three variable function $f(x, y, z)$, you would have

$$\mathcal{H} = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix}$$

To find all of the local extrema of f, you find all of the critical points by taking all of the first order partial derivatives and setting them equal to 0 simultaneously. For each critical point, you plug it into all of the functions in the Hessian matrix, and then do the following:
If $\det \mathcal{H} = 0$, the test is inconclusive.
If $\det \mathcal{H} \neq 0$, then determine the signs of the eigenvalues of $\mathcal{H}$
- If all of the eigenvalues are positive, then f has a local minimum at this critical point.
- If all of the eigenvalues are negative, then f has a local maximum at this critical point.
- If $\mathcal{H}$ has positive *and* negative eigenvalues, then f has a saddle point at this critical point.

For an example, suppose you have $f(x, y, z) = x^2 + (y - 1)^2 + (z + 1)^2$. Then $f_x = 2x$, $f_y = 2y - 2$, $f_z = 2z + 2$. Setting all of these equal to 0, we get the critical point $(0, 1, -1)$. The 2nd order partial derivatives are $f_{xx} = 2$, $f_{yy} = 2$, $f_{zz} = 2$ and all mixed partials are 0:

$$\mathcal{H} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

One can check that $\det \mathcal{H} = 8 \neq 0$, and that all of the eigenvalues of $\mathcal{H}$ are 2, so they are all positive. Therefore, $f(x, y, z) = x^2 + (y - 1)^2 + (z + 1)^2$ has a local minimum at $(0, 1, -1)$.

It takes a bit of linear algebra knowledge to show, but the 2nd derivative test we saw this semester is a special case of this more general second derivative test. In particular, the discriminant function is the determinant of the 2×2 Hessian matrix. And one can show that for a 2×2 matrix A: if $\det A > 0$ and the diagonal entries are positive, then the eigenvalues of A are positive; if $\det A > 0$ and the diagonal entries are negative, then the eigenvalues of A are negative; and if $\det A < 0$, then the eigenvalues of A are opposite signs.

Diagonalization of Matrices: A digonal matrix is a matrix in which all entries off of the diagonal are 0 (the entries on the diagonal can be anything). The following 3 matrices are all diagonal matrices:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Multiplying by a diagonal matrix is very simple. For example, if you were to multiply AM for any 3×3 matrix M, you would simply multiply all of the entries in row 1 of M by 1, multiply all of the entries in row 2 of M by -2, and multiply all of the entries in row 3 of M by 4.

In particular, it is very easy to find powers of diagonal matrices: all you have to do is raise each diagonal entry to that power! So, for example,

$$A^5 = \begin{bmatrix} 1^5 & 0 & 0 \\ 0 & (-2)^5 & 0 \\ 0 & 0 & 4^5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -32 & 0 \\ 0 & 0 & 1024 \end{bmatrix}$$

As such, it is often nice if we can "diagonalize" a matrix. That is, if we could write a matrix A as $P^{-1}DP$ where D is a diagonal matrix and P is a nonsingular matrix. Of course, this would allow us to find high powers of A much more easily. If $A = P^{-1}DP$ for a diagonal matrix D, then $A^n = P^{-1}D^nP$, and it is easy to find D^n since you just have to raise each diagonal entry to the power n.
Why is this true? Let's look at A^4. $A^4 = (P^{-1}DP)^4 = (P^{-1}DP)(P^{-1}DP)(P^{-1}DP)(P^{-1}DP) = P^{-1}DPP^{-1}DPP^{-1}DPP^{-1}DP = P^{-1}DDDDP$ (since PP^{-1} is the identity matrix), which is $P^{-1}D^4P$.

A theorem in linear algebra tells us that if a matrix A is digonalizable, then the diagonal matrix D has the eigenvalues of A as its diagonal entries, and we can choose P to be the matrix whose columns are eigenvectors corresponding to the eigenvalues of A.

For an example, let's find A^{99} where $A = \begin{bmatrix} -3 & -4 \\ 2 & 3 \end{bmatrix}$

First, we find the eigenvalues of A: $0 = \det \begin{bmatrix} \lambda + 3 & 4 \\ -2 & \lambda - 3 \end{bmatrix} = (\lambda + 3)(\lambda - 3) - (-2)(4) = \lambda^2 - 9 + 8 = \lambda^2 - 1 = (\lambda + 1)(\lambda - 1)$. So the eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = 1$. We can now find corresponding eigenvectors:

$$\lambda_1 = -1 : \left[\begin{array}{cc|c} -1+3 & 4 & 0 \\ -2 & -1-3 & 0 \end{array}\right] = \left[\begin{array}{cc|c} 2 & 4 & 0 \\ -2 & -4 & 0 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Hence, $\vec{x_1} = \begin{bmatrix} -2t \\ t \end{bmatrix}$ for any $t \neq 0$. Choosing $t = 1$, we get $\vec{x_1} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

$$\lambda_1 = 1 : \left[\begin{array}{cc|c} 1+3 & 4 & 0 \\ -2 & 1-3 & 0 \end{array}\right] = \left[\begin{array}{cc|c} 4 & 4 & 0 \\ -2 & -2 & 0 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Hence, $\vec{x_1} = \begin{bmatrix} -t \\ t \end{bmatrix}$ for any $t \neq 0$. Choosing $t = 1$, we get $\vec{x_1} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Therefore, we get

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad P = [\vec{x_1}\vec{x_2}] = \begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix}$$

Computing $\det P = -2 + 1 = -1$, we get that

$$P^{-1} = \frac{1}{-1}\begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & 2 \end{bmatrix}$$

Hence,

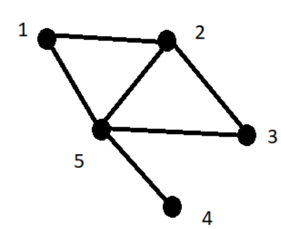
$$\begin{bmatrix} -3 & -4 \\ 2 & 3 \end{bmatrix} = A = P^{-1}DP = \begin{bmatrix} -1 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix}$$

Thus,

$$A^{99} = \begin{bmatrix} -1 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} (-1)^{99} & 0 \\ 0 & 1^{99} \end{bmatrix}\begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -3 & -4 \\ 2 & 3 \end{bmatrix}$$

Although this took some work, this general method could be much faster than multiplying A by itself 99 times.

Why care about powers of matrices?: Powers of matrices can show up in lots of situations. One possible situation comes from graph theory. A graph is a collection of vertices and edges. Consider the graph below, with its verties labeled 1 through 5, made in MS Paint:



An important tool in analyzing a graph is the adjacency matrix. The (i, j)-entry of the adjacency matrix is 0 if there is no edge connecting vertex i and j and is 1 if there is an edge connecting vertex i and j. So the adjacency matrix for the above graph is

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Powers of the adjacency matrix give you information about the graph. The (i, j)-entry of A^n tells you the number of distinct paths of length n in the graph that start at vertex i and end at vertex j.

Adjacency matrices are always diagonalizable, so it is relatively easy (assuming you can find eigenvalues and eigenvectors of A!) to find information about length n paths in the graph, by diagonalizing A, then taking the appropriate power. You can imagine that this information might be useful to companies that are concerned with data analysis of networks. Facebook, Google, Amazon, Netflix, YouTube, etc. care about this sort of information in terms of connectivity of friends, related searches, products one might be interested in based on other products they've viewed, movies/shows one might be interested in based on other movies/shows they've watched and rated, etc. Of course, the bigger the data set is, the harder it is to actually compute these things!

Systems of differential equations: We saw earlier in the semester that differential equations are very useful for modeling change, since it is often easy to identify how a quantity changes. But sometimes, you may care about different quantities that are related to each other.

For example, you may care about the populations of a predator and prey in a particular region. The population of the predator will change based on the population of both the predator and the prey, and the population of the prey will change based on the population of both the predator and the prey. So you can get a system of differentiation equations like

$$x' = f(t)\,x + g(t)\,y, \qquad y' = h(t)\,x + j(t)\,y$$

which can be represented in the matrix format

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} f(t) & g(t) \\ h(t) & j(t) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Knowing the eigenvalues and eigenvectors of the coefficient matrix $\begin{bmatrix} f(t) & g(t) \\ h(t) & j(t) \end{bmatrix}$ can help you find solutions to the system of differential equations.