**Linear Algebra Packet**  Name:

**Part I: Gaussian Elimination**

1. Solve the following system using Gaussian elimination and then use the Grapher utility on a Mac to visualize the solution point:



2. Solve the following system using Gaussian elimination and then use the Grapher utility on a Mac to visualize the solution line:



3. Solve the following system using Gaussian elimination and then use the Grapher utility on a Mac to visualize the solution plane:



4. (Try) to solve the following system using Gaussian elimination and then use the Grapher utility on a Mac to visualize why there is no solution:



Homework:

1. Solve:



2. Solve:



**Part II: Matrix Multiplication & Elementary Matrices**

1. Find the **inner product** (aka dot product) of the row vector  and the column vector .

2. Consider the system of equations:



We can think of the left hand side as a matrix multiplication in two different ways.

A. Write the left hand side as three rows where each row includes an inner product.

B. Write the left hand side as a linear combination of three column vectors.

3. Write out what the general ith row of the product Ax is using sigma notation where:



4. What are the requirements on the size of matrices A and B in order to multiply A\*B? What is the size of the resulting matrix?

5. Find the following products:

a.  b. 

c. 

6. More generally, write the entry in the ith row and jth column, denoted (AB)ij, of this product:



7. **Matrix multiplication is associative**, meaning that (AB)C=A(BC). Show that this is true for:



8. **In general, matrix multiplication is NOT commutative**, meaning that AB does not usually equal BA. Show that AB and BA are not equal for .

Some definitions:

The  **identity matrix** has 1’s along its diagonal and zero’s everywhere else. For any matrix A,  when the dimensions match up appropriately.

The elementary matrices  are the identity matrices with an extra –L term in the ij*th* entry. Multiplying an elementary matrix on the left has the effect of subtracting L times row j from row i. For example:

 has the effect of subtracting L times row 1 from row 3.

Gaussian elimination is done by multiplying by elementary matrices.

9. Consider the system:



a. Use an elementary matrix in which you multiply row 1 by 2 and subtract it from row 2.

b. From there, use an elementary matrix in which you add row 1 to row 3.

c. From there, use an elementary matrix in which you add row 2 to row 3.

d. You should now have an upper triangular matrix. Solve for x,y,z.

Homework:

Use elementary matrices to solve:

a.



b.



**Part III: Inverses**

Def: A matrix A is **invertible** if there exists a matrix A-1 such that .

1. Prove this theorem: If a matrix does have an inverse, then it is unique.

2. The inverse of a 2x2 matrix, if it exists, is given by the formula:



We’ll soon get to that the determinant of matrix A is ad-bc.

3. What is the inverse of a diagonal matrix ?

4. Prove this theorem: .

Gauss-Jordan Substitution: To solve for the inverse of a matrix, we can start out with this equation AX=I below:



And then perform Gaussian elimination until the A matrix is transformed into the identity matrix. Whatever is left in the X position will be in the inverse matrix.

5. Explain why this process works.

6. Use Gauss-Jordan elimination to find the inverse of:



Homework:

1. Find the inverse of:

a. 

b.



c. 

d. Explain why  is a rotation matrix in 2D using 90 degrees as an example.

Explain why  is its inverse using:

a.) Your 2x2 formula.

b.) Geometric intuition of what would “undo” a rotation by an angle. Hint: You’ll need to use even and odd properties of sine and cosine.

**Part IV: Determinants**

Recall our 2x2 formula for the determinant:



The following properties will extend to higher order matrices, but let’s gain some intuition with the 2D case.

1. Show that when rows are exchanged, the determinant changes sign.

2. Show that multiplying a row by a scalar multiple of the original row multiples the determinant by that amount.

3. Show that if two rows are equal, then the determinant is zero.

4. Show that if the matrix has a row (or column) or zeros, then the determinant is zero.

5. Show that subtracting a multiple of one row from another row leaves the same determinant (so that Gaussian elimination does not affect the determinant).

6. Show that if A is diagonal, then the determinant is the product of its diagonal elements.

7. Show that if A is triangular, then the determinant is the product of its diagonal elements.

8. Show that A is invertible if and only if the determinant does not equal zero.

9. Show that .

10. The **transpose** of a matrix A, denoted AT, is obtained by putting the ijtj elements of A in the jith position in AT. For example, the transpose of  is .

Show that .

11. Show that  in the general case (non 2x2). To do this, use property #9 and the definition of inverse.

**Higher Order Determinant Formula:**

Det(A) is equal to a linear combination of any row i (or column j) times its cofactors:

,

where the cofactor, Cij, is the determinant of Mij with the correct sign:

,

where Mij is the minor matrix formed by deleting the ith row and jth column of A.

Here’s a 3x3 example:



Note: Because you’ll multiply each determinant by a coefficient aij, you’ll want to choose the row or column with the most zeros in order to simplify your calculations.

Example:



**Cramer’s Rule:**

The solution of Ax = b can be found by computing , where Bj is the A matrix with its jth column replaced by the b vector.

Example:

The solution to  is:

 and 

**Homework:**

Calculate the determinant of the following matrices. Try using determinant properties to simplify your calculations on a few of them.

a. 

b. 

c. 

d. 

2. Solve  using Cramer’s rule.

**Part V: Projections and Least Squares**

Consider systems with more equations than unknowns:



1. For what values of b’s will this system be solvable?

Suppose the b’s are not in that special form. Then this system is inconsistent, which arises all the time in real life, and still must be solved. We can determine x by solving part of the system and ignoring the rest of the equations, but this is hard to justify. Instead, let’s choose x that minimizes the average error E from all n equations. Meaning, let’s minimize:



2. Why do we square each term above instead of just adding up the differences?

3. Take the derivative of E(x) with respect to x and set it equal to zero in order to solve for your approximation, .

4. Show that . This is called our **least squares solution**.

In our case, if we were trying to solve , then our least squares solution would be , which is in between ¼ and ½ (the solutions to the first and last equations individually).

How close does this approximation get to the b vector (1,1,1)? The **projection** is given by .

More generally, let’s try to solve . We’ll need find an approximation  that minimizes the error of . The proof for the formula for  is a bit beyond the scope of this class, as we haven’t talked about column and null space, but the general idea is that you need to solve .

5. Show that  can be transformed to , our **least squares solution** formula.

How well does our solution do at approximating b? The closest we get is the **projection** .

6. Show that if A is invertible then the above formula  simplifies to the exact solution .

For example, to solve:

, we could calculate , where , to find that  and the projection is: 

Homework:

Find the a.) least squares solution and b.) projection of the following system:

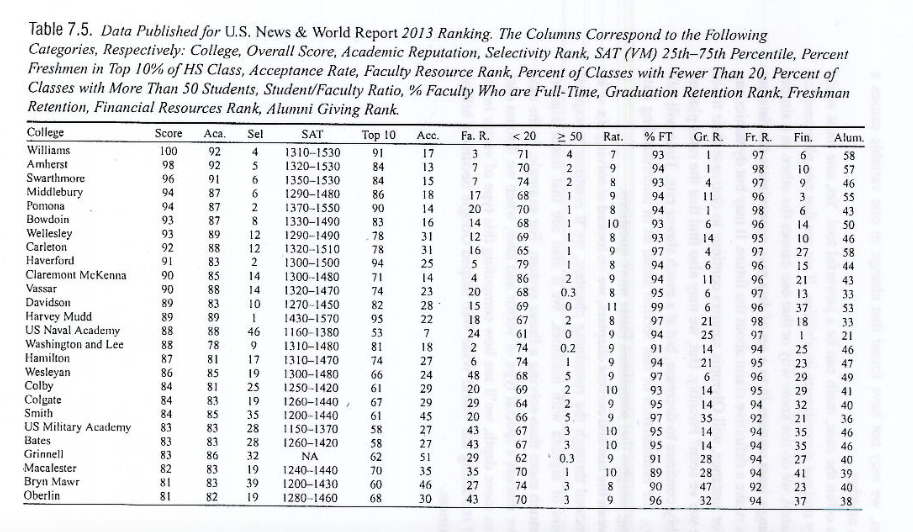


2. The goal of regression is to fit a mathematical model to a set of observed points. Say we’re collecting data on the number of machine failures per day in some factory. Imagine we’ve got three data points: (day, number of failures) (1,1) (2,2) (3,2).

The goal is to find a linear equation that fits these points. We believe there’s an underlying mathematical relationship that maps “days” uniquely to “number of machine failures” form , where x is the day and y is the number of failures.

Put this data into a matrix system and then use least squares to solve it.

3. Open up the College Rankings Jupyter Notebook to find the following table from Life is Linear:



Create a Python program that does the following:

1. Uses *all* of the colleges in the above list except for Grinnell (who doesn’t publish SATs) and all of the variables above (Academic Reputation through Alumni Giving Rank) to create a matrix, A. In addition, create a vector, b, containing all of the colleges’ (except for Grinnell’s) US News World & Report Score (2nd column).

2. Solves for the weight of each ranking using the formula .

3. What variables affect a school’s ranking most in a positive way? In a negative way?

4. Use the weights to calculate the US News World & Report score for Colby. (Don’t enter Colby’s data manually but rather use the row in your matrix above).

5. Read about the similarities and differences between the Ordinary Least Squares and Gradient Descent methods.

**Part VI: Eigenvalues and Eigenvectors**

Consider the following system of differential equations:

,

also written as: .

Note that if this was just a one dimensional system, then by separation of variables we would find that  has the general exponential solution .

We shall take a direct approach and look for solutions with the same exponential dependence on t as in the one dimensional case.

1. Assume that  and  and then plug these formulas and their derivatives into the original system above to obtain .

Note that

 can be written in matrix form as , where .

2. Show that solving  is equivalent to solving .

If we are assuming that v is a nonzero eigenvector, then we must have the **characteristic** equation .

In the 2x2 case,

 .

Once we solve this quadratic equation for , we will have our **eigenvalues  and **, and then we can plug each value into  to solve for each **eigenvector** v1 and v2. Note that  and  are the pure exponential solutions to  and the superposition of both of them gives us the most general purely exponential solution: 

Note: As this is not a full linear algebra course, we will only deal with real, unique roots (not repeated roots or imaginary roots). Their most general solution formulas are a bit different. A few things to know about eigenvalues:

a. Their form greatly affects the type of solution behavior. Purely imaginary eigenvalues lead to strictly periodic (sine and cosine) behavior. Whether or not the real part of the root is positive or negative affects whether there is exponential growth or decay.

b. Eigenvalues are related to the frequency of oscillations of the solutions. As a historical note, soldiers do not march in step as they go across a bridge because if they happen to march at the same frequency as one of the eigenvalues of the bridge, then the bridge begins to oscillate. (Just as a child’s swing, you soon notice the natural frequency of the swing, and by matching it you go higher). An engineer tries to keep natural frequencies of his bridge away from those of wind. The Tacoma Narrows Bridge actually crashed in 1940 due to wind and the Broughton Bridge collapsed in 1831 due to soldiers marching.

Let’s return to solving this example:



3. Find the eigenvalues.

4. Find each eigenvector.

5. Thus, the general solution is . Use the initial conditions to show that the particular solution is .

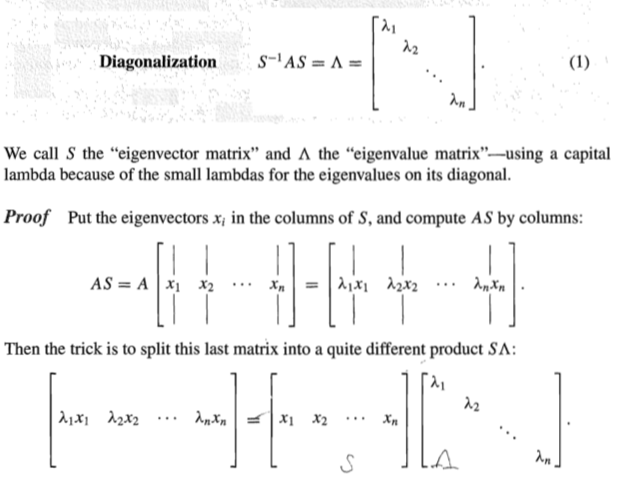
Homework: Find the particular solution to:

a. 

b. 

**Part VII: Singular Value Decomposition**

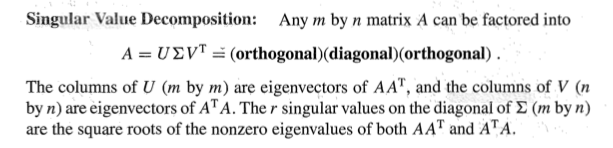
Suppose the nxn matrix A has n distinct, real eigenvalues. If their corresponding eigenvectors are the columns of the matrix S, then S-1AS is a diagonal matrix, with eigenvalues along the diagonal:



1. Show that  implies  and .

2. Recall our example from before , where  with eigenvalues -1 and 2 and corresponding eigenvectors  and . Show that .

The above formula works if A is a square matrix. What if A is not a square matrix? Well,  and  are always square. The diagonal matrix  has eigenvalues from . Those positive entries are called the singular values and will be placed along the diagonal.



We won’t prove this theorem but we will use it.

3. We’ll decompose  into this form using the following steps:

a. Find the eigenvalues of .

b. Find the eigenvalues of .

c. The only positive eigenvalues are 25 and 9. Stack the square roots of these eigenvalues together to get . **Note that the size of this matrix will always have the number of rows as  and the number of columns of .**

d. Find the normalized eigenvectors of . Stack these to form U: 

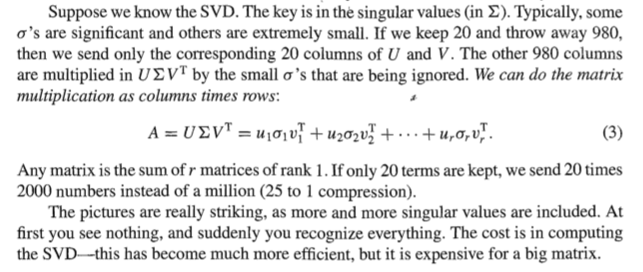
e. Find the normalized eigenvectors of . Stack these to form V: 

f. Transpose V and we now have our decomposition of A:



**Application of SVD: Image compression**

Suppose a satellite takes a picture and wants to send it back to earth. The picture may contain 1000x1000 pixels – a million little squares, each with color. We can code the colors and send back 1,000,000 numbers. It is better to find the *essential* information inside the 1000x1000 matrix and send only that.



**Homework:**

1. Explain the previous claim that only 20 times 2000 numbers will need to be sent.

2. Open the image compression Python notebook to view SVD image compression. Describe your observations. What image characteristics were picked up using only two eigenvalues? What were the maximum number of eigenvalues that you could have used? How many eigenvalues were needed to make a reasonably good picture?

2. Show that SVD decomposition of  is given by:



3. Show that the SVD decomposition of  is given by:

