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:

$$3x + y + z = 3$$
$$12x - 2y + 3z = 3$$
$$-6x - 2y - 4z = 12$$

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

$$2x + 3y + 4z = 12$$

$$x + y + z = 6$$

$$4x + 6y + 8z = 24$$

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

$$3x+2y+4z=12$$
$$6x+4y+8z=24$$
$$12x+8y+16z=48$$

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:

$$x+y+z=2$$
$$2x+2y+2z=6$$
$$4x+6y+8z=10$$

Homework:

1. Solve:

$$x + y + z = 2$$
$$x + 3y + 3z = 0$$
$$x + 3y + 5z = 2$$

2. Solve:

$$x + 2y + z = 6$$
$$3x + 6y + 3z = 18$$
$$x + y + z = 12$$

Part II: Matrix Multiplication & Elementary Matrices

1. Find the **inner product** (aka dot product) of the row vector $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ and the column vector $\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$.

2. Consider the system of equations:

$$\begin{aligned} 2x + y + z &= 5 \\ 4x - 6y \quad &= -2 \\ -2x + 7y + 2z &= 9 \end{aligned}$$

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 i^{th} row of the product Ax is using sigma notation where:

$$\begin{bmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{bmatrix}$$

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. $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \\ 4 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -2 & -2 \\ -0 & -1 \end{bmatrix}$

b. $\begin{bmatrix} 1 & 2 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \\ 0 \\ 3 \end{bmatrix}$

c. $\begin{bmatrix} -1 \\ 1 \\ 0 \\ 3 \end{bmatrix}\begin{bmatrix} 1 & 2 & 0 & 1 \end{bmatrix}$

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

$$\begin{bmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} b_{11} & \cdots & \cdots & b_{1p} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ b_{n1} & \cdots & \cdots & b_{np} \end{bmatrix}$$

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

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

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 $A=\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, B=\begin{bmatrix} 2 & 1 \\ 3 & 0 \end{bmatrix}$.

Some definitions:

The I_{nxn} **identity matrix** has 1's along its diagonal and zero's everywhere else. For any matrix A, $A = AI = IA$ when the dimensions match up appropriately.

The elementary matrices E_{ij} 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:

$$E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -l & 0 & 1 \end{bmatrix}$$ 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:

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}$$

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.

$$3x + y + z = 3$$
$$12x - 2y + 3z = 3$$
$$-6x - 2y - 4z = 12$$

b.

$$x + y + z = 2$$
$$x + 3y + 3z = 0$$
$$x + 3y + 5z = 2$$

Part III: Inverses

Def: A matrix A is **invertible** if there exists a matrix A^{-1} such that $AA^{-1} = A^{-1}A = I$.

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:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

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

3. What is the inverse of a diagonal matrix $\begin{bmatrix} d_1 & \cdots & \cdots & 0 \\ \vdots & d_2 & & \vdots \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & d_n \end{bmatrix}$?

4. Prove this theorem: $(AB)^{-1} = B^{-1}A^{-1}$.

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

$$\begin{bmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_{11} & \cdots & \cdots & x_{1n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ x_{n1} & \cdots & \cdots & x_{nn} \end{bmatrix} = \begin{bmatrix} 1 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 1 \end{bmatrix}$$

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:

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}$$

Homework:

1. Find the inverse of:

a. $\begin{bmatrix} 5 & 6 \\ -1 & 2 \end{bmatrix}$

b.

$$\begin{bmatrix} -1 & 2 & -1 \\ 2 & -1 & 0 \\ 0 & -1 & 2 \end{bmatrix}$$

c. $$\begin{bmatrix} 1 & -2 & 1 \\ -1 & 2 & 2 \\ 1 & 0 & 1 \end{bmatrix}$$

d. Explain why $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is a rotation matrix in 2D using 90 degrees as an example.

Explain why $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$ 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:

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

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 $\det(AB) = \det(A)\det(B)$.

10. The **transpose** of a matrix A, denoted A^T, is obtained by putting the ij^{tj} elements of A in the ji^{th} position in A^T. For example, the transpose of $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ is $A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$.
Show that $\det A = \det A^T$.

11. Show that $\det A^{-1} = \dfrac{1}{\det A}$ 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:

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \ldots + a_{in}C_{in},$$

where the cofactor, C_{ij}, is the determinant of M_{ij} with the correct sign:

$$C_{ij} = (-1)^{i+j}\det M_{ij},$$

where M_{ij} is the minor matrix formed by deleting the ith row and jth column of A.

Here's a 3x3 example:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}(-1)^{1+1}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{21}(-1)^{2+1}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}(-1)^{3+1}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

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

Example:

$$\begin{vmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 1 & 0 & 1 & 3 \\ 2 & 4 & 1 & 2 \end{vmatrix} = 2\begin{vmatrix} 2 & -1 & 0 \\ 0 & 1 & 3 \\ 4 & 1 & 2 \end{vmatrix} + (-1)(-1)\begin{vmatrix} -1 & -1 & 0 \\ 1 & 1 & 3 \\ 2 & 1 & 2 \end{vmatrix}$$

$$= 2\left(2\begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix} + (-1)(-1)\begin{vmatrix} 0 & 3 \\ 4 & 2 \end{vmatrix}\right) + 1\left((-1)\begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix} + (-1)(-1)\begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix}\right)$$

$$= 2(2(-1) + (1)(-12)) + ((-1)(-1) + (1)(-4))$$

$$= -28 - 3 = -31$$

Cramer's Rule:

The solution of Ax = b can be found by computing $x_j = \frac{\det B_j}{\det A}$, where B_j is the A matrix with its jth column replaced by the b vector.

Example:

The solution to $\begin{aligned} x_1 + 3x_2 &= 0 \\ 2x_1 + 4x_2 &= 6 \end{aligned}$ is:

$$x_1 = \frac{\begin{vmatrix} 0 & 3 \\ 6 & 4 \end{vmatrix}}{\begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix}} = 9 \quad \text{and} \quad x_2 = \frac{\begin{vmatrix} 1 & 0 \\ 2 & 6 \end{vmatrix}}{\begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix}} = -3$$

Homework:

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

a. $\begin{bmatrix} 2 & 1 & 1 \\ 3 & 2 & -1 \\ 0 & 3 & 4 \end{bmatrix}$

b. $\begin{bmatrix} 2 & 1 & 5 & 1 \\ 3 & 1 & 1 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 3 & 0 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & -2 & 1 \\ 0 & 3 & 0 & 1 \end{bmatrix}$

d. $\begin{bmatrix} 0 & 2 & 2 & 2 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & -2 & 1 \\ 0 & 3 & 0 & 1 \end{bmatrix}$

2. Solve $\begin{array}{l} 3x_1 + x_2 = 2 \\ x_1 - x_2 = 6 \end{array}$ using Cramer's rule.

Part V: Projections and Least Squares

Consider systems with more equations than unknowns:

$$2x = b_1$$
$$3x = b_2$$
$$4x = b_3$$

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:

$$E = (2x - b_1)^2 + (3x - b_2)^2 + (4x - b_3)^2$$

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, $\hat{x}$.

4. Show that $\hat{x} = \dfrac{a^T b}{a^T a}$. This is called our **least squares solution**.

$$2x = 1$$
$$3x = 1$$

In our case, if we were trying to solve $4x = 1$, then our least squares solution would be $\hat{x} = \frac{2+3+4}{2^2+3^2+4^2} = \frac{9}{29}$, 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

$$A\hat{x} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \frac{9}{29} = \begin{bmatrix} 18/29 \\ 27/29 \\ 36/29 \end{bmatrix}.$$

More generally, let's try to solve $A_{mxn} x_{nx1} = b_{mx1}$. We'll need find an approximation $\hat{x}$ that minimizes the error of $E = \| Ax - b \|$. The proof for the formula for $\hat{x}$ 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 $A^T(b - A\hat{x}) = 0$.

5. Show that $A^T(b - A\hat{x}) = 0$ can be transformed to $\hat{x} = (A^T A)^{-1} A^T b$, our **least squares solution** formula.

How well does our solution do at approximating b? The closest we get is the **projection** $A\hat{x} = A(A^T A)^{-1} A^T b$.

6. Show that if A is invertible then the above formula $\hat{x} = (A^T A)^{-1} A^T b$ simplifies to the exact solution $\hat{x} = A^{-1} b$.

For example, to solve:

$$x = 0$$
$$x + y = 8$$
$$x + 3y = 8$$
$$x + 4y = 20$$

, we could calculate $\hat{x} = (A^T A)^{-1} A^T b$, where $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$, to find that

$\hat{x} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ and the projection is: $A\hat{x} = \begin{bmatrix} 1 \\ 5 \\ 13 \\ 17 \end{bmatrix}$

Homework:

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

$$\begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

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:

Table 7.5. *Data Published for* U.S. News & World Report *2013 Ranking. The Columns Correspond to the Following Categories, Respectively: College, Overall Score, Academic Reputation, Selectivity Rank, SAT (VM) 25th–75th Percentile, Percent Freshmen in Top 10% of HS Class, Acceptance Rate, Faculty Resource Rank, Percent of Classes with Fewer Than 20, Percent of Classes with More Than 50 Students, Student/Faculty Ratio, % Faculty Who are Full-Time, Graduation Retention Rank, Freshman Retention, Financial Resources Rank, Alumni Giving Rank.*

College	Score	Aca.	Sel	SAT	Top 10	Acc.	Fa. R.	< 20	≥ 50	Rat.	% FT	Gr. R.	Fr. R.	Fin.	Alum.
Williams	100	92	4	1310–1530	91	17	3	71	4	7	93	1	97	6	58
Amherst	98	92	5	1320–1530	84	13	7	70	2	9	94	1	98	10	57
Swarthmore	96	91	6	1350–1530	84	15	7	74	2	8	93	4	97	9	46
Middlebury	94	87	6	1290–1480	86	18	17	68	1	9	94	11	96	3	55
Pomona	94	87	2	1370–1550	90	14	20	70	1	8	94	1	98	6	43
Bowdoin	93	87	8	1330–1490	83	16	14	68	1	10	93	6	96	14	50
Wellesley	93	89	12	1290–1490	78	31	12	69	1	8	93	14	95	10	46
Carleton	92	88	12	1320–1510	78	31	16	65	1	9	97	4	97	27	58
Haverford	91	83	2	1300–1500	94	25	5	79	1	8	94	6	96	15	44
Claremont McKenna	90	85	14	1300–1480	71	14	4	86	2	9	94	11	96	21	43
Vassar	90	88	14	1320–1470	74	23	20	68	0.3	8	95	6	97	13	33
Davidson	89	83	10	1270–1450	82	28	15	69	0	11	99	6	96	37	53
Harvey Mudd	89	89	1	1430–1570	95	22	18	67	2	8	97	21	98	18	33
US Naval Academy	88	88	46	1160–1380	53	7	24	61	0	9	94	25	97	1	21
Washington and Lee	88	78	9	1310–1480	81	18	2	74	0.2	9	91	14	94	25	46
Hamilton	87	81	17	1310–1470	74	27	6	74	1	9	94	21	95	23	47
Wesleyan	86	85	19	1300–1480	66	24	48	68	5	9	97	6	96	29	49
Colby	84	81	25	1250–1420	61	29	20	69	2	10	93	14	95	29	41
Colgate	84	83	19	1260–1440	67	29	29	64	2	9	95	14	94	32	40
Smith	84	85	35	1200–1440	61	45	20	66	5	9	97	35	92	21	36
US Military Academy	83	83	28	1150–1370	58	27	43	67	3	10	95	14	94	35	46
Bates	83	83	28	1260–1420	58	27	43	67	3	10	95	14	94	35	46
Grinnell	83	86	32	NA	62	51	29	62	0.3	9	91	28	94	27	40
Macalester	82	83	19	1240–1440	70	35	35	70	1	10	89	28	94	41	39
Bryn Mawr	81	83	39	1200–1430	60	46	27	74	3	8	90	47	92	23	40
Oberlin	81	82	19	1280–1460	68	30	43	70	3	9	96	32	94	37	38

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 (2^{nd} column).

2. Solves for the weight of each ranking using the formula $\hat{x} = (A^T A)^{-1} A^T b$.

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:

$$\frac{dx_1}{dt} = 4x_1 - 5x_2$$
$$\frac{dx_2}{dt} = 2x_1 - 3x_2$$
$$x_1(0) = 8, x_2(0) = 5$$,

also written as: $\frac{dx}{dt} = Ax$.

Note that if this was just a one dimensional system, then by separation of variables we would find that $\frac{dx}{dt} = ax$ has the general exponential solution $x(t) = ce^{at}$.

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 $x_1(t) = c_1 e^{\lambda t}$ and $x_2(t) = c_2 e^{\lambda t}$ and then plug these formulas and their derivatives into the original system above to obtain $\begin{aligned} \lambda c_1 &= 4c_1 - 5c_2 \\ \lambda c_2 &= 2c_1 - 3c_2 \end{aligned}$.

Note that

$\begin{aligned} \lambda c_1 &= 4c_1 - 5c_2 \\ \lambda c_2 &= 2c_1 - 3c_2 \end{aligned}$ can be written in matrix form as $Av = \lambda v$, where $v = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$.

2. Show that solving $Av = \lambda v$ is equivalent to solving $(A - \lambda I)v = 0$.

If we are assuming that v is a nonzero eigenvector, then we must have the **characteristic** equation $\det(A-\lambda I)=0$.

In the 2x2 case,

$$\det\begin{bmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{bmatrix} = (a_{11}-\lambda)(a_{22}-\lambda)-a_{12}a_{21}=0$$

.

Once we solve this quadratic equation for λ, we will have our **eigenvalues** λ_1 **and** λ_2, and then we can plug each value into $(A-\lambda I)v=0$ to solve for each **eigenvector** v_1 and v_2. Note that $x(t)=e^{\lambda_1 t}v_1$ and $x(t)=e^{\lambda_2 t}v_2$ are the pure exponential solutions to $\frac{dx}{dt}=Ax$ and the superposition of both of them gives us the most general purely exponential solution:

$$x(t)=c_1 e^{\lambda_1 t}v_1 + c_2 e^{\lambda_2 t}v_2$$

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:

$$\frac{dx_1}{dt} = 4x_1 - 5x_2$$
$$\frac{dx_2}{dt} = 2x_1 - 3x_2$$
$$x_1(0) = 8, x_2(0) = 5$$

3. Find the eigenvalues.

4. Find each eigenvector.

5. Thus, the general solution is $x(t) = c_1 e^{-1t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{2t} v_2 \begin{bmatrix} 5 \\ 2 \end{bmatrix}$. Use the initial conditions to show that the particular solution is $\begin{matrix} x_1 = 3e^{-t} + 5e^{2t} \\ x_2 = 3e^{-t} + 2e^{2t} \end{matrix}$.

Homework: Find the particular solution to:

a.
$$x' = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} x,$$
$$x(0) = \begin{bmatrix} 0 \\ -5 \end{bmatrix}$$

b.
$$x' = \begin{bmatrix} -5 & 1 \\ 4 & -2 \end{bmatrix} x,$$
$$x(0) = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

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 $S^{-1}AS$ is a diagonal matrix, with eigenvalues along the diagonal:

$$\textbf{Diagonalization} \qquad S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}. \qquad (1)$$

We call S the "eigenvector matrix" and Λ the "eigenvalue matrix"—using a capital lambda because of the small lambdas for the eigenvalues on its diagonal.

Proof Put the eigenvectors x_i in the columns of S, and compute AS by columns:

$$AS = A\begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \\ | & | & & | \end{bmatrix}.$$

Then the trick is to split this last matrix into a quite different product $S\Lambda$:

$$\begin{bmatrix} \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \end{bmatrix} = \underbrace{\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}}_{S} \underbrace{\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}}_{\Lambda}.$$

1. Show that $AS = S\Delta$ implies $S^{-1}AS = \Delta$ and $A = S\Delta S^{-1}$.

2. Recall our example from before

$$\frac{dx_1}{dt} = 4x_1 - 5x_2$$
$$\frac{dx_2}{dt} = 2x_1 - 3x_2$$
$$x_1(0) = 8, x_2(0) = 5$$

, where $A = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}$ with eigenvalues -1 and 2 and corresponding eigenvectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$. Show that $A = S\Delta S^{-1}$.

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

Singular Value Decomposition: Any m by n matrix A can be factored into

$$A = U\Sigma V^{\mathrm{T}} = \textbf{(orthogonal)(diagonal)(orthogonal)}.$$

The columns of U (m by m) are eigenvectors of AA^{T}, and the columns of V (n by n) are eigenvectors of $A^{\mathrm{T}}A$. The r singular values on the diagonal of Σ (m by n) are the square roots of the nonzero eigenvalues of both AA^{T} and $A^{\mathrm{T}}A$.

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

3. We'll decompose $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$ into this form using the following steps:

a. Find the eigenvalues of AA^T.

b. Find the eigenvalues of $A^T A$.

c. The only positive eigenvalues are 25 and 9. Stack the square roots of these eigenvalues together to get $\Sigma = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$. **Note that the size of this matrix will always have the number of rows as AA^T and the number of columns of $A^T A$.**

d. Find the normalized eigenvectors of AA^T. Stack these to form U: $U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$

e. Find the normalized eigenvectors of $A^T A$. Stack these to form V:

$$V = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \\ 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \end{bmatrix}$$

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

$$A = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & -2/3 & -1/3 \end{bmatrix}$$

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.

Suppose we know the SVD. The key is in the singular values (in Σ). Typically, some σ's are significant and others are extremely small. If we keep 20 and throw away 980, then we send only the corresponding 20 columns of U and V. The other 980 columns are multiplied in $U\Sigma V^{\mathrm{T}}$ by the small σ's that are being ignored. *We can do the matrix multiplication as columns times rows*:

$$A = U\Sigma V^{\mathrm{T}} = u_1\sigma_1 v_1^{\mathrm{T}} + u_2\sigma_2 v_2^{\mathrm{T}} + \cdots + u_r\sigma_r v_r^{\mathrm{T}}. \tag{3}$$

Any matrix is the sum of r matrices of rank 1. If only 20 terms are kept, we send 20 times 2000 numbers instead of a million (25 to 1 compression).

The pictures are really striking, as more and more singular values are included. At first you see nothing, and suddenly you recognize everything. The cost is in computing the SVD—this has become much more efficient, but it is expensive for a big matrix.

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 $A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$ is given by:

$$A = \begin{bmatrix} -1/3 & 2/\sqrt{5} & 2/\sqrt{5} \\ 2/3 & 0 & 1/\sqrt{5} \\ 2/3 & 1/\sqrt{5} & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} [1]$$

3. Show that the SVD decomposition of $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ is given by:

$$A = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$