Chapter 6: Eigenvalues and Eigenvectors

Chapter 6.1: Introduction to Eigenvalues

- Linear equations $Ax=b$ come from steady state problems. Eigenvalues have their greatest importance in **dynamic problems**. The solution of $\frac{du}{dt}=Au$ is changing with time — growing, decaying, or oscillating. We can't find it by elimination.
- This chapter enters a new part of linear algebra, based on $Ax=\lambda x$. All matrices in this chapter are square.
- A good model comes from the powers $A, A^2, A^3, \dots$ of a matrix. Suppose you need the hundredth power A^{100}

$$\underset{A}{\begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}} \quad \underset{A^2}{\begin{bmatrix} .70 & .45 \\ .30 & .55 \end{bmatrix}} \quad \underset{A^3}{\begin{bmatrix} .650 & .525 \\ .350 & .475 \end{bmatrix}} \quad \dots, \quad \underset{A^{100}}{\begin{bmatrix} .60000 & .60000 \\ .40000 & .4000 \end{bmatrix}}$$

  - A^{100} approaches $[.6, .6; .4, .4]$
  - A^{100} is found by using the eigenvalues of A, not by multiplying 100 matrices
- To explain eigenvalues, we must first explain eigenvalues. Almost all vectors change direction when multiplied by A. Certain exceptional vectors x remain in the same direction; that is $Ax=\lambda x$, where λ is a scalar.
  - The vector x is an "eigenvector"
  - The scalar λ (lambda) is an "eigenvalue". It tells how much a specific eigenvector is scaled in Ax. It can be any real number
- If $A=I$, then all vectors are eigenvectors of I, and all eigenvalues are 1
- We can find eigenvalues by using $\det(A-\lambda I)=0$ *
- Example:

$A=\begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}$ has eigenvalues $\lambda_1=1$, $\lambda_2=\frac{1}{2}$

$\det\left(\begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}-\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)=\det\left(\begin{bmatrix} -.2 & .3 \\ .2 & -.3 \end{bmatrix}\right)=0$

$\det\left(\begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}-\frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)=\det\left(\begin{bmatrix} .3 & .3 \\ .2 & .2 \end{bmatrix}\right)=0$

$x_1=\begin{bmatrix} .6 \\ .4 \end{bmatrix}$ and $Ax_1=\begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}\begin{bmatrix} .6 \\ .4 \end{bmatrix}=x_1 \quad (\lambda_1=1)$

$x_2=\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $Ax_2=\begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}=\begin{bmatrix} .5 \\ -.5 \end{bmatrix}=\frac{1}{2}x_2 \quad (\lambda_2=\frac{1}{2})$

Solving for λ:

"characteristic polynomial"

$\det\left(\begin{bmatrix} .8-\lambda & .3 \\ .2 & .7-\lambda \end{bmatrix}\right)=\lambda^2-\frac{3}{2}\lambda+\frac{1}{2}=(\lambda-1)(\lambda-\frac{1}{2})$

The eigenvectors x_1 and x_2 are in the nullspace of $A-I$ and $A-\frac{1}{2}I$

$(A-I)x_1=0 \Rightarrow Ax_1=x_1 \Rightarrow x_1=(.6, .4)$

$(A-\frac{1}{2}I)x_2=0 \Rightarrow Ax_2=\frac{1}{2}x_2 \Rightarrow x_2=(1,-1)$

*$\boxed{\begin{array}{l} Ax=\lambda x \Rightarrow Ax-\lambda x=0 \Rightarrow (A-\lambda I)x=0 \\ (A-\lambda I)x=0 \text{ has a nonzero sol. iff} \\ \det(A-\lambda I)=0 \end{array}}$

  - Apply A over and over on the eigenvectors do not change its direction. The eigenvalues are exponentiated. $AAx=A\lambda x=\lambda^2 x$, etc
- All other vectors are combinations of the eigenvectors's

- The first column of the example matrix A can be split into $x_1 + (.2)x_2$

$$\begin{bmatrix} .8 \\ .2 \end{bmatrix} = x_1 + (.2)x_2 = \begin{bmatrix} .6 \\ .4 \end{bmatrix} + \begin{bmatrix} .2 \\ -.2 \end{bmatrix}$$

Multiplying by A gives $(.7, .3)$, the first column of A^2. Do it separately for x_1 and $(.2)x_2$. A multiplies each eigenvector by its eigenvalue

$$A\begin{bmatrix} .8 \\ .2 \end{bmatrix} = A\left(\begin{bmatrix} .6 \\ .4 \end{bmatrix} + \begin{bmatrix} .2 \\ -.2 \end{bmatrix}\right) = A\begin{bmatrix} .6 \\ .4 \end{bmatrix} + (0.2)A\begin{bmatrix} 1 \\ -1 \end{bmatrix} = (1)\begin{bmatrix} .6 \\ .4 \end{bmatrix} + (0.2)(0.5)\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} .7 \\ .3 \end{bmatrix}$$

Then,

$$A^{99}\begin{bmatrix} .8 \\ .2 \end{bmatrix} \text{ is really } \overbrace{x_1}^{\text{multiply by } \lambda_1^{99} = 1^{99} = 1} + \underbrace{(.2)\left(\tfrac{1}{2}\right)^{99} x_2}_{\text{multiply by } \lambda_2^{99} = (\frac{1}{2})^{99}} = \begin{bmatrix} .6 \\ .4 \end{bmatrix} + \begin{bmatrix} \text{very small} \\ \text{vector} \end{bmatrix}$$

- The eigenvector x_1 is a "steady state" that doesn't change (b/c $\lambda_1 = 1$), while the eigenvector x_2 is a "decaying mode" that virtually disappears (b/c $\lambda_2 = .5$)

- This particular matrix A is a Markov matrix. Its entries are positive and each column sums to 1. These facts guarantee the largest eigenvalue as 1. Its eigenvector $x_1 = (.6, .4)$ is the steady state that all columns of A^k will approach.
- For projections we can spot the steady state ($\lambda = 1$) and the nullspace ($\lambda = 0$)
- Example:

The projection matrix $P = \begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix}$ has eigenvalues $\lambda = 1$ and $\lambda = 0$

Its eigenvectors are $x_1 = (1, 1)$ and $x_2 = (1, -1)$. For those vectors, $Px_1 = x_1$ (steady state) and $Px_2 = 0$ (nullspace)

1. Each column of P sums to 1, so $\lambda = 1$ is an eigenvalue
2. P is singular so $\lambda = 0$ is an eigenvalue
3. P is symmetric, so its eigenvectors $(1,1)$ and $(1,-1)$ are perpendicular

The eigenvectors for $\lambda = 0$ (which means $Px = 0x$) fill up the nullspace. The eigenvectors for $\lambda = 1$ (which means $Px = x$) fill up the column space

The nullspace is projected to 0. The column space projects to itself. The projection keeps the column space and destroys the nullspace

$$v = \overbrace{\begin{bmatrix} 1 \\ -1 \end{bmatrix}}^{x_2} + \overbrace{\begin{bmatrix} 2 \\ 2 \end{bmatrix}}^{2x_1} \Rightarrow Pv = \overbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}^{\text{multiply by } \lambda_1 = 0} + \overbrace{\begin{bmatrix} 2 \\ 2 \end{bmatrix}}^{\text{multiply by } \lambda_2 = 1}$$

- Permutation matrices have all $|\lambda| = 1$
- Example:

$R = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -1$ and eigenvectors $x_1 = (1,1)$ and $x_2 = (1,-1)$

- The eigenvalues for R are the same as P because reflection = 2(projection) − I

$$R = 2P - I \Rightarrow \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = 2\begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

- If $Px = \lambda x$, then $2Px = 2\lambda x$. The eigenvalues are doubled when the matrix is doubled. Subtracting $Ix = x$, we get $(2P-I)x = 2Px - Ix = 2\lambda x - x = (2\lambda - 1)x$
  When a matrix is shifted by I, each λ is shifted by 1.

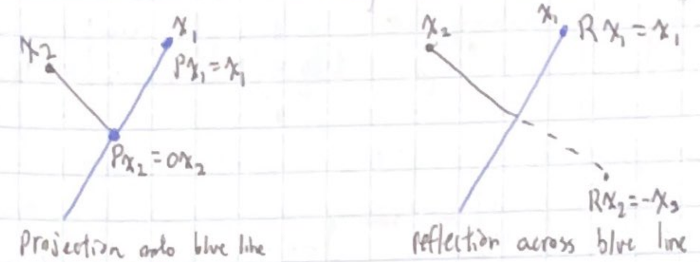

fig 6.1.1
Projections P have eigenvalues 0 and 1
Reflections R have eigenvalues −1 and 1

- Key idea: The eigenvalues of R and P are related exactly as the matrices are related.
  - The eigenvalues of $R = 2P - I$ are $2(1) - 1 = 1$ and $2(0) - 1 = -1$
  - The eigenvalues of R^2 are λ^2. In this case $R^2 = I$. Check $(1)^2 = 1$ and $(-1)^2 = 1$

The Equation for the Eigenvalues

- Start with $Ax = \lambda x$ and move λx to the left side

$$Ax - \lambda x = 0 \Rightarrow (A - \lambda I)x = 0$$

- The eigenvectors x are in the nullspace of $A - \lambda I$. When we know an eigenvalue λ, we find an eigenvector by solving $(A - \lambda I)x = 0$.
- We want a nonzero solution to $(A - \lambda I)x = 0$ (0 is not considered an eigenvector), which occurs if and only if $\det(A - \lambda I) = 0$ (i.e. the matrix is singular)
- Using the big formula for determinants, we get a "characteristic polynomial" in λ, whose roots give us the eigenvalues. Then,

  For each eigenvalue λ, solve $(A - \lambda I)x = 0$ or $Ax = \lambda x$ to find an eigenvector x

- Example:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \Rightarrow A - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix} \Rightarrow \det(A - \lambda I) = (1-\lambda)(4-\lambda) - (2)(2) = \lambda^2 - 5\lambda$$

$$\lambda^2 - 5\lambda = 0 \Rightarrow \lambda(\lambda - 5) = 0 \Rightarrow \lambda = 0, 5$$

  - Note: If A is already singular, the $\lambda = 0$ is an eigenvalue

The matrices $A - 0I$ and $A - 5I$ are singular

$$(A - 0I)x = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ yields an eigenvector } \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \text{ for } \lambda_1 = 0$$

$$(A - 5I)x = \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ yields an eigenvector } \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ for } \lambda_2 = 5$$

The eigenvectors $(2,-1)$ and $(1,2)$ are in the nullspace of $(A - \lambda I)x = 0$

- Summary:
    1. Compute $\det(A-\lambda I)$, which is a polynomial in λ of degree n
    2. Find the roots of this "characteristic polynomial" $\det(A-\lambda I)=0$. The n roots are the n eigenvalues of A. They make $A-\lambda I$ singular.
    3. For each eigenvalue λ, solve $(A-\lambda I)x=0$ to find an eigenvector x
- A note on the eigenvectors of 2 by 2 matrices. When $A-\lambda I$ is singular, both rows are multiples of a vector (a,b). The eigenvector is any multiple of $(b,-a)$.
    - $\lambda=0$: rows of $A-0I$ in the direction $(1,2)$; eigenvector in the direction of $(2,-1)$
    - $\lambda=5$: rows of $A-5I$ in the direction $(-4,2)$; eigenvector in the direction of $(2,4)$
- Note that if x is an eigenvector, so is cx, where c is a constant. There is a whole line of eigenvectors in the direction of x.
- Some 2 by 2 matrices only have one line of eigenvectors, which happens only if two eigenvalues are equal.
- n by n matrices don't always have n independent eigenvectors, which doesn't give us a basis. We can't write every v as a combination of eigenvectors.

Good News, Bad News

- Elimination does not preserve eigenvalues.

  $U=\begin{bmatrix}1&3\\0&0\end{bmatrix}$ has $\lambda=0,1$ $\qquad$ $A=\begin{bmatrix}1&3\\2&6\end{bmatrix}$ has $\lambda=0,7$
    - Row exchanges and adding rows usually changes eigenvalues
- The products and sums of eigenvalues:
    - The product of the n eigenvalues equals the determinant of the matrix
    - The sum of the n eigenvalues equal the sum of the n diagonal entries.
        - The sum of the entries on the main diagonal is called the trace of A.
- The above are useful as checks but not for computation
- The determinant test makes the product of the λ's equal to the product of the pivots (assuming no row exchanges). But the sum of the λ's is not the sum of the pivots. The individual λ's have almost nothing to do with the pivots. In this new part of linear algebra, the key equation is really nonlinear: λ multiplies x.

Imaginary Eigenvalues

Eigenvalues are generally complex numbers.

- Example

  $Q=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$ has no real eigenvectors. $(-\lambda)(-\lambda)-(-1)(1)=0 \Rightarrow \lambda^2+1=0$
- Its eigenvalues are $\lambda=i$ and $\lambda=-i$, which matches with the sum and product checks for eigenvalues.
- Q is a $90°$ rotation matrix. No vector Qx keeps its direction after a rotation besides 0. We must go into the complex numbers.
- If we look at Q^2, which is $-I$. Its eigenvalues are -1 and -1. These values are the squares of the eigenvalues of Q, and i^2 and $(-i)^2=-1$ and -1.

- Somehow, these complex vectors $x_1 = (1, i)$ and $x_2 = (i, 1)$ keep their directions upon rotation.
- Q is an orthonormal matrix, so the absolute value of each λ is $|\lambda| = 1$
- Q is a skew-symmetric matrix so each λ is pure imaginary.
- The symmetric matrix ($A^T = A$) can be compared to a real number and a skew-symmetric matrix ($A^T = -A$) can be compared to an imaginary number.
  An orthogonal matrix ($A^TA = I$) can be compared to a complex number with $|\lambda| = 1$

Problem set 6.1

1

a)

$$A = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}$$

$$A - \lambda I = \begin{bmatrix} .8-\lambda & .3 \\ .2 & .7-\lambda \end{bmatrix}$$

$$\det(A - \lambda I) = \left(\tfrac{4}{5} - \lambda\right)\left(\tfrac{7}{10} - \lambda\right) - \left(\tfrac{1}{5}\right)\left(\tfrac{3}{10}\right)$$

$$= \tfrac{14}{25} - \tfrac{3}{2}\lambda + \lambda^2 - \tfrac{3}{50}$$

$$= \lambda^2 - \tfrac{3}{2}\lambda + \tfrac{1}{2}$$

Set $\det(A - \lambda I) = 0$

$\lambda^2 - \frac{3}{2}\lambda + \frac{1}{2} = 0$

$2\lambda^2 - 3\lambda + 1 = 0$

$(\lambda - 1)(2\lambda - 1) = 0$

$\lambda - 1 = 0$ or $2\lambda - 1 = 0$ (zpp)

$\lambda = 1$ or $\lambda = \frac{1}{2}$

$$A' = \begin{bmatrix} .2 & .7 \\ .8 & .3 \end{bmatrix}$$

$$A' - \lambda I = \begin{bmatrix} .2-\lambda & .7 \\ .8 & .3-\lambda \end{bmatrix}$$

$$\det(A' - \lambda I) = \left(\tfrac{1}{5} - \lambda\right)\left(\tfrac{3}{10} - \lambda\right) - \left(\tfrac{4}{5}\right)\left(\tfrac{7}{10}\right)$$

$$= \tfrac{3}{50} - \tfrac{1}{2}\lambda + \lambda^2 - \tfrac{14}{25}$$

$$= \lambda^2 - \tfrac{1}{2}\lambda - \tfrac{1}{2}$$

Set $\det(A - \lambda I) = 0$

$\lambda^2 - \frac{1}{2}\lambda - \frac{1}{2} = 0$

$2\lambda^2 - \lambda - 1 = 0$

$(\lambda - 1)(2\lambda + 1) = 0$

$\lambda - 1 = 0$ or $2\lambda + 1 = 0$ (zpp)

$\lambda = 1$ or $\lambda = -\frac{1}{2}$

b)

The $\lambda = 0$ eigenvalue leads to the equation

$(A - 0I)x = 0$

$Ax = 0$

elimination does not change the nullspace of A

2.

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \quad A + I = \begin{bmatrix} 2 & 4 \\ 2 & 4 \end{bmatrix} = A'$$

$$A - \lambda I = \begin{bmatrix} 1-\lambda & 4 \\ 2 & 3-\lambda \end{bmatrix}$$

$$\det(A - \lambda I) = (1-\lambda)(3-\lambda) - (4)(2)$$

$$= 3 - 4\lambda + \lambda^2 - 8$$

$$= \lambda^2 - 4\lambda - 5$$

Set $\det(A - \lambda I) = 0$

$\lambda^2 - 4\lambda - 5 = 0$

$(\lambda - 5)(\lambda + 1) = 0$

$\lambda - 5 = 0$ or $\lambda + 1 = 0$ (zpp)

$\lambda = 5$ or $\lambda = -1$

Case $\lambda = 5$

$(A - 5I)x_1 = 0$

$$\begin{bmatrix} -4 & 4 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$x_1 = (1, 1)$

Case $\lambda = 1$:

$(A + I)x_2 = 0$

$$\begin{bmatrix} 2 & 4 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$x_2 = (2, -1)$

Same, increased

$$A' - \lambda I = \begin{bmatrix} 2-\lambda & 4 \\ 2 & 4-\lambda \end{bmatrix}$$

$$\det(A' - \lambda I) = (2-\lambda)(4-\lambda) - (4)(2)$$

$$= 8 - 6\lambda + \lambda^2 - 8$$

$$= \lambda^2 - 6\lambda$$

Set $\det(A' - \lambda I) = 0$

$\lambda^2 - 6\lambda = 0$

$\lambda(\lambda - 6) = 0$

$\lambda = 0, 6$

$(A' - 0I)x_1 = 0$

$x_1 = (2, -1)$

$(A' - 6I)x_2 = 0$

$$\begin{bmatrix} -4 & 4 \\ 2 & -2 \end{bmatrix} x_2 = 0$$

$x_2 = (1, 1)$

3.

$A = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}$

$A - \lambda I = \begin{bmatrix} -\lambda & 2 \\ 1 & 1-\lambda \end{bmatrix}$

$\det(A - \lambda I) = (-\lambda)(1-\lambda) - (2)(1)$
$= \lambda^2 - \lambda - 2$
$= (\lambda - 2)(\lambda + 1)$

Set $\det(A - \lambda I) = 0$
$(\lambda - 2)(\lambda + 1) = 0$
$\lambda - 2 = 0$ or $\lambda + 1 = 0$ (zpp)
$\lambda = 2$ or $\lambda = -1$

Case $\lambda = 2$
$(A - 2I)x_1 = 0$
$\begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
$\begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
$x_1 = (1, 1)$

Case $\lambda = -1$
$(A + I)x_2 = 0$
$\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = 0$
$x_2 = (2, -1)$

$A^{-1} = \begin{bmatrix} -\frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{bmatrix}$

$A^{-1} - \lambda I = \begin{bmatrix} -\frac{1}{2} - \lambda & 1 \\ \frac{1}{2} & -\lambda \end{bmatrix}$

$\det(A^{-1} - \lambda I) = (-\frac{1}{2} - \lambda)(-\lambda) - (\frac{1}{2})(1)$
$= \lambda^2 + \frac{1}{2}\lambda - \frac{1}{2}$

set $\det(A^{-1} - \lambda I) = 0$
$0 = 2\lambda^2 + \lambda - 1$
$0 = (\lambda + 1)(2\lambda - 1)$
$\lambda = -1, \frac{1}{2}$
$\frac{1}{\lambda_1}, \frac{1}{\lambda_2}$

7.
determinant, 1, 0

26.
$Ax = \begin{bmatrix} Bx & Cy \\ 0 & Dy \end{bmatrix} =$

27.
A.
rank = 1

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

$A - \lambda I = \begin{bmatrix} 1-\lambda & 1 & 1 & 1 \\ 1 & 1-\lambda & 1 & 1 \\ 1 & 1 & 1-\lambda & 1 \\ 1 & 1 & 1 & 1-\lambda \end{bmatrix}$

9.
a) $AAx = \lambda Ax = \lambda\lambda x$
b) $Ax = \lambda x$
$x = \lambda A^{-1}x$
$A^{-1}x = \frac{1}{\lambda}x$
c) $(A + I)x = Ax + Ix$
$= \lambda x + x$
$= (\lambda + 1)x$

16.
$\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\dots(\lambda_n - \lambda)$
$\det(A - 0I) = \lambda_1\lambda_2\lambda_3 \dots \lambda_n$

17.
$\det(A - \lambda I) = \lambda^2 - (a+d)\lambda + ad - bc = 0$

$\lambda = \dfrac{(a+d) \pm \sqrt{(a+d)^2 - 4(1)(ad - bc)}}{2(1)}$

$\lambda_{sum} = (a+d) = \text{trace}$

Chapter 6.2: Diagonalizing a Matrix

- When x is an eigenvector, multiplication by A is just multiplication by a number λ: $Ax = \lambda x$. We can follow the eigenvectors separately instead of handling an interconnected system.
- It's like having a diagonal matrix, with no off-diagonal interconnections. The 100th power of a diagonal matrix is easy.
- The matrix A turns into a diagonal matrix Λ (uppercase lambda) when we use the eigenvectors properly
- Diagonalization: Suppose the n by n matrix A has n linearly independent eigenvectors $x_1, \dots, x_n$. Put them into the columns of an eigenvector matrix S. Then $S^{-1}AS$ is the eigenvalue matrix Λ:

Matrix A is "diagonalized"

$$S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

- Proof: Multiply A by its eigenvectors, which are the columns of S. The first column of AS is Ax_1, which is $\lambda_1 x_1$. Each column of S is multiplied by its eigenvalue λ_i:

$$AS = A\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} = \begin{bmatrix} Ax_1 & \cdots & Ax_n \end{bmatrix} = \begin{bmatrix} \lambda_1 x_1 & \cdots & \lambda_n x_n \end{bmatrix}$$

The trick is to split this matrix AS into S times Λ.

$$S\Lambda = \begin{bmatrix} \lambda_1 x_1 & \cdots & \lambda_n x_n \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Then:

$AS = S\Lambda$ or $\Lambda = S^{-1}AS$ or $\underbrace{A = S\Lambda S^{-1}}$ ← this is called the eigendecomposition of A

S^{-1} exists because we assume S contains n independent eigenvectors as columns.
Without n independent eigenvectors, we can't diagonalize.

- A and Λ have the same eigenvalues $\lambda_1, \dots, \lambda_n$. The eigenvectors are different.
The job of the original eigenvectors $x_1, \dots, x_n$ was to diagonalize A. Those eigenvectors in S produce $A = S\Lambda S^{-1}$.
- Example

Eigenvectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\quad \underbrace{\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}}_{S^{-1}} \underbrace{\begin{bmatrix} 1 & 5 \\ 0 & 6 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{S} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}}_{\Lambda}$

$A = S\Lambda S^{-1}$ so $A^2 = S\Lambda S^{-1} S\Lambda S^{-1} = S\Lambda^2 S^{-1}$

In general

$A^n = S\Lambda^n S^{-1}$ (Same eigenvectors in S) (exponentiated eigenvalues in Λ)

For this A,

$$A^k = \begin{bmatrix} 1 & 5 \\ 0 & 6 \end{bmatrix}^k = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1^k & \\ & 6^k \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 6^k - 1 \\ 0 & 6^k \end{bmatrix}$$

$K = 1 \Rightarrow \begin{bmatrix} 1 & 5 \\ 0 & 6 \end{bmatrix}$

$K = 0 \Rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$K = -1 \Rightarrow \begin{bmatrix} 1 & -\frac{5}{6} \\ 0 & \frac{1}{6} \end{bmatrix} = A^{-1}$

- Notes:
  - If all eigenvalues are unique, then it is automatic that the eigenvectors are independent.
  Any matrix with unique eigenvalues is diagonalizable
  - Eigenvectors can be scaled arbitrarily without changing their eigenvector status.
  - The eigenvalues in Λ come in the same order as the eigenvector in S.
  - Matrices with too few eigenvectors cannot be diagonalized.
  - There is no connection between diagonalizability and invertibility.
    - Invertibility is concerned with the eigenvalues ($\lambda = 0$ or $\lambda \neq 0$)
    - Diagonalizability is concerned with the eigenvectors (too few or too many for S)
- Eigenvectors $x_1, \dots, x_n$ that correspond to distinct eigenvalues are linearly independent.
  - Proof: Suppose $c_1 x_1 + c_2 x_2 = 0$. Multiply by A to get $c_1\lambda_1 x_1 + c_2 \lambda_2 x_2 = 0$. Multiply the original equation by λ_2 to get $c_1 \lambda_2 x_1 + c_2 \lambda_2 x_2 = 0$. Subtract them,
  we get $(\lambda_1 - \lambda_2) c_1 x_1 = 0 \Rightarrow c_1 = 0$ since λ are all different and c_2 also equals 0.
  No other combination gives $c_1 x_1 + c_2 x_2 = 0$ so x_1 and x_2 must be independent.
  This proof extends to j eigenvectors.

- Suppose $c_1 x_1 + \ldots + c_j x_j = 0$. Multiply by A and λ_j separately and subtract. This removes x_j. Repeat until $(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_j) c_1 x_1 = 0 \Rightarrow c_1 = 0$.
- Example:

$$A = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & .5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ .4 & -.6 \end{bmatrix} = S \Lambda S^{-1}$$

$$A^k = S \Lambda^k S^{-1} = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix} \begin{bmatrix} 1^k & 0 \\ 0 & .5^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ .4 & -.6 \end{bmatrix}$$

$$\lim_{k \to \infty} A^k = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ .4 & -.6 \end{bmatrix} = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} .6 & .6 \\ .4 & .4 \end{bmatrix}$$

- $\lim_{k \to \infty} A^k = 0$ when all $|\lambda| < 1$

Fibonacci Numbers

- Challenge: Find F_{100}, the 100th Fibonacci number.
- Start with the equation $u_{k+1} = A u_k$

  Let $u_k = (F_{k+1}, F_k)$. The rule $\begin{cases} F_{k+2} = F_{k+1} + F_k \\ F_{k+1} = F_{k+1} \end{cases}$ is $u_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} u_k$
- Each step multiplies by $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, $u_{100} = A^{100} u_0$

$$u_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad u_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad u_3 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \quad \ldots \quad u_{100} = \begin{bmatrix} F_{101} \\ F_{100} \end{bmatrix}$$

$$A - \lambda I = \begin{bmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{bmatrix} \Rightarrow \det(A - \lambda I) = (1-\lambda)(-\lambda) - (1)(1) = \lambda^2 - \lambda - 1$$

Using the quadratic formula, $\lambda = \frac{1+\sqrt{5}}{2}$, $\lambda = \frac{1-\sqrt{5}}{2}$

- Then our eigenvectors are $x_1 = (\lambda_1, 1)$ and $x_2 = (\lambda_2, 1)$.
- Then we find the combination of eigenvectors which produce $u_0 = (1, 0)$

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\lambda_1 - \lambda_2} \left(\begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} - \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix} \right) \text{ or } u_0 = \frac{x_1 - x_2}{\lambda_1 - \lambda_2}$$

Then

$$u_{100} = A^{100} u_0 = \frac{A^{100} x_1 - A^{100} x_2}{\lambda_1 - \lambda_2} = \frac{(\lambda_1)^{100} x_1 - (\lambda_2)^{100} x_2}{\lambda_1 - \lambda_2}$$

We want F_{100} the second component of u_{100}. The second components of x_1 and x_2 are 1. Then

$$F_{100} = \frac{1}{\frac{1+\sqrt{5}}{2} - \frac{1-\sqrt{5}}{2}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^{100} - \left(\frac{1-\sqrt{5}}{2} \right)^{100} \right] = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^{100} - \left(\frac{1-\sqrt{5}}{2} \right)^{100} \right]$$

Matrix Powers of A^k

- Fibonacci's example is a typical difference equation $u_{k+1} = Au_k$. Each iterative step multiplies by A. The solution is $u_k = A^k u_0$. Diagonalizing matrices allow computation as $S\Lambda^k S^{-1} u_0 = u_k$
- Procedure
  1. Write u_0 as a combination $c_1 x_1 + \dots + c_n x_n$ of the eigenvectors. Then $c = S^{-1} u_0$
  2. Multiply each eigenvector x_i by $(\lambda_i)^k$. Now we have $\Lambda^k S^{-1} u_0$
  3. Add up all the pieces $c_i (\lambda_i)^k x_i$ to find the solution $u_k = A^k u_0$. This is $S\Lambda^k S^{-1} u_0$

  $u_{k+1} = Au_k \Rightarrow u_k = A^k u_0 = c_1(\lambda_1)^k x_1 + \dots + c_n(\lambda_n)^k x_n$
- Example 1

  Start from $u_0 = (1,0)$ with

  $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ has $\lambda_1 = 2$ and $x_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\lambda_2 = -1$ and $x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

  $u_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ so $c_1 = c_2 = \frac{1}{3}$ (1)

  Multiply the two parts by $(\lambda_1)^k = 2^k$ and $(\lambda_2)^k = (-1)^k$

  $\frac{1}{3}(2)^k \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{1}{3}(-1)^k \begin{bmatrix} 1 \\ -1 \end{bmatrix} = u_k$

Nondiagonalizable Matrices

- Eigenvalues may be repeated and we want to know the multiplicity. There's two different ways to count these: ($GM \leq AM \; \forall \lambda$)
  1. Geometric multiplicity = GM
     - Count the independent eigenvectors for λ. This is the dimension of $N(A - \lambda I)$
  2. Algebraic multiplicity = AM
     - Count the repetitions of λ in the roots of $\det(A - \lambda I) = 0$
- If A has $\lambda = 4, 4, 4$, then it has $AM = 3$, $GM = 1, 2,$ or 3, depending on whether each eigenvalue corresponds to an independent eigenvector or not.
- For

  $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2 \Rightarrow \lambda = 0, 0$ but only 1 eigenvector $\begin{pmatrix} AM = 2 \\ GM = 1 \end{pmatrix}$

  $A = \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}$ has $\det(A - \lambda I) = (\lambda - 5)^2 \Rightarrow \lambda = 5, 5$ but $A - 5I$ has only rank 1 $\begin{pmatrix} AM = 2 \\ GM = 1 \end{pmatrix}$

Eigenvalues of AB and $A+B$

- The following is false because we are assume A and B share the same eigenvalue x

  $ABx = ABx = BAx = B\lambda x$

- Eigenvalues do not generally add either for $A+B$
- When all n eigenvectors are shared we can multiply eigenvalues
- Suppose both A and B are diagonalizable. They share S iff $AB = BA$

Problem Set 6.2

1.

a)

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

$A - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 0 & 3-\lambda \end{bmatrix}$

$\det(A-\lambda I) = (1-\lambda)(3-\lambda) - (2)(0)$

$= (\lambda-3)(\lambda-1)$

Set $\det(A-\lambda I) = 0$

$(\lambda-3)(\lambda-1) = 0$

$\lambda = 1, 3$

case $\lambda = 1$: $(A-I)x = 0$ $\begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix} x_1 = 0$ $x_1 = (1, 0)$

case $\lambda = 3$: $(A-3I)x = 0$ $\begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix} x_2 = 0$ $x_2 = (1, 1)$

$S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $S^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$

$A = S \Lambda S^{-1}$

$= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$

$A = \begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix}$

$A - \lambda I = \begin{bmatrix} 1-\lambda & 1 \\ 3 & 3-\lambda \end{bmatrix}$

$\det(A-\lambda I) = (1-\lambda)(3-\lambda) - (1)(3)$

$= \lambda^2 - 4\lambda + 3 - 3$

$= \lambda^2 - 4\lambda$

$= \lambda(\lambda-4)$

Set $\det(A-\lambda I) = 0$

$\lambda(\lambda-4) = 0$

$\lambda = 0, 4$

Case $\lambda = 0$: $(A - 0I)x_1 = 0$ $\begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix} x_1 = 0$ $x_1 = (1, -1)$

Case $\lambda = 4$: $(A-4I)x_2 = 0$ $\begin{bmatrix} -3 & 1 \\ 3 & -1 \end{bmatrix} x_2 = 0$ $x_2 = (1, 3)$

$S = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}$ $S^{-1} = \begin{bmatrix} \frac{3}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix}$

$A = S\Lambda S^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} \frac{3}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix}$

b)

$A^3 = S\Lambda^3 S^{-1}$

$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 27 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$

$= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 27 \end{bmatrix}$

$= \begin{bmatrix} 1 & 26 \\ 0 & 27 \end{bmatrix}$

$A^{-1} = S\Lambda^{-1}S^{-1}$

$= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$

$= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & \frac{1}{3} \end{bmatrix}$

$= \begin{bmatrix} 1 & -\frac{2}{3} \\ 0 & \frac{1}{3} \end{bmatrix}$

2.

$A = S\Lambda S^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$

$= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & -2 \\ 0 & 5 \end{bmatrix}$

$= \begin{bmatrix} 2 & 3 \\ 0 & 5 \end{bmatrix}$

4.

a) Maybe

b) True

c) True

d) Maybe

8.

$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\begin{bmatrix} \frac{\sqrt{5}}{5} & -\frac{\lambda_2}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{\lambda_1}{\sqrt{5}} \end{bmatrix}$

15.

$|\lambda| < 1$

$A_1 - \lambda I = \begin{bmatrix} \frac{3}{5}-\lambda & \frac{9}{10} \\ \frac{2}{5} & \frac{1}{10}-\lambda \end{bmatrix}$

$\det(A_1 - \lambda I) = (\frac{3}{5}-\lambda)(\frac{1}{10}-\lambda) - (\frac{9}{10})(\frac{2}{5})$

$= \lambda^2 - \frac{7}{10}\lambda - \frac{3}{10} = 0$

$10\lambda^2 - 7\lambda - 3 = 0$

$(\lambda-1)(10\lambda+3) = 0 \implies \lambda = 1, -\frac{3}{10}$

11.

a) F b) F c) T $\rightarrow \begin{bmatrix} 3 & -3 \\ 3 & -3 \end{bmatrix} x_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

12.

a) F b) T c) T $x_1 = (1, 1)$

13.

$A = \begin{bmatrix} 8 & -3 \\ 3 & 2 \end{bmatrix}$

$Ax = 5x$

$(A - 5I)x$

14.

$A - 3I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ $\begin{bmatrix} 3 & 1 \\ 0 & 4 \end{bmatrix} \implies A - 3I = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ A_2

$r = 0$

16.

$A_1 = \begin{bmatrix} .6 & .9 \\ .4 & .1 \end{bmatrix}$

$\lambda_1 = 1, \lambda_2 = -\frac{3}{10}$

$(A - I)x_1 = 0$ | $(A + \frac{3}{10}I)x_2 = 0$

$\begin{bmatrix} -.4 & .9 \\ .4 & -.9 \end{bmatrix} x_1 = 0$ $\begin{bmatrix} .9 & .9 \\ .4 & .4 \end{bmatrix} x_2 = 0$

$x_1 = (.9, .4)$ $x_2 = (1, -1)$

$S = \begin{bmatrix} .9 & 1 \\ .4 & -1 \end{bmatrix}$, $S^{-1} = \begin{bmatrix} \frac{10}{13} & \frac{10}{13} \\ \frac{4}{13} & -\frac{9}{13} \end{bmatrix}$

$\lim_{k\to\infty} \Lambda^k = \begin{bmatrix} 1^k & 0 \\ 0 & (-\frac{3}{10})^k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

$\lim_{k\to\infty} A^k = \lim_{k\to\infty} S\Lambda^k S^{-1}$

$= \begin{bmatrix} .9 & 1 \\ .4 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{10}{13} & \frac{10}{13} \\ \frac{4}{13} & -\frac{9}{13} \end{bmatrix}$

$= \begin{bmatrix} .9 & 1 \\ .4 & -1 \end{bmatrix}\begin{bmatrix} \frac{10}{13} & \frac{10}{13} \\ 0 & 0 \end{bmatrix}$

$= \begin{bmatrix} \frac{9}{13} & \frac{9}{13} \\ \frac{4}{13} & \frac{4}{13} \end{bmatrix}$

17.

$A_2 = \begin{bmatrix} .6 & .9 \\ .1 & .6 \end{bmatrix}$

$A - \lambda I = \begin{bmatrix} .6-\lambda & .9 \\ .1 & .6-\lambda \end{bmatrix}$

$\det(A - \lambda I) = (.6-\lambda)(.6-\lambda) - (.1)(.9)$
$= \lambda^2 - \frac{6}{5}\lambda + \frac{9}{25} - \frac{9}{100}$
$= \lambda^2 - \frac{6}{5}\lambda + \frac{27}{100}$

Set $\det(A - \lambda I) = 0$

$\lambda_1 = \frac{9}{10}, \lambda_2 = \frac{3}{10}$

Case $\lambda_1 = \frac{9}{10}$ | Case $\lambda_2 = \frac{3}{10}$

$(A - \frac{9}{10}I)x_1 = 0$ | $(A - \frac{3}{10}I)x_2 = 0$

$\begin{bmatrix} -.3 & .9 \\ .1 & -.3 \end{bmatrix} x_1 = 0$ | $\begin{bmatrix} .3 & .9 \\ .1 & .3 \end{bmatrix} x_2 = 0$

$x_1 = (3, 1)$ | $x_2 = (-3, 1)$

$S = \begin{bmatrix} 3 & -3 \\ 1 & 1 \end{bmatrix}$ $S^{-1} = \begin{bmatrix} \frac{1}{6} & \frac{1}{2} \\ -\frac{1}{6} & \frac{1}{2} \end{bmatrix}$ $\Lambda = \begin{bmatrix} \frac{9}{10} & \\ & \frac{3}{10} \end{bmatrix}$

a)

$u_0 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1x_2 + 0x_2 = x_1$

$(A_2)^{10} u_0 = (A_2)^{10}(x_1) = (\lambda_1)^{10}(x_1) = \left(\frac{9}{10}\right)^{10} x_1 = \begin{bmatrix} 3(\frac{9}{10})^{10} \\ (\frac{9}{10})^{10} \end{bmatrix}$

b)

$u_0 = \begin{bmatrix} 3 \\ -1 \end{bmatrix} = -x_2$

$(A_2)^{10} u_0 = (\lambda_2)^{10}(-x_2) = (\frac{3}{10})^{10}(-x_2) = \begin{bmatrix} -3(\frac{3}{10})^{10} \\ (\frac{3}{10})^{10} \end{bmatrix}$

c)

$u_0 = \begin{bmatrix} 6 \\ 0 \end{bmatrix} = x_1 + x_2$

$A^{10} u_0 = A^{10}(x_1 + x_2) = A^{10}x_1 + A^{10}x_2 = \lambda_1^{10}x_1 + \lambda_2^{10}x_2$
$= (\frac{9}{10})^{10}x_1 + (\frac{3}{10})^{10}x_2$
$= \begin{bmatrix} 3(\frac{9}{10})^{10} + 3(\frac{3}{10})^{10} \\ (\frac{9}{10})^{10} - (\frac{3}{10})^{10} \end{bmatrix}$

20.

$\det A = (\det S)(\det \Lambda)(\det S^{-1})$
$= (\det S)(\prod \lambda_i)(\frac{1}{\det S})$
$= \prod \lambda_i$

25.

$N(A - I)$ or $C(A), N(A)$

26.

$\lambda \neq 0$ / ~~doesn't occupy the full column space~~

$C(A) \neq N(A)$ ∴ overlap

6.3: Applications to Differential Equations

- Eigen -values, -vectors, and -decompositions are also perfect for differential equations $du/dt = Au$. Our most important derivative will be $\frac{d}{dt}(e^{\lambda t}) = \lambda e^{\lambda t}$
- $\frac{du}{dt} = \lambda u$ is solved by $u = Ce^{\lambda t}$. We can solve for C given an initial condition
- Now we want to solve n equations

  $\frac{du}{dt} = Au$ with initial condition $u(0)$ at $t=0$

- These differential equations are linear. If $u(t)$ and $v(t)$ are solutions, so is $Cu(t) + Dv(t)$
- We will need n constants like C and D to match the n components of $u(0)$
- Note that A is a constant matrix. We usually have A change as t changes (linear) or A change when u changes (nonlinear). Here we have $\frac{du}{dt} = Au$ as "linear with constant coefficients"
- Our main point will be:

<u>Solve linear constant coefficient equations by exponentials $e^{\lambda t}x$ when $Ax = \lambda x$.</u>

Solution of $du/dt = Au$

- Our pure exponential solution will be $e^{\lambda t}$ times a fixed vector x. λ is an eigenvalue of A and x is the eigenvector.

$$\frac{d}{dt}(e^{\lambda t}x) = A(e^{\lambda t}x) \Rightarrow \lambda e^{\lambda t}x = \lambda e^{\lambda t}x$$

- All components of this solution $u = e^{\lambda t}x$ share the same $e^{\lambda t}$ part. The solution grows when $\lambda > 0$, decays if $\lambda < 0$. If λ is complex, its real parts decides growth or decay, and its imaginary part gives oscillation.
- Example

Solve $\frac{du}{dt} = Au = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ with initial condition $u(0) = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$

$\frac{du}{dt} = Au \quad \frac{d}{dt}\begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} y \\ z \end{bmatrix}$ means that $\frac{dy}{dt} = z$ and $\frac{dz}{dt} = y$

The idea of eigenvectors is to combine those equations in a way that gets it to a 1 equation ("1 by 1") problem

The matrix A has eigenvalues 1 and -1 and eigenvectors $(1,1), (1,-1)$. The pure exponential solutions takes the form

$$u_1(t) = e^{\lambda_1 t}x_1 = e^{t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad u_2(t) = e^{\lambda_2 t}x_2 = e^{-t}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

- Note: These u's are eigenvectors. They satisfy $Au_1 = u_1$ and $Au_2 = -u_2$ just like x_1 and x_2. The factors e^{t} and e^{-t} change with time. These factors give $du_1/dt = u_1 = Au_1$ and $du_2/dt = -u_2 = Au_2$. We now have 2 solutions. To get all solutions, take all linear combinations of those 2

$$u(t) = Ce^{t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + De^{-t}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} Ce^{t} + De^{-t} \\ Ce^{t} - De^{-t} \end{bmatrix}$$

$u(0)$ gives C, D

- Procedure, summarized:

1. Write $u(0)$ as a combination $C_1x_1 + \dots + C_nx_n$ of the eigenvectors of A
2. Multiply each eigenvector x_i by $e^{\lambda_i t}$.
3. The solution is the combination of pure solutions $e^{\lambda t}x$.

$$u(t) = C_1e^{\lambda_1 t}x_1 + \dots + C_ne^{\lambda_n t}x_n$$

- If the two λ's are equal, with only one eigenvector, another solution is needed (It will be $te^{\lambda t}x$). Step 1 needs $A = S\Lambda S^{-1}$ to be diagonalizable
- Example

Solve $du/dt = Au$ knowing the eigenvalues $\lambda = 1, 2, 3$ of A

$$\frac{du}{dt} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix} u \text{ with } u(0) = \begin{bmatrix} 9 \\ 7 \\ 4 \end{bmatrix}$$

The eigenvectors are $x_1 = (1,0,0)$, $x_2 = (1,1,0)$, and $x_3 = (1,1,1)$

$$u(0) = 2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 4\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 2x_1 + 3x_2 + 4x_3$$. Thus $(c_1, c_2, c_3) = (2, 3, 4)$

Then,

The pure exponential solutions are $e^{t}x_1$, $e^{2t}x_2$, and $e^{3t}x_3$

The solution is

$$u(t) = 2e^{t}x_1 + 3e^{2t}x_2 + 4e^{3t}x_3$$

Second order Equations

- The equation $my'' + by' + ky = 0$ is very important in mechanics.
  - The first term $my'' = ma$ balances the force F. The force includes the damping $-by'$ and the elastic restoring force $-ky$.
- In a differential equations course, the method of solution is to substitute $y = e^{\lambda t}$

Each derivative brings a factor λ. We want $y = e^{\lambda t}$

$$m\frac{d^2y}{dt^2} + b\frac{dy}{dt} + ky = 0 \quad \text{becomes} \quad (m\lambda^2 + b\lambda + k)e^{\lambda t} = 0$$

Everything depends on $m\lambda^2 + b\lambda + k = 0$. The equation for λ has 2 roots λ_1 and λ_2.
Then the equation for y has 2 pure solutions $y_1 = e^{\lambda_1 t}$ and $y_2 = e^{\lambda_2 t}$.
Their combinations $c_1y_1 + c_2y_2$ give the complete solution unless $\lambda_1 = \lambda_2$

- We can convert this equation into a vector one. Suppose $m = 1$

$$\begin{cases} \dfrac{dy}{dt} = y' \\ \dfrac{dy'}{dt} = -ky - by' \end{cases} \quad \text{converts to} \quad \frac{d}{dt}\begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -k & -b \end{bmatrix}\begin{bmatrix} y \\ y' \end{bmatrix}$$

  - The second equation connects y'' to y' and y. Together, the equation connects u' to u.
  - $A - \lambda I = \begin{bmatrix} -\lambda & 1 \\ -k & -b-\lambda \end{bmatrix}$ has determinant $\lambda^2 + b\lambda + k = 0$. The equation for the λ are the same. The roots λ_1 and λ_2 are now the eigenvalues of A.
  - The eigenvectors and the solution are

$$x_1 = \begin{bmatrix} 1 \\ \lambda_1 \end{bmatrix} \quad x_2 = \begin{bmatrix} 1 \\ \lambda_2 \end{bmatrix} \quad u(t) = c_1e^{\lambda_1 t}\begin{bmatrix} 1 \\ \lambda_1 \end{bmatrix} + c_2e^{\lambda_2 t}\begin{bmatrix} 1 \\ \lambda_2 \end{bmatrix}$$

- Example 3 Motion around a circle with $y'' + y = 0$ and $y = \cos t$
  $y'' + y = 0 = my'' + by' + ky$ with $m = 1, b = 0, k = 1$
  Substitute $y = e^{\lambda t}$
  $\lambda^2 e^{\lambda t} + e^{\lambda t} = 0$

  $e^{\lambda t}(\lambda^2 + 1) = 0$
  The roots are $\lambda = i$ and $\lambda = -i$. Then half of $e^{it} + e^{-it}$ gives the solution $y = \cos t$
  As a matrix system
  $y(0) = 1$, $y'(0) = 0$ go into $u(0) = (y(0), y'(0)) = (1, 0)$

  use $y'' = -y$ $\quad \frac{du}{dt} = \frac{d}{dt}\begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} y \\ y' \end{bmatrix} = Au$

  The eigenvalues of A are again $\lambda = i$ and $\lambda = -i$. A is anti-symmetric with the eigenvectors $x_1 = (1, i)$ and $x_2 = (1, -i)$. The combination that matches $u(0) = (1, 0)$ is $\frac{1}{2}x_1 + \frac{1}{2}x_2$.
  $u(t) = \frac{1}{2}e^{it}\begin{bmatrix} 1 \\ i \end{bmatrix} + \frac{1}{2}e^{-it}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} \cos t \\ -\sin t \end{bmatrix} = \begin{bmatrix} y(t) \\ y'(t) \end{bmatrix}$
  The vector $u = (\cos t, -\sin t)$

Stability of 2 by 2 Matrices

- For the solution $du/dt = Au$, there is a fundamental question. Does the solution approach $u = 0$ as $t \to \infty$. Is the problem stable? The examples in previous sections included e^t (unstable). Stability depends on the eigenvalues of A.
- The complete solution $u(t)$ is built from pure solutions $e^{\lambda t}x$. If $\lambda \in \mathbb{R}$, $e^{\lambda t}$ will approach zero if λ is negative
- If $\lambda \in \mathbb{C}$, $\lambda = r + is$, the real part r must be negative. When $e^{\lambda t}$ splits into $e^{rt}e^{ist}$, the factor e^{ist} has absolute value fixed at 1.
  $e^{ist} = \cos st + i\sin st$ has $|e^{ist}|^2 = \cos^2 st + \sin^2 st = 1$

  The factor e^{rt} controls growth ($r > 0$ is instability) or decay ($r < 0$ is stability)
- The question is when are the real parts of λ all negative
- Stability
  A is stable and $\lim_{t \to \infty} u(t) = 0 \iff \mathrm{Re}(\lambda) < 0 \ \forall \lambda_i$
  The 2 by 2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ must pass two tests:
  1. $\lambda_1 + \lambda_2 < 0$. The trace $a + d$ must be negative
  2. $\lambda_1\lambda_2 > 0$. The determinant must be positive

The Exponential of A Matrix

- We want to write the solution $u(t)$ in a new form $e^{At}u(0)$
- To do this, we use the Taylor series

$$e^{x} = \sum_{n=1}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

$$e^{x} = \sum_{1} \frac{x^n}{n!} \quad \text{(Maclaurin)}$$

$$e^{At} = I + At + \tfrac{1}{2}(At)^2 + \tfrac{1}{6}(At)^3 + \ldots$$

Its derivative is

$$Ae^{At} = A + A^2t + \tfrac{1}{2}A^3t^2 + \ldots$$

Its eigenvalues are $e^{\lambda t}$

$$(I + At + \tfrac{1}{2}(At)^2 + \ldots)x = (1 + \lambda t + \tfrac{1}{2}(\lambda t)^2 + \ldots)x$$

- $e^{At}u(0)$ solves the differential equation even if there is a shortage of eigenvectors
- Assume A is diagonalizable, then

$$e^{At} = e^{S\Lambda S^{-1}t} = I + S\Lambda S^{-1}t + \tfrac{1}{2}(S\Lambda S^{-1}t)(S\Lambda S^{-1}t) + \ldots$$

$$= S[I + \Lambda t + \tfrac{1}{2}(\Lambda t)^2 + \ldots]S^{-1}$$

$$= Se^{\Lambda t}S^{-1}$$

- e^{At} equals $Se^{\Lambda t}S^{-1}$. Then Λ is a diagonal matrix and so is $e^{\Lambda t}$. The numbers $e^{\lambda_i t}$ are on its diagonal. Multiply by $u(0)$ to get $u(t)$

$$e^{At}u(0) = Se^{\Lambda t}S^{-1}u(0) = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

- Example

Solve $du/dt = Au = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} u$ starting from $u(0) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$\lambda_1 = 1 \Rightarrow x_1 = (1,0)$

$\lambda_2 = 2 \Rightarrow x_2 = (1,1)$

$u(0) = x_1 + x_2 \Rightarrow u(t) = e^{t}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + e^{2t}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

For every $u(0)$

$$u(t) = Se^{\Lambda t}S^{-1}u(0) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} e^{t} & \\ & e^{2t} \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}u(0) = \begin{bmatrix} e^{t} & e^{2t} - e^{t} \\ 0 & e^{2t} \end{bmatrix}u(0)$$

Chapter 6.4: Symmetric Matrices

- For projection onto a plane in $\mathbb{R}^3$, the plane is full of eigenvectors (where $Px = x$). The other eigenvectors are perpendicular to the plane (where $Px = 0$)
- The eigenvalues $\lambda = 1, 1, 0$ are real. Three eigenvectors can be chosen perpendicular to each other
- What is special about $Ax = \lambda x$ when A is symmetric?

  If we diagonalize $A = S\Lambda S^{-1}$,

  $A^T = (S^{-1})^T (\Lambda)^T (S)^T = (S^{-1})^T (\Lambda)(S^T)$.

  Since

  $A^T = A$,

  $A = (S^{-1})^T \Lambda S^T = S\Lambda S^{-1} \Rightarrow S^{-1} = S^T$ and $S^T S = I$, which makes S orthonormal
- Fact:
  1. A symmetric matrix has only real eigenvalues
  2. The eigenvalues can be chosen orthonormal
- Those n orthonormal eigenvectors go into the columns of S. Every symmetric matrix can be diagonalized. Its eigenvector matrix S becomes an orthonormal matrix Q.
- We "chose" because any scalar multiple of an eigenvector is itself an eigenvector. We pick eigenvectors of unit length. Then for symmetric matrices, $S\Lambda S^{-1}$ is in its special form $Q\Lambda Q^T$
- Spectral Theorem:

  Every symmetric matrix has the factorization $A = Q\Lambda Q^T$ with real eigenvalues in Λ and orthonormal eigenvectors in $S = Q$:

  Symmetric diagonalization $\quad A = Q\Lambda Q^{-1} = Q\Lambda Q^T$ with $Q^{-1} = Q^T$
- It's easy to see $Q\Lambda Q^T$ is symmetric $(Q\Lambda Q^T)^T = (Q^T)^T \Lambda^T Q^T = Q\Lambda Q^T$.

  The harder part is to prove every symmetric matrix has real λ's and orthonormal eigenvectors.
- First, an example!

  Find the λ's and x's when $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

  $\det(A - \lambda I) = \lambda^2 - 5\lambda = 0$

  $\lambda = 0, 5$ (both real)
- Two eigenvectors are $(2, -1)$ and $(1, 2)$ – orthogonal but not yet orthonormal.
- The eigenvector for $\lambda = 0$ is in the nullspace of A. The eigenvector for $\lambda = 5$ is in the column space. The fundamental Theorem says the row space and nullspace are orthogonal. Since $A^T = A$, the row space is equivalent to the column space, so our 2 eigenvectors must be orthogonal.
- To make them orthonormal, we just divide each one by its length ($\sqrt{5}$)
- Then, to get Λ

$$\Lambda = Q^{-1} A Q = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} = \Lambda$$

- Proof: All eigenvalues of a real symmetric matrix are real.

Suppose that $Ax = \lambda x$, with $\lambda \in \mathbb{C}$. Its complex conjugate is $\bar{\lambda} = a - ib$. x is also a complex vector ($\mathbb{C}^n$), and taking the complex conjugate of each component gives $\bar{x}$. $\bar{\lambda}\bar{x} = \overline{(\lambda x)}$. We take the conjugate of both sides, with A being a real symmetric matrix.

$Ax = \lambda x$ leads to $A\bar{x} = \bar{\lambda}\bar{x}$. Transpose to $\bar{x}^T A = \bar{x}^T \bar{\lambda}$.

Now take the dot product of the first equation with $\bar{x}$ and the second equation with x:

$$\bar{x} A x = \bar{x}^T \lambda x \quad \text{and} \quad \bar{x}^T A x = \bar{x}^T \bar{\lambda} x$$

The left sides are equal so the right sides are also equal. One equation has λ and another has $\bar{\lambda}$. They multiply $\bar{x}^T x = |x_1|^2 + |x_2|^2 =$ length squared which is not zero. Therefore λ must equal $\bar{\lambda}$ and $b = 0$ (QED)

- Orthogonal Eigenvectors: Eigenvectors of a real symmetric matrix (when they correspond to different eigenvalues) are always perpendicular
  - Proof: Suppose $Ax = \lambda_1 x$ and $Ay = \lambda_2 y$. We are assuming that $\lambda_1 \neq \lambda_2$. Take dot products of the first equation with y and the second with x.

$$(\lambda_1 x)^T y = (Ax)^T y = x^T A^T y = x^T (Ay) = x^T \lambda_2 y$$

The left side is $x^T \lambda_1 y$, the right side is $x^T \lambda_2 y$. Since $\lambda_1 \neq \lambda_2$, then $x^T y = 0$, which means the two eigenvectors x and y are orthogonal.

- Example 2:

The eigenvectors of a 2 by 2 symmetric matrix have a special form

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \text{ has } x_1 = \begin{bmatrix} b \\ \lambda_1 - a \end{bmatrix} \text{ and } x_2 = \begin{bmatrix} \lambda_2 - c \\ b \end{bmatrix}$$

$$x_1^T x_2 = \begin{bmatrix} b & \lambda_1 - a \end{bmatrix} \begin{bmatrix} \lambda_2 - c \\ b \end{bmatrix} = b(\lambda_2 - c) + b(\lambda - a) = b(\lambda + \lambda_2 - a - c) = 0$$

$\lambda_1 + \lambda_2$ is the trace, which is equal to $a + c$

- From $A = Q\Lambda Q^T$ with $Q^T Q = I$, every 2 by 2 symmetric matrix looks like

$$A = Q\Lambda Q^T = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} \begin{bmatrix} x_1^T \\ x_2^T \end{bmatrix}$$

The columns x_1 and x_2 multiply the rows $\lambda_1 x_1^T$ and $\lambda_2 x_2^T$ to produce A. We can write A as the sum of rank one matrices

$$A = \lambda_1 x_1 x_1^T + \lambda_2 x_2 x_2^T$$

- When the symmetric matrix is n by n, there are n columns in Q multiplying n rows in Q^T. The n products $x_i x_i^T$ are projection matrices

- Including the eigenvalues the spectral theorem $A = Q\Lambda Q^T$ for symmetric matrices says that A is a combination of projection matrices

$$A = \lambda_1 P_1 + \dots + \lambda_i P_i + \dots + \lambda_n P_n \quad \lambda_i = \text{eigenvalue} \quad P_i = \text{projection onto eigenspace}$$

Complex Eigenvalues of Real Matrices

- If A is symmetric, then its eigenvalues are real. A nonsymmetric matrix can easily produce complex eigenvalues and eigenvectors. In this case $A\bar{x} = \bar{\lambda}\bar{x}$ is different from $Ax = \lambda x$. It gives us a new eigenvalue $(\bar{\lambda})$ and a new eigenvector $(\bar{x})$.
- For real matrices complex λ's and x's come in conjugate pairs.
- Example:

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ has } \lambda_1 = \cos\theta + i\sin\theta \text{ and } \lambda_2 = \cos\theta - i\sin\theta$$

$$\lambda x: \quad Ax = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = (\cos\theta + i\sin\theta)\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

$$\bar{\lambda}\bar{x}: \quad A\bar{x} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 1 \\ i \end{bmatrix} = (\cos\theta - i\sin\theta)\begin{bmatrix} 1 \\ i \end{bmatrix}$$

The eigenvectors $(1, -i)$ and $(1, i)$ are complex conjugates because A is real.
- $|\lambda| = 1$ for the eigenvalues of every orthogonal matrix

Eigenvalues versus Pivots

- The eigenvalues of A are very different from pivots. The only connection we have this far is:

  product of pivots = determinant = product of eigenvalues

  - We are assuming n pivots and n real eigenvalues.
- For symmetric matrices the pivots and the eigenvalues have the same sign
  - The number of positive eigenvalues of $A = A^T$ equals the number of positive pivots
  - Special case: A has all $\lambda_i > 0$ if and only if all pivots are positive
- Example

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \text{ has } \begin{array}{ll} \text{Pivots} & 1 \text{ and } -8 \\ \text{Eigenvalues} & 4 \text{ and } -2. \end{array}$$

- Proof:

  If $A^T = A$ and A is invertible, A can be factorized as $A = LDL^T$

$$\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -8 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$

  - The eigenvalues of LDL^T are 4 and -2. The eigenvalues of IDI^T are 1 and -8 (the pivots). The eigenvalues are changing, as the "3" in L moves to zero, but to change sign, a real eigenvalue would have to cross 0, making it singular. Our matrix always has pivots 1 and -8 so it is never singular. The signs cannot change.

All Symmetric Matrices are Diagonalizable

- A repeated eigenvalue can lead to a shortage of eigenvectors for diagonalization. This happens sometimes for nonsymmetric matrices. It never happens for symmetric matrices. There are always enough eigenvectors to diagonalize $A = A^T$
- Proof:
  - Every square matrix factors into $A = QTQ^{-1}$ where T is upper triangular and $\bar{Q}^T = Q^{-1}$
  - If A has real eigenvalues then Q and T can chosen real: $Q^TQ = I$
  - This is Schur's Theorem. We are looking for $AQ = QT$. The first q_1 of Q must be a unit eigenvector of A. Then the first columns of AQ and QT are Aq_1 and $t_1 q_1$. But the other columns of Q need not be eigenvectors when T is only triangular.
  - So use any $n-1$ columns that complete q_1 to a Q_1 with orthonormal columns. At this point only the first columns of Q and T are set, where $Aq_1 = t_1 q_1$.

$$\bar{Q}_1^T A Q_1 = \begin{bmatrix} \bar{q}_1^T \\ \vdots \\ \bar{q}_n^T \end{bmatrix} \begin{bmatrix} Aq_1 & \cdots & Aq_n \end{bmatrix} = \begin{bmatrix} t_{11} & \cdots & \\ 0 & A_2 & \\ 0 & & \end{bmatrix}$$

  - By induction, we can repeat this for submatrix A_2
  - When A is symmetric, T is the diagonal Λ when A is symmetric. This is because T is both symmetric and triangular, so it is diagonal.

Problem set 6.4

1.

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 4 & 3 & 0 \\ 8 & 6 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 6 \\ 3 & 3 & 3 \\ 6 & 3 & 5 \end{bmatrix} + \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -3 \\ 2 & 3 & 0 \end{bmatrix}$$

2.

$(A^TCA)^T = A^TC^T(A^T)^T = A^TCA$

$C = 3$ by 3

$(A^TCA) = 6$ by 6

3.

$$A - \lambda I = \begin{bmatrix} 2-\lambda & 2 & 2 \\ 2 & -\lambda & 0 \\ 2 & 0 & -\lambda \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} \end{bmatrix}$$

$$\begin{aligned} \det(A - \lambda I) &= (2-\lambda)(-\lambda)(-\lambda) - (2)(2)(-\lambda) - (2)(2)(-\lambda) \\ &= -\lambda^3 + 2\lambda^2 + 8\lambda \\ &= -\lambda(\lambda^2 - 2\lambda - 8\lambda) \\ &= -\lambda(\lambda - 4)(\lambda + 2) \end{aligned}$$

set $\det(A - \lambda I) = 0$

$\lambda = 0, -2, 4$

Case $\lambda = 0$

$(A - 0I)x_1 = 0$

$$\begin{bmatrix} 2 & 2 & 2 \\ 2 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix} x_1 = 0$$

$x_1 = (0, -1, 1)$

$q_1 = (0, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$

case $\lambda = -2$:

$(A + 2I)x_2 = 0$

$$\begin{bmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix} x_2 = 0$$

$x_2 = (-1, 1, 1)$

$q_2 = (-\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3})$

case $\lambda = 4$

$(A - 4I)x_3 = 0$

$$\begin{bmatrix} -2 & 2 & 2 \\ 2 & -4 & 0 \\ 2 & 0 & -4 \end{bmatrix} x_3 = 0$$

$x_3 = (2, 1, 1)$

$q_3 = (\frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{6})$

4.

$$\begin{aligned} \det(A - \lambda I) &= (-2-\lambda)(7-\lambda) - (6)(6) \\ &= \lambda^2 - 5\lambda - 14 - 36 \\ &= \lambda^2 - 5\lambda - 50 \\ &= (\lambda - 10)(\lambda + 5) \end{aligned}$$

$\lambda = -5, 10$

Case $\lambda = 10$:

$$\begin{bmatrix} -12 & 6 \\ 6 & -3 \end{bmatrix} x_1 = 0$$

$x_1 = (1, 2)$

$q_1 = (\frac{\sqrt{5}}{5}, \frac{2\sqrt{5}}{5})$

Case $\lambda = -5$:

$$\begin{bmatrix} 3 & 6 \\ 6 & 12 \end{bmatrix} x_2 = 0$$

$x_2 = (2, -1)$

$q_2 = (\frac{2\sqrt{5}}{5},$

$$Q = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$$

11.

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

$$\det(A - \lambda I) = (3-\lambda)(3-\lambda) - (1)(1)$$
$$= 9 - 6\lambda + \lambda^2 - 1$$
$$= \lambda^2 - 6\lambda + 8$$
$$= (\lambda - 4)(\lambda - 2)$$

$\det(A - \lambda I) = 0 \Rightarrow \lambda = 2, 4$

Case $\lambda = 2$:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} x_1 = 0$$

$x_1 = (1, -1)$

$q_1 = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$

Case $\lambda = 4$:

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} x_2 = 0$$

$x_2 = (1, 1)$

$q_2 = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$

$$A = Q \Lambda Q^T = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

$$= 2 \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix} + 4 \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

19.

orthonormal, $i, i, -i, -i$

21.

a) False

b) True

c) True

d) False

23.

A: Invertible, Permutation

B:

26.

$$\begin{bmatrix} 1 & 10^{-15} \\ 0 & 1+10^{-15} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

$\lambda = 1, \ \lambda = 1 + 10^{-15}$

$$\begin{bmatrix} 10^{-15} & 10^{-15} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 0$$

$$\theta = \cos^{-1}\left(\frac{(1)}{\sqrt{2}\sqrt{1}}\right) = 45°$$

Chapter 6.5: Positive Definite Matrices

- This chapter focuses on symmetric matrices that have positive eigenvalues.
- Symmetric matrices with positive eigenvalues are called positive definite matrices.
- We can determine if a matrix is positive definite without calculating eigenvalues.
- Remember, the eigenvalues are real because the matrix is symmetric.
- Start with 2 by 2. When does $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ have $\lambda_1 > 0$ and $\lambda_2 > 0$
  - The eigenvalues of A are positive if and only if $a > 0$ and $ac - b^2 > 0$

    $A_1 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ is not positive definite because $ac - b^2 = (1)(1) - (2)^2 = -3 < 0$

    $A_2 = \begin{bmatrix} 1 & -2 \\ -2 & 6 \end{bmatrix}$ is positive definite because $ac - b^2 = (1)(6) - (-2)^2 = 2 > 0$

    $A_3 = \begin{bmatrix} -1 & 2 \\ 2 & -6 \end{bmatrix}$ is not positive definite because $a = -1$
  - Proof: The determinant $(ac - b^2)$ is the product of the pivots. If it is positive, λ_1 and λ_2 must be the same sign. The trace $a + c$ is the sum of the eigenvalues is also positive and a, c are both positive as well (or else $ac - b^2$ fails)
  - This test uses the 1 by 1 determinant and 2 by 2 determinant. A 3 by 3 matrix will use the 1 by 1, 2 by 2, and 3 by 3 determinant.

- Another test:
  - The eigenvalues of $A = A^T$ are positive iff the pivots are all positive

    $a > 0$ and $\frac{ac-b^2}{a} > 0$
  - This is a corollary of the first test, but we should recognize a and $\frac{ac-b^2}{a}$ are pivots

    $$\begin{bmatrix} a & b \\ b & c \end{bmatrix} \longrightarrow \begin{bmatrix} a & b \\ 0 & -\frac{b^2}{a}+c \end{bmatrix} \quad -\frac{b^2}{a}+c = \frac{ac-b^2}{a}$$
  - Pivots are a lot easier and faster to compute than eigenvalues

Energy-based Definition

- From $Ax = \lambda x$, multiply by x^T to get $x^T Ax = \lambda x^T x$. The right is a positive λ times a positive number $x^T x = \|x\|^2$. So $x^T Ax$ is positive for any eigenvector.
- The new idea is that $x^T Ax$ is positive for all nonzero vectors x, not just eigenvectors
- In many applications, this number $x^T Ax$ (or $\frac{1}{2}x^T Ax$) is the energy in the system. The requirement of positive energy is another definition of a positive definite matrix.
  - Eigenvalues and pivots are two equivalent ways to test the new requirements $x^T Ax > 0$
  - A is positive definite if $x^T Ax > 0$ for every nonzero vector x:

    $$x^T Ax = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} ax+by \\ bx+cy \end{bmatrix} = ax^2 + 2bxy + cy^2 > 0$$
  - If A and B are symmetric positive definite, so is $A+B$
    - $x^T(A+B)x = x^T Ax + x^T Bx$. The energies add, so the result has positive energy as well
- Start any matrix R, possibly rectangular. We know that $A = R^T R$ is square and symmetric.
  - If the columns of R are independent, then $A = R^T R$ is positive definite.

    $x^T Ax = x^T R^T R x = (Rx)^T (Rx)$. The vector Rx is not zero when $x \neq 0$

    (b/c independent columns). Then $x^T Ax = \|Rx\|^2 > 0$ and A is positive definite.
- When a symmetric matrix has one of these five properties, it has them all
  1. All n pivots are positive
  2. All k by k, $k \in \mathbb{Z}$, $k \in [0, n]$ upper left determinants are positive
  3. All n eigenvalues are positive
  4. $x^T Ax$ is positive for all x except $x = 0$
  5. A equals $R^T R$ for a matrix R with independent columns
- Example

  $$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$
  - Pivots: $2, \frac{3}{2}, \frac{4}{3}$
  - Upper left determinants: $2, 3, 4$
  - Eigenvalues: $2-\sqrt{2}, 2, 2+\sqrt{2}$
  - $x^T Ax = 2(x_1^2 - x_1x_2 + x_2^2 + x_2x_3 + x_3^2) = 2(x_1 - \frac{1}{2}x_2)^2 + \frac{3}{2}(x_2 - \frac{2}{3}x_3)^2 + \frac{4}{3}(x_3)^2$
    - All squares, so $x^T Ax > 0$
  - $R = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$

- We can also choose R using $A = LDL^T$

$$LDL^T = \begin{bmatrix} 1 & & \\ -\frac{1}{2} & 1 & \\ 0 & -\frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & & \\ & \frac{3}{2} & \\ & & \frac{4}{3} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ & 1 & -\frac{2}{3} \\ & & 1 \end{bmatrix} = (L\sqrt{D})(L\sqrt{D})^T = R^T R$$

- This R has square roots, but it is the only R that is 3 by 3 and upper triangular. It is the "Cholesky factor" of A
- In applications, the rectangular R is how we build A and this Cholesky R is how we break it apart
- Eigenvalues give the symmetric choice $R = Q\sqrt{\Lambda}Q^T$

Positive Semidefinite Matrices

- Often we are at the edge of positive definiteness. The determinant is zero. The smallest eigenvalue is zero. The energy in its eigenvector is $x^T A x = 0$. These matrices are called <u>positive semidefinite</u>
  - Here are two examples:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

A has eigenvalues 0 and 5. Its upper left determinants are 1 and 0. Its rank is only 1. This matrix A factors in $R^T R$ with dependent columns in R.

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$$

- Positive semidefinite matrices have all $\lambda \geq 0$ and $x^T A x \geq 0$.

First Application: The Ellipse $ax^2 + 2bxy + cy^2 = 1$

- Think of a tilted ellipse $x^T A x = 1$. Its center is (0,0) as in fig 6.5.1. Turn it to align it with the x and y axes. These 2 ellipses show the geometry behind $A = Q \Lambda Q^T$.
  1. The tilted ellipse is associated with A. Its equation is $x^T A x = 1$.
  2. The lined-up ellipse is associated with Λ. Its equation is $X^T \Lambda X = 1$
  3. The rotation matrix that lines up the ellipse is the eigenvector matrix Q
- Example
  - Find the axes of the tilted ellipse $5x^2 + 8xy + 5y^2 = 1$
  - Start with the positive definite matrix that matches this equation

$$[x \; y]\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 1 \Rightarrow \text{the matrix is } A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$$

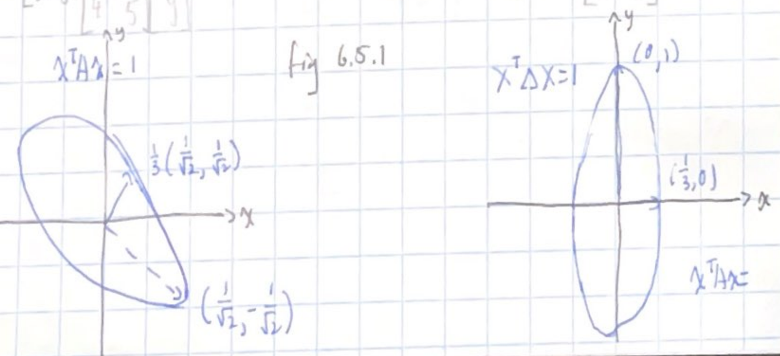

fig 6.5.1

- The eigenvectors are $\begin{bmatrix}1\\1\end{bmatrix}$ and $\begin{bmatrix}1\\-1\end{bmatrix}$. Divide by $\sqrt{2}$ to get unit vectors. Then $A=Q\Lambda Q^T$

$$\begin{bmatrix}5&4\\4&5\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\1&-1\end{bmatrix}\begin{bmatrix}9&0\\0&1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\1&-1\end{bmatrix}$$

- Now multiply by $[x\ y]$ on the left and $\begin{bmatrix}x\\y\end{bmatrix}$ on the right to get to x^TAx

$$[x\ y]\begin{bmatrix}5&4\\4&5\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=5x^2+8xy+5y^2=9\left(\frac{x+y}{\sqrt{2}}\right)^2+1\left(\frac{x-y}{\sqrt{2}}\right)^2$$

- These coefficients are from the eigenvalues 9 and 1 from Λ. Inside the squares are the eigenvectors $\frac{(1,1)}{\sqrt{2}}$ and $\frac{(1,-1)}{\sqrt{2}}$.

- The axes of the tilted ellipse point along the eigenvectors. This $A=Q\Lambda Q^T$ is called the "Principal axis theorem" It also gives the axis lengths (from the eigenvalues).
  - To align the ellipse,

    Let $\frac{x+y}{\sqrt{2}}=X$ and $\frac{x-y}{\sqrt{2}}=Y$ and $9X^2+Y^2=1$
  - The largest value of $X^2=\frac{1}{9}$. The endpoint of the shorter axis has $X=\frac{1}{3}$ and $Y=0$. The bigger eigenvalue gives the shorter axis, of half-length $1/\sqrt{\lambda_1}=\frac{1}{3}$. The smaller eigenvalue $\lambda_2=1$ gives the greater length $1/\sqrt{\lambda_2}=1$.
  - In the xy system, the axes are along the eigenvectors of A. In the XY system, the axes are along the eigenvectors of Λ, the coordinate axes.
- Suppose $A=Q\Lambda Q^T$ is positive definite, so $\lambda_i>0$. The graph of $x^TAx=1$ is an ellipse:

$$[x\ y]Q\Lambda Q^T\begin{bmatrix}x\\y\end{bmatrix}=[X\,Y]\Lambda\begin{bmatrix}X\\Y\end{bmatrix}=\lambda_1X^2+\lambda_2Y^2=1$$

The axes point along eigenvectors. The half lengths are $1/\sqrt{\lambda_1}$ and $1/\sqrt{\lambda_2}$

Problem Set 6.5

2.

A_4

3.

$$A=\begin{bmatrix}1&2\\2&9\end{bmatrix}=\begin{bmatrix}1&\\2&1\end{bmatrix}\begin{bmatrix}1&\\&4\end{bmatrix}\begin{bmatrix}1&2\\0&1\end{bmatrix}$$

4.

$$A=\begin{bmatrix}1&2\\2&9\end{bmatrix}$$

$$x^TAx=[x\,y]\begin{bmatrix}1&2\\2&9\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$$

$$=[x\ y]\begin{bmatrix}x+2y\\2x+9y\end{bmatrix}$$

$$x^2+4xy+9y^2$$

6.

$f(x,y)=2xy$

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

$\lambda=-1,1\qquad(-\lambda)(-\lambda)-1=\lambda^2-1$

8.

$$A=\begin{bmatrix}3&6\\6&16\end{bmatrix}$$

9.

$r=1$

$p=4$

a.

$x_1^2-2x_1(x_2+2x_3)+(x_2-2x_3)^2$

$x_1^2-2x_1x_2+4x_1x_3$

15

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 4 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 4x_1 + x_2 + x_3 \\ x_1 + 2x_3 \\ x_1 + 2x_2 + 5x_3 \end{bmatrix}$$

$= 4x_1^2 + x_1x_2 + x_1x_3 + x_1x_2 + 2x_2x_3 + x_1x_3 + 2x_2x_3 + 5x_3^2$

$= 4x_1^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3 + 5x_3^2$

choose $x = (1, -5, 0)$

$x^TAx = 4(1)^2 + 2(1)(-5) = -6$

18.

$Ax = \lambda x$

$x^TAx = x^T(\lambda x) = \lambda x^Tx = \lambda\|x\|^2$

$\mathrm{sgn}(\lambda) = \mathrm{sgn}(x^TAx)$

20.

a) All n pivots $\quad a^2 - b c < 0$

b)

P is symmetric, $\quad ax^2 + 2bxy + cy^2 > 0$

$P = Q$ $\quad$ iff $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is positive definite

19

$x = c_1x_1 + c_2x_2 + \dots + c_nx_n$

$x^TAx = (c_1x_1 + c_2x_2 + \dots + c_nx_n)A(c_1x_1 + c_2x_2 + \dots + c_nx_n)$

$(c_1x_1 + c_2x_2 + \dots + c_nx_n)(c_1\lambda_1x_1 + c_2\lambda_2x_2 + c_n\lambda_nx_n)$

$= c_1^2\lambda_1x_1^Tx_1 + c_1c_2\lambda_2x_1^Tx_2 + \dots$

Since A is symmetric, all eigenvectors are orthogonal

$x^TAx = \sum_1 c_i^2\lambda_i x_i^Tx_i > 0$

28

$A = Q\Lambda Q^T$

a) 10 $\quad$ $\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

b) 2, 5

a) All eigenvalues > 0

22.

$A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$

$\lambda = 1, 9$

$x_1 = (1, -1)$

$x_2 = (1, 1)$

$q_1 = (\frac{1}{\sqrt2}, -\frac{1}{\sqrt2})$

$q_2 = (\frac{1}{\sqrt2}, \frac{1}{\sqrt2})$

$A = Q\Lambda Q^T$

$= \begin{bmatrix} \frac{1}{\sqrt2} & -\frac{1}{\sqrt2} \\ \frac{1}{\sqrt2} & \frac{1}{\sqrt2} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt2} & \frac{1}{\sqrt2} \\ -\frac{1}{\sqrt2} & \frac{1}{\sqrt2} \end{bmatrix}$

$A^{\frac12} = \begin{bmatrix} \frac{1}{\sqrt2} & -\frac{1}{\sqrt2} \\ \frac{1}{\sqrt2} & \frac{1}{\sqrt2} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt2} & \frac{1}{\sqrt2} \\ -\frac{1}{\sqrt2} & \frac{1}{\sqrt2} \end{bmatrix}$

23.

$\frac{1}{a^2} = \lambda_1 \Rightarrow a = \frac{1}{\sqrt{\lambda_1}}, \; b = \frac{1}{\sqrt{\lambda_2}}$

$a = \frac13, \; b = \frac12$

24.

$A = \begin{bmatrix} 1 & \frac12 \\ \frac12 & 1 \end{bmatrix}$ $\quad a = \frac{1}{\sqrt{\lambda_1}} = \frac{1}{\sqrt{\frac12}} = \sqrt2$

$\lambda = \frac12, \frac32$ $\quad b = \frac{1}{\sqrt{\lambda_2}} = \frac{1}{\sqrt{\frac32}} = \frac{\sqrt6}{3}$

25.

$A = C^TC$

$= \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}$

$= \begin{bmatrix} 10 & 2 \\ 2 & 4 \end{bmatrix}$

$A = LDL^T$

$\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

$= \left(\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\right)\left(\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\right)$

$C = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

$= \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix}$

26

$A = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 9 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

$C = \sqrt{D}L^T = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

$= \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}$

Chapter 6.6: Similar Matrices

- Matrices with n unique eigenvalues can be diagonalized as $S\Lambda S^{-1}$. However, we can't do this for all matrices. Some matrices have too few eigenvectors.
- In this section, the eigenvector matrix S remains the choice when it is available, but we now allow any invertible matrix M.
- Starting from A we go to $M^{-1}AM$. No matter which matrix M we choose, the eigenvalues remain the same. The matrices A and $M^{-1}AM$ are called similar.
- A typical matrix A is similar to a whole family of other matrices because M can be any invertible matrix.
- Let M be any invertible matrix. Then $B = M^{-1}AM$ is similar to A.
  - If $B = M^{-1}AM$ then immediately $A = MBM^{-1}$. If B is similar to A, A is similar to B.
- A diagonalizable matrix is similar to Λ. In that special case M is S.
- The combination $M^{-1}AM$ appears when we change variables in a differential equations. Start with an equation for u and set $u = Mv$

$$\frac{du}{dt} = Au \text{ becomes } M\frac{dv}{dt} = AMv \text{ which is } \frac{dv}{dt} = M^{-1}AMv$$

- When $M = S$ the system is diagonal - the maximal simplicity. Other choices of M could make a system triangular and easier to solve. Since we can always go back to u, similar matrices must give the same growth and decay. More precisely, the eigenvalues of A and B are the same
- Similar matrices A and $M^{-1}AM$ have the same eigenvalues. If x is an eigenvector of A, then $M^{-1}x$ is an eigenvector of $M^{-1}AM$.
  - Proof: Since $B = M^{-1}AM$ gives $A = MBM^{-1}$. Suppose $Ax = \lambda x$

$$MBM^{-1}x = \lambda x \text{ means } B(M^{-1}x) = \lambda(M^{-1}x)$$

The eigenvalue of B is the same λ. The eigenvector has changed to $M^{-1}x$
  - Two matrices can have the same repeated λ, and fail to be similar
- Example

The projection $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ is similar to $\Lambda = S^{-1}AS = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

Now choose $M = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$, The similar matrix $M^{-1}AM$ is $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

Also choose $M = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, The similar matrix $M^{-1}AM$ is $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$

- Example 2 (repeated eigenvalues)

  $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is similar to $\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$ and all $B = \begin{bmatrix} cd & d^2 \\ -c^2 & -cd \end{bmatrix}$ except $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

  - The above matrices are all of determinant 0 and rank 1.
  - These matrices can't be diagonalized. $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is the closest we can get. It is the "Jordan normal form" for the family of matrices B.

- What M changes
  - Not changed by M
    - Eigenvalues
    - Trace and determinant
    - Rank
    - Number of independent eigenvectors
    - Jordan form
  - Changed by M
    - Eigenvectors
    - Nullspace
    - Column Space
    - Row Space
    - Left Nullspace

    $\Big\}$ All 4 Fundamental Subspaces
    - Singular values

Examples of the Jordan Form

- Example

  $$J = \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix} \text{ has } \lambda = 5,5,5 \text{ and } x = (1,0,0).\ AM = 3,\ GM = 1$$

  $$J - 5I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \text{ has rank 2}$$

  - Every similar matrix $B = M^{-1}JM$ has $\lambda = 5,5,5$. Also $B - 5I$ must have the same rank 2. Its nullspace has dimension 1. So every B that is similar to this "Jordan block" J has only one independent eigenvector $M^{-1}x$.
  - The transpose matrix J^T has the same eigenvalues 5, 5, 5 and $J^T - 5I$ has the same rank 2. Jordan's theorem says J^T is similar J. The matrix M that produces this similarity happens to be the reverse identity.

  $$J^T = M^{-1}JM \text{ is } \begin{bmatrix} 5 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 1 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

  - The eigenvector of J^T is $M^{-1}(1,0,0) = (0,0,1)$
  - This matrix J is similar to every matrix A with eigenvalues 5,5,5 and one line of eigenvectors.

- Example
  - Since J is not diagonal, the equation $du/dt = Ju$ cannot be simplified by changing variables.

  $$\frac{du}{dt} = Ju = \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ is } \begin{matrix} dx/dt = 5x + y \\ dy/dt = 5y + z \\ dz/dt = 5z \end{matrix}$$

- We solve this by back substitution

$\frac{dz}{dt} = 5z$ yields $z = z(0)e^{5t}$

$\frac{dy}{dt} = 5y + z$ yields $y = (y(0) + tz(0))e^{5t}$

$\frac{dx}{dt} = 5x + y$ yields $x = (x(0) + ty(0) + \frac{1}{2}t^2 z(0))e^{5t}$

- The missing eigenvectors are responsible for the te^{5t} and t^2e^{5t} in y and x.

The Jordan Form

- For every A, we want to choose M so that $M^{-1}AM$ is as nearly diagonal as possible.
- When A has a full set of n eigenvectors, they go into the columns of $M = S$. Then the matrix $S^{-1}AS$ is diagonal, period.
  - The matrix Λ is the Jordan Form of A - when A can be diagonalized.
- In general, suppose A has s independent eigenvectors. Then it is similar to a matrix with s blocks. Each "Jordan block" has an eigenvalue on the diagonal and 1's above it. This block accounts for one eigenvector of A. When there are n eigenvectors and n blocks, each block is 1 by 1 and J is Λ.
- Jordan Form
  - If A has s independent eigenvectors, it is similar to a matrix J that has s Jordan blocks on its diagonal. Some matrix M puts A into Jordan form.

$$M^{-1}AM = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix} = J$$

  - Each block in J has one eigenvalue λ_i, one eigenvector, and 1's above the diagonal.

$$J_i = \begin{bmatrix} \lambda_i & 1 & \\ & \ddots & 1 \\ & & \lambda_i \end{bmatrix}$$

  - The Jordan block has one off-diagonal 1 for each missing eigenvector.
  - In each family of similar matrices, we have a Jordan Form which is diagonal or nearly so. For that J, we can solve $du/dt = Ju$, we can take powers J^k. Every other matrix in the family has the form $A = MJM^{-1}$
  - A is similar to B iff they have the same Jordan Form.
  - The proof for Jordan's theorem is very intricate and computations with the Jordan Form is not stable.
  - Example

If A has $\lambda = 4, 2, 3, 3$, its Jordan Form is

$$J = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Problem Set 6.6

1.

$C = F^{-1}AF$

$B = GCG^{-1}$

$B = GF^{-1}AFG^{-1}$

$= GF^{-1}A(GF^{-1})^{-1}$

$M = GF^{-1}$

3.

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

$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \left(\frac{1}{ad-bc}\right)$

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

$= \frac{1}{ad-bc} \begin{bmatrix} ad+bd & -ab-b^2 \\ cd+d^2 & -bc-bd \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

$M = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$

7.

a)

$Ax = 0$

$(M^{-1}AM)(M^{-1}x) = 0$

$M^{-1}Ax = 0$

$Ax = 0$

b) dimension

9.

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

$A^2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

$A^3 = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$

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

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

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

10.

$J = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$

$J^2 = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{bmatrix}$

$J^3 = \begin{bmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{bmatrix} \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} \lambda^3 & 3\lambda^2 \\ 0 & \lambda^3 \end{bmatrix}$

$J^k = \begin{bmatrix} \lambda^k & k\lambda^{k} \\ 0 & \lambda^k \end{bmatrix}$

$J^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$

$J^{-1} = \begin{bmatrix} \frac{1}{\lambda} & -\frac{1}{\lambda} \\ 0 & \frac{1}{\lambda} \end{bmatrix}$

11.

$\frac{du}{dt} = Ju = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} u$

$y = y(0)e^{\lambda t} = 5e^{\lambda t}$

$x = (x(0) + ty(0))e^{\lambda t} = (5+2t)e^{\lambda t}$

$u = \begin{bmatrix} (5+2t)e^{\lambda t} \\ 5e^{\lambda t} \end{bmatrix} \Rightarrow \frac{du}{dt} = \begin{bmatrix} 2e^{\lambda t} + \lambda e^{\lambda t}(5+2t) \\ 5\lambda e^{\lambda t} \end{bmatrix}$

13.

$J_1 = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}$ $J_2 = \begin{bmatrix} 0 & & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}$ $J_3 = \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}$

$J_4 = \begin{bmatrix} 0 & 1 & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}$ $J_5 = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & \\ & & & 0 \end{bmatrix}$

15.

$\det(M^{-1}AM - \lambda M^{-1}M)$

$= \det[M^{-1}(AM - \lambda M)]$

$= \det[M^{-1}M(A - \lambda I)]$

$= \det(A - \lambda I)$

d)

$\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$ has $\lambda = 3, 3$

20.

a) $A = M^{-1}BM$

$A^2 = M^{-1}BMM^{-1}BM$

$= M^{-1}B^2M$

c) $\begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$ has $\lambda = 3, 4$

$\begin{bmatrix} 3 & 1 \\ 0 & 4 \end{bmatrix}$ has $\lambda = 3, 4$

$J = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$

Chapter 6.7: Singular Value Decomposition

- A is any m by n matrix. Its rank is r. We diagonalize this matrix A, but not by $S^{-1}AS$.
- The eigenvectors in S have three big problems
  1. They are usually not orthogonal
  2. There may be insufficient eigenvectors
  3. $Ax = \lambda x$ requires A to be square
- Instead, we use the singular vectors of A.
- The price we pay is to have two sets of singular vectors, u's and v's. The u's are eigenvectors of AA^T and the v's are the eigenvectors of A^TA. Since those matrices are both symmetric, their eigenvectors can be chosen orthonormal.
- The fact that $A(A^TA)$ is the same as $(AA^T)A$ will lead to a remarkable property of u and v

"A is diagonalized" $Av_1 = \sigma_1 u_1 \quad Av_2 = \sigma_2 u_2 \; \ldots, \; Av_r = \sigma_r u_r$

- The singular vectors $v_1, v_2, v_3, \ldots$ are in the row space of of A. The outputs $u_1, \ldots, u_r$ are in the column space of A. The singular values $\sigma_1, \sigma_2, \ldots$ are positive scalars. When the v's and u's go into the columns of U and V respectively, orthogonality dictates $V^TV = I$ and $U^TU = I$. The σ's go into a diagonal matrix Σ.
- Just as $Ax_i = \lambda_i x_i$ led to the diagonalization $AS = S\Lambda$, the equations $Av_i = \sigma_i u_i$ tell us column by column that $AV = U\Sigma$

(m by n)(n by r)
equals
(m by r)(r by r)

$$A\begin{bmatrix} v_1 & \cdots & v_r \end{bmatrix} = \begin{bmatrix} u_1 & \cdots & u_r \end{bmatrix}\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

- The v's and u's account for the row and column space of A. We need $n-r$ more v's and $m-r$ more u's for the nullspace and left nullspace. They can be orthonormal bases for those two nullspaces (and then automatically orthogonal to the first r v's and u's.
- Include all v's and all u's in V and U so these matrices become square. We still have $AV = U\Sigma$.

(m by n)(n by n)
equals
(m by m)(m by n)

$$A\begin{bmatrix} v_1 & \cdots & v_r & \cdots & v_n \end{bmatrix} = \begin{bmatrix} u_1 & \cdots & u_r & \cdots & u_m \end{bmatrix}\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & & \end{bmatrix}$$

- The new Σ is m by n. It is just the old r by r matrix (Σ_r) with $m-r$ new zero rows and $n-r$ new zero columns. The real change is in the shapes U and V and Σ. Still $V^TV = I$ and $U^TU = I$, with sizes n and m.
- V is now a square orthogonal matrix, with inverse $V^{-1} = V^T$. So $AV = U\Sigma$ can become $A = U\Sigma V^T$. This is the Singular Value Decomposition

SVD $\quad A = U\Sigma V^T = u_1\sigma_1 v_1^T + \ldots + u_r\sigma_r v_r^T = \sigma_1 u_1 v_1 + \ldots + \sigma_r u_r v_r$

- $A = U_r \Sigma_r V_r^T$ is equally true and is the "reduced SVD". It gives the same splitting of A as a sum of rank 1 matrices.
- We will see that $\sigma_i^2 = \lambda_i$ is an eigenvalue of A^TA and AA^T. When we put the singular values in descending order $\sigma_1 \geq \sigma_2 \geq \dots \sigma_r > 0$, it gives the r rank one pieces of A in order of importance
- Example 1
  - When is $U\Sigma V^T$ (singular values) the same as $S\Lambda S^{-1}$ (eigenvalues)
    We need orthonormal eigenvectors in $S = U$. We need non negative eigenvalues in $\Lambda = \Sigma$, so A must be a positive semidefinite (or definite) symmetric matrix $Q\Lambda Q^T$
- Example 2
  - If $A = xy^T$ with unit vectors x and y, what is the SVD of A?
    The reduced SVD is exactly $A = xy^T$ with rank $r=1$. It has $u_1 = x$, $v_1 = y$ and $\sigma_1 = 1$. For the full SVD, complete $u_1 = x$ to an orthonormal basis of u's and $v_1 = y$ to an orthonormal basis of v

Image Compression
- An image (specifically a raster image) can be described as a matrix of colors. Each color is a vector (RGB).
- Image compression is very useful. Here is a SVD approach
  - Say we had a 256 by 512 image.
  - We replace the 256 by 512 image by a matrix of rank 1. That is just a column times a row, which is only 256 + 512 pixels.
    This is a $(256)(512)/(256+512) \sim 170:1$ compression ratio.
  - This is our best case. Instead we could use a linear combination of 5 rank one matrices and still get a $\sim 34:1$ compression ratio.
  - Where does SVD come in? The best rank one approximation is given by $\sigma_1 u_1 v_1^T$, where σ_1 is the largest singular value. The best rank 5 approximation is $\sigma_1 u_1 v_1^T + \dots + \sigma_5 u_5 v_5^T$, where the σ's are in descending order.
- A library compresses a different matrix, where rows correspond to key words and columns to titles in the library. The entry in this word-title matrix is $a_{ij} = 1$ if word i is in title j (otherwise it is zero). We normalize columns so long titles don't get an advantage
  - Once the indexing matrix is created, it has to be compressed, perhaps with SVD.

The Bases and the SVD
- Start with a 2 by 2 matrix. Let its rank be $r=2$, so A is invertible. We want v_1 and v_2 to be perpendicular unit vectors. We also want $Av_1 = \sigma_1 u_1$ and $Av_2 = \sigma_2 u_2$ to be perpendicular. Then the corresponding unit vectors will be orthonormal

- Example

$$A = \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix}$$

  - No orthogonal matrix Q will make $Q^{-1}AQ$ diagonal. We need $\Sigma_1 = U^{-1}AV$. The two bases will be different. The singular values are the lengths $\|Av_1\|$ and $\|Av_2\|$

$AV = U\Sigma$ $\qquad A[v_1 \ v_2] = [\sigma_1 u_1 \ \ \sigma_2 u_2] = [u_1 \ u_2]\begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix}$

$A = U\Sigma_1 V^T$

  - To get V by itself, take $A^TA = (U\Sigma_1 V^T)^T(U\Sigma_1 V^T) = V\Sigma_1^T\Sigma_1 V^T$
    Multiplying those diagonal Σ_1^T and Σ_1 gives σ_1^2 and σ_2^2. Then
    Eigenvalues: σ_1^2, σ_2^2 $\qquad A^TA = V\begin{bmatrix} \sigma_1^2 & \\ & \sigma_2^2 \end{bmatrix}V^T$
    Eigenvectors: v_1, v_2
  - This is exactly like $A = Q\Lambda Q^T$
    - The symmetric matrix is A^TA
    - <u>The columns of V are the eigenvectors of A^TA</u>
- Likewise, the columns of U are the eigenvectors of AA^T. We can also compute them by:
  $Av_i = \sigma_i u_i \implies u_i = Av_i / \sigma_i$

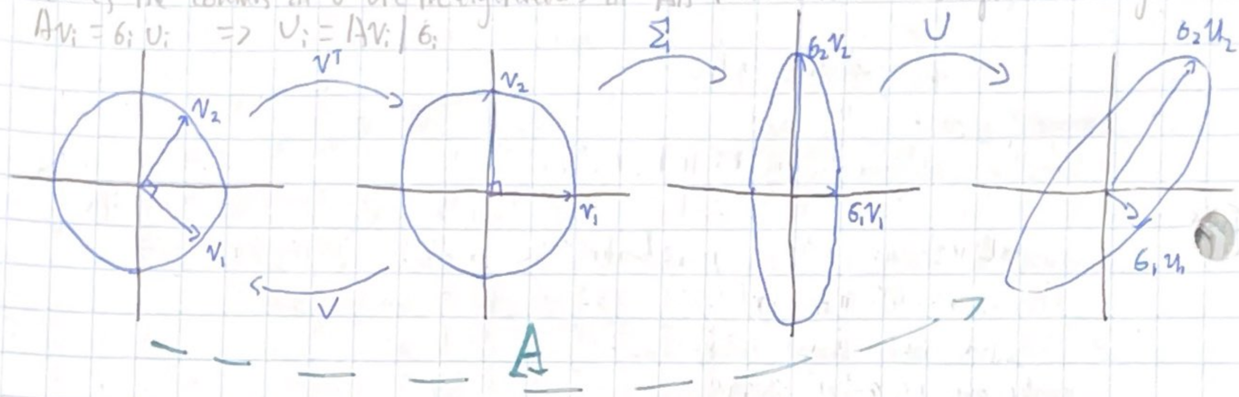

fig 6.7.1: U and V are rotations and reflections, Σ_1 stretches the circle than ellipse.

- Example
  - Find the SVD of $A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$
  - Compute A^TA and its orthonormal eigenvectors

$$A^TA = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} \text{ has unit eigenvectors } v_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \text{ and } v_2 = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

  - The eigenvalues are 8 and 2. The v's are orthogonal because A^TA is automatically symmetric and symmetric matrices have perpendicular eigenvectors.
  - We find u_i, we normalize Av_i

$$Av_1 = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix} \implies u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

The singular value is $2\sqrt{2}$.
Note $\sigma_1^2 = (2\sqrt{2})^2 = 8 = \lambda_1$

$$Av_2 = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix} \implies u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The singular value is $\sqrt{2}$
Note $\sigma_2^2 = (\sqrt{2})^2 = 2 = \lambda_2$

  - Then,

$$A = U\Sigma_1 V^T \text{ is } \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}\begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

- We could also have found the u's as the eigenvectors of AA^T

  Use $V^TV = I$ $\quad AA^T = (U\Sigma_1 V^T)(V\Sigma_1^T U^T) = U\Sigma_1\Sigma_1^T U^T$

  $AA^T = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix} \Rightarrow u_1 = (1,0),\ u_2 = (0,1),\ \sigma_1^2 = 8,\ \sigma_2^2 = 2$

- Example
  - Find the SVD of $A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$
  - The row space has only one basis vector $v_1 = \frac{1}{\sqrt{2}}(1,1)$. Same with the column space; $u_1 = \frac{1}{\sqrt{5}}(2,1)$. Then $Av_1 = (4,2)/\sqrt{2} = \sigma_1 u_1$. It does, with $\sigma_1 = \sqrt{10}$

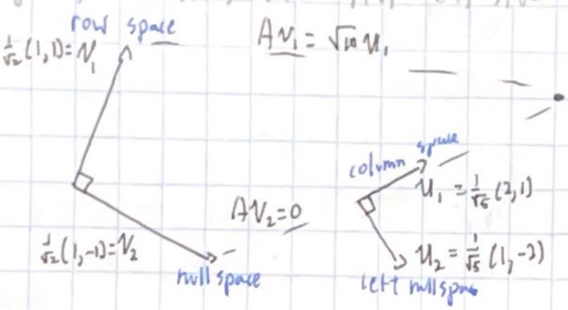

Fig 6.7.2: The SVD chooses orthonormal bases for 4 subspaces so that $Av_i = \sigma_i u_i$

  - The SVD could stop after the column and row space but it is customary for U and V to be square.
  - The vector v_2 is in the null space. It is perpendicular to v_1 in the row space, so we can chose $\frac{1}{\sqrt{2}}(1,-1)$. Likewise for u_2, which we can choose as $\frac{1}{\sqrt{5}}(1,-2)$ to make it perpendicular to u_1. We set σ to zero to get

    $A = U\Sigma V^T$ Full size $\quad \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

- The matrices U and V contain orthonormal bases for all 4 fundamental subspaces
  - first r columns of V: row space of A
  - last $n-r$ columns of V: nullspace of A
  - first r columns of U: column space of A
  - last $m-r$ columns of U: left nullspace of A
- As long as the last v's and u's in the nullspaces are orthonormal, the SVD will be correct.
- Proof of the SVD
  - Start from $A^TAv_i = \sigma_i^2 v_i$, which gives the v's and σ's. Multiplying by v_i^T leads to $v_i^TA^TAv_i = v_i^T\sigma_i^2 v_i \Rightarrow (Av_i)^T(Av_i) = \sigma_i^2 v_i^Tv_i \Rightarrow \|Av_i\|^2 = \sigma_i^2\|v_i\|^2 = 1$

    $\|Av_i\|^2 = \sigma_i^2$ so $\|Av_i\| = \sigma_i$

    Multiply by A

    $AA^TAv_i = \sigma_i^2 Av_i$ gives $(AA^T)(Av_i) = \sigma_i^2(Av_i) \Rightarrow Av_i$ as an eigenvector of AA^T.

    Then $u_i = Av_i/\sigma_i$ as a unit eigenvector of AA^T

Problem Set 6.7

1.

$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$

$A^TA = \begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix}$

$\det(A^TA - \lambda I) = (10-\lambda)(40-\lambda) - (20)(20)$
$= \lambda^2 - 50\lambda + 400 - 400$
$= \lambda^2 - 50\lambda$
$= \lambda(\lambda - 50)$

Set $\det(A^TA - \lambda I) = 0$

$\lambda = 0, 50$

case $\lambda = 0$
$(A^TA - 0I)x_1 = 0$
$\begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix} x_1 = 0$
$x_1 = (2, -1)$
$v_1 = \frac{1}{\sqrt{5}}(2, -1)$

case $\lambda = 50$:
$(A^TA - 50I)x_2 = 0$
$\begin{bmatrix} -40 & 20 \\ 20 & -10 \end{bmatrix} x_2 = 0$
$x_2 = (1, 2)$
$v_2 = \frac{1}{\sqrt{5}}(1, 2)$

$Av_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \sigma_1 = 0,\ u_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$Av_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix} 5 \\ 15 \end{bmatrix} = \sqrt{5}\begin{bmatrix} 1 \\ 3 \end{bmatrix}$

$u_2 = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \sigma_2 = \sqrt{50}$

$\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{10}}{10} & \frac{3\sqrt{10}}{10} \\ \frac{3\sqrt{10}}{10} & -\frac{\sqrt{10}}{10} \end{bmatrix}\begin{bmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & -\frac{\sqrt{5}}{5} \end{bmatrix}$

choose to be orthogonal to u_1 / v_1

2.

$C(A^T) = c\left(\frac{\sqrt{5}}{5}, \frac{2\sqrt{5}}{5}\right)$
$C(A) = c\left(\frac{\sqrt{10}}{10}, \frac{3\sqrt{10}}{10}\right)$
$N(A) = c\left(\frac{2\sqrt{5}}{5}, -\frac{\sqrt{5}}{5}\right)$
$N(A^T) = c\left(\frac{3\sqrt{10}}{10}, -\frac{\sqrt{10}}{10}\right)$

7.

Largest $\sigma_i = \sqrt{3}$

$u_1 \sigma_1 v_1^T = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)(\sqrt{3})\begin{bmatrix} \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{6} \end{bmatrix}$

$= \sqrt{3}\begin{bmatrix} \frac{\sqrt{3}}{6} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{6} \\ \frac{\sqrt{3}}{6} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{6} \end{bmatrix}$

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

4.

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

$A^TA = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$

$\det(A^TA - \lambda I) = (2-\lambda)(1-\lambda) - (1)(1)$
$= \lambda^2 - 3\lambda + 2 - 1$
$= \lambda^2 - 3\lambda + 1$

$\lambda_1 = \frac{3-\sqrt{5}}{2}, \ \lambda_2 = \frac{3+\sqrt{5}}{2}$

case $\lambda_1 = \frac{3-\sqrt{5}}{2}$
$\begin{bmatrix} 2-\lambda_1 & 1 \\ 1 & 1-\lambda_1 \end{bmatrix} x_1 = 0$
$x_1 = \left(\frac{1-\sqrt{5}}{2}, 1\right)$
$a_1 =$

6.

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

$A^TA = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

$\det(A^TA - \lambda I) = (1-\lambda)(2-\lambda)(1-\lambda) - (1)(1)(1-\lambda) - (1-\lambda)(1)(1)$
$= (1-\lambda)^2(2-\lambda) - 2(1-\lambda)$
$= (\lambda^2 - 2\lambda + 1)(2-\lambda) - 2 + 2\lambda$
$= 2\lambda^2 - 4\lambda + 2 - \lambda^3 + 2\lambda^2 - \lambda - 2 + 2\lambda$
$= -\lambda^3 + 4\lambda^2 - 3\lambda$
$= -\lambda(\lambda^2 - 4\lambda + 3)$
$= -\lambda(\lambda - 3)(\lambda - 1)$

Set $\det(A^TA - \lambda I) = 0$

$\lambda = 0, 1, 3$

case $\lambda = 0$:
$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} x_1 = 0$
$x_1 = (1, -1, 1)$
$v_1 = \left(\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right)$

case $\lambda = 1$
$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} x_2 = 0$
$x_2 = (1, 0, -1)$
$v_2 = \left(\frac{\sqrt{2}}{2}, 0, -\frac{\sqrt{2}}{2}\right)$

case $\lambda = 3$
$\begin{bmatrix} -2 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & -2 \end{bmatrix} x_3 = 0$
$x_3 = (1, 2, 1)$
$v_3 = \left(\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6}\right)$

$AA^T = \begin{bmatrix}1&1&0\\0&1&1\end{bmatrix}\begin{bmatrix}1&0\\1&1\\0&1\end{bmatrix} = \begin{bmatrix}2&1\\1&2\end{bmatrix}$

$\det(AA^T - \lambda I) = (2-\lambda)(2-\lambda) - (1)(1)$
$= \lambda^2 - 4\lambda + 4 - 1$
$= \lambda^2 - 4\lambda + 3$
$= (\lambda - 3)(\lambda - 1)$

Set $\det(AA^T - \lambda I) = 0$
$\lambda = 1, 3$

Case $\lambda = 1$: $\begin{bmatrix}1&1\\1&1\end{bmatrix} x_1 = 0$ Case $\lambda = 3$: $\begin{bmatrix}-1&1\\1&-1\end{bmatrix} x_2 = 0$

$x_1 = (1, -1)$ $x_2 = (1, 1)$
$u_2 = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$ $u_1 = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$

$Av_3 = \sigma_1 u_1 \Rightarrow \begin{bmatrix}1&1&0\\0&1&1\end{bmatrix}\begin{bmatrix}\frac{\sqrt{6}}{6}\\ \frac{\sqrt{6}}{3}\\ \frac{\sqrt{6}}{6}\end{bmatrix} = \begin{bmatrix}\frac{\sqrt{6}}{2}\\ \frac{\sqrt{6}}{2}\end{bmatrix} = \sqrt{3}\begin{bmatrix}\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\end{bmatrix}$

$\sigma_1 = \sqrt{3}$

$Av_2 = \sigma_2 u_2 \Rightarrow \begin{bmatrix}1&1&0\\0&1&1\end{bmatrix}\begin{bmatrix}\frac{\sqrt{2}}{2}\\ 0\\ -\frac{\sqrt{2}}{2}\end{bmatrix} = \begin{bmatrix}\frac{\sqrt{2}}{2}\\ -\frac{\sqrt{2}}{2}\end{bmatrix}$

$\sigma_2 = 1$

$A = U\Sigma V^T \Rightarrow \begin{bmatrix}1&1&0\\0&1&1\end{bmatrix} = \begin{bmatrix}\frac{\sqrt{2}}{2}&\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}&-\frac{\sqrt{2}}{2}\end{bmatrix}\begin{bmatrix}\sqrt{3}&0&0\\0&1&0\end{bmatrix}\begin{bmatrix}\frac{\sqrt{6}}{6}&\frac{\sqrt{6}}{3}&\frac{\sqrt{6}}{6}\\ \frac{\sqrt{2}}{2}&0&-\frac{\sqrt{2}}{2}\\ \frac{\sqrt{3}}{3}&-\frac{\sqrt{3}}{3}&\frac{\sqrt{3}}{3}\end{bmatrix}$

SVD of A:

$A = \begin{bmatrix}-1&3\\2&-6\end{bmatrix} = U\Sigma V^T = \begin{bmatrix}\frac{\sqrt{5}}{5}&\frac{2\sqrt{5}}{5}\\ \frac{2\sqrt{5}}{5}&\frac{\sqrt{5}}{5}\end{bmatrix}\begin{bmatrix}5\sqrt{2}&0\\0&0\end{bmatrix}\begin{bmatrix}\frac{\sqrt{10}}{10}&-\frac{3\sqrt{10}}{10}\\ \frac{3\sqrt{10}}{10}&\frac{\sqrt{10}}{10}\end{bmatrix}$

$A^TA = \begin{bmatrix}-1&2\\3&-6\end{bmatrix}\begin{bmatrix}-1&3\\2&-6\end{bmatrix} = \begin{bmatrix}5&-15\\-15&45\end{bmatrix}$

$\det(A^TA - \lambda I) = (5-\lambda)(45-\lambda) - (-15)(-15)$
$= \lambda^2 - 50\lambda + 225 - 225$
$= \lambda^2 - 50\lambda$
$= \lambda(\lambda - 50)$

set $\det(A^TA - \lambda I) = 0 \Rightarrow \lambda = 0, 50$

case $\lambda = 50$:
$(A^TA - 50I)x_1 = 0$
$\begin{bmatrix}-45&-15\\-15&-5\end{bmatrix} x_1 = 0$
$x_1 = (1, -3)$
$v_1 = (\frac{\sqrt{10}}{10}, -\frac{3\sqrt{10}}{10})$

case $\lambda = 0$
$(A^TA - 0I)x_2 = 0$
$\begin{bmatrix}5&-15\\-15&45\end{bmatrix} x_2 = 0$ $v_1^T v_2 = 0$
$x_2 = (3, 1)$
$v_2 = (\frac{3\sqrt{10}}{10}, \frac{\sqrt{10}}{10})$

$Av_1 = \sigma_1 u_1$:
$\begin{bmatrix}-1&3\\2&-6\end{bmatrix}\begin{bmatrix}\frac{\sqrt{10}}{10}\\ -\frac{3\sqrt{10}}{10}\end{bmatrix} = \begin{bmatrix}-\sqrt{10}\\ 2\sqrt{10}\end{bmatrix} = 5\sqrt{2}\begin{bmatrix}-\frac{\sqrt{5}}{5}\\ -\frac{2\sqrt{5}}{5}\end{bmatrix}$

$\sigma_1 = 5\sqrt{2}$, $u_1 = (\frac{\sqrt{5}}{5}, \frac{2\sqrt{5}}{5})$

v_2 is chosen to be orthogonal to v_1
u_2 is chosen to be orthogonal to u_1
$u_2 = (-\frac{2\sqrt{5}}{5}, \frac{\sqrt{5}}{5})$

SVD of A:

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

Procedure:
Find v_i and then
u from Av_i;
easier to number

$A^TA = \begin{bmatrix}1&0&1\\0&1&1\end{bmatrix}\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix} = \begin{bmatrix}2&1\\1&2\end{bmatrix}$

$\det(A^TA - \lambda I) = (2-\lambda)(2-\lambda) - (1)(1)$
$= \lambda^2 - 4\lambda + 4 - 1$
$= \lambda^2 - 4\lambda + 3$
$= (\lambda - 3)(\lambda - 1)$

Set $\det(A^TA - \lambda I) = 0 \Rightarrow \lambda = 1, 3$

Case $\lambda = 1$
$(A^TA - I)x_1 = 0$
$\begin{bmatrix}1&1\\1&1\end{bmatrix} x_1 = 0$
$x_1 = (1, -1)$
$v_1 = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$

case $\lambda = 3$
$(A^TA - 3I)x_2 = 0$
$\begin{bmatrix}-1&1\\1&-1\end{bmatrix} x_2 = 0$
$x_2 = (1, 1)$
$v_2 = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$

Find u and σ

$Av_1 = \sigma_1 u_1$:
$\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix}\begin{bmatrix}\frac{\sqrt{2}}{2}\\ -\frac{\sqrt{2}}{2}\end{bmatrix} = \begin{bmatrix}\frac{\sqrt{2}}{2}\\ -\frac{\sqrt{2}}{2}\\ 0\end{bmatrix} \Rightarrow \sigma_1 = 1,\ u_1 = \begin{bmatrix}\frac{\sqrt{2}}{2}\\ -\frac{\sqrt{2}}{2}\\ 0\end{bmatrix}$

$Av_2 = \sigma_2 u_2$:
$\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix}\begin{bmatrix}\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\end{bmatrix} = \begin{bmatrix}\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\\ \sqrt{2}\end{bmatrix} \Rightarrow \sigma_2 = \sqrt{3},\ u_2 = \begin{bmatrix}\frac{\sqrt{6}}{6}\\ \frac{\sqrt{6}}{6}\\ \frac{\sqrt{6}}{3}\end{bmatrix}$

Choose u_3 to be orthonormal to u_1 and u_2
$u_3 = \frac{1}{\sqrt{3}}(1, 1, -1) = (\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3})$

$A = U\Sigma V^T$:

$\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix} = \begin{bmatrix}\frac{\sqrt{2}}{2}&\frac{\sqrt{6}}{6}&\frac{\sqrt{3}}{3}\\ -\frac{\sqrt{2}}{2}&\frac{\sqrt{6}}{6}&\frac{\sqrt{3}}{3}\\ 0&\frac{\sqrt{6}}{3}&-\frac{\sqrt{3}}{3}\end{bmatrix}\begin{bmatrix}1&0\\0&\sqrt{3}\\0&0\end{bmatrix}\begin{bmatrix}\frac{\sqrt{2}}{2}&-\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}&\frac{\sqrt{2}}{2}\end{bmatrix}$

Table of Eigenvalues and Eigenvectors

Symmetric $A^T = A$	real λ's	orthogonal $x_i^T x_j = 0$
Orthogonal $Q^T = Q^{-1}$	all $\lvert\lambda\rvert = 1$	orthogonal $\bar{x}_i^T x_j = 0$
Skew-Symmetric $A^T = -A$	imaginary λ's	orthogonal $\bar{x}_i^T x_j = 0$
Complex Hermitian $\bar{A}^T = A$	real λ's	orthogonal $\bar{x}_i^T x_j = 0$
Positive Definite $x^T A x > 0 \ \forall x \neq 0$	all $\lambda > 0$	orthogonal since $\bar{A}^T = A$
Markov $m_{ij} > 0$, $\sum_{i=1}^{n} m_{ij} = 1$	$\lambda_{max} = 1$	Steady state $x > 0$
Similar $B = M^{-1} A M$	$\lambda(B) = \lambda(A)$	$x(B) = M^{-1} x(A)$
Projection $P = P^2 = P^T$	$\lambda = 1; 0$	column space; nullspace
Plane Rotation	$e^{i\theta}$ and $e^{-i\theta}$	$x = (1, i)$ and $(1, -i)$
Reflection $I - 2uu^T$	$\lambda = -1; 1, \dots, 1$	u; whole plane $u_\perp$
Rank One uv^T	$\lambda = v^T u; 0, \dots, 0$	u; whole plane $v^\perp$
Inverse A^{-1}	$1/\lambda(A)$	Same eigenvectors
Shift $A + cI$	$\lambda(A) + c$	Same eigenvectors
Stable powers $A^n \to 0$	all $\lvert\lambda\rvert < 1$	any eigenvectors
Stable exponential $e^{At} \to 0$	all $Re(\lambda) < 0$	any eigenvectors
Cyclic Permutation row 1 of I last	$\lambda_k = e^{2\pi i k/n}$	$x_k = (1, \lambda_k, \dots, \lambda_k^{n-1})$
Tridiagonal $-1, 2, -1$ on diagonals	$\lambda_k = 2 - 2\cos\frac{k\pi}{n+1}$	$x_k = (\sin\frac{k\pi}{n+1}, \sin\frac{2k\pi}{n+1}, \dots)$
Diagonalizable $A = S\Lambda S^{-1}$	diagonal of Λ	columns of S are independent
Symmetric $A = Q\Lambda Q^T$	diagonal of Λ (real)	columns of Q are orthonormal
Schur $A = QTQ^{-1}$	diagonal of T	columns of Q if $A^T A = AA^T$
Jordan $J = M^{-1} A M$	diagonal of J	each block gives $x = (0, \dots, 1, \dots, 0)$
Rectangular $A = U\Sigma V^T$	rank(A) = rank(Σ)	eigenvectors of $A^T A$, AA^T in V, U

Chapter 7: Linear Transformations

Chapter 7.1: The Idea of a Linear Transformation

- When the matrix A multiplies a vector v, it "transforms" into another vector Av. In goes v, out comes Av. $T(v) = Av$
  - A transformation is very much like a function. To find a specific output, we simply evaluate Av.
  - The deeper goal is to see all v's at once. We are transforming the entire space V when we multiply all vectors v.
- Start with a matrix A. It transforms vector v into Av and vector w into Aw. Then we know what it does to $u = v + w$; it equals $Av + Aw$. Matrix multiplication $T(v) = Av$ gives a linear transformation.
- A transformation T assigns an output $T(v)$ to each input v in V. It is linear if it meets the below requirements for all v and w

a) $T(v + w) = T(v) + T(w)$ b) $T(cv) = cT(v)$ for all c