

AM 120 Notes

Bannus Van der Kloot

November 13, 2012

Contents

1	September 6 Lecture	3
1.1	Geometric interpretation	3
1.2	Algorithmic approach	4
2	September 11 Lecture	6
2.1	General Example	8
3	September 13 Lecture	9
4	September 18 Lecture	12
5	September 20 Lecture	14
5.1	Finite precision	15
6	September 25 Lecture	16
6.1	Floating point nonsense	16
6.2	Catastrophic cancellation	17
7	September 27 Lecture	18
7.1	Ill-conditioned Matrices	18
8	October 2 Lecture	20
8.1	Over/underconstrained Systems	20
8.2	Vector Spaces	21
9	October 4 Lecture	23
10	October 9 Lecture	26
11	October 11 Lecture	30
11.1	Fundamental Theorem of Linear Algebra	30

12 October 16 Lecture	33
12.1 Fourier series	34
13 October 18 (Guest) Lecture	36
13.1 Face Recognition	36
14 October 23 Lecture	38
14.1 Discrete Fourier Series	38
14.2 Least-Squares Minimization	39
14.3 Physics Example	39
15 October 25 Lecture	41
16 October 30 Lecture	42
17 November 1 Lecture	46
17.1 Eigenvectors	48
18 November 6 Lecture	49
18.1 Eigenvalues and Eigenvectors	49
19 November 8 Lecture	53
20 November 13 Lecture	55
20.1 Quiz Questions	55
20.2 Spectral Theorem	55
20.3 Singular Value decomposition	55
21 Midterm	56

1 September 6 Lecture

There are two main problems that we will learn how to handle in this class.

1. Find $x \in R^n$ such that $Ax = b$. A is m by n matrix, $b \in R^n$ vector
2. Find x and λ such that $Ax = \lambda x$

Example
$$\begin{aligned} x + 2y &= 3 \\ 4x + 5y &= 6 \end{aligned}$$

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

There are 3 ways to solve:

1. $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \end{bmatrix} \Rightarrow y = 2, x = -1$
2. $A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} 5 & -2 \\ -4 & -1 \end{bmatrix} \det A = -3$
 $x = A^{-1}b \Rightarrow \frac{1}{-3} \begin{bmatrix} +3 \\ -6 \end{bmatrix}$
3. Kramer's rule

Summary Topics covered in next 3 classes:

1. Geometric interpretation of solving linear systems
2. Matrix notation (LU factorization)
3. Singular cases (no solution, multiple soln's)
4. Efficient way to solve $Ax = b$ using computers

1.1 Geometric interpretation

Example Graphical method:

Row interpretation (plot lines on coordinate system):

$$\begin{aligned} 2x - y &= 1 \\ x + y &= 5 \end{aligned} \quad \text{Solution: } x = 2, y = 3$$

Column interpretation:

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} x + \begin{bmatrix} -1 \\ 1 \end{bmatrix} y = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

Example 3 by 3 system:

Each row represents a plane:

$$2u + v + w = 5$$

$$4u - 6v + 0 = -2$$

$$-2u + 7v + 2w = 9$$

Remember: inner product of vector with another vector equals 0 \Rightarrow orthogonal.

Column interpretation:

$$\begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix} u + \begin{bmatrix} 1 \\ -6 \\ 7 \end{bmatrix} v + \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} w = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}$$

Example Overdetermined system:

$$\begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 7 \end{bmatrix}$$

Solution: $c = 1, d = 1$

In 4 dimensions, the rows represent 3-spaces, which are ‘flat’ relative to 4 dimensional space. If we intersect $(x, y, z, t = 0)$ with $(x, y, z = 0, t)$, two three spaces, we get (x, y) plane.

$$a_1u + a_2v + a_3w + a_4z = b$$

$$A = (a_1|a_2|a_3|a_4)$$

1.2 Algorithmic approach

Generalizing to n by n . How to solve $Ax = b$ in a way that scales well? Gaussian elimination (row reduction).

$$2u + v + w = 5$$

$$4u - 6v + 0 = -2$$

$$-2u + 7v + 2w = 9$$

$$\hline 2u + v + w = 5$$

$$\Rightarrow -8v - 2w = -12$$

$$8v + 3w = 14$$

$$\hline 2u + v + w = 5$$

$$\Rightarrow -8v - 2w = -12$$

$$w = 2$$

$$\hline \Rightarrow v = 1, u = 1$$

We need a process that takes:

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

$$x = \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

$$b = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}$$

... this $Ax = b$ problem and transforms it to a $Ux = \hat{b}$ problem. We can get an upper triangular matrix, and obtain solution by back substitution.

Problems One issue that could arise is if the bottom row is all 0s: infinitely many solutions.

2 September 11 Lecture

Last class:

- Introduced first central problem of linear algebra: solving linear equations
- Studied column and row interpretation of linear systems
- Introduced Gaussian elimination

Example (from previous class) Row/Column interpretation.

$$2u + v + w = 5$$

$$4u - 6v + 0 = -2$$

$$-2u + 7v + 2w = 9$$

Row: Three planes intersecting. Column: linear combination of three vectors

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

We were trying to figure out how to transform matrix A into an upper triangular matrix.

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix} = \begin{bmatrix} 5 \\ -12 \\ 9 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ -2 & 7 & 2 \end{bmatrix}$$

Matrix operations Addition is associative: $A+B+C = (A+B)+C = A+(B+C)$

Multiplication: dimension $m \times n$ multiplied by $n \times p$ results in $m \times p$ matrix.
 $AB \neq BA$.

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 7 & 8 \\ 2 & 3 \end{bmatrix}$$
$$\begin{bmatrix} 2 & 3 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 8 & 7 \end{bmatrix}$$

Matrix multiplication:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & & & \ddots & \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n a_{1i}x_i \\ \sum_{i=1}^n a_{2i}x_i \\ \vdots \\ \sum_{i=1}^n a_{ni}x_i \end{bmatrix}$$

$$\begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

$$\begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} | & | \\ b_1 & b_2 \\ | & | \end{bmatrix} = \begin{bmatrix} | & | \\ Ab_1 & Ab_2 \\ | & | \end{bmatrix}$$

Row reduction In matrix form

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

1. Subtract 2 times row 1 to row 2

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}_{E_{21}} \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}_A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ -2 & 7 & 2 \end{bmatrix}$$

2. Subtract -1 times row 1 to row 3

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}_{E_{31}} E_{21}A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 8 & 3 \end{bmatrix}$$

3. Subtract -1 times row 2 to row 3

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}_{E_{32}} E_{31}E_{21}A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

Originally we wanted to solve $Ax = b$. Now we have:

$$E_{32}E_{31}E_{21}A = U$$

where U is an upper triangular matrix.

$$E_{32}E_{31}E_{21}Ax = Ux$$

Let's let $E_{32}E_{31}E_{21} = L$. Then, we have

$$\begin{aligned} L^{-1}A &= U \\ A &= LU \\ Ux &= C = E_{32}E_{31}E_{21}b \end{aligned}$$

Now we can solve by back substitution.

$$L^{-1} = E_{32}E_{31}E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}$$

Matrix inverse properties:

$$(AB)^{-1} = B^{-1}A^{-1}$$

$$(A_1A_2 \dots A_n)^{-1} = A_n^{-1} \dots A_2^{-1}A_1^{-1}$$

So we have:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} = E_{21}^{-1}E_{31}^{-1}E_{32}^{-1}$$

Row reduction matrices A matrix that subtracts l times row j from row i is such that it includes $-l$ in row i , column j .

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}_L \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}_U$$

L is lower triangular and U is upper triangular.

1. Compute LU factorization
2. Solve for c in $Lc = b$ (forward substitution)
3. Solve for x in $Ux = c$ (back substitution)

We want to solve $Ax = b$. We factor to get $LUx = b$. First we find c such that $Lc = b$

2.1 General Example

$$\begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \Rightarrow \begin{aligned} c_1 &= b_1/l_{11} \\ c_2 &= b_2 - b_1l_{21}/l_{11} \\ c_3 &= b_3 - l_{31}b_1 - l_{32}(b_2 - b_1l_{21}) \end{aligned}$$

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \Rightarrow \begin{aligned} x_3 &= c_3/u_{33} \\ x_2 &= \frac{1}{u_{22}}(c_2 - u_{23}c_3/u_{33}) \\ x_1 &= \dots \end{aligned}$$

3 September 13 Lecture

Announcements

- Matlab tutorials (sections)
- Final projects
 - Adjustment based on class size
 - Pairs
- Assignment 1 due Fri @ 7pm in Pierce 303
- Collaboration policy

From last time:

- Linear equations \rightarrow Matrix notation
- Column j of $AB = Ab_j$

$$A \begin{bmatrix} | & | & \cdots & | \\ b_1 & b_2 & \cdots & b_n \\ | & | & \cdots & | \end{bmatrix} = [Ab_1 \quad Ab_2 \quad \cdots \quad Ab_n]$$

- Introduced the LU factorization of square matrix A (see general example at end of last lecture)

$$Ax = b \Rightarrow LUx = b$$

1. Find LU
2. Solve for c in $Lc = b$
3. Solve for x in $Ux = c$

Example LU factorization

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

1. Subtract 2 times row 1 to row 2

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}_{E_{21}} A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 4 & 5 \end{bmatrix}$$

$$E_{21}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2. Subtract 3 times row 1 to row 2

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}_{E_{31}} E_{21} A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 4 & 2 \end{bmatrix}$$

3. Subtract 2 times row 2 to row 3

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}_E 32 E_{31} E_{21} A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}_U$$

$$L^{-1} = E_{32} E_{31} E_{21} = \begin{bmatrix} 1 & 0 & 0 & -2 & 1 & 0 & -3 & -2 & 1 \end{bmatrix}$$

$$\begin{aligned} L^{-1} A &= U \\ L &= E_{21}^{-1} E_{31}^{-1} E_{32}^{-1} \\ L &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 0 \end{bmatrix} \end{aligned}$$

Generalizing LU factorization To $n \times n$ matrix:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \ddots & & & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}$$

1. Introduce zeros below a_{11} by subtracting multiples of row 1
2. Use multipliers $l = \frac{a_{i1}}{a_{11}}$
3. Repeat 1 and 2 for $a_{22}^*, a_{33}^*, \dots$

Step 1:

$$\begin{array}{c|cccc}
 a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\
 \hline
 0 & a_{22}^* & a_{23}^* & \dots & a_{2n}^* \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 0 & a_{n2}^* & a_{n3}^* & \dots & a_{nn}^*
 \end{array}$$

Step 2:

$$\begin{array}{c|cc|cc}
 a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\
 \hline
 0 & a_{22}^* & a_{23}^* & \dots & a_{2n}^* \\
 \vdots & 0 & a_{33}^* & \dots & a_{3n}^* \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 0 & 0 & a_{n3}^* & \dots & a_{nn}^*
 \end{array}$$

How many operations does this algorithm use?

$$\sum_{k=1}^n k^2 - \sum_{k=1}^n k = \frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2}$$

4 September 18 Lecture

To review: solving $Ax = b$:

1. Find $LU = A$
2. Solve for c in $Lc = b$ (forward substitution)
3. Solve for x in $Ux = c$ (back-subst)

Multipliers to find U are entries of L .

What is the # of operations needed to get LU factorization?

$$\approx \frac{n^3 - n}{3}$$

Forward substitution Number of operations:

$$(n-1) + (n-2) + \dots + 1 \approx O(n^2)$$

Back substitution is similar process (also $O(n^2)$). Most time consuming place is step 1.

Algorithm Failure This $Ax = b$:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

has solution $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. However, our algorithm won't find the answer because it can't switch rows. If the algorithm fails we have two options:

1. We need to rearrange rows
2. No solution
3. Infinitely many solutions

Example of (2):

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Example of (3):

$$\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Fact $\det(A) = \det(LU) = \det(L)\det(U)$

$$\det(U) = \prod_{i=1}^n u_{ii}$$

Example Consider this:

$$\begin{aligned} \begin{bmatrix} 0.0001 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ \begin{bmatrix} 0.0001 & 1 \\ 0 & -9999 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 1 \\ -9998 \end{bmatrix} \\ \Rightarrow x_2 &= \frac{9998}{9999} \\ 0.0001x_1 + \frac{9998}{9999} &= 1 \\ \Rightarrow x_1 &= \frac{10000}{9999} \end{aligned}$$

If we do all of this with limited precision (say 3 digits), we do the following:

$$\begin{aligned} \begin{bmatrix} 0.0001 & 1 \\ 0 & -10^4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 1 \\ -10^4 \end{bmatrix} \\ \Rightarrow x_2 &= 1 \end{aligned}$$

Then if we use the first equation, we get

$$\Rightarrow x_1 = 0$$

This is called **catastrophic cancellation**.

5 September 20 Lecture

First part of AM120 is to solve $Ax = b$ for arbitrary $\|A\| = n$.

$$u_{11} = a_{11}, u_{22} = a_{22}$$

Pseudocode did not have 0s in L and U . Second part of code is given L and b , should output c . Third part takes U and c and outputs x .

Assignment 2 Due on Monday morning (9am).

This Doolittle algorithm can fail:

1. If there is a pivot = 0
 - (a) System is singular $\Rightarrow \det(A) = 0$. This means there is no solution or infinitely many solutions.
 - (b) We can exchange rows and 'cure' system.

$$\det(A) = \det(L) \det(U) = 1 \prod_{k=1}^n u_{kk}$$

Example From last class:

$$\begin{bmatrix} 0.0001 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

This had true solution:

$$x_1 = \frac{10000}{9999}, x_2 = \frac{9998}{9999}$$

But with limited precision (three digit arithmetic), we got'

$$x_1 = 0, x_2 = 1$$

What if we switch the rows?

$$\begin{bmatrix} 1 & 1 \\ 0.0001 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 10^{-4} & 1 \end{bmatrix}_L \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 - 2 \cdot 10^{-4}t \end{bmatrix}$$

5.1 Finite precision

The computer represents a floating point number with a sign, exponent, and digits for the value itself. When we are talking about n -digit arithmetic, we are referring to the number of digits storing the value.

$$\begin{aligned} U &= \text{max exponent} \\ L &= \text{lowest exponent} \\ P &= \text{mantissa number of digits} \\ \beta &= \text{base} \end{aligned}$$

For example, for $L = -1, U = 1, p = 2$ and $\beta = 10$

$$(\text{sign})d_0.d_1d_2 \dots d_p \times 10^{\text{exponent}}$$

$$\begin{array}{ll} \text{largest} & 9.9 \times 10^1 = 99 \\ \text{smallest (non-zero)} & 1.0 \times 10^{-1} = 0.1 \end{array}$$

Examples of real values in floating point systems:

	β	P	L	U	
IEEE	2	24	-126	123	single
	2	53	-1022	1023	double
HP	10	12	-499	499	

What is the total number of floating point numbers?

$$2(\beta - 1)\beta^{p-1}(u - l + 1) + 1$$

Largest representable number:

$$(\beta - 1).(\beta - 1) \dots (\beta - 1) \cdot \beta^U$$

Smallest number (absolute value):

$$\beta^L(\text{underflow})$$

Machine precision Note that the difference between the real number and the floating point number chosen depends on exponent.

$$\forall x \in R, \exists \text{fl}(x) = \hat{x} \text{ such that } |x - \hat{x}| \leq \sum |x|$$

Note that this is not really true for all $x \in R$ – only within a certain range. ϵ_{mach} is the largest number s.t. $\text{fl}(1 + \epsilon_{\text{mach}}) > 1$.

Floating point numbers are not associative:

$$A + (B + C) \neq (A + B) + C$$

6 September 25 Lecture

Need to solve $Ax = b$. In theory, we can find A^{-1} . In practice, this is not necessary – it takes too many operations!

$$A \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ e_1 & e_2 & \cdots & e_n \\ | & | & \cdots & | \end{bmatrix} =$$

$$AA^{-1} = I$$

$$x = A^{-1}b$$

In practice, we find LU . Then we solve $Lc = b$

$$Lc = b$$

$$c = L^{-1}b$$

Then, we solve $Ux = c$:

$$Ux = cL^{-1}Ax = L^{-1}b$$

This algorithm fails if any of the pivots are 0.

Singular matrix A matrix is singular if $\det(A) \neq 0 \iff Ax = b$ has a unique solution.

6.1 Floating point nonsense

$$fl(1 + \epsilon) > 1 \Rightarrow \epsilon = \frac{1}{2}\beta^{1-\rho}$$

Machine precision (ϵ_{mach}): single 10^{-8} ; double 10^{-16} .

$$a = 1.23456 \times 10^2 a = -1.23455 \times 10^{-2}a = -1.11123 \times 10^{-3}a + b + c = (1 \times 10^{-3}) - 1.11123 \times 10^{-3}$$

This is the result if we calculate $(a+b)+c$. Notice that if we use 6-digit arithmetic and calculate $a + (b + c)$, we get 0! Floating point arithmetic is not associative.

$$\sum_{n=1}^{\infty} \frac{1}{n} \rightarrow$$

1. $\frac{1}{n}$ i UFL
2. $\sum_{n=1}^{k_2} \frac{1}{n}$ i OFL
3. $\text{fl}\left(\left(\sum_{n=1}^{k_3} \frac{1}{n}\right) + \frac{1}{k_3+1}\right) = \sum_{n=1}^{k_3} \frac{1}{n}$

6.2 Catastrophic cancellation

If we add a very small number to a big number, sometimes we get the same big number! Like $1.0 \times 10^8 + 1.0 \times 10^{-9} = 1.0 \times 10^8$.

$$A + b = A$$

$$A - b = A$$

Partial pivoting minimized catastrophic cancellation. *Bunch of matrices about the steps of partial pivoting, that I don't even think were accurate*

Theorem For a non-singular and square matrix A , $\exists P$ (permutation matrix) that reorders rows of A to avoid zeros in the pivot positions. $Ax = b$ has a unique solution and with the rows ordered "in advance."

$$PA = LU \text{ and } L \text{ and } U \text{ are unique}$$

7 September 27 Lecture

Last Class Theory: For a non-singular and square matrix A , $\exists P$ (permutation matrix) that reorders rows of A to avoid zeroes in the pivot positions

$Ax = b$ has a unique solution, and with the rows ordered “in advance.”

$$PA = LU \text{ where } L \text{ and } U \text{ are unique.}$$

Note: If A is singular, no P can produce a full set of pivots and elimination fails. Gaussian elimination with partial pivoting: if a pivot is zero, then A is singular

$$Ax = \begin{bmatrix} 0 & 4 & 1 \\ 1 & 1 & 3 \\ 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 6 \\ -1 \end{bmatrix} = b$$

$$PA = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 4 & 1 \\ 1 & 1 & 3 \end{bmatrix}, P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\tilde{A} = PA = LU = \begin{bmatrix} 1 & 0 & 00 & 1 & 01/2 & 1/2 & 1 \end{bmatrix}_{\tilde{L}} \begin{bmatrix} 2 & -2 & 10 & 4 & 10 & 0 & 6 \end{bmatrix}_{\tilde{U}}$$

7.1 Ill-conditioned Matrices

The presence of round-off error makes it difficult to identify singular matrices.

$$A = \begin{bmatrix} 1000 & 999 \\ 999 & 998 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ .999 & 1 \end{bmatrix}_L \begin{bmatrix} 1000 & 999 \\ 0 & -.001 \end{bmatrix}_U$$

$$Ax = b = \begin{bmatrix} 1999 \\ 1997 \end{bmatrix} \rightarrow x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

With limited precision (5-digit arithmetic), A appears singular, because $0.999(999)=998.00$.

$$Ax = \hat{b} = \begin{bmatrix} 1998.99 \\ 1997.01 \end{bmatrix} = b + \delta b = b + 10^{-2} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \Rightarrow x = \begin{bmatrix} 20.97 \\ -18.99 \end{bmatrix}$$

Small change in b and same A , x changes a lot. We call A ‘ill-conditioned,’ meaning that its ‘condition number,’ $k(A)$ is big. This is the definition of condition number:

$$\frac{\|\delta x\|}{\|x\|} \leq k(A) \frac{\|\delta b\|}{\|b\|}$$

Hilbert matrices are of this kind: changing b a little bit, x changes a lot (they are ill conditioned). Condition number of a singular matrix A is ∞ .

If $\frac{\|\delta x\|}{\|x\|} > 1$ we don’t expect to find a solution close to the one we were looking for.

In numerical calculations, singular matrices are indistinguishable from ill-conditioned matrix.

$$\begin{aligned} -\frac{d^2u}{dx^2} &= f(x), 0 \leq x \leq 1 \\ u(0) &= c_1 u(1) = c_2 \end{aligned}$$

$$u(x+h) = u(x) + hu'(x) + h^2 \frac{u''(x)}{2} + \dots + h^k \frac{d^k u}{dx^k}$$

Discretize interval into points x_i , solve for value of u at each point. At each point we can solve the problem as a linear algebra problem. We ignore higher terms, and say:

$$u'(x_i) \approx \frac{u(x_i+h) - u(x_i)}{h}$$

$$-u''(x_i) \approx -\frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1}))}{h^2} = f(x_i)$$

This results in a large matrix:

$$\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ 0 & -1 & 2 & -1 & \\ & & & \ddots & \\ & & & & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} = \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_m) \end{bmatrix}$$

8 October 2 Lecture

Last class Condition numbers. Blah blah blah blah.

We have implemented solving $Ax = b$ the same way that MATLAB's "\ " function works. Now we move on to other things.

8.1 Over/underconstrained Systems

$$-\frac{d^2u}{dx^2} = f(x), u(0) = \alpha, u(1) = \beta$$

Approximation of second derivative at discrete point x_i :

$$-\frac{d^2u(x_i)}{dx^2} \approx \frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1}))}{h^2}$$

$$h^2 f(x_i) \approx u(x_{i+1}) - 2u(x_i) + u(x_{i-1}))$$

At the boundary:

$$h^2 f(x_1) \approx u(x_2) - 2u(x_1) + \alpha$$

We got this from this, ignoring smaller terms:

$$u(x+h) = u(x) + \frac{du}{dx}(x)h + \frac{d^2u}{dx^2} \frac{h^2}{2} + \dots$$

This yields:

$$\begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ 0 & -1 & 2 & -1 \\ & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} u(x_1) \\ u(x_2) \\ \vdots \\ u(x_m) \end{bmatrix} = \begin{bmatrix} -\alpha + f(x_1)h^2 \\ f(x_2)h^2 \\ \vdots \\ -\beta + f(x_m)h^2 \end{bmatrix}$$

Need to solve $Ax = b$. Now we are studying methods that have to do with matrices that are not square! First, we'll say $m < n$:

$$\begin{bmatrix} & & & \\ & A & & \\ & & & \end{bmatrix}_{m \times n} \begin{bmatrix} x \\ \\ \\ \end{bmatrix}_{n \times 1} = \begin{bmatrix} b \\ \\ \end{bmatrix}_{m \times 1}$$

The left side of A has an $m \times m$ square section. This is underdetermined, so there are many solutions. What is $m > n$:

$$\begin{bmatrix} A \end{bmatrix}_{m \times n} \begin{bmatrix} x \end{bmatrix}_{n \times 1} = \begin{bmatrix} b \end{bmatrix}_{m \times 1}$$

This is an overconstrained system, which may not have a solution. A real example of this is fitting a line to points in a least-squares sense. If each point is at (t_i, y_i) , and we want to find a line $y = mt + b$, then we solve:

$$\begin{bmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{bmatrix}_A \begin{bmatrix} m \\ b \end{bmatrix}_x = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}_b$$

8.2 Vector Spaces

Vector spaces have two operations:

[Example: \mathbb{R}^n]

1. if $x, y \in \mathbb{R}^n$, $x + y = z \in \mathbb{R}^n$
2. if $x \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$, $\alpha x \in \mathbb{R}^n$

These properties must hold:

Addition: for $x, y, z \in \mathbb{R}^n$

1. Commutativity: $x + y = y + x$
2. Associativity: $x + (y + z) = (x + y) + z$
3. $\exists!$ zero vector s.t. $x + 0 = x \forall x \in \mathbb{R}^n$
4. $\exists! -x \in \mathbb{R}^n$ s.t. $x + (-x) = 0 \forall x \in \mathbb{R}^n$

Scalar multiplication

5. $1x = x$: 1 is scalar
6. $(c_1 c_2)x = c_1(c_2 x)$

¹! indicates there exists some unique zero vector

$$7. \ c(x + y) = cx + cy$$

$$8. \ (c_1 + c_2)x = c_1x + c_2x$$

A **subspace** is a non-empty subset of a vector space that satisfies all these properties and all linear combinations stay in the subspace.

Example

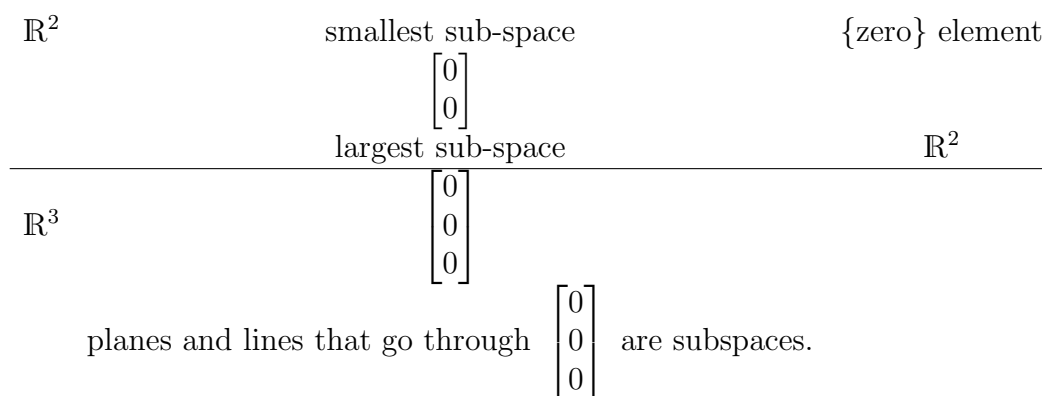
$$\begin{bmatrix} 1 & 0 \\ 5 & 4 \\ 4 & 4 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 4 \end{bmatrix} u + \begin{bmatrix} 0 \\ 4 \\ 4 \end{bmatrix} v$$

9 October 4 Lecture

Use ‘\’ operator in MATLAB from now on unless we specify otherwise. Assignment 4 is now due Tuesday in class.

Last class We discussed vector spaces, with the intention of being able to solve any linear algebra problem, whether it be overdetermined or underdetermined. Vector spaces have two operations, addition and scalar multiplication, with 8 axioms. Note: no notion of proximity or distance (no topology).

Subspace A non-empty subset of a vector space. Closed under addition and scalar multiplication. ($x + y \in \text{subspace}$, $ax \in \text{subspace}$).



$$A = \begin{bmatrix} 1 & 0 \\ 5 & 4 \\ 2 & 4 \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, Ax = \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix} x_1 + \begin{bmatrix} 0 \\ 4 \\ 4 \end{bmatrix} x_2$$

Column spaces of A (denoted $C(A)$) is the spaces that contains all linear combinations of the columns of A .

If we have a matrix $[A]_{M \times N}$ (with m rows, n columns), then $C(A) \in \mathbb{R}^m$.

b and $\tilde{b} \in C(A)$, $\exists x$ and \tilde{x}

$$\begin{aligned} Ax &= b & A(x + \tilde{x}) &= b + \tilde{b} \\ A\tilde{x} &= \tilde{b} \\ cb & & A(cx) &= cAx = cb \end{aligned}$$

Null space:

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

Null space of A consists of all vectors x such that $Ax = 0$, denoted $N(A) \in \mathbb{R}^n$.

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

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in N(A), \begin{bmatrix} 0 \\ 0 \end{bmatrix} \in N(A)$$

Theorem If zero is the only element of $N(A) \Rightarrow$ columns of A are linearly independent.

If $N(A) = \{0\}$ and A is a square matrix $\Rightarrow \exists! x$ such that $Ax = b$ for any b .

Basis for a **vector space** $V\{v_k\}$.

1. v_k 's are linearly independent
2. they span V (any $v \in V$ is a linear combination of the basis vectors $\{v_k\}$)

$\Rightarrow \exists!$ way to represent any element of V

$\text{dim}(V) = \#$ of basis vectors

The **complete solution** of a linear system of equations

$Ax = b$ is given by $x = x_p + x_n$ (if it exists)

where

$$Ax_p = b \text{ and } Ax_n = 0$$

$$A(x_p + x_n) = b + 0 = b$$

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad C(A) = k \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ line}$$

$$N(A) = d \begin{bmatrix} 2 \\ -1 \end{bmatrix} \text{ line}$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix} = Ax = b = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

$$x = \begin{bmatrix} 3 \\ 0 \end{bmatrix}_{x_p} + d \begin{bmatrix} 2 \\ -1 \end{bmatrix}_{x_n}$$

Theorem : For any $m \times n$ matrix $A \exists P$ (permutation) and L lower triangular matrix and an $m \times n$ Echelon matrix U such that $PA = LU$

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix}$$

Let's find LU :

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix}, U = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Note that L is $m \times m$ square, and U is also $m \times n$. Also, we see that U has 2 LI columns.

We already knew that the column space of A lives in \mathbb{R}^3 , so one column had to be dependent. Now we know that only two vectors in the 4 columns are LI, so the column space is a plane. Null space is also a plane, but it lives in \mathbb{R}^4 .

Another example:

$$\begin{aligned} (t_1, y_1)' &= (0, 0) \\ (t_2, y_2) &= (2, 1) \\ \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}_A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} y &= P(t) = x_1 + x_2 t \\ y_1 &= P(t_1) = x_1 + x_2 t_1 \\ y_2 &= P(t_2) = x_1 + x_2 t_2 \end{aligned}$$

We get $P = \frac{1}{2}t$

10 October 9 Lecture

The transpose of a matrix A (denoted A^T) is a matrix with columns directly from rows of A (the i th row becomes the i th column of A^T). $(AB)^T = B^T A^T$.

Column space of A Denoted $C(A)$. Contains all linear combinations of the columns of A . $C(A)$ is a subspace of \mathbb{R}^m .

Complete solution to the problem $Ax = b$ can be expressed as $x = x_p + x_n$. $x_n \in N(A)$, x_p a particular solution.

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

$$Ax_p = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

Should know

1. What is a vector space?
2. When a set of vectors are linearly independent?
3. Dimension of a vector space
4. Basis for a vector space

Theorem $N(A)$ contains only the vector iff zero the columns of A are linearly independent. Therefore, if A is square, $\exists! x$ for any b .

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix}_L \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}_U$$

The *row space* of A is the column space of A^T ($C(A^T)$). It is a subspace of \mathbb{R}^n .

The left null space of A is the null space of A^T . $y \in N(A^T)$ if $A^T y = 0$ iff $y^T A = 0$.

Let's calculate the null space of A . We've broken A into LU . Since U was obtained by adding and subtracting rows of A , it follows that $N(A) = N(U)$.

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}_U \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned}
3w + 3y &= 0 \Rightarrow w = -y \\
u + 3v - 3y + 2y &= 0 \\
u &= y - 3v
\end{aligned}$$

$$x \in N(A) = \begin{bmatrix} y - 3v \\ v \\ -y \\ y \end{bmatrix} = y \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix} + v \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

The linear combination of these two vectors generates a plane in \mathbb{R}^4 , the null space of A .

$$Ax = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}_U \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 - 2b_1 \\ (b_3 + b_1) - 2(b_2 - 2b_1) \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 - 2b_2 + 5b_1 \end{bmatrix}$$

$$\Leftrightarrow b_3 - 2b_2 + 5b_1 = 0 \text{ solvability condition}$$

Now, we choose:

$$\begin{aligned}
b &= \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix}, \tilde{b} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} \\
\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}_U \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} &= \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
3w + 3y &= 3 \Rightarrow w = 1 - y \\
u + 3v + 3(1 - y) + 2y &= 1 \Rightarrow u = -2 - 3v + y
\end{aligned}$$

Now we can write

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}_U \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} \Rightarrow x = \underbrace{\begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}}_{x_p} + v \underbrace{\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}}_{x_n} + y \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

Remarks:

- The null space of A , $N(A)$, and the row space of A , $C(A^T)$, are subspaces of \mathbb{R}^n .
- The left null space, $N(A^T)$, and column space of A , $C(A)$ are subspace of \mathbb{R}^m .

Transform $A \xrightarrow[\text{using Gauss elim}]{\quad} U$ we can immediately identify a basis for $C(A^T)$.

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

$$C(A) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ a line in } \mathbb{R}^2$$

$$\text{Row space of } A = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \text{ c lives in } \mathbb{R}^3$$

$N(A)$ lives in $\mathbb{R}^3 : Ax = 0$:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \vec{0}$$

We see that x_1 must be 0, but the other values are free:

$$x_1 = 0$$

$$x_2 = a$$

$$x_3 = b$$

$$x = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} a + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} b \text{ a plane in } \mathbb{R}^3$$

Left null space $A^T y = 0$:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \vec{0}$$

Dimension of row space corresponds with # of linearly independent rows. $C(A^T)$ dimension is r (r linearly independent rows). If we have $A_{m \times n}$, then $r \leq m$ and $r \leq n$. $C(A)$ dimension is r (even though they live in different spaces).

The dimension of $N(A)$ is $n - r$. Null space lives in \mathbb{R}^n

The dimension of $N(A^T)$ is $m - r$.

11 October 11 Lecture

Final projects: think about what brought you to study Applied Math. What problems do you like to solve? We'll find some linear algebra component to it.

Last class:

Row space of A $A \rightarrow U$, the ' r ' non zero rows are a basis for the row space $C(A^T)$. It has dimension r and it is a subspace of \mathbb{R}^n . The row space of U is the same as the row space of A , since they only differ by linear combinations of rows.

The row space of A and U have the same basis.

Null space of A $N(A) = N(U)$. If r rows are linearly independent \Rightarrow there are $(n - r)$ free variables the dimension of $N(A) = n - r$.

Null space definition:

$$x \in \mathbb{R}^n \text{ s.t. } Ax = 0$$

Column space of A $C(A)$. When we transform A to U , the first non-zero elements' index in each row determines which variable of x will be a pivot variable, suggesting that those indices determine the column space of A . The column space of A is *not* the same column space of U . Dimension of $C(A) = r$. $C(A)$ is a subspace of \mathbb{R}^n .

Left null space

$$\begin{bmatrix} y_1 & y_2 & \cdots & y_m \end{bmatrix} \begin{bmatrix} A^T \\ \end{bmatrix}_{m \times n} = \underbrace{\begin{bmatrix} 0 & \cdots & 0 \end{bmatrix}}_n \text{ it is a subspace of } \mathbb{R}^m$$

$N(A^T)$ is $A^T y = 0$. If y is in the null space of A^T , then y^T is in the left null space of A ($y^T A = 0$).

11.1 Fundamental Theorem of Linear Algebra

- $\dim(C(A)) = r$
- $\dim(C(A^T)) = r$
- $\dim(N(A)) = n - r$
- $\dim(N(A^T)) = m - r$

$Ax = b$. Case: $(m \leq n)$

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}_{m \times n} \begin{bmatrix} \\ x \\ \end{bmatrix}_{n \times 1} = \begin{bmatrix} \\ b \\ \end{bmatrix}_{m \times 1}$$

Existence: If A has the maximum number of linearly independent rows($= m$) A is said to have full “row” rank. There exists at least one solution for any b

In this case, A has a right inverse. An example:

$$A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix} C = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/5 \\ \alpha & \beta \end{bmatrix}$$

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

In this example, the right inverse is not unique! Any values of α and β will work.

$$ACb = b$$

$$Ax = bx \qquad \qquad \qquad = Cb$$

Uniqueness $m \geq n$

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}_{m \times n} \begin{bmatrix} \\ x \\ \end{bmatrix}_{n \times 1} = \begin{bmatrix} \\ b \\ \end{bmatrix}_{m \times 1}$$

If all columns of A are linearly independent, A is said to be full “column” rank.

If $b \in C(A) \exists$ 1 unique solution. If $b \notin C(A)$, then no solution.

\mathbb{R}^n : normed vector space (with an inner product)

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i$$

$$||x||^2 = \langle x, x \rangle = x^T x$$

Interesting property:

$$\langle x, y \rangle = ||x|| ||y|| \cos \theta$$

x and y are said to be orthogonal if $\langle x, y \rangle = 0$. If we have k non-zero vectors (v_1, \dots, v_k) are mutually orthogonal, they are linearly independent. Then we can say that there is only one combination that satisfies the following:

$$c_1 v_1 + c_2 v_2 + \dots + c_k v_k = 0$$

All c_i must be 0. Proof:

$$v_1^T (c_1 v_1 + c_2 v_2 + \dots + c_k v_k) = 0$$

$$v_1^T c_1 v_1 + v_1^T c_2 v_2 + \dots + v_1^T c_k v_k = 0$$

$$c_1 ||v_1||^2 = 0$$

Repeat for all $v_i \Rightarrow c_i = 0 \forall i \Rightarrow \{v_k\}$ are linearly independent.

$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad \text{1 at the } i\text{th element}$$

$\{e_i\}$ is an orthonormal basis of \mathbb{R}^n . Any vector in \mathbb{R}^n can be generated as a linear combination of them. If every vector in a subspace V is orthogonal to every vector in subspace $W \Rightarrow V$ and W are said to be orthogonal subspaces.

1. The row space of A , $C(A^T)$ is the orthogonal complement to $N(A)$.
2. The column space of A , $C(A)$ is the orthogonal complement to the left null space $N(A^T)$.
1. Why is this true? By definition, the null space are vectors x such that $Ax = 0$. If A is $m \times n$. This means that the inner product of every row of A with x must be 0. In other words, x is orthogonal to every row of A . The rows define the row space $C(A^T)$, so (1) is true.
2. The left null space is defined as y such that $y^T A = 0$. If A is $m \times n$, then the inner product of y and each of the n columns of A must be 0.

12 October 16 Lecture

Fundamental theorem of Linear Algebra

- The null space of A , $N(A)$, is the orthogonal complement of the row space of A (living in \mathbb{R}^n)
- The left null space $N(A^T)$ is the orthogonal complement of the column space of A

$$\min_{\hat{x} \in \mathbb{R}} \|a\hat{x} - b\|$$

$$\begin{aligned} b - \hat{x}a &\perp a \\ \langle a, b - \hat{x}a \rangle &= 0 \\ a^t(b - \hat{x}a) &= 0 \\ \hat{x} &= \frac{a^t b}{a^t a} \\ p &= \hat{x}a \end{aligned}$$

In \mathbb{R}^n :

$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad 1 \text{ at the } i\text{th element}$$

These $\{e_k\}$ form an orthonormal basis for \mathbb{R}^n .

$$\begin{aligned} \mathbb{R}^5 : v = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} &= 1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + 5 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \\ v &= c_1 e_1 + \cdots + c_5 e_5 \text{ where } c_i = \frac{\langle v, e_i \rangle}{\|e_i\|^2} \end{aligned}$$

12.1 Fourier series

Similarly, we can express an arbitrary function $f(x)$ as a sum:

$$f(x) = a_0 + \sum_{k=1}^{\infty} a_k \cos kx + b_k \sin kx$$

The basis functions are $\sin(kx)$ and $\cos(kx)$.

Define $f(x) : [-\pi, \pi] \rightarrow \mathbb{R}$.

$$\int_{-\pi}^{\pi} |f(x)|^2 < M$$

M is finite. If $f, g \in L^2([-\pi, \pi]) \Rightarrow \langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$

Functions that map from finite intervals to the real numbers form a vector space.

What should a_k, b_k and a_o be?

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) dx$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) dx$$

Euler's formula:

$$e^{ikx} = \cos kx + i \sin kx$$

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx} \text{ with } c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{ikx} dx$$

How do we bring this problem to finite dimension, so we can solve a problem of the form $Ax = b$?

Suppose you have a discrete signal on a fixed interval.

$$f(x) = c_0 + c_1 e^{ikx} + c_2 e^{i2kx} + c_3 e^{i3kx}$$

$$f(0) = f_0 = c_0 + c_1 + c_2 + c_3$$

$$f(\pi/2) = f_1 = c_0 + c_1 i - c_2 - c_3 i$$

$$f(\pi) = f_2$$

$$f(3\pi/2) = f_3$$

This gives us matrix A :

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix}$$

For use in problem $Ac = f$:

$$c = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}, f = \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \end{bmatrix}$$

Sunspots and crazy stuff.

13 October 18 (Guest) Lecture

J. Nathan Kutz from University of Washington.

If the determinant is 0 – BAD.

Forearm shiver. This is how you win a wrestling match.

1. $\det A - \lambda I \neq 0$. $\vec{X} = (A - \lambda I)^{-1} \cdot 0 = 0$
2. $\det A - \lambda I = 0$

$$A = \begin{bmatrix} 1 & 3 \\ -1 & 5 \end{bmatrix}$$

$$\begin{bmatrix} 1 - \lambda & 3 \\ -1 & 5 - \lambda \end{bmatrix} \vec{x} = 0$$

$$(1 - \lambda)(5 - \lambda) + 3 = 0\lambda^2 - 6\lambda + 8 = 0(\lambda - 2)(\lambda - 4) = 0\lambda = 2, 4$$

Protip: don't do algebra or spell in public.

$$\lambda = 2 : \begin{bmatrix} -1 & 3 \\ -1 & 3 \end{bmatrix} \vec{x} = 0$$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$x_1 = 3x_2$$

One equation, two unknowns \rightarrow infinite solutions.

$$\lambda = 4 : \begin{bmatrix} -3 & 3 \\ -1 & 1 \end{bmatrix} \vec{x} = 0$$

$$-x_1 + x_2 = 0$$

Dot products! Geometrically, indicates how much one is projected onto the other.
Help

13.1 Face Recognition

We can read images with the followig command:

```
A = imresize(double(rgb2gray(imread('pic','jpeg'))),[120 80])
```

Then, we reshape them into a single row:

```
a = reshape(A,1,120*80)
```

We use each of these rows to assemble an $n \times 120 \times 80$ matrix B . Then, we calculate $C = B^T B$.

We use the command `[V,D] = eigs(C,20,'lm')` to get the first 20 largest magnitude values.

The eigenvectors define a basis of faces. Each face has a different (hopefully unique) set of coefficients that can be used for face recognition.

14 October 23 Lecture

Announcements:

- Graded assignments (Rebecca at Pierce 318)
- No assignment this week

Last class we were talking about Fourier decomposition.

$$\begin{aligned}f(x) &= a_0 \sum_{k=1}^{\infty} a_k \cos(kx) + b_k \sin(kx) \\a_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx \\a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) dx \\b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) dx\end{aligned}$$

14.1 Discrete Fourier Series

When $n = 4$:

$$f = c_0 + c_1 e^{ix} + c_2 e^{2ix} + c_3 e^{3ix}$$

We saw that this becomes an interpolation problem:

$$\begin{aligned}A &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} \\Ac &= f \\A^{-1} &= \frac{1}{4} \bar{A} \\c &= A^{-1} f\end{aligned}$$

Note that \bar{A} is the conjugate transpose.

14.2 Least-Squares Minimization

If we have a group of points (t_i, y_i) , we want to find a line that fits these ‘best.’ If the line is given by $y = at + b$ A basic idea is to minimize $\sum |y_i - (at_i + b)|$ (the vertical error).

If we only consider lines running through the origin:

$$\begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} x = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

Find x such that $Ax = b$. If b is in $C(A)$, there is a unique solution. This would be the case in which all points are already on a line.

If this is not the case, what can we do? Say we have $(2, y_1), (3, y_2), (4, y_3)$. We want to minimize:

$$\|r\|^2 = (2x - y_1)^2 + (3x - y_2)^2 + (4x - y_3)^2 = J \quad = \|Ax - b\|^2$$

How do we maximize/minimize?

$$\begin{aligned} \frac{dJ}{dx} &= 0 = 4(2x - y_1) + 6(3x - y_2) + 8(4x - y_3) \\ x &= \frac{2y_1 + 3y_2 + 4y_3}{2^2 + 3^2 + 4^2} \end{aligned}$$

Alternately we can write this as the norm of the error. We note that $x^T = x$ since it just a scalar value.

$$\begin{aligned} J &= \langle ax - y, ax - y \rangle = (ax - y)^T(ax - y) \\ &= (xa^T - y^T)(ax - y) \\ &= a^T ax^2 - 2a^T yx + y^T y \\ \frac{dJ}{dx} &= 2xa^T a - 2a^T y = 0 \\ x &= \frac{a^T y}{a^T a} = \frac{\langle a, y \rangle}{\langle a, a \rangle} = \frac{\langle a, y \rangle}{\|a\|^2} \end{aligned}$$

We see that x is the projection of y onto the column space of A .

14.3 Physics Example

For a flying ball:

$$\begin{aligned}\frac{d^2y}{dt^2} &= -g \\ y &= \frac{-gt^2}{2} + at + b \\ y_i &= c_2 t_i^2 + c_1 t_i + c_0\end{aligned}$$

Say we had sampled many (t_i, y_i) points that we expect to be close to a polynomial path. If we have more than 3 measurements, we have an overdetermined system. This is how it will look in matrix form:

$$\begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_1 & t_1^2 \\ & \vdots & \\ 1 & t_n & t_n^2 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

We look for $x \in \mathbb{R}^m$ such that $\|Ax - b\|^2 = \langle Ax - b, Ax - b \rangle$ is minimized ($\delta J = 0$).

$$A^T Ax = A^T b \text{ "Normal equations"}$$

Simple example:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Something about residuals being perpendicular to something. $A^T r = 0$ means that r is perpendicular to every column of A . $A^T(b - Ax) = 0$ is the same expression as $A^T Ax = A^T b$.

This problem can be solved if the columns of A are linearly independent. We'll prove this next class.

15 October 25 Lecture

Final project ideas.

16 October 30 Lecture

Last class: least square problem. We have a set of m measurements and would like to represent them by a lower dimensional approximation. This lower dimensional approximation is typically proposed on theoretical (or practical) reasons

Today:

- Noisy data \longrightarrow smooth (polynomial representation)
- Complicated function \longrightarrow simple approx.

In the language of linear algebra: need to find a projection of higher dimension onto a lower-dimensional space.

Examples: last class, linear fit, parabolic fit

$$\begin{bmatrix} A \end{bmatrix}_{m \times n} \begin{bmatrix} x \end{bmatrix}_{n \times 1} \approx \begin{bmatrix} b \end{bmatrix}_{m \times 1}$$

$$\|r\|^2 = \|Ax - b\|^2$$

$$\Delta J = 0 : \hat{x} \text{ such that}$$

$$\|r\|^2 \text{ is minimized}$$

$$A^t A \hat{x} = A^T b \text{ Normal equations}$$

$$r \perp C(A)$$

$$A^T(r) = 0 \iff A^T(Ax - b) = 0$$

The (least squares) solution \hat{x} exists and is unique when $(A^T A)$ is invertible.

$$\hat{x} = (A^T A)^{-1} A^T b$$

We know that $A^T A$ is of size $n \times n$. How can we show that the null space of A is the same as the null space of $A^T A$?

Proof. If $x \in N(A)$, by definition $Ax = 0$. Therefore, $A^T Ax = A^T 0 = 0$.

If $x \in N(A^T A)$, by definition $A^T Ax = 0$. Therefore, $x^T A^T Ax = x^T 0 = 0$. This is by definition $\|Ax\|^2 = 0 \iff Ax = 0$. \square

If A is not well-conditioned (rows are close to linearly dependent), then the normal equations are very ill-conditioned.

Simple example Suppose I want to solve this problem:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Seems that there is no solution, since $b_3 = 1$. Suppose we calculate the residual $\|r\|^2 = \|Ax - b\|^2 = (x_1 - 1)^2 + (x_2 - 1)^2 + (0 - 1)^2$. Part of the residual must be 1, so to minimize we find \hat{x}_1 and \hat{x}_2 to 1.

If we were to use the normal equations:

$$A^T A \hat{x} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \hat{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

This yields the solutions we found above.

$$\begin{bmatrix} R_{n \times n} \\ 0 \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \begin{matrix} n \times 1 \\ m - n \times 1 \end{matrix}$$

If we have a problem with A in this form:

$$\begin{aligned} \|r\|^2 &= \|Rx - b_1\|^2 + \|0 - b_2\|^2 \\ \hat{x} &\text{ is a minimum if} \\ R\hat{x} &= b_1 \Rightarrow \|r\|^2 = \|b_2\|^2 \end{aligned}$$

When we wanted to solve $Ax = b$ previously, our method was to take A and transform it into LU . Then, we solved the system $Ux = L^{-1}b = \tilde{b}$.

When solving the least squares problem

$$\begin{aligned} A\hat{x} &\approx b \\ Q^T A\hat{x} &\approx Q^T b \\ \|r\|^2 &= \|Ax - b\|^2 = \|Q^T(Ax - b)\|^2 = \|Q^T Ax - Q^T b\|^2 = \|\tilde{r}\|^2 \end{aligned}$$

If we multiply the original problem by Q^T such that $Q^T Q = I$, then multiplying the problem solves the equivalent problem.

$$\begin{aligned} \|Qv\|^2 &= \langle Qv, Qv \rangle \\ &= (Qv)^T Qv \\ &= v^T Q^T Qv \\ &= v^T v = \|v\|^2 \end{aligned}$$

So we can transform any arbitrary A using Q (orthogonal transformation) like so:

$$Q^T A = \begin{bmatrix} R_{n \times n} \\ 0 \end{bmatrix}$$

Note that if we designate R as upper triangular, we can just do back substitution. Why are the normal equations not good to be solved by a computer?

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

We choose $0 < \epsilon < \sqrt{\epsilon_{\text{mach}}}$ and $\epsilon_{\text{mach}} < \epsilon$

This is ill conditioned, because the rows and columns are close to being linearly dependant.

$$A^T A = \begin{bmatrix} 1 + \epsilon^2 & 1 \\ 1 & 1 + \epsilon^2 \end{bmatrix}$$

Since ϵ^2 is less than machine precision, all of a sudden we have a singular matrix.

Q is **orthogonal** if $Q^T Q = I$.

- Rotations
- Reflections
-

Suppose I want to find the QR factorization of this matrix:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \rightarrow \begin{bmatrix} a_{11}^* & a_{12}^* \\ 0 & a_{22}^* \\ 0 & 0 \end{bmatrix}$$

Simple example:

$$Q \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \rightarrow \begin{bmatrix} \alpha \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \rightarrow \begin{bmatrix} \alpha \\ 0 \end{bmatrix}$$

Now he's making a giant matrix that I don't want to write out. Step 1 is $Q_1^T A$ to get a matrix that has all 0s below the first diagonal. Then the next step has $Q_2^T Q_1^T A$ which has all 0s below the second diagonal.

17 November 1 Lecture

Least squares problem Find \hat{x} such that $Ax \approx b$. A is an $m \times n$ matrix ($m \geq n$), x is n -vector, b is m -vector. In other words, $J(x) = \|Ax - b\|^2 = \|r\|^2$ (residual) is minimized. If A is full rank (column) then r must be perpendicular to $C(A)$.

$$\Rightarrow A^T r = A^T (Ax - b) = 0$$

$$\Rightarrow A^T A \hat{x} = A^T b$$

$$\Rightarrow \hat{x} = (A^T A)^{-1} A^T b$$

Normal equations

are the coeff (theory)

$P = A\hat{x}$ is the projection of b onto $C(A)$. If P' is the projection matrix, then $A\hat{x} = P'x = \underbrace{A(A^T A)^{-1} A^T}_{P'} b$.

Since $C(A)$ is the orthogonal complement of $N(A^T)$ (left-null space) in R^m :

$$b = P'b + (I - P')b$$

$$P'b \in C(A)$$

$$(I - P)b \in N(A^T)$$

We want to transform A into a matrix that has $R_{n \times n}$ at the top and $0_{m \times m-n}$ at the bottom. Q is an orthogonal transformation if $Q^T Q = I \Rightarrow Q^T = Q^{-1}$.

For any vector $v \in \mathbb{R}^m$:

$$\|Qv\|^2 = (Qv)^T Qv = v^T Q^T Qv = v^T v = \|v\|^2$$

Householder transformations

$$Ha = H \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

If we have A :

$$\begin{aligned}
A &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \ddots & \\ a_{m1} & & & a_{mn} \end{bmatrix} \\
H_1 A &= \begin{bmatrix} \alpha & * & \cdots & * \\ 0 & a_{22} & \cdots & a_{2n} \\ 0 & \ddots & & \\ \vdots & & & \\ 0 & & & a_{mn} \end{bmatrix} \\
H_2 H_1 A &= \begin{bmatrix} \alpha & * & \cdots & * \\ 0 & * & \cdots & * \\ 0 & 0 & \ddots & a_{3n} \\ \vdots & & & \\ 0 & 0 & & a_{mn} \end{bmatrix}
\end{aligned}$$

How to find H ?

$$\begin{aligned}
H &= I - 2 \frac{vv^T}{||v||^2} \text{ where } a = \alpha e_1 \\
\alpha &= -\text{sign} a_1 ||a||
\end{aligned}$$

Simple example:

$$\begin{aligned}
a &= \begin{bmatrix} 4 \\ 3 \end{bmatrix} \\
v &= \begin{bmatrix} 4 \\ 3 \end{bmatrix} + 5 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 9 \\ 3 \end{bmatrix} \\
Ha &= \left[\underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_I - 2 \frac{\begin{bmatrix} 9 \\ 3 \end{bmatrix} \begin{bmatrix} 9 & 3 \end{bmatrix}}{90} \right] \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix} - \frac{2(45)}{90} \begin{bmatrix} 9 \\ 3 \end{bmatrix} \\
&= \begin{bmatrix} -5 \\ 0 \end{bmatrix}
\end{aligned}$$

So how does this help? We're looking to transform:

$$\begin{aligned}
A &\rightarrow \begin{bmatrix} R \\ 0 \end{bmatrix} \\
Q^T A &= \begin{bmatrix} R \\ 0 \end{bmatrix} \Rightarrow A = Q \begin{bmatrix} R \\ 0 \end{bmatrix} \\
A &= \left[\underbrace{Q_1}_n \quad \underbrace{Q_2}_{m-n} \right]_{m \times m} \begin{bmatrix} R \\ 0 \end{bmatrix} = Q_1 R + Q_2 0 = Q_1 R
\end{aligned}$$

17.1 Eigenvectors

A is a square matrix.

Find x and λ s.t. $Ax = \lambda x$. The x are called eigenvectors and λ are eigenvalues.

Suppose $\exists n\{x_i\}$ and $\{\lambda_i\} \Rightarrow$ any $y \in R^n, y = \sum_{i=1}^n c_i x_i$.

$$Ay = A \sum_{i=1}^n c_i x_i = c_1 \lambda_1 x_1 + c_2 \lambda_2 x_2 + \cdots + c_n \lambda_n x_n$$

Suppose we have S :

$$\begin{aligned}
S &= \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix} \\
AS &= \begin{bmatrix} | & | & \cdots & | \\ Ax_1 & Ax_2 & \cdots & Ax_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \\ | & | & & | \end{bmatrix} \\
&= \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}
\end{aligned}$$

So we can write:

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

This is another factorization – a diagonalization process. If A is symmetric and positive, then S and S^{-1} are orthogonal.

18 November 6 Lecture

18.1 Eigenvalues and Eigenvectors

Find x and λ s.t. $Ax = \lambda x$.

A is a square ($n \times n$) matrix.

In the language of linear algebra, find λ s.t. $(A - \lambda I)$ has a non-zero null space.

$\iff (A - \lambda I)$ has linearly dependent columns

$\iff A - \lambda I$ has at least a zero pivot

$\iff A - \lambda I$ is singular

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

Recall: If we have a set of n linearly independent eigenvectors of A , the A can be diagonalized:

$$\begin{aligned} S &= \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix} \text{ eigenvectors are columns of } S \\ AS &= \begin{bmatrix} | & | & \cdots & | \\ Ax_1 & Ax_2 & \cdots & Ax_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \lambda x_1 & \lambda x_2 & \cdots & \lambda x_n \\ | & | & & | \end{bmatrix} \\ &= \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \end{aligned}$$

Where Λ is a diagonal matrix:

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

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

Example Find eigenvalues:

$$\begin{aligned}
A &= \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} \\
\lambda_1 &= 3, \lambda_2 = 3 \\
(A - \lambda I)x &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x = 0 \\
x &= \begin{bmatrix} \alpha \\ 0 \end{bmatrix} \\
p(\lambda) &= (3 - \lambda)^2 = 0
\end{aligned}$$

How would we find λ for (real-life) matrices A ?

Fundamental theorem of algebra: any polynomial of degree n has n roots.

(Abel, 1824) proved that the roots of a polynomial of degree greater than 4 cannot always be expressed by a closed-form involving the coefficients of the polynomial

1. For any matrix A there is an associated polynomial whose roots are the eigenvalues of A
2. For any polynomial there is a matrix whose eigenvalues are the roots of the polynomial

THIS COULD BE IN THE MIDTERM:

$$\begin{aligned}
A &= \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix} \\
\det(A - \lambda I) &= \begin{vmatrix} 4 - \lambda & -5 \\ 2 & -3 - \lambda \end{vmatrix} \\
p(\lambda) &= (4 - \lambda)(-3 - \lambda) + 10 \\
&= (\lambda + 1)(\lambda - 2) \\
\lambda_1 = -1 &\Rightarrow (A - \lambda I)x = \begin{bmatrix} 5 & -5 \\ 2 & -2 \end{bmatrix} x = 0 \Rightarrow x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
\lambda_2 = 2 &\Rightarrow (A - \lambda I)x = \begin{bmatrix} 2 & -5 \\ 2 & -5 \end{bmatrix} x = 0 \Rightarrow x = \begin{bmatrix} 5 \\ 2 \end{bmatrix}
\end{aligned}$$

This means we can write A like so:

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

Important results

1. If a matrix A has no repeated (distinct) eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$
 - Its eigenvalues are linearly independent
 - Any matrix with distinct eigenvalues can be diagonalized
2. The diagonalizing matrix S is not unique

$$\begin{aligned}\det(A) &= \det(S\Lambda S^{-1}) \\ &= \det(S) \underbrace{\det(\Lambda)}_{\prod_{i=1}^n \lambda_i} \det(S^{-1})\end{aligned}$$

Spectral theorem: ($A = A^T$)

1. A real symmetric matrix A can be diagonalized
2. The eigenvalues of A are real

$$A = Q\Lambda Q^T$$

where Q is orthogonal i.e. $Q^T Q = I$. The columns of Q are an orthonormal basis of $C(A)$.

$$\begin{aligned}A &= Q\Lambda Q^T \\ &= \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} | & | & \cdots & | \\ \lambda x_1 & \lambda x_2 & \cdots & \lambda x_n \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} - & x_1^T & - \\ \vdots & \vdots & \vdots \\ - & x_n^T & - \end{bmatrix}\end{aligned}$$

Eigenvalue decomposition:

$$A = \lambda_1 x_1 x_1^T + \cdots + \lambda_n x_n x_n^T$$

Suppose we have $\lambda_i \gg \lambda_{i+1}$. We can then approximate:

$$A \approx \lambda_1 x_1 x_1^T + \cdots + \lambda_i x_i x_i^T$$

Suppose $Ax = \lambda x$, with $\{\lambda_i\}$ the eigenvalues of A . What are the eigenvalues of A^2 ?

$$A^2 = A Ax = A \lambda x = \lambda Ax = \lambda \lambda x = \lambda^2 x$$

This implies that the eigenvalues of A^2 are $\{\lambda_i^2\}$

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

Eigenvectors of A are the same eigenvectors of A^2

Remember: We're trying to find a numerical algorithm to find eigenvectors and eigenvalues for $n > 4$ (characteristic polynomial approach may fail).

Main result The eigenvalues of A^k are $\{\lambda_i^k\}$ if $\{\lambda_i\}$ are the eigenvalues of A .

$$\begin{aligned} A^k &= (S\Lambda S^{-1})(S\Lambda S^{-1}) \cdots (S\Lambda S^{-1}) \\ A^k &= S\Lambda^k S^{-1} \end{aligned}$$

Power iteration Pseudocode:

- x_0 = arbitrary vector
- for $k = 1, 2, \dots$
 - $x_k = Ax_{k-1}$
- end
- $x_k = Ax_{k-1} = AAx_{k-2} = A^k x_0$

Assume that a full set of eigenvectors $\{v_i\}$ and a unique eigenvalue λ_1 of maximum “modulus”. The pseudocode converges in the long run to the first eigenvector of A . Once you’ve found that value, you can use QR factorization to remove the component along that direction, and then repeat to find the second largest eigenvalue, etc.

19 November 8 Lecture

Last class, we show that the eigenvalues of A^k are $\{\lambda_i^k\}$.

Interesting fact: the eigenvalues of A^{-1} are $\{1/\lambda_i\}$.

Assuming A has a full set of eigenvectors $\{v_i\}$ and a unique λ_1 of maximum modulus with corresponding eigenvectors v_1 , then the power iteration is defined as:

- $x_0 =$ arbitrary vector

- for $k = 1, 2, \dots$

- $x_k = Ax_{k-1}$

- $\lambda_k = \frac{(x_k)_i}{(x_{k-1})_i}$

- end

$$x_1 = Ax_0$$

$$x_2 = Ax_1 = A^2x_0$$

$$\vdots$$

$$x_k = Ax_{k-1} = A^kx_0$$

Any vector in \mathbb{R} can be decomposed as a linear combination of a basis:

$$x_0 = \sum_{j=1}^n \alpha_j v_j$$

$$x_k = A^k \sum_{j=1}^n \alpha_j v_j$$

According to the power iteration. Interesting fact about this sum is that since $\{v_j\}$ are eigenvectors, it can be expressed as:

$$\begin{aligned} x_k &= \sum_{j=1}^n \alpha_j A^k v_j \\ &= \sum_{j=1}^n \alpha_j \lambda_j^k v_j \\ &= \lambda_1^k \left(\alpha_1 v_1 + \sum_{j=2}^n \alpha_j \left(\frac{\lambda_j}{\lambda_1} \right)^k v_j \right) \end{aligned}$$

If we have $\left| \frac{\lambda_j}{\lambda_1} \right| < 1$, then $\frac{\lambda_j}{\lambda_1}^k \rightarrow 0$ as $k \rightarrow \infty$. Therefore:

$$x_k \rightarrow \beta v_1 \text{ as } k \rightarrow \infty$$

Limitations:

1. if x_0 is such that $\alpha_1 = 0 \Rightarrow x_k$ will not converge to v_1
2. Convergence is slow
3. Method may overflow

Aside: this power iteration is the method used to find steady state of Markov chain.

Once you have found λ_1, v_1 , the matrix $A - \lambda_1 v_1 v_1^T$ has $0, \lambda_2, \dots, \lambda_n$ as its eigenvalues.

The QR method (Spectral Theorem) If A is symmetric and real, A can be diagonalized by an orthogonal matrix Q . The columns of Q contain the eigenvectors of A .

How can we generalize this eigenvalue decomposition?

Three things to remember about this class:

1. LU factorization
2. QR factorization
3. UDV factorization (SVD)

Generalization of Eigenvalue Decomposition A general $A_{m \times n}$ has no determinant defined, so it has no characteristic polynomial, and no eigenvalues.

We can decompose into $A = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T$ where U and V^T are orthogonal, and Σ is diagonal.

If we have AA^T , the columns of this product are linear combinations of A . The columns space of this product are embedded in the columns space of A . It could be smaller in general, but since it's A^T , and assuming A is full column rank, $C(AA^T) = C(A)$. This matrix AA^T is symmetric, as is $n \times n$ matrix $A^T A$.

$$\begin{aligned} AA^T &= (U \Sigma V^T)(V \Sigma^T U^T) \\ &= U \Sigma \Sigma^T U^T \\ &= Q \Lambda Q^T \end{aligned}$$

20 November 13 Lecture

20.1 Quiz Questions

- Actual midterm will only differ from practice midterm by changing a few values²
- No SVD
- No QR within the context of finding eigenvalues/eigenvectors.
- No need to know IEEE standards for machine precision, max values
- Exam Location: maybe not here
- No calculator needed
- $n \times n$ matrix with n non-zero eigenvalues has rank n
- Pseudoinverse was not yet discussed

20.2 Spectral Theorem

Every real symmetric ($A^T = A$) matrix A can be diagonalized by an orthogonal matrix Q . (The eigenvalues of A are real in this case). Q contains the orthonormal eigenvectors of A .

$$A = Q\Lambda Q^T = \sum_{i=1}^n \lambda_i x_i x_i^T$$

20.3 Singular Value decomposition

Any $m \times n$ matrix A can be factored into $A = U\Sigma V^T = (\text{orthogonal})(\text{“diagonal”})(\text{orthogonal})$

$$\begin{aligned} AA^T(m \times m) &= (U\Sigma V^T)(V\Sigma^T U^T) = U\Sigma\Sigma^T U^T \\ A^T A(n \times n) &= (V\Sigma^T U^T)(U\Sigma V^T) = V\Sigma^T \Sigma V^T \end{aligned}$$

In the first line, U contains the eigenvectors of AA^T in its columns, and in the second line V contains the eigenvectors of $A^T A$ in its columns. $\Sigma\Sigma^T$ is the eigenvalue matrix for AA^T , size $m \times m$ with $\{r_i^2\}$ on the diagonal.

AA^T is a covariance matrix. If one finds the eigenvector that corresponds with the largest eigenvalue, we've found the direction of the largest variance (?).

²Hi Jimmy.

$$Av = U\Sigma \Rightarrow A \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ 0 & & \sigma_r & \\ & & & 0 \\ & & & & \ddots \end{bmatrix}$$

Theorem U and V produce an orthonormal basis for all four fundamental spaces.

1. First r columns of U : column space of A
2. First $m - r$ columns of U : left null space of A
3. First r columns of V : row space of A
4. Last $n - r$ columns of U : null space of A

21 Midterm

- LU factorization (multipliers $\rightarrow L$ with $(-)$ sign)
- $LU \rightarrow$ solve $Ax = b$
- Finite precision (machine precision)
 - relevance \Rightarrow to what extent can we identify matrices that are singular?
 - Gaussian elimination \Rightarrow how ill-conditioned
- Computer 4 fundamental spaces
- Particular solution based on if on column space
- Least squares using QR , normal equations
- Fundamental theorem of linear algebra
- Householder
- Computing eigenvectors and eigenvalues
- How to solve least squares solution using calculus (setting derivative equal to 0) or some other approach involving something perpendicular to something else (??)