MAT 281 - Linear Algebra and App

Matrix: An array of numbers arranged in rows and columns

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 3 \\ 7 & 6 \end{bmatrix} \quad \text{a } 3\times 2 \text{ matrix}$$

$(A)_{2,1} = a_{21}$ $\qquad$ $(A)_{3,2} = a_{32}$

Transpose of A: $(A^T)_{i,j} = (A)_{j,i}$

$$A^T = \begin{bmatrix} 1 & 5 & 7 \\ 2 & 3 & 6 \end{bmatrix}$$

Scalar-Matrix multiplication: $(cA)_{i,j} = c \cdot (A)_{i,j}$

$$B = 2A = \begin{bmatrix} 2 & 4 \\ 10 & 6 \\ 14 & 12 \end{bmatrix}$$

Matrix addition: $(A+B)_{i,j} = (A)_{i,j} + (B)_{i,j}$

$$A+B = \begin{bmatrix} 3 & 6 \\ 15 & 9 \\ 21 & 18 \end{bmatrix}$$

A column vector: ($n\times 1$ matrix) $v = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ → v_1, → v_2

A row vector: ($1\times n$ matrix)

$x = [2, 1, 0]$ $\quad v^T = [2 \;\; -1 \;\; 1]$

Dot product of two vectors: $v \cdot x = \sum_{i=1}^{n} v_i x_i$

$= 4 + (-1) + 0 = 3$

Matrix-vector multiplication:

$$c = A \underbrace{\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{y} = \begin{bmatrix} 2 + 2 \\ 10 + 3 \\ 14 + 6 \end{bmatrix} = \begin{bmatrix} 4 \\ 13 \\ 20 \end{bmatrix}$$

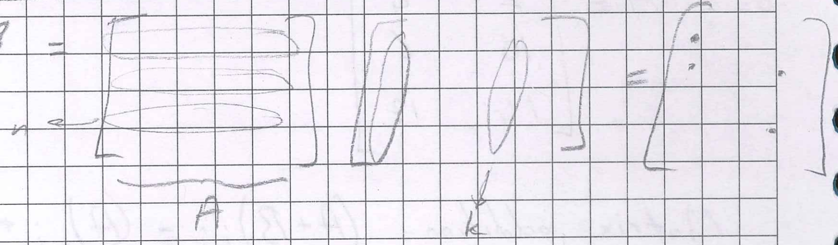
$$c_i = \sum_{j=1}^{n} a_{ij} y_j$$

Another view:

$$c = 2 \cdot \begin{bmatrix} 1 \\ 5 \\ 7 \end{bmatrix} + 1 \cdot \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix}$$

Matrix-Matrix Multiplication

$C = A \cdot B =$



$$c_{mk} = \sum_{j=1}^{n} a_{mj} b_{jk}$$

Solving Linear Equations

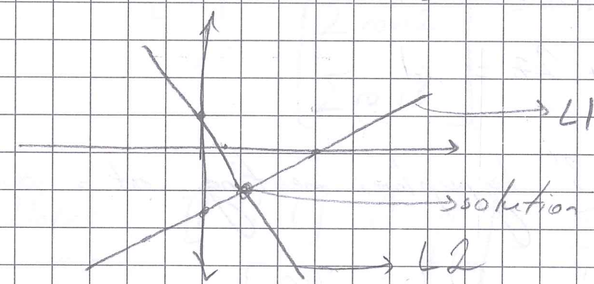
"Row Picture"

$x - 2y = 3$

$2x + y = 1$

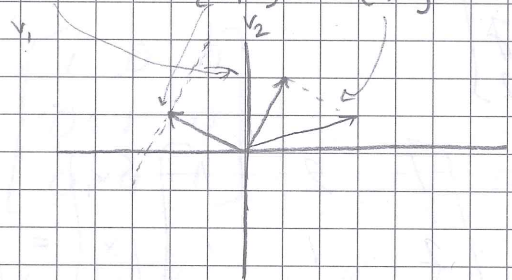
Solution of the 1st equation is a line. - L1

" 2nd " - L2



"Column Picture"

$$x \underset{v_1}{\begin{bmatrix} 1 \\ 2 \end{bmatrix}} + y \underset{v_2}{\begin{bmatrix} -2 \\ 1 \end{bmatrix}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$



Solving the system of lin. eqns boils down to finding the right linear comb. of v_1 & v_2.

$\Rightarrow$ We are essentially solving

$$\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

$\rightarrow$ coefficient matrix.

3 Eqns, 3 Unknowns

$$x + 2y + z = 2$$
$$2x + y + z = 1$$
$$-x + y + 2z = 1$$

Row Picture: 3 hyperplanes meeting at a point.

Column Picture: $x \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + y \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + z \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$

Soln: $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

Matrix form: $\underbrace{\begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ -1 & 1 & 2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}}_{b}$

Multiplication by rows: $Ax = \begin{bmatrix} (\text{row } 1) \cdot x \\ (\text{row } 2) \cdot x \\ (\text{row } 3) \cdot x \end{bmatrix}$

Multiplication by columns: $Ax = x_1 (\text{col } 1) + x_2 (\text{col } 2) + x_3 \cdot (\text{col } 3)$

Ex: $A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \sum \text{row } 1 \\ \sum \text{row } 2 \\ \sum \text{row } 3 \end{bmatrix} = \text{col } 1 + \text{col } 2 + \text{col } 3$

Notice $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \sum_i \sum_j a_{ij}$

A special matrix: Identity matrix, denoted by I.
The matrix that satisfies $Ix = x$ for any x.

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

Has 1's on the main diagonal and 0's elsewhere.

Idea of Elimination

Elimination will give a systematic way to solve linear equations

$$\begin{aligned} x + 3y &= 5 \\ 2x - y &= 3 \end{aligned} \quad \longrightarrow \quad \begin{aligned} x + 3y &= 5 \\ -7y &= -7 \end{aligned}$$

$\Downarrow$ subtract 2·eqn 1 from eqn 2 (elimination)

Solve the 2^{nd} eqn $\Rightarrow y = 1$, (back-subst)

Back-substitute $y = 1$ into 1^{st} eqn $\Rightarrow x + 3 = 5$

$\Rightarrow$ Soln: $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Take a look at the coefficient matrix:

$$\begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 3 \\ 0 & -7 \end{bmatrix}$$

$\Downarrow$ full matrix $\qquad$ $\Downarrow$ upper triangular system

For elimination, we used a_{11} to cancel a_{21}.

a_{11} is the first pivot.

Pivot: first entry in the row that does the elimination.

The pivot is important because it determines the multiplier.

Multiplier: $\dfrac{\text{entry to eliminate}}{\text{pivot}}$

Breakdown of elimination

2 cases lead to permanent failure:

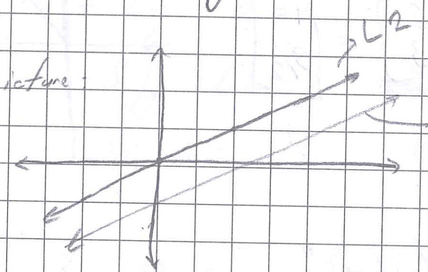
(1) No solution

(2) Infinitely many solutions

Ex 1: $\begin{aligned} 3x + 6y &= 1 \\ x + 2y &= 0 \end{aligned} \xrightarrow{\text{Elimination}} \begin{aligned} 3x + 6y &= 1 \\ 0 &= -1/3 \end{aligned}$ (3: pivot)

$\boxed{\text{No solution}}$

row picture:

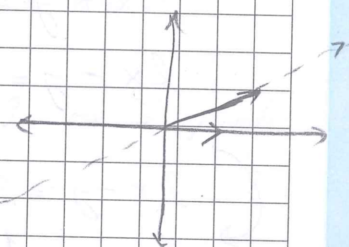


parallel lines do not meet

column picture:

$$x\begin{bmatrix} 3 \\ 1 \end{bmatrix} + y\begin{bmatrix} 6 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$\underbrace{\qquad = 2y\begin{bmatrix} 3 \\ 1 \end{bmatrix}}_{(x+2y)\begin{bmatrix} 3 \\ 1 \end{bmatrix}}$

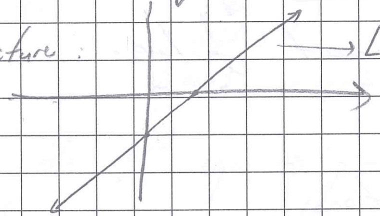


Ex 2:

$$\begin{aligned} 2x - 2y &= 1 \\ 3x - 3y &= 3/2 \end{aligned} \implies \begin{aligned} 2x - 2y &= 1 \\ 0 &= 0 \end{aligned}$$

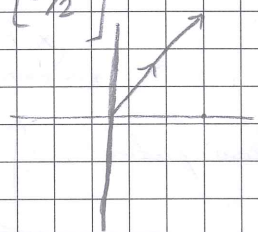
Again 1 pivot but now no inconsistensies

$\Rightarrow$ infinitely many solutions

row picture:  $\rightarrow$ L1, L2 coincide

col picture: $x\begin{bmatrix}2\\3\end{bmatrix} + y\begin{bmatrix}-2\\-3\end{bmatrix} = \begin{bmatrix}1\\3/2\end{bmatrix}$

$$(x-y)\begin{bmatrix}2\\3\end{bmatrix} = \begin{bmatrix}1\\3/2\end{bmatrix}$$



Ex 3: "Temporary failure"

$$\begin{aligned} 0x + y &= 1 \\ 2x - y &= 2 \end{aligned}$$

zero in pivot!

Remedy: exchange rows.
(Permutation)

$$\begin{aligned} 2x - y &= 2 \\ 0x + y &= 1 \end{aligned}$$

Coefficient matrices in Ex 1 & 2 are singular (no second pivot). Whenever we end up with a full set of pivots, we will call the coef. matrix non-singular (in which case we will be able to solve the system.)

3 equations 3 unknowns:

$$\begin{aligned} x + 2y - z &= 1 \\ 3x + y + z &= 6 \\ 2x - y - z &= -1 \end{aligned} \implies \begin{aligned} x + 2y - z &= 1 \\ -5y + 4z &= 3 \\ 2x - y - z &= -1 \end{aligned}$$

$$\implies \begin{aligned} x + 2y - z &= 1 \\ -5y + 4z &= 3 \\ -5y + z &= -3 \end{aligned} \implies \begin{aligned} x + 2y - z &= 1 \\ -5y + 4z &= 3 \\ -3z &= -6 \end{aligned}$$

$\Rightarrow z = 2 \Rightarrow -5y + 8 = 3 \Rightarrow x + 2 - 2 = 1$

$\Rightarrow y = 1 \qquad \Rightarrow x = 1$

sol: $\begin{bmatrix}1\\1\\2\end{bmatrix}$ pivots: $1, -5, -3 \Rightarrow 3$ pivots.

Elimination Using Matrices

The previous example can be written in matrix form as,

$$\underbrace{\begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 1 \\ 2 & -1 & -1 \end{bmatrix}}_{A_1} \Rightarrow \underbrace{\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 4 \\ 2 & -1 & -1 \end{bmatrix}}_{A_2} \Rightarrow \begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 4 \\ 0 & -5 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 4 \\ 0 & 0 & -3 \end{bmatrix}$$

(started with a full matrix, ended up with an upper-triangular)

Can we perform these steps using Matrices?

Recall:

$$\begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 1 \\ 2 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

$\Rightarrow$ a linear combination of columns of A.

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 1 \\ 2 & -1 & -1 \end{bmatrix} = x_1 \cdot \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} + x_2 \cdot \begin{bmatrix} 3 & 1 & 1 \end{bmatrix} + x_3 \cdot \begin{bmatrix} 2 & -1 & -1 \end{bmatrix}$$

$\Rightarrow$ a linear combination of rows of A.

If I want to subtract $3 \cdot$ row 1 from row 2, I'd set $x_1 = -3$, $x_2 = 1$.

This is the 2^{nd} row of A_2

$$\Rightarrow \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{B} \Rightarrow$$

$B \cdot A$ keeps the 1^{st} & 3^{rd} rows of A the same, places in the 2^{nd} row: row 2 $- 3 \cdot$ row 1

Can we undo this? Yes: add back $3 \times$ row 1 to row 2 using

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

$\Rightarrow C \cdot B \cdot A = A.$

2 Rules about Matrix Mult:

Associative law: $A(B \cdot C) = (AB)C$

Commutative law does not hold in general: $AB \neq BA$.

Permutation matrices:

Exchange rows $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ exchanges eqn 2 with eqn 1.

Augmented Matrix: Since we apply the same operations to the rhs, we augment it as an additional column, forming the augmented matrix: $[A \; b]$

$$\begin{bmatrix} 1 & 2 & -1 & 1 \\ 3 & 1 & 1 & 6 \\ 2 & -1 & -1 & -1 \end{bmatrix} \xrightarrow{B} \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & -5 & 4 & 3 \\ 2 & -1 & -1 & -1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & -5 & 4 & 3 \\ 0 & -5 & 1 & -3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & -5 & 4 & 3 \\ 0 & 0 & -3 & -6 \end{bmatrix}$$

Inverse Matrices

Defn

Notes ① Inverse exists iff elimination produces n pivots.

② Inverse is unique

③ $Ax = b \Rightarrow x = A^{-1}b$

④ If $\exists\, x \neq 0$ s.t. $Ax = 0 \Rightarrow A$ is not invertible.

Inverse of a Product

Inverses of elim. matrices.

Rules for Matrix Operations

Addition Laws:

$A + B = B + A$ (commutative law)

$c(A+B) = cA + cB$ (distributive law)

$A + (B+C) = (A+B) + C$ (associative law)

Multiplication Laws

$AB \neq BA$ (in general com. law does not hold)

$C(A+B) = CA + CB$ (dist. from left)

$(A+B)C = AC + BC$ (right)

$A(BC) = (AB)C$ (associative law)

$\Rightarrow$ parentheses are not needed.

Block Matrices & Block Multiplication

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

Block Multiplication: Suppose we can cut both A & B into blocks, which have sizes that allow multiplication. We can treat blocks like regular entries of a matrix

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

$$= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \\ A_{31}B_{11} + A_{32}B_{21} & A_{31}B_{12} + A_{32}B_{22} \end{bmatrix}$$

$$\begin{bmatrix} I & I & I \\ 2I & 2I & 2I \end{bmatrix} \left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \\ \hline 0 & 1 \\ 0 & 0 \\ \hline 0 & 0 \\ 1 & 1 \end{array}\right] = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

4 Different Ways to see Matrix Multiplication

$$\underset{n\times m}{A} \cdot \underset{m\times k}{B} = \underset{n\times k}{C}$$

(1) Regular (row-column inner product)

c_{ij} = (row i of A · row j. of B)

Decompose A into rows, B into columns

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \cdots & b_k \end{bmatrix} = \begin{bmatrix} a_1 \cdot b_1 & a_1 \cdot b_2 & \cdots & a_1 \cdot b_k \\ a_2 \cdot b_1 & \cdots & & \\ a_n \cdot b_1 & \cdots & & a_1 \cdot b_k \end{bmatrix}$$

(2) Decompose B into columns:

$$A \cdot B = A\begin{bmatrix} b_1 & b_2 & \cdots & b_k \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \cdots & Ab_k \end{bmatrix}$$

(3) Decompose A into rows:

$$A \cdot B = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} B = \begin{bmatrix} a_1 B \\ a_2 B \\ \vdots \\ a_n B \end{bmatrix}$$

(4) Decompose A into columns, B into rows.

$$A \cdot B = \begin{bmatrix} \tilde{a}_1 & \tilde{a}_2 & \tilde{a}_3 & \cdots & \tilde{a}_m \end{bmatrix} \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \tilde{b}_m \end{bmatrix} = \tilde{a}_1 \cdot \tilde{b}_1 + \tilde{a}_2 \cdot \tilde{b}_2 + \cdots + \tilde{a}_m \tilde{b}_m$$

Ex: $\begin{bmatrix}1&0\\2&-1\end{bmatrix}\begin{bmatrix}3&1\\0&2\end{bmatrix}=\begin{bmatrix}3&1\\6&0\end{bmatrix}$

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

Inverse Matrices

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

A is said to be invertible if there exists a matrix A^{-1} s.t

$$A \cdot A^{-1} = I \quad \text{and} \quad A^{-1} \cdot A = I.$$

If A is invertible, A^{-1} is called the inverse of A.

Fact: There are matrices which are not invertible.

Ex: $\begin{bmatrix}1&0\\0&0\end{bmatrix}\begin{bmatrix}a&b\\c&d\end{bmatrix}=\begin{bmatrix}a&b\\0&0\end{bmatrix}\neq\begin{bmatrix}1&0\\0&1\end{bmatrix}$

Notes: 1) If inverse exists, then elimination produces n pivots.

Converse, iff.

2) If inverse exists, solution of $Ax=b$ is $x=A^{-1}b$.

3) Inverse is unique. if $AB=I$ & $CA=I \Rightarrow B=C$.

4) If there is a non-zero vector x s.t. $Ax=0$ then A is not invertible.

$$x \neq A^{-1} 0 = 0$$

5) Inverse of a diagonal matrix.

Inverse of Products: $(AB)^{-1} = B^{-1}A^{-1}$

If: $(AB)\cdot(B^{-1}A^{-1}) = A(B\,B^{-1})A^{-1} = A \cdot I \cdot A^{-1} = I$.

In general, $(A \cdot B \cdot C)^{-1} = (C)^{-1}(B)^{-1}(A)^{-1}$

Calculating A^{-1} by Gauss-Jordan Elimination

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

We want to find the matrix A^{-1} s.t.

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

Decompose A^{-1} into columns $\Rightarrow A \cdot A^{-1} = A\begin{bmatrix}x_1 & x_2 & x_3\end{bmatrix}$

$= \begin{bmatrix}Ax_1 & Ax_2 & Ax_3\end{bmatrix} \Rightarrow Ax_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} \quad Ax_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$

$Ax_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}.$

For $n \times n$ A, $\Rightarrow$ n X {n equations in n unknowns}

For $Ax = b$, we form the Augmented Matrix $[A \; b]$ and operate on the rows. For our case, we will form a new augmented matrix as $[A \; I]$.

Gauss-Jordan solves the equations together.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 3 & -1 & 0 & 1 & 0 \\ -1 & -3 & 0 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ -1 & -3 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & 1 & 3 & -1 & 1 \end{array}\right]$$ upper-triangular

Continue with elimination $\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 & -1 & 1 \end{array}\right]$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 0 & 2 \\ 0 & -1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 & -1 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 0 & 2 \\ 0 & 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 3 & -1 & 1 \end{array}\right]$$

Multiply by a diagonal matrix.

In summary: $A^{-1}[A \; I] = [I \; A^{-1}]$

How to Solve $AX = B$? Augmented matrix $[A \; B]$

$\Rightarrow$ In order for the inverse to exist elimination must produce n pivots.

LU Decomp: Factorization of A into lower and upper triangular matrices.

Ex: Let $A = \begin{bmatrix} 2 & 3 \\ 8 & 1 \end{bmatrix}$

Apply elimination: $\underbrace{\begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix}}_{E_{21}} \underbrace{\begin{bmatrix} 2 & 3 \\ 8 & 1 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 2 & 3 \\ 0 & -11 \end{bmatrix}}_{U}$

$E_{21}^{-1} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \Rightarrow E_{21}^{-1} E_{21} A = \underbrace{\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 2 & 3 \\ 0 & -1 \end{bmatrix}}_{U}$

In general we apply a number of elimination matrices to A to reach an upper-triangular U:

$$E_n \cdots E_3 E_2 E_1 A = U$$

$$A = \underbrace{E_1^{-1} E_2^{-1} \cdots E_n^{-1}}_{L} U$$

Why is $(E_n \cdots E_1)^{-1}$ lower triangular?

1. Notice that every E_k and E_k^{-1} is lower diagonal

Ex: E: subtract $3 \cdot$ row 1 from row 2:

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

2. Product of two lower-triangular matrices is also lower triangular.

Ex:
$$\begin{bmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ 4 & 5 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 2 \end{bmatrix}$$
$$\ c_1 \quad c_2 \quad c_3$$

$$= \begin{bmatrix} 3 \cdot c_1 + c_2 + c_3 & 1 \cdot c_2 + 2 \cdot c_3 & 2 \cdot c_3 \end{bmatrix} = \begin{bmatrix} x & 0 & 0 \\ x & x & 0 \\ x & x & x \end{bmatrix}$$

Combining 1 & 2 $\Rightarrow E_1^{-1} E_2^{-1} E_3^{-1} \cdots E_n^{-1}$ is lower triang.

Notice that all the diagonal entries of L are 1. This is not the case for U. If we force the diagonal entries of U to be 1 as well, we get $LD\tilde{U}$, where $D\tilde{U} = U$, $\tilde{U}$ has 1 on the diagonal.

LU Decomposition for solving systems of eqns:

$Ax = b \Rightarrow LUx = b$

1. Solve $Ly = b$ (*)
2. Solve $Ux = y$ (back-subst.)

For 1. Notice that $L = E_1^{-1} E_2^{-1} \cdots E_n^{-1}$

3×3 case: E_1: $\begin{bmatrix} 1 & 0 & 0 \\ x & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ E_2: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ x & 0 & 1 \end{bmatrix}$

$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & x & 1 \end{bmatrix}$

$\Rightarrow L =$

Transpose: $(A^T)_{ij} = (A)_{ji}$ rows of A = columns of A^T.

Rules: 1) $(A+B)^T = A^T + B^T$ 2) $(AB)^T = B^T A^T$

3) $(A^{-1})^T = (A^T)^{-1}$

Pf 2) $(Ax)^T = (x_1 a_1 + x_2 a_2 + \dots + x_n a_n)^T = x_1 a_1^T + \dots + x_n a_n^T$

$= x^T A^T$

$\Rightarrow \left(A\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}\right)^T = \begin{bmatrix} Ax_1 & Ax_2 & \cdots & Ax_n \end{bmatrix}^T$

$= \begin{bmatrix} x_1^T A^T \\ x_2^T A^T \\ \vdots \\ x_n^T A^T \end{bmatrix} = \begin{bmatrix} x_1^T \\ x_2^T \\ x_n^T \end{bmatrix} A^T$

3) $A \cdot A^{-1} = I \Rightarrow (A^{-1})^T \cdot A^T = I$

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

~~Inner Product.~~

Symmetric Matrices: A is symmetric if $A^T = A$.

$R^T R$ is symmetric for any R.

Permutation Matrices: A permutation matrix has the rows of identity in any order.

$n!$ distinct perm matrices

$P^T = P^{-1}$. $PA = LU$

Chp 2 Vector Spaces & Subspaces

(3.1) Spaces of Vectors

$\mathbb{R}^n$: The set of all column vectors with n components

Vector space: A set of vectors with two operations addition & sc. mult. that satisfy the following properties

(1) $x + y = y + x$

(2) $x + (y + z) = (x + y) + z$

(3) There is a unique "zero vector" s.t. $x + 0 = x$ $\forall x$.

(4) For each x, there is a unique $-x$ s.t. $x + (-x) = 0$.

(5) $1 \cdot x = x$

(6) $(c_1 c_2)x = c_1(c_2 x)$

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

(8) $(c_1 + c_2)x = c_1 x + c_2 y$

Two essential properties:

1) We can add any two vectors and the result will be in the space.

2) We can multiply a vector with a scalar and the result will be in the space.

GIPTA

Ex: M = All real 2×2 matrices

F = All real functions $f(x)$

Z = The zero vector in $\mathbb{R}^n$.

Subspace: Let V be a vector space. $S \subset V$ is a subspace of V if

1) Addition of any two vectors from S lies in S. ($x_1, x_2 \in S \Rightarrow x_1 + x_2 \in S$)

2) $x \in S \Rightarrow \alpha \cdot x \in S$

In other words, S is "closed" under addition and multiplication by a scalar.

FACT: Every subspace contains the zero vector.

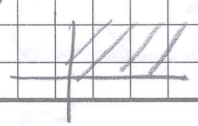
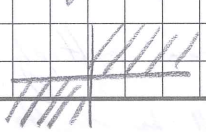
Why?

Subspaces of $\mathbb{R}^3$: 1) Any line through 0

2) Any plane through 0

3) $\mathbb{R}^3$ 4) $\{0\}$.

Examples which are not subspaces of $\mathbb{R}^2$:

1)  2) 

Column Space of A:

Consider $Ax = b$. For what values of 'b' can we find a solution x?

Trivial answer: Those 'b' which are linear combinations of the columns of A.

This is called the 'column space of A'.

This is a space! Check the requirements:

1) let $b_1, b_2 \in C(A) \Rightarrow \exists\, x_1, x_2$ s.t.

$Ax_1 = b_1$, $Ax_2 = b_2$ $\Rightarrow$ $A(x_1 + x_2) = \underbrace{(b_1 + b_2)}_{\text{in the space}}$

2) $b_1 \in C(A) \Rightarrow Ax_1 = b \Rightarrow A(\alpha x_1) = \alpha b_1$

$\Rightarrow \alpha b_1 \in C(A)$.

If A is $m \times n$, $C(A)$ is a subspace of $\mathbb{R}^m$.

Q: Is 0 in $C(A)$? (i.e. can we solve $Ax = 0$)

Span of a set of vectors:

S = Set of vectors in a space V.

SS = all linear combinations of vectors in S

SS = the subspace of V spanned by S.

Is this a subspace? Why?

The column space of A is the 'span' of the columns of A.

Example: What are the column spaces of

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & -3 \\ 2 & -6 \end{bmatrix}, \quad \begin{bmatrix} 1 & -3 & 2 \\ 2 & -6 & 1 \end{bmatrix}$$

Nullspace of A

$N(A)$: The set of vectors x s.t. $Ax = 0$.

if A is $m \times n$, $N(A) \subset \mathbb{R}^n$

$N(A) = \{ x \in \mathbb{R}^n : Ax = 0 \}$.

Is $N(A)$ a subspace? Check the conditions

i) $x_1, x_2 \in N(A) \Rightarrow \left.\begin{matrix} Ax_1 = 0 \\ Ax_2 = 0 \end{matrix}\right\} \Rightarrow A(x_1 + x_2) = 0$

$\Rightarrow x_1 + x_2 \in N(A)$

ii) $x_1 \in A \Rightarrow \alpha x_1 \in A$.

$N(A)$: All the solutions of $Ax = 0$.

Can we change the rhs to obtain new subspaces?

Is the set of all solutions of $Ax = b$ a subspace? Why?

Ex: $x + y + 2z = 0 \Rightarrow \begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$

$\Rightarrow$ The set of all solutions is a plane in $\mathbb{R}^3$ that passes through the origin.

Ex: A is $n \times n$, invertible. $N(A) = \{0\}$. Why?

Ex: A: $m \times n$, $m < n$. Can $N(A) = \{0\}$? No.

Ex: Describe systematically the nullspace of

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 8 & 11 & 8 \end{bmatrix}$$

We are trying to solve $Ax = 0$. Do elimination!

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 8 & 11 & 8 \end{bmatrix} x = 0 \Rightarrow \underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 2 & -4 \end{bmatrix}}_{\text{Echelon form}} x = 0$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 1 & 8 \\ 0 & 2 & 2 & -4 \end{bmatrix} \rightarrow \underbrace{\begin{bmatrix} 1 & 0 & 1 & 8 \\ 0 & 1 & 1 & -2 \end{bmatrix}}_{\text{Reduced row echelon form}}$$

$$\begin{bmatrix} 1 & 0 & 1 & 8 \\ 0 & 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

pivot variables: x_1, x_2

free variables: x_3, x_4

Once you select the free variables, the pivot variables are determined

Special solution are obtained by setting one of the free variables to 1, rest to zero.

$$s_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} \qquad s_2 = \begin{bmatrix} 1/8 \\ 1/2 \\ 0 \\ 1 \end{bmatrix}$$

$$s = \alpha_1 s_1 + \alpha_2 s_2$$

Idea: I) Forward elimination $A \to$ triangular U or r-r-e R

II) Back substitution to solve $Ux = 0$

$\Rightarrow$ # of special solutions = # of free variables.

$N(A)$ = all linear combinations of the special soln.

Another ex:
$$\begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 4 & 7 & 4 \\ -1 & -2 & -2 & 1 \end{bmatrix}$$

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

pivot var: x_1, x_3

free var: x_2, x_4

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

$$s_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \begin{pmatrix} x_2 = 0 \\ x_4 = 0 \end{pmatrix}$$

$$s_2 = \begin{bmatrix} 5 \\ 0 \\ -8 \\ 1 \end{bmatrix}$$

$$\Rightarrow x = \alpha_1 s_1 + \alpha_2 s_2$$

$$= \begin{bmatrix} -2\alpha_1 + 5\alpha_2 \\ \alpha_1 \\ -8\alpha_2 \\ \alpha_2 \end{bmatrix}$$

If $m < n$ an $m \times n$ matrix has a non trivial nullspace.

GIPTA

Echelon Matrix: The "staircase" matrix that remains after elimination

$$\begin{bmatrix} p & x & x & x & x \\ 0 & 0 & p & x & x \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

pivot variables: x_1, x_3

free variables: x_2, x_4, x_5.

pivot columns

free columns

Obtain the special solution by setting one of the free variables to 1, rest to 0 and then determining the pivot variables.

How many special soln. in this case? 3.

The nullspace is the span of these 3 solutions.

Reduced Row Echelon Matrix

$$\begin{bmatrix} 1 & x & 0 & x & x \\ 0 & 0 & 1 & x & x \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Rank

Given an $m \times n$ matrix A, we have at hand m equations in n unknowns. Some of these equations are combinations of other equations and could have been eliminated. The number of "independent" equations will be called the rank. More precisely

Rank: The rank of A is the number of pivots. This number is r.

FACT: Pivot Columns are not combinations of earlier columns. The free columns are combinations of earlier columns. Those combinations are the special sol.

Ex:

$$E \cdot \underbrace{\begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 4 & 7 & 4 \\ -1 & -2 & -2 & 1 \end{bmatrix}}_{A} = \begin{array}{c} \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ \begin{array}{cccc} c_1 & c_2 & c_3 & c_4 \end{array} \end{array}$$

$$= \begin{bmatrix} c_1 & 2 \cdot c_1 & c_3 & 2 \cdot c_3 - 5 \cdot c_1 \end{bmatrix}$$

$$= \begin{bmatrix} c_1 & c_3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & -5 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$$A = E^{-1} \cdot \begin{bmatrix} c_1 & c_3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & -5 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 3 \\ 2 & 7 \\ -1 & -2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 0 & -5 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

A special solution is obtained by setting one of the free variables to 1, the rest of the free variables to 0 and computing the pivot variables.

\# of special soln = \# of free variables

$= \#$ of columns of $A - \underbrace{\#\text{ of pivot var.}}_{\text{rank}}$

$= n - r.$

Nullspace Matrix: $N = \begin{bmatrix} y_1 & y_2 & \cdots & y_{n-r} \end{bmatrix}$

$$\Rightarrow AN = \begin{bmatrix} Ay_1 & Ay_2 & \cdots & Ay_{n-r} \end{bmatrix} = 0.$$

The Complete Solution to $Ax = b$.

The null-space is the set of solution to $Ax=0$. We saw that the set of solution for $Ax=b$ with $b \neq 0$ is not a subspace. But how do we obtain this set?

Ex: $$\underbrace{\begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 2 & -2 & 5 & 8 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \\ 4 \\ 6 \end{bmatrix}}_{b}$$

Augmented Matrix: $$\begin{bmatrix} A & b \end{bmatrix} = \begin{bmatrix} 1 & -1 & 2 & 3 & 1 \\ 0 & 0 & 1 & 2 & 4 \\ 2 & -2 & 5 & 8 & 6 \end{bmatrix}$$

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

$$\rightarrow \begin{bmatrix} \underbrace{\begin{matrix} 1 & -1 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{matrix}}_{R} & \underbrace{\begin{matrix} -2 \\ 4 \\ 0 \end{matrix}}_{d} \end{bmatrix}$$

$$\Rightarrow I \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} + \begin{bmatrix} -1 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2 \\ +4 \end{bmatrix}$$

$Rx = d$

To obtain a "particular solution" set the free variables to 0.

$$\Rightarrow \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} -7 \\ +4 \end{bmatrix} \Rightarrow x_p = \begin{bmatrix} -7 \\ 0 \\ 4 \\ 0 \end{bmatrix} \text{ solves the system.}$$

This is only one of the solutions. How do we obtain the whole set of solutions?

Claim: Set of solutions $= x_p + N(A)$.

For this case, we have $\overset{=2}{n-r}$ special solutions for $N(A)$:

$$y_{s_1} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad y_{s_2} = \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

$N(A)$ = All vectors of the form $\alpha_1 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix}$.

Solution set for $Ax=b$ =

All vectors of the form $y_p + \alpha_1 y_{s_1} + \alpha_2 y_{s_2}$

Pf of claim: First notice that

$$A(y_p + \alpha_1 y_{s_1} + \alpha_2 y_{s_2} + \dots + \alpha_n y_{s_r})$$

$$= Ay_p + \alpha_1 Ay_{s_1} + \alpha_2 Ay_{s_2} + \dots + \alpha_n y_{s_r} = Ay_p = d.$$

Conversely if $Ax_0 = d$, then we should be able to express x_0 as $x_0 = y_p + y_n$ where $y_n \in N(A)$.

$$x_0 = y_p + (x_0 - y_p)$$

We need to show that $(x_0 - y_p) \in N(A)$.

$$\left.\begin{aligned} Ax_0 &= d \\ Ay_p &= d \end{aligned}\right\} \Rightarrow A(x_0 - y_p) = 0 \Rightarrow (x_0 - y_p) \in N(A).$$

Procedure for obtaining the solution set of $Ax=b$.

1. Find a particular soln. y_p by solving $Ax=b$ (using elimination).
2. Determine $N(A)$
3. Solution set $= y_p + y_n$ where $y_n \in N(A)$

Q: Can we solve $Ax=b$ for all b?

A: No, we can solve it only when $b \in C(A)$

Special Case: Full-column rank

$A: m \times n$, rank $= n$.

Ex: $\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$

Elimination on $[A \ b] \rightarrow \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

$\rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

No free variables $\Rightarrow N(A) = \{0\}$.

Particular solution: $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Full Column Rank

- $Ax=b$ has $\begin{cases} 1 \text{ soln. if } b \in C(A) \\ 0 \text{ soln. if } b \notin C(A) \end{cases}$
- $N(A) = \{0\}$ (because no free variables)

Full-Row Rank: $A = m \times n$, $r = m$.

of pivot columns = # of rows.

Reduced-Row Echelon Form

$$\begin{bmatrix} 1 & d & 0 & 0 & e \\ 0 & 0 & 1 & 0 & f \\ 0 & 0 & 0 & 1 & g \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_5 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$I \begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix} + \begin{bmatrix} d & e \\ 0 & f \\ 0 & g \end{bmatrix} \begin{bmatrix} x_2 \\ x_5 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

We can always find a solution (set $x_2 = x_5 = 0$,

$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$

$\Rightarrow C(A) = \mathbb{R}^m$.

If $n > m$, then # of free variables > 0

$\Rightarrow N(A)$ is larger than $\{0\}$.

$\Rightarrow Ax=b$ has infinitely many solutions.

Independence, Basis, Dimension

Def'n: The columns of A are linearly independent when the only vector that solves $Ax = 0$ is $x = 0$.
(In other words $N(A) = \{0\}$).

$$Ax = [a_1 \; a_2 \; \dots \; a_n] = x_1 a_1 + x_2 a_2 + \dots + x_n a_n = 0$$

only when $x = 0$.

Def'n: A sequence of vectors $v_1, v_2, \dots, v_n$ is said to be linearly independent if the only linear combination that gives the zero vector is $0 \cdot v_1 + 0 \cdot v_2 + \dots + 0 \cdot v_n$.

If a non-zero combination produces zero, we will say that the vectors are dependent.

Ex: $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$; $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}$; $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\begin{pmatrix} c \\ c \end{pmatrix}$;

$\begin{pmatrix} a \\ b \end{pmatrix}$ $\begin{pmatrix} c \\ d \end{pmatrix}$ $\begin{pmatrix} e \\ f \end{pmatrix}$

Fact: Any set of n vectors in $\mathbb{R}^m$ with $n > m$ are linearly dependent

Fact: Full-column rank $\Rightarrow N(A) = \{0\}$

$\Rightarrow$ columns are independent.

Defn: A set of vectors *span* a space if any element of the space can be expressed as a linear combination of the elements from the set

In particular columns of A span $C(A)$.

Ex: $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$ span $\mathbb{R}^2$

$\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\-1\end{pmatrix}$ span the line $\alpha\begin{pmatrix}1\\1\end{pmatrix}$.

$\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$ span $\mathbb{R}^2$

Defn: A basis for a vector space is a sequence of vectors that

(i) span the vector space

(ii) are linearly independent.

Claim: Given an element v of a vector space and a basis $v_1, \dots, v_n$, there's a unique linear comb. that gives $v = \alpha_1 v_1 + \dots + \alpha_n v_n$

Pf: ...

Notice that there are inf. many different bases for a vector space.

Ex: $A: n\times n$, invertible.

$C(A) = \mathbb{R}^n$ because for any $b \in \mathbb{R}^n$, we can find x s.t. $Ax = b \Rightarrow x = A^{-1}b$.

Fact: For $\mathbb{R}^n$, $v_1, \dots, v_n$ form a basis if the matrix $[v_1\ v_2 \cdots v_n]$ is invertible.

Fact: Pivot columns of A form a basis for $C(A)$.

$\Downarrow$

Bunu echelon matrislerle göster

Dimension:

Claim: Let V be a space. If $v_1, v_2, \dots, v_n$ and $w_1, \dots, w_m$ are both bases for V, then $n = m$.

Pf: $W = VA$ with wide $A \Rightarrow \exists\, x \neq 0$ s.t. $Ax = 0$.

$\Rightarrow Wx = 0$

Defn: The dimension of a space is the number of vectors in every basis.

Ex: Dimension of $C(A)$ is its rank.

Dimensions of the Four Subspaces

For an $m \times n$ matrix A,
We've seen two subspaces so far,

(I) The column space $C(A)$.

(II) The null-space $N(A)$.

We can add two more

(III) The row space of A, all linear combinations of the rows of A, equivalent to $C(A^T)$.

(IV) The left null-space: $N(A^T)$.

The Four Subspace for R

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

Notice that $\operatorname{rank}(R) = 3$.

(1) The row space of R has dimension $3 = \text{rank}$.
It is obviously spanned by the 3 rows and these rows are independent.

In general, $\dim(R(A)) = r$. The non-zero rows form a basis

(2) The column space has dimension $3 = \text{rank}$.
Same argument.

In general, $\dim(C(A)) = r$. The pivot columns form a basis.

(3) The null-space has dimension $= n - r = 5 - 3 = 2$.
The special solutions form a basis for the null-space.

For this matrix, the free variables are x_2, x_4
the pivot variables are x_1, x_3, x_5

$$R \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_5 \end{bmatrix} = I \begin{bmatrix} x_1 \\ x_3 \\ x_5 \end{bmatrix} + \begin{bmatrix} 3 & 5 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} \quad (*)$$

$$\Rightarrow s_2 = \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad s_4 = \begin{bmatrix} -5 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}$$

(1) Notice that s_2 & s_4 are independent.

If we can show that they also span $N(A)$, then we can say that they form a basis for $N(A)$.

We need to show that the set of all linear comb of s_2 & s_4 is equal to $N(A)$

First notice that $R \cdot (\alpha_1 s_1 + \alpha_2 s_2) = 0$

$\Rightarrow$ Set of all linear comb. is a subset of $N(A)$.

i.e. ($\text{span}(s_2, s_4) \subset N(A)$).

If we can show that any vector from $N(A)$ also lies in $\text{span}(s_1, s_2)$ then that would imply $N(A) \subset \text{span}(s_1, s_2)$.

Taken together, "$\text{span}(s_1, s_2) \subset N(A)$" and "$N(A) \subset \text{span}(s_1, s_2)$" imply that "$N(A) = \text{span}(s_1, s_2)$".

Now let $y \in N(A)$. Because of (*), we see that once the free variables are chosen, the pivot variables are determined. But we see that "$y_2 \cdot s_2 + y_4 \cdot s_4$" is an element of the null-space and its free variables are equal to those of y.

so it must hold that $y = y_2 \cdot s_2 + y_4 \cdot s_4$

Thus y can be expressed as a linear comb. of the sp. solutions. $\Rightarrow y \in \text{span}(s_2, s_4)$.

Since y was an arbitrary element of $N(A)$, this implies that $N(R) \subset \text{span}(s_2, s_4)$

$\Rightarrow N(R) = \text{span}(s_2, s_4)$.

In general, $N(A) = n - r$, the special solutions form a basis for $N(A)$.

(4) The left null-space of R has dimension $m-r$.

Easier to see. These are all y that give

$y \cdot R = 0 \Rightarrow y_1 \cdot r_1 + y_2 \cdot r_2 + y_3 \cdot r_3 + y_4 \cdot r_4 = 0$

y_1, y_2, y_3 must be zero $\Rightarrow N(A) = \alpha \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$.

In general, $\dim(N(A)) = m - r$.

Let us now consider a more general matrix A.

For ex:

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

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{E_1} A = R$$

$\Rightarrow A = E^{-1} R$ where $E = PE_1$.

(1.) A has the same row space as R. The same dimension r & the same basis.

Because $A = E^{-1}R$, every row of A is a lin. comb. of rows of R. $\Rightarrow C(A^T) \subset C(R^T)$.

$C(R^T) \subset C(A^T)$ because $R = EA$

$\Rightarrow$ First r rows of R form a basis for $C(A^T)$ & $C(R^T)$.

(2) The colum space of A has dimension r.

The number of indep. columns = # indep. rows

Is $C(A) = C(R)$ in general? No. Why not?

We see that if $x \in C(R) \Rightarrow x_4 = 0$ but this need not be the case for a vector from $C(A)$.

Defn: Pivot columns of A are those columns that turn into the pivot columns of R.

For our ex, these are the 1^{st}, 3^{rd} & 5^{th} columns.

Claim: Pivot columns form a basis for $C(A)$

Pf for our ex.: We need to show that

(i) Pivot columns are indep.

(ii) span (piv. columns) = $C(A)$

(i) Suppose this is not true, i.e. we can find a non-zero linear comb. (x_1, x_3, x_5) of pivot column that give 0.

Then we have $c_1 \cdot x_1 + c_3 x_5 + c_5 \cdot x_5 = 0$

But we have $\underbrace{[c_1 \ c_3 \ c_5]}_{C} = E^{-1}(I)$

$\Rightarrow Cx = E^{-1}x = 0$

But if $E^{-1}x = 0 \Rightarrow x = E \cdot 0 = 0 \Rightarrow$ a contradiction

$\Rightarrow$ pivot columns must be independent.

(ii) First notice $sp(\text{piv. columns}) \subset sp(\text{all columns}) = C(A)$.

If we can show $C(A) \subset sp(\text{piv.})$ we are done

Take $b \in C(A)$. $\Rightarrow \exists\, x$ s.t. $Ax = b$.

If we can find y with $y_{free} = 0$ and $Ay = b$, this would imply that $b \in sp(\text{piv. col})$

Now let c be the vector from $N(A)$ with $c_{free} = -x_{free}$ $\Rightarrow$ $A(x+c) = b$ &

for $y = x + c$, $y_{free} = 0$. We are done.

(3.) $N(A) = N(R)$, Same dimension $n-r$, same basis.

(4.) $N(A^T)$ has dim $m-r$

Important facts: $\dim(C(A)) = \dim(C(A^T)) = \text{rank}$.

$\dim(C(A)) + \dim(N(A)) = n$

Rank one Matrices

$$A = \begin{bmatrix} 1 & -2 \\ 2 & -4 \\ 4 & -8 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & -2 \end{bmatrix}$$

Every rank-1 matrix has the form

$A = u \cdot v^T$ where u, v are length-m & length-n vectors.

Orthogonality

Two vectors are said to be orthogonal when their dot product is zero:

$$v \cdot w = v^T w = \sum v_i w_i = 0.$$

Recall that $\|v\|^2 = v^T v > 0$ ($=0$ only if $v=0$)

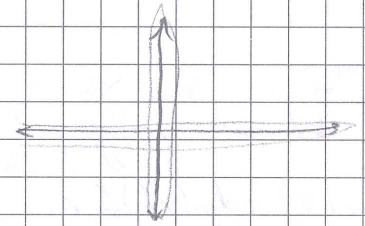
For orthogonal vectors, we have

$\|v+w\|^2 = (v+w)^T(v+w) = \|v\|^2 + \|w\|^2$.

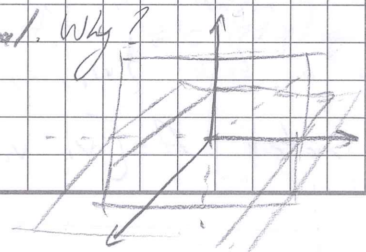
Let V, W be two subspaces of R^n. We say that V and W are orthogonal if every vector $v \in V$ is orthogonal to every vector $w \in W$.

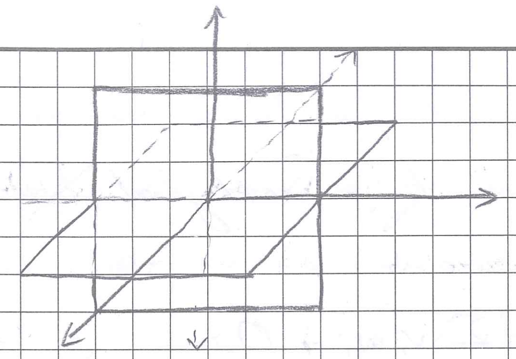
$$\boxed{v^T w = 0 \text{ for all } v \in V \text{ \& all } w \in W.}$$

Ex:



Ex: In R^3, $z=0$ (x-y plane) and $y=0$ (x-z plane) are not orthogonal. Why?





Intersection is the x-axis.

ie $V \cap W = \alpha \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

Claim: If two vector spaces are orthogonal, their intersection is the zero vector.

Pf: Let $v \in V$, $v \in W$,

Then $v^T v$ must be 0. This is only possible if $v = 0$.

Claim: Suppose V and W are orthogonal subspaces of $\mathbb{R}^n$. Then $\dim(V) + \dim(W) \leq n$.

Pf: Let $k = \dim V$, $m = \dim W$.

Then V has a basis $\{v_1, \dots, v_k\}$

W has a basis $\{w_1, \dots, w_m\}$

Consider the matrix $A = [v_1 \ v_2 \dots v_k \ w_1 \dots w_m]$

A is $n \times (m+k)$.

We will show that if $(m+k) > n$, V & W have a non-zero intersection.

The collection $\{v_1, \dots, v_k, w_1, \dots, w_m\}$ contains more than n elements. Therefore they cannot be linearly independent. In particular, considering the null-space of A, it has to be non-empty.

Let $y \in N(A)$, $y \neq 0$

$$\Rightarrow \sum_{i=1}^{k} y_i v_i = -\sum_{i=1}^{m} y_{i+k} w_i = x$$

$x \neq 0$ (Why?)

$x \in V$, $x \in W$! $\Rightarrow$ V & W cannot be orthogonal.

Prop: $N(A)$ and $C(A^T)$ are orthogonal subspaces of $\mathbb{R}^n$.

Pf. $x \in N(A) \Rightarrow Ax = 0$

$y \in C(A^T) \Rightarrow y = A^T c$ for some c.

$y^T x = (c^T A) x = c^T (Ax) = 0$.

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

$N(A) = \alpha \cdot \begin{bmatrix}1\\-1\\1\end{bmatrix}$

$y^Tx = \left(\beta_1 \begin{bmatrix}1 & 2 & 1\end{bmatrix} + \beta_2 \begin{bmatrix}2 & 3 & 1\end{bmatrix}\right) \cdot \alpha \begin{pmatrix}1\\-1\\1\end{pmatrix}$

$= 0$

Prop $N(A^T)$ and $C(A)$ are orthogonal subspaces.

Take a look at the dimension of the four fundamental spaces.

For $m \times n$ A, $N(A)$ and $C(A^T)$ are subspace of $\mathbb{R}^n$. $\Rightarrow \dim(N(A)) + \dim(C(A^T))$ must not exceed n. In fact, $\underbrace{\dim(N(A))}_{n-r} + \underbrace{\dim(C(A^T))}_{r} = n$

Similarly, $N(A^T)$, $C(A)$ are subspaces of $\mathbb{R}^m$ and $\underbrace{\dim(N(A^T))}_{m-r} + \underbrace{\dim(C(A))}_{r} = m$.

Defn: The orthogonal complement of a subspace V contains every vector that is perpendicular to V. This subspace is denoted by $V^{\perp}$.

1

A^{+1a}

Lemma: Let $\{v_1, \dots, v_k\}$ be a set of linearly independent vectors. If $v \neq 0$ is orthogonal to $\{v_1, \dots, v_k\}$ then the set of vectors $\{v_1, \dots, v_k, v\}$ are linearly independent.

Pf: Suppose not. Then,

$v = a_1 v_1 + a_2 v_2 + \dots + a_k v_k$

But then $v^Tv = a_1 v^Tv_1 + a_2 v^Tv_2 + \dots + a_k v^T v_k = 0$

which is a contradiction since $v \neq 0$.

Prop $C(A^T) = N(A)^{\perp}$ (Prop 1 kullanarak).

First, since $C(A^T)$ is orthogonal to $N(A)$, we have, $C(A^T) \subset N(A)^{\perp}$.

Now let $x \in N(A)^{\perp}$. Then if $x \notin C(A^T)$ we can form the matrix

$B = \begin{bmatrix}A\\ x^T\end{bmatrix}$.

We have $N(B) = N(A)$, but

$\dim(C(B^T)) = \dim(C(A^T)) + 1$.

$\Rightarrow \dim(C(B^T)) + \dim(N(B)) = n+1$; a contradiction.

---

* We note that every $x \in \mathbb{R}^n$ can be written uniquely as $x = x_r + x_n$ where $x_r \in C(A^T)$, $x_n \in N(A)$.

* For every $b \in C(A)$ we can find a unique $x_r \in C(A^T)$

---

(2) Prop: Any k linearly indep vectors in a k-dim space V, must span V.

Pf: Let $\{v_1, \dots, v_k\} \in V$, be a basis for V. and $\{d_1, \dots, d_k\} \in V$ be lin. indep.

Then $\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_k \end{bmatrix} = A \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}$

A is $k \times k$. Suppose $d \notin \text{span}\{d_1, \dots, d_k\}$.

$\Rightarrow \begin{bmatrix} d_1 \\ \vdots \\ d_k \\ d \end{bmatrix} = \underbrace{\begin{bmatrix} A \\ x \end{bmatrix}}_{B} \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}$

$B : (k+1) \times k \Rightarrow N(B^T)$ contains non-zero vectors.

$\exists$ a vector $c \neq 0$ s.t. $c^T \cdot B = 0$

$\Rightarrow c^T \begin{bmatrix} d_1 \\ \vdots \\ d_k \\ d \end{bmatrix} = c^T B \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix} = 0$

$\Rightarrow c_{k+1} d = c_1 d_1 + c_2 d_2 + \dots + c_k d_k$

since d_i's are lin. indep, $c_{k+1} \neq 0$

$\Rightarrow d = \frac{c_1}{c_{k+1}} d_1 + \dots + \frac{c_k}{c_{k+1}} d_k \Rightarrow$ a contradiction.

In general, if V is a subspace of $\mathbb{R}^n$, any $x \in \mathbb{R}^n$ can be written uniquely as

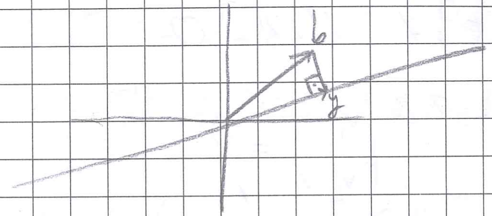
$x = x_1 + x_2$ where $x_1 \in V$, $x_2 \in V^\perp$.

Let us now see how this can be done

Projections:

Let p be a given vector, and let L be a line through the origin. (L is a subspace in this case)

How can we find the point on L, closest to p?



Geometrically it is the point y on L such that $p-y$ is orthogonal to L.

In general, for given b, L is described as αx where $\alpha \in \mathbb{R}$, and x is a vector on L.

Projection yields a $\hat{\alpha}$ such that $\|b - \hat{\alpha} x\|^2$ is minimized. Using our geometric intuition, we must have that the error vector $e = b - \hat{\alpha} x$ is orth. to L (i.e. to αx). That is

$$e^T \alpha x = \alpha (b - \hat{\alpha} x)^T x = 0$$

$$\Rightarrow b^T x = \hat{\alpha} x^T x \Rightarrow \hat{\alpha} = \frac{b^T x}{x^T x}.$$

Algebraically to minimize $\|b - \hat{\alpha} x\|^2$ we have

$$(b - \hat{\alpha} x)^T (b - \hat{\alpha} x) = b^T b - 2 \hat{\alpha} b^T x + \hat{\alpha}^2 x^T x$$

Set the derivative wrt $\hat{\alpha}$ to zero $\Rightarrow$

Notice that we can write the projection as

$$\hat{b} = x \hat{\alpha} = \frac{x x^T}{x^T x} b$$

$\Downarrow$

This is the projection matrix (rank-1).

Ex: Find the projection of $b = \begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix}$ onto the line through $x = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$.

$$P = \frac{\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \end{bmatrix}}{\begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}} = \frac{\begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix}}{6}$$

$$\Rightarrow Pb = \begin{bmatrix} 12 \\ 12 \\ 24 \end{bmatrix} / 6 = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}$$

Check: is $(b - Pb)^T x = 0$?

$$(b - Pb)^T x = \begin{bmatrix} 2 & 2 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = 0$$

What is the projection of $c = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}$ onto the same line?

$Pc = c$. (It's already on the line).

$\Rightarrow$ We have, actually $P^2 = P$.

Projection onto a Subspace:

For a subspace, V, the projection of a point b onto V is the point $p \in V$ that is closest to b.

Assuming we have a basis for V as $\{v_1, v_2, \dots, v_k\}$, the problem becomes

Problem: Find the combination

$p = x_1 v_1 + x_2 v_2 + \dots + x_k v_k$ that is closest to b.

We will once again make use of the geometric result: $(b-p)$ is orthogonal to the subspace.

$$\Rightarrow p = \underbrace{\begin{bmatrix} v_1 & v_2 & v_k \end{bmatrix}}_{V} \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}$$

$(b - Vx)^T v_1 = 0$

$(b - Vx)^T v_2 = 0 \qquad \Rightarrow \qquad V^T(b - Vx) = 0$

$(b - Vx)^T v_k = 0$

$\Rightarrow V^T V x = V^T b$

$V^T V$ is invertible if v_i's are l-i.

Why? Suppose not. Then $\exists\ y \neq 0$ s.t. $V^T V y = 0$

$\Rightarrow y^T V^T V y = 0$

But $Vy = x \neq 0 \Rightarrow y^T V^T V y = x^T x \neq 0$

A contradiction. $\Rightarrow V^T V$ must be invertible.

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

$$\Rightarrow p = Vx = \underbrace{V (V^T V)^{-1} V^T}_{\text{projection matrix} = P} b$$

Notice $P^2 = P$.

Compare this with the 1-D case!

3-Aralık, Bunadan başlar.

Given a vector space S, suppose we have the projection matrix P. Then given x, we can decompose it as

$$x = \underbrace{Px}_{x_s} + \underbrace{(I-P)x}_{x_{s'}}$$

$x_s \in S$ and $x_{s'} \in S^\perp$

To see this, multiply $x_{s'}$ by $P \Rightarrow P x_{s'} = 0$.

This decomposition is unique. That is, if

$x = x_S + x_{S'}$ with $x_S, y_S \in S$

$x = y_S + y_{S'}$ $\quad x_{S'}, y_{S'} \in S^{\perp}$

then $x_S = y_S$, $x_{S'} = y_{S'}$.

Because $\underbrace{x_S - y_S}_{\in S} = \underbrace{x_{S'} - y_{S'}}_{\in S'}$

The only vector common to S & S' is 0.

Review:

$C(A) \perp N(A^T) \Rightarrow C(A) \cap N(A^T) = \{0\}$

$N(A) \perp C(A^T) \Rightarrow N(A) \cap C(A^T) = \{0\}$

Let us consider $Ax = b$.

If $b \in C(A)$, we can find an x that satisfies the equation.

If $b \notin C(A)$, we can try to approximate b as much as possible, using the columns of A, i.e. minimize $\|Ax - b\|^2$

A vector x for which $\|Ax - b\|^2$ is minimum is called a least-squares solution.

How to find it? (2 cases): (1) A has indep columns
(2) Not

(2) Suppose that $b \notin C(A)$.

Since $C(A) = N(A^T)^{\perp}$, we can decompose b as

$b = b_1 + b_2$ where $b_1 \in C(A)$

$b_2 \in N(A^T)^{\perp}$.

Let us define $e = b - Ax$. Our task is to minimize $\|e\|^2$.

Now since $Ax \in C(A)$, we have that

$b_1 - Ax \in C(A)$

$\Rightarrow e = \underbrace{(b_1 - Ax)}_{\in C(A)} + \underbrace{b_2}_{\in C(A)^{\perp}}$

$\Rightarrow \|e\|^2 = \|b_1 - Ax\|^2 + \|b_2\|^2$

Minimizing $\|e\|^2$ is equivalent to minimizing $\|b_1 - Ax\|^2$.

If the columns of A are independent this is easy to do. $\underbrace{b_1 = A(A^TA)^{-1}A b}_{\text{solves the problem.}}$

What to do when the columns of A are not independent?

Consider $A^T b$.

Notice that $A^T b = A^T b_1$ → has a solution $\in C(A)$

$\Rightarrow A^T A x = A^T b_1 \rightarrow \in C(A^T)$ → Q: Does this system always have soln?

Suppose y solves this system of equations.

Does it also solve $Ay = b_1$?

Suppose not ie $Ay = c \neq b_1$, $c \in C(A)$

$\Rightarrow A^T(Ay - c) = A^T Ay - A^T c = A^T b_1 - A^T c$

$= A^T(b_1 - c) = 0 \Rightarrow (b_1 - c) \in C(A)^\perp$

$\in C(A)$

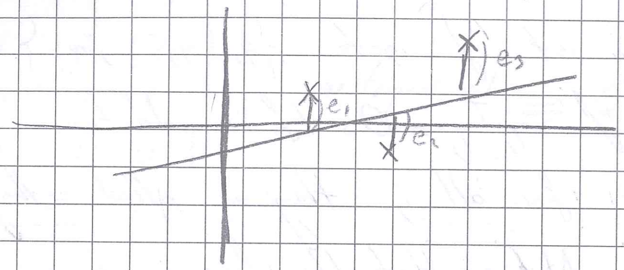
$\Rightarrow b_1 - c = 0 \Rightarrow b_1 = c$.

Recipe: To minimize $\|Ax - b\|^2$,

solve $A^T A x = A^T b$

(This system always has a solution)

Ex: Given the points on the plane



find the line that minimizes the error

$e_1^2 + e_2^2 + e_3^2$.

Let the points be $(2, 1)$, $(3, -1)$, $(4, 3)$.

A line is described by the equation $Cx + D$

$\Rightarrow e_1 = 1 - (C \cdot 2 + D)$

$e_2 = -1 - (C \cdot 3 + D)$

$e_3 = 3 - (C \cdot 4 + D)$

$$\Rightarrow \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$$

$\Rightarrow$ minimize $\|Ax - b\|^2$ where $A = \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$

For this, solve $A^T A x = A^T b$

Orthogonal Bases & Gram-Schmidt

We call a set of vectors $\{q_1, q_2, \dots, q_n\}$ orthogonal if $q_i^T q_j = 0$. If in addition, $q_i^T q_i = 1$ for all i, they are called orthogonal unit vectors. Notice that they form a basis for their span — this basis is said to be orthonormal.

Consider the matrix $Q = \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix}$.

We have $Q^T Q = \begin{bmatrix} q_1^T \\ \vdots \\ q_n^T \end{bmatrix} \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix} = I$

In particular, if Q is a square matrix, we have

$Q^T = Q^{-1}$.

In this case Q is called an orthogonal matrix.

Ex 1: 'Rotation Matrix'

$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$Q^T Q = \begin{bmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \cos^2\theta + \sin^2\theta \end{bmatrix}$$

Ex 2: Permutation Matrix:

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$ orthonormal columns $\Rightarrow Q^T Q = I$.

Fact ① If Q has orthonormal columns, then it doesn't change the length of a vector

i.e. $\|x\|^2 = \|Qx\|^2$.

② It also preserves inner products: $x^T y = (Qx)^T (Qy)$

Converses of these statements are also valid:

If Q doesn't change the lengths of vectors $\Rightarrow Q$ is orth.

Projections using orthogonal bases.

Suppose that our basis vectors are orthonormal.

Recall that the projection matrix is

$P = A(A^T A)^{-1} A^T \Rightarrow$ Since $Q^T Q = I$

We've $P = Q(Q^T Q)^{-1} Q^T = Q Q^T$

$$= \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix} \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix} b = q_1 \cdot (q_1^T b) + \cdots + q_n (q_n^T b)$$

This gives a decomposition of b in terms of the columns of q, which

<u>Ex</u>: $Q = \begin{bmatrix} -1 & +2 \\ 2 & -1 \\ 2 & 2 \end{bmatrix} / 3$ $\qquad \left(N(Q^T) = \alpha \cdot \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \right)$

Suppose $b = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

$\Rightarrow$ Projection of b onto the column-space of Q is:

$$Pb = q_1 \cdot (q_1^T b) + q_2 \cdot (q_2^T \cdot b) = +\frac{1}{9} \begin{bmatrix} +1 \\ -2 \\ -2 \end{bmatrix} + \frac{2}{9} \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$$

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

$$Pb - b = \frac{1}{9} \begin{bmatrix} -4 \\ -4 \\ 2 \end{bmatrix} = -\frac{2}{9} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \in N(Q)$$

<u>Gram-Schmidt Process</u>

Given a set of linearly indep. vectors $a_1, \dots, a_n$

Obtain a set of orth. vectors $q_1, \dots, q_n$ that span the same space.

Let's do this for 4 vectors, a, b, c, d.

We want to find A, B, C, D, orthogonal.

Once we get A, B, C, D it's easy to obtain orthonormal vectors as $\frac{A}{\|A\|}, \frac{B}{\|B\|}, \frac{C}{\|C\|}, \frac{D}{\|D\|}$.

Set $A = a$.

$$B = b - \frac{A^T b}{A^T A} \qquad \left(B = b - Proj_A(b) \right)$$

$$C = c - \frac{B^T c}{B^T B} - \frac{A^T c}{A^T A} \qquad \left(C = c - Proj_{A,B}(c) \right)$$

<u>Ex</u>: $$\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix} / 9 \cdot \begin{bmatrix} 9 & 9 & 18 \\ 0 & 18 & 0 \\ 0 & 0 & 9 \end{bmatrix} = \begin{array}{ccc} \left[\begin{matrix} -1 \\ 2 \\ 2 \end{matrix}\right. & \begin{matrix} 3 \\ 0 \\ 6 \end{matrix} & \left.\begin{matrix} 0 \\ 6 \\ 3 \end{matrix}\right] \\ \shortparallel & \shortparallel & \shortparallel \\ a & b & c \end{array}$$

QR Factorization

Suppose A has independent columns $a_1, \dots, a_n$ $\Rightarrow$ from these we can obtain orthonormal vectors $q_1, \dots, q_n$ that span the same space. Moreover, Gram-Schmidt gives

* a_1 spans the same space as q_1
* a_1, a_2 " " q_1, q_2
* a_1, a_2, a_3 " " q_1, q_2, q_3.

$$\Rightarrow a_1 = q_1 (q_1^T a_1)$$

$$\Rightarrow a_2 = q_1 (q_1^T a_2) + q_2 (q_2^T a_2)$$

$$\Rightarrow a_3 = q_1 (q_1^T a_3) + q_2 (q_2^T a_3) + q_3 (q_3^T a_3)$$

$$\Rightarrow \underbrace{\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} q_1^T a_1 & q_1^T a_2 & q_1^T a_3 \\ 0 & q_2^T a_2 & q_2^T a_3 \\ 0 & 0 & q_3^T a_3 \end{bmatrix}}_{R}$$

$\Rightarrow A = QR$

$$A^T A x = A^T b \Rightarrow R^T R x = R^T Q^T b \Rightarrow Rx = Q^T b.$$

Properties of Determinants

Determinant assigns a number to a square matrix. We denote it by $\operatorname{Det}(A)$ or $|A|$.

Properties ① The determinant of the n by n Identity matrix is 1.

② The determinant changes sign when two rows are exchanged.

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

③ The determinant is a linear function of each row separately.

$$\begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

$$\begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}$$

(In particular $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ c & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & d \end{vmatrix}$

The rest of the properties can be follow from these.

4. If two rows of A are equal, then $\det A = 0$

(Follows from Rule 2) $\left(\begin{vmatrix} a & b \\ a & b \end{vmatrix} = - \begin{vmatrix} a & b \\ a & b \end{vmatrix} \right)$

5. Subtracting a multiple of one row from another row leaves $|A|$ unchanged

$$\begin{vmatrix} a & b \\ c-\alpha a & d-\alpha b \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} - \alpha \begin{vmatrix} a & b \\ a & b \end{vmatrix}$$

6. A matrix with a row of zeros has $|A|=0$.

$$\begin{vmatrix} a & b \\ 0 & 0 \end{vmatrix} = \begin{vmatrix} a & b \\ a & b \end{vmatrix}$$

7. If A is triangular, $|A| = \prod a_{ii}$

(product of diagonal entries)

$$\begin{vmatrix} a_{11} & b \\ 0 & a_{12} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 \\ 0 & a_{12} \end{vmatrix} = a_{11} a_{12} \quad \text{if } a_{12} \neq 0$$

if $a_{12} = 0 \Rightarrow \det A = 0$ anyway.

8. If A is singular, then $|A| = 0$.

If A is invertible, $|A| \neq 0$.

(Reduce A to U by elimination & use Prop 7)

$$|A| = \pm |U|$$

9. $|AB| = |A|\,|B|$

10. $|A^T| = |A|$

All of the rules that apply to the rows also apply to the columns because of Prop 10.

Computing The Determinant Using Rules 1, 2, 3.

"The big formula".

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$\begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \\ 0 & 0 & a_{23} \\ a_{31} & 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{vmatrix}$$

$$\begin{vmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ 0 & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ 0 & a_{22} & 0 \\ a_{31} & 0 & 0 \end{vmatrix}$$

$$= a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32}$$

$$- a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} - a_{13} a_{22} a_{31}$$

Eigenvalues and Eigenvectors

Given a square matrix A, a vector 'x' is said to be an eigenvector of A with eigenvalue λ if $Ax = \lambda x$.

Ex: $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ satisfies

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 4 \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{x_1}, \quad A \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \end{bmatrix} = -1 \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

Ex 2: $B = \dfrac{A}{4} \Rightarrow Bx_1 = \dfrac{A}{4} x_1 = \dfrac{4}{4} x_1 = x_1$

What is B^{100}?

$$Bx_2 = -\frac{1}{4} x_2$$

Instead of B^{100}, Compute $B^{100} \cdot [x_1 \ \ x_2] = D$

$$D = \left[1^{100} \cdot x_1 \quad \left(-\frac{1}{4}\right)^{100} x_2 \right]$$

$$\Rightarrow B^{100} = D \cdot [x_1 \ \ x_2]^{-1}$$

How do we compute the eigenvalues?

$Ax = \lambda x$ can be written as

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

$\Rightarrow \det(A - \lambda I) = 0$.

$\Rightarrow \det(A - \lambda I)$ is a polynomial of λ of degree n ($A: n \times n$). The roots of this polynomial give the eigenvalues. Once the eigvalues are determined, solve $(A - \lambda I)x = 0$ to find an eigenvector x.

Remarks: (1.) Elimination does not preserve the eigenvalues.

(2) $\det A = \lambda_1 \cdot \lambda_2 \cdots \lambda_n$

(3) Trace $A = \sum \lambda_i$

(4) The eigenvalues may be imaginary.

(5) For Q orthogonal, $Qx = \lambda x \Rightarrow |\lambda| = 1$.

(6) If $Ax = \lambda x \Rightarrow A(\alpha x) = \lambda(\alpha x)$

$\Rightarrow$ One can multiply an eigvector by α but the eigvalue doesn't change.

Diagonalizing a Matrix:

Suppose A has n linearly independent eigenvectors $x_1, \ldots, x_n$. Construct

$S = [x_1 \cdots x_n]$.

Then $AS = [\lambda_1 x_1 \; \lambda_2 x_2 \cdots \lambda_n x_n]$

$$= S \underbrace{\begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_n \end{bmatrix}}_{\Lambda}$$

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

S^{-1} exists because $x_1, \ldots, x_n$ linearly indep.

Without n indep. eigvectors, we cannot diagonalize.

How to Compute Powers of A?

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

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

How to Compute $A^k x$?

Write $x = c_1 x_1 + \dots + c_n x_n$

$\Rightarrow A^k x = c_1 \lambda_1^k x_1 + \dots + c_n \lambda_n^k x_n$

Remarks on Diagonalization

1. If the eigenvalues are different, then there exist n linearly indep. eigenvectors.

2. Some matrices do not have enough eigenvectors.

Ex: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ $\Rightarrow$ both eigenvalues $= 0$, only 1 eigenvectors

Pf. of 1: 3×3 case

$A x_1 = \lambda_1 x_1$

$A x_2 = \lambda_2 x_2$

$A x_3 = \lambda_3 x_3$

if not lin. indep. $\Rightarrow$ $\exists$ a linear combination

$c_1 x_1 + c_2 x_2 + c_3 = 0$

Difference Equations:

Suppose we define a system as

$$y(n) = a\,y(n-1) + b\,y(n-2).$$

Such an equation is called a linear difference equation. Given $y(0), y(1)$, we can compute

$$y(2) = a\,y(1) + b\,y(0)$$

$$y(3) = a\,y(2) + b\,y(1)$$

Q: Can we find a closed form expression for $y(n)$? Yes, if the "incidence matrix" can be diagonalized.

Let us define $u_n = \begin{bmatrix} y(n) \\ y(n-1) \end{bmatrix}$

Then we have,

$$u_n = \begin{bmatrix} a\,y(n-1) + b\,y(n-2) \\ y(n-1) \end{bmatrix} = \underbrace{\begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} y(n-1) \\ y(n-2) \end{bmatrix}}_{u_{n-1}}$$

$\Rightarrow u_n = A\,u_{n-1}$

$\Rightarrow u_n = A \cdot u_{n-1} = A^2 u_{n-2} \cdots = A^{n-1} u_1 = A^{n-1}\begin{bmatrix} y(1) \\ y(0) \end{bmatrix}$

$\Rightarrow$ Power of A become useful here.

Ex: $y(n) = y(n-1) + 2y(n-2)$; $y(0)=0$, $y(1)=1$
Compute $y(100)$.

$\Rightarrow A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ eigenvalues: solve $(1-\lambda)(-\lambda) - 2 = 0$

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

eigenvectors: for $\lambda = -1$,

$\begin{bmatrix} 2 & 2 \\ 1 & +1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow y = 1, \; x = -1$

$\Rightarrow A\begin{bmatrix} -1 \\ 1 \end{bmatrix} = -1\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

for $\lambda = 2$, $\begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

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

$\Rightarrow u_0 = \underbrace{\begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}}_{S}\begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$

$\Rightarrow \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix} = S^{-1} u_0$

$\Rightarrow A \cdot S = S\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$

$\Rightarrow u_{100} = A^{99} \cdot u_0 = S\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}^{99} S^{-1} u_0$

$= S$

Ex: $y(n) = a\,y(n-1) + b\,y(n-2) + c\,y(n-3) \cdots$

Symmetric Matrices

In general, a matrix may have complex eigenvalues and its eigenvectors may not be able to span the whole space. For symmetric matrices (i.e. $A = A^T$), the situation is more favorable.

For symmetric A.

1. All eigenvalues are real.
2. The eigenvectors can be chosen orthonormal
   ie $A = S\Lambda S^{-1}$ with S: orthonormal, $S^T = S^{-1}$
   $\Rightarrow A = S\Lambda S^T$

Spectral Thm

Ex: $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$

eigenvalues come from $(a-\lambda)(c-\lambda)-b^2=0$

$\Rightarrow \lambda^2-(a+c)\lambda+ac-b^2=0$

$$\lambda^2-(a+c)\lambda+\left(\frac{a+c}{2}\right)^2=b^2-ac+\left(\frac{a+c}{2}\right)^2$$

$$=b^2+\left(\frac{a}{2}\right)^2+\left(\frac{c}{2}\right)^2$$

$$\Rightarrow \lambda_{1,2}=(a+c)\pm\sqrt{b^2+\left(\frac{a}{2}\right)^2+\left(\frac{c}{2}\right)^2}$$

$\Rightarrow$ eigenvalues are real.

Proof that the eigenvalues are always real for symmetric A:

Suppose $Ax=\lambda x$. Because A is real, we have

$A\bar{x}=\bar{\lambda}\bar{x} \Rightarrow \bar{x}^T A^T=\bar{x}^T A=\bar{\lambda}\bar{x}^T$

$\Rightarrow \bar{x}^T A x=\lambda \bar{x}^T x$

$=\bar{\lambda}\bar{x}^T x \Rightarrow \bar{\lambda}=\lambda$; λ has to be real

The proof the second part of the spectral thm is easier if it's known that all eigenvalues are distinct.

Orthogonal eigenvectors when $\lambda_j\neq\lambda_j$: Pf:

Let $Ax_j=\lambda_j x_j$, $Ax_k=\lambda_k x_k$.

Then $x_j^T A x_k = x_j^T \lambda_k x_k$

$= \lambda_j x_j^T x_k$

$\Rightarrow x_j^T x_k = \langle x_j, x_k\rangle$

$=0$

Sum of Rank 1 matrices:

$$\Rightarrow A=Q\Lambda Q^T=\begin{bmatrix} q_1 & q_2 & \cdots & q_n\end{bmatrix}\begin{bmatrix}\lambda_1 & & 0\\ & \lambda_2 & \\ 0 & & \lambda_n\end{bmatrix}\begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T\end{bmatrix}$$

$$=\lambda_1 q_1 q_1^T+\lambda_2 q_2 q_2^T+\cdots+\lambda_n q_n q_n^T$$

$\Rightarrow A$ is a linear combination of Projection matrices onto orthogonal subspaces.

For the general proof that A (symmetric) can be decomposed as $A = Q\Lambda Q^T$ with $Q^T = Q^{-1}$, Λ diagonal, we need the so-called Schur Decomposition:

– Every square matrix can be decomposed as $A = QTQ^{-1}$ where T is upper triangular and $Q^T = Q^{-1}$!
If A has real eigenvalues, then Q & T can be chosen real:

We will prove the statement by induction. Let A: $n \times n$
Assume that this is possible for matrices of size $(n-1)\times(n-1)$.
Pick an eigenvector of A, namely q_1

$$\Rightarrow A q_1 = \lambda_1 q_1$$

Complete $[q_1]$ to an orthonormal matrix Q

$$A[q_1 \dots q_n] = Q \begin{bmatrix} \lambda_1 & b \\ 0 & \\ \vdots & T_2 \\ 0 & \end{bmatrix}$$

Now T_2 can be decomposed as $T_2 = Q_2 T_3 Q_2^T$

$$\Rightarrow AQ_1 = \underbrace{Q_1 \begin{bmatrix} \lambda_1 & b \\ 0 & Q_2 T_3 Q_2^T \end{bmatrix}}$$

$$= \underbrace{\begin{bmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & Q_2 \\ 0 & \end{bmatrix}}_{Q_3} \underbrace{\begin{bmatrix} \lambda_1 & b \\ 0 & \\ \vdots & T_3 \\ 0 & \end{bmatrix}}_{T} \underbrace{\begin{bmatrix} 1 & 0 \cdots 0 \\ \vdots & \\ 0 & Q_2^T \end{bmatrix}}_{Q_3^T}$$

$$\Rightarrow AQ_1 = Q_1 Q_3 \, T \, Q_3^T \qquad Q_1 Q_3 = Q$$

$$\Rightarrow A = QTQ^T \text{ with } T \text{ upper triangular}$$

For $A = A^T$

$$\Rightarrow (QTQ^T)^T = QT^TQ^T \Rightarrow T = T^T$$

$\Rightarrow T$ diagonal

$\Rightarrow Q$ must contain the eigenvectors, which are orthonormal!!

<u>Application</u>: Maximize $\dfrac{x^T A x}{\|x\|^2}$

Let $H = A^T A$.

Notice that $H = Q \Lambda Q^T$

where Q is orthonormal and Λ is diagonal.

Therefore $H = \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T + \dots + \lambda_n q_n q_n^T$

where $q_i q_i^T = P_i$ is a projection matrix to S_i, ($S_i \perp S_j$)

Also, $P_i^2 = P_i$ and $P_i^T = P_i$

$\Rightarrow x^T H x = \sum_i \lambda_i x^T P_i P_i x = \sum_i \lambda_i \|x_{P_i}\|^2$

where $\sum_i \|x_{P_i}\|^2 = 1$.

This is maximized if $\|x_{P_i}\| = 1$ where $i = \arg\max_i \lambda_i$

Therefore the solution is given by q_i.

For an alternative approach, consider the columns of Q: they form an orth. basis and any $x = \sum c_i q_i$ where $c = Q^T x$.

We also have that $\|Q^T x\| = \|x\|$.

Now $\max\limits_{x, \|x\|=1} x^T Q^T \Lambda Q x$ is the same as $y^T \Lambda y$

$= \sum_i \lambda_i y_i^2$

where $\sum y_i^2 = 1$

Positive Definite Matrices

We know that $\|Ax\|^2 \geq 0$ for any A & x.

But this means that

$x^T A^T A x \geq 0 \quad \forall x$.

a Ha

For $H = Q \Lambda Q^T$, if $x = q_i$, then we have,

$x^T H x = \lambda_i \Rightarrow \lambda_i \geq 0$.

All of the eigenvalues of H are non-negative.

Defn. A symmetric real H is called positive semi-defn. if $x^T H x \geq 0 \ \forall x$.

According to this defn. if $H = A^T A$ for some A, then it is p.d.

What if we do not have such an A at hand?

Write $H = Q \Lambda Q^T$.

Take $x = q_i \Rightarrow x^T H x = \lambda_i \Rightarrow \lambda_i \geq 0 \ \forall i$.

Write $\sqrt{\Lambda} = \begin{bmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{bmatrix} \Rightarrow \Lambda = \sqrt{\Lambda} \cdot \sqrt{\Lambda}$

$\Rightarrow H = \underbrace{(Q\sqrt{\Lambda})}_{A} \underbrace{(\sqrt{\Lambda} \cdot Q^T)}_{A^T}$

We have shown that the following are equivalent for symmetric H:

- $x^T H x \geq 0$
- $H = A^T A$ for some A.
- All of the eigenvalue are nonnegative.

SVD

For a non-symmetric (possibly rectangular A), we cannot have $A = Q \Lambda Q^{-1}$.

Instead of $Av = \lambda v_i$, we will have

$A v_i = \sigma u_i$.

Specifically, if A is $m \times n$, we will write $A = U \Sigma V$ where U: $m \times m$ orth.

V: $n \times n$ orth,

Σ: $m \times n$ diag.

Specifically u_i's & v_i's come from AA^T and $A^T A$ respectively.

In particular, let $v_1, \ldots, v_n$ ~~be~~ the eigvectors of $A^T A$ with $\|v_i\| = 1$, $v_i \perp v_j$.

Then, $A^T A v_i = \lambda_i v_i$ for $\lambda_i \geq 0$

$$AA^T (Av_i) = \lambda_i (Av_i)$$

$\Rightarrow Av_i$ is an eigenvector of AA^T with eigvalue λ_i. Moreover, $\|Av_i\|^2 = v_i^T A^T A v_i = \lambda_i$

Set $u_i = Av_i / \|Av_i\| = Av_i / \sqrt{\lambda} \Rightarrow Av_i = \sqrt{\lambda_i}\, u_i$

Then $u_i \perp u_j$ since $(v_i^T A^T) A v_j = 0$

If r of λ_i are positive, this gives us r u_i's. with

$$A[v_1 \ldots v_r \mid v_{r+1} \ldots v_n] = [\sigma_1 u_1 \ldots \sigma_r u_r \mid 0 \ldots 0]$$

Complete u_i's to an orth. basis using Gram-Sch. to obtain

$$A[v_1 \ldots v_n] = [u_1, \ldots, u_m] \left[\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \sigma_2 & & 0 \\ 0 & & \sigma_m & \end{array}\right]$$

Note: $\begin{cases} v_1, \ldots, v_r & \text{span } C(A^T) \\ v_{r+1}, \ldots, v_n & \text{span } N(A) \\ u_1, \ldots, u_r & \text{span } C(A) \\ u_{r+1}, \ldots, u_m & \text{span } N(A^T) \end{cases}$

Application: minimize $\|Ax-b\|$

$\|Ax-b\| = \|U\Sigma V x - b\| = \|\Sigma V x - U^T b\|$

$\Rightarrow$ solve $\min_y \|\Sigma y - U^T b\|$, set $x = V^T y$.

Norm of a Matrix

We define the norm of A : $\|A\| = \max \frac{\|Ax\|}{\|x\|}$.

Ex. $\|I\| = 1$.

Ex: If H is hermitian $\|H\| = \max_i |\lambda_i|$.

* In general, $\|A\| \neq \max_i |\lambda_i|$ but $\|A\| \geq \max_i |\lambda_i|$

* $\|A\| = \sqrt{\max \text{eig}\ (A^T A)}$

Ex: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ $\Rightarrow \|A\| = 1$

but all its eigenvalues $= 0$.

Note: $\|A\| = \max_i \sigma_i$ where $A = U \Sigma V$ with $\Sigma = \text{diag}(\sigma_i)$.

Note 2: $\|AB\| \leq \|A\| \|B\|$

Iterative Methods for Solving Linear Systems

Wish to solve $Ax = b$ (assume A is invertible).

Suppose $A = S - T$

$\Rightarrow Sx = Tx + b$

Define the iterations:

$$x_{n+1} = S^{-1} T x_n + S^{-1} b. \qquad (*)$$

Does this method converge?

Define $e_n = x - x_n$

Since $x = S^{-1} T x + S^{-1} b$, subtracting $(*)$, we find $e_{n+1} = S^{-1} T e_n$

$\Rightarrow e_{n+1} = (S^{-1} T)^n e_1$.

If $\underbrace{\|S^{-1} T\|}_{\rho =} < 1 \Rightarrow \|e_{n+1}\| \leq \|e_1\| \rho^n$

$\Rightarrow \|e_{n+1}\| \to 0 \Rightarrow x_n$'s converge to the solution.