Lecture 1: Course Overview and Background

Prof. Krishna R. Pattipati

Dept. of Electrical and Computer Engineering
University of Connecticut
Contact: krishna@engr.uconn.edu (860) 486-2890

ECE 6435

Adv Numerical Methods in Sci Comp

Fall 2008
August 27, 2008

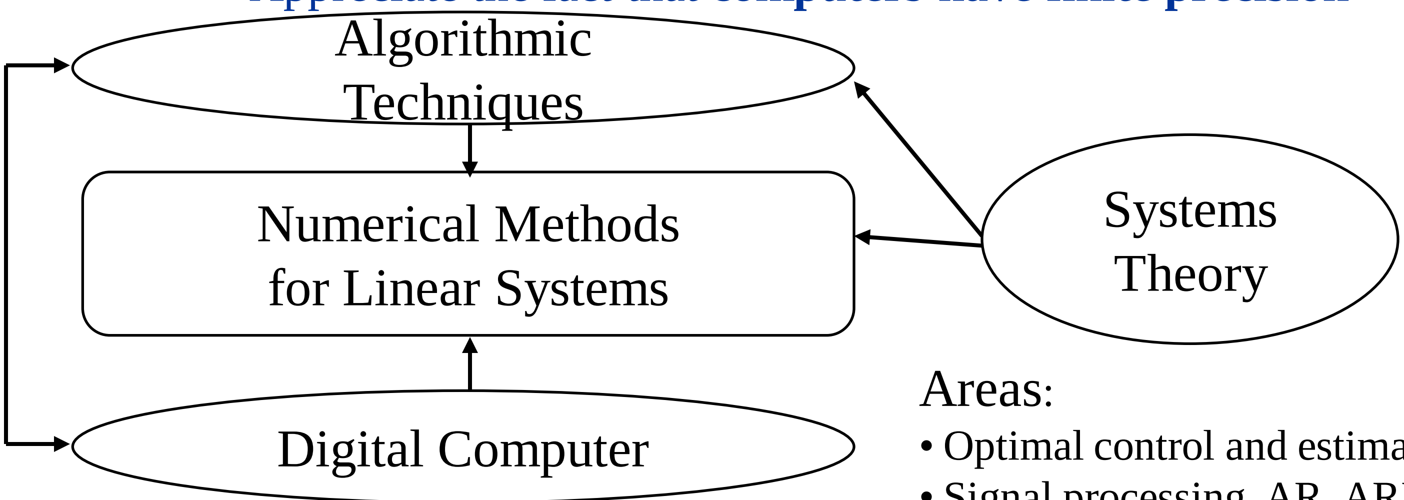
Outline of Lecture 1

- Course Objectives
- Round-off Errors
  - Computer representation of numbers
  - Machine accuracy
  - Illustration of round-off error problems via examples
- Background on Matrix Algebra
  - Matrix-vector notation
  - Matrix-vector product
  - Linear spaces associated with $A\underline{x}=\underline{b}$
  - Matrix inverse and pseudo inverse
  - Eigen values and Eigen vectors
  - Vector and Matrix Norms
  - Singular value decomposition (SVD)

Course Objectives

- ❖ Provide Systems analyst with effective software tools
- ❖ Use skills from Math. and CS to solve systems theory problems
- ❖ Three recurring Themes
  - Need to **understand mathematically the problem to be solved** (i.e., systems theory)
  - Express the problem **algorithmically**
  - Appreciate the fact that **computers have finite precision**

} CS



-Round-off Errors (finite precision)
Word length 32 bits (≈ 6-7 digit accuracy)
Word length 64 bits (≈ 13-14 digit accuracy)
- Truncation Errors (e.g., infinite summations truncated.)

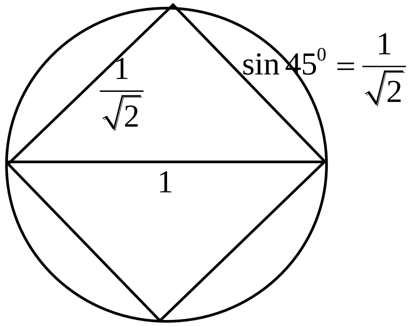
Areas:
- Optimal control and estimation
- Signal processing, AR, ARMA & LS
- Statistics, Multivariate Analysis
- Communication Theory

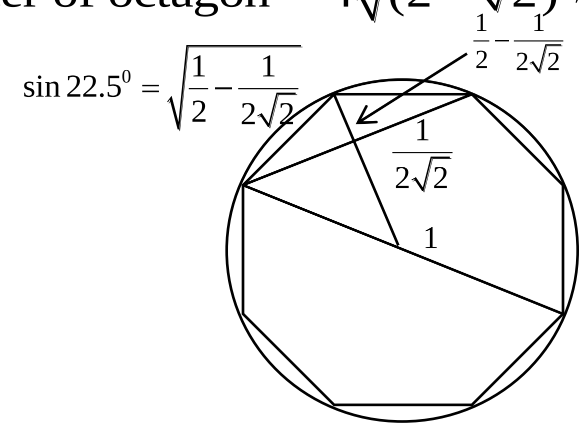
Round-Off Errors - 1

- **Computer solution is not the same as Hand calculation**
  - Humans have been computing for thousands of years
  - Famous examples are Pythagorean formula and Archimedes's approximation of π.

Example 1: Archimedes approximation of π

- Take a circle with diameter 1 $\Rightarrow$ circumference = π
- Idea: Approximate the circumference by the perimeters p_n of inscribed polygons with 2^n sides, n = 2, 3, ...
- $n = 2 \Rightarrow$ square $\Rightarrow$ perimeter of square $= 2\sqrt{2} = 2.828 < \pi$
- $n = 3 \Rightarrow$ octagon $\Rightarrow$ perimeter of octagon $= 4\sqrt{(2-\sqrt{2})} \approx 3.08 < \pi$





Round-Off Errors - 2

- Computer Algorithm:

Start with $p_2 = 2\sqrt{2}$
For $n = 3, 4, \ldots, 60$ DO

$$p_{n+1} = 2^n \sqrt{2\left(1 - \sqrt{1 - \left(\frac{p_n}{2^n}\right)^2}\right)}$$

End DO

- DIASTER!!!

Combination of underflow and catastrophic cancellations

- Solution is to rewrite the recursion (due to W. Kahan)

Let

$$p_{n+1} = 2^n \sqrt{r_{n+1}}; \quad r_{n+1} = 2\left(1 - \sqrt{1 - \left(\frac{p_n}{2^n}\right)^2}\right); \quad r_3 = \frac{2}{(2+\sqrt{2})}$$

Stable Recursion:

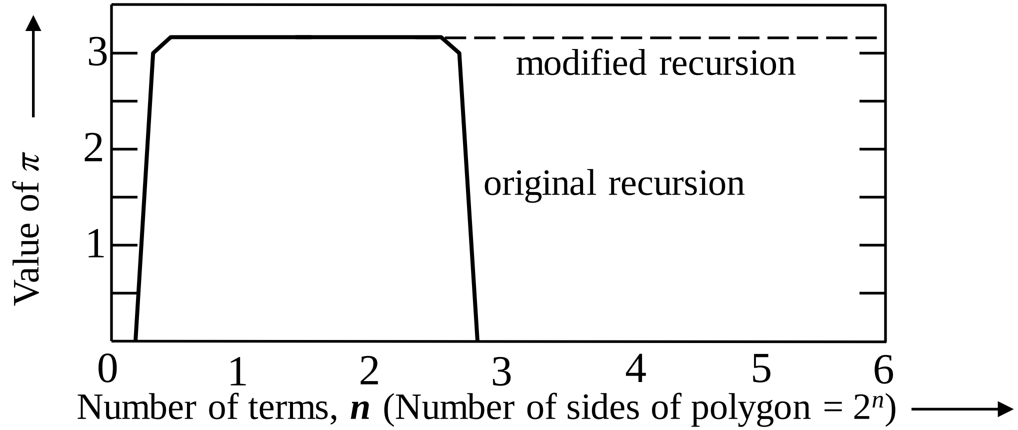
$$r_{n+1} = \frac{r_n}{2 + \sqrt{4 - r_n}}$$

$$p_{n+1} = 2^n \sqrt{r_{n+1}}; \quad r_{n+1} = 2\left(1 - \sqrt{1 - \frac{r_n}{4}}\right)$$

$$r_{n+1} = 2 - \sqrt{4 - r_n} = \frac{r_n}{2 + \sqrt{4 - r_n}}$$

Round-Off Errors - 3

Approximation of π by the Perimeter of Inscribed Polygons



- Alternatively, can show that for an 2^n-sided inscribed polygon, the perimeter is: $p_n = 2^n \sin(180°/2^n)$

Example 2: Taylor series for e^x

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

Let S_n = sum of first n terms = $\sum_{k=0}^{n} \frac{x^k}{k!}$

S_{n+1} = sum of first $(n+1)$ terms = $S_n + \frac{x^{n+1}}{(n+1)!}$

- How far do we go?
  - We sum terms until $S_{n^*} \approx S_{n^*+1} = S$

Round-Off Errors - 4

- Results on a VAX 780 (# bits per floating point word= 32)

x	S	e^x
1	2.718282	2.718282
5	148.4132	148.4132
10	22026.47	22026.46
15	3269017	3269017
20	4.8516531×10^{8}	4.8516531×10^{8}
-1	0.3678794	0.3678795
-5	6.7377836×10^{-3}	6.7379470×10^{-3}
-10	-1.6408609×10^{-4}	4.5399930×10^{-5}
-15	-2.2377001×10^{-2}	3.0590232×10^{-7}
-20	1.202966	2.0611537×10^{-9}

Works for $x>0$; but fails miserably for $x<0$

- Solutions: 1) Use $e^{-x} = 1/e^x$
  2) Integer x, use $e^{-x} = (e^{-1})^x$
  3) Much better methods than Taylor series (e.g., Chebyshev approximation, Pade approximation)
- These two examples illustrate the need to understand computer arithmetic & their effects on computation.

Computer Representation of Numbers

- Computer representation of numbers
  - Integers
    - On a 32 bit computer: Largest integer = $2^{31} - 1 = 2{,}147{,}483{,}647$
    - Smallest integer = -2^{31}
  - Floating point (real) arithmetic
    - On a 32 bit computer (IEEE standard): 24 bits for mantissa & 8 bits for exponent. One bit from each for sign
    - Largest floating point #: $+\,111\ldots1_2 \times 2^{127} \approx 1.7 \times 10^{38}$;
    - Smallest positive floating point # $\approx 10^{-38}$
    - Double precision $\Rightarrow$ Twice as many bits to represent each number.
- Machine Accuracy
  - Smallest (in magnitude) floating point number which when added to the floating point number 1.0 produces a floating point number different from 1 (also called macheps or ε_m or machine constant).

$\Rightarrow$Smallest $\varepsilon_m \ni 1 + \varepsilon_m \neq 1 \Rightarrow \varepsilon_m$ = machine accuracy (function of the # of bits in the mantissa)

$\Rightarrow$32-bit floating point arithmetic $\Rightarrow \varepsilon_m \approx 2^{-23} \approx 1.2 \times 10^{-7}$

$\Rightarrow$6-7 digit accuracy.

Illustration of Round-Off Error Problems - 1

Example 3: Addition of two positive floating point numbers x, y

- $x + y = x(1+y/x)$
- Note that if $y/x < \varepsilon_m$, $fl(x+y) \approx x$
- Suppose $y/x > \varepsilon_m$. what is the relative error?

$$\text{relative error} = \frac{|\text{true value} - \text{computed value}|}{|\text{true value}|}$$

- The mere act of reading a floating point number (e.g., 0.3) into the computer causes an error. so, $x_{stored} = x(1+\delta_x)$ or $x_{stored} - x = x\,\delta_x$; $\delta_x \le \varepsilon_m$

$$\text{relative error} = \frac{|(x+y)-(x_{stored}+y_{stored})|}{|(x+y)|}$$

$$= \frac{|(x\delta_x + y\delta_y)|}{|(x+y)|} \le \max(\delta_x, \delta_y) \le \varepsilon_m$$

$\Rightarrow$ Every floating point operation introduces a fractional round-off error of as much as ε_m

Round-off Error →
- $\sqrt{N}\varepsilon_m$ irregular computations; up or down equal prob., mean = 0, $\text{sd} = \sqrt{N}\,\frac{\varepsilon_m}{2}$
- $N\,\varepsilon_m$ regular N = # of operations
- also depends on implementation

Illustration of Round-Off Error Problems - 2

Example 4: Catastrophic cancellation

- Suppose $y - x \approx 2\,\varepsilon_m$. then

$$\text{relative error} = \frac{|(y-x)-(y_{stored}-x_{stored})|}{|(y-x)|} = \frac{|\delta_y - \delta_x|}{|y-x|} \le 1$$

$\Rightarrow$ Difference can be wrong in every digit

- Solutions:
  - Use double precision arithmetic $(\varepsilon_m = 2^{-52} = 2.22\,\text{x}\,10^{-16})$
  - Modify algorithm to minimize catastrophic cancellations (e.g., approximation to π and e^x for x<0 from $1/e^{|x|}$)

Example 5: Roots of a quadratic equation

- Subtraction of nearly equal numbers is dangerous!!

$$x = \frac{-b \pm \sqrt{b^2-4ac}}{2a} \quad \text{if } ac << b^2 \text{ trouble!!}$$

$$\text{alternative } x = \frac{2c}{-b \mp \sqrt{b^2-4ac}} \quad \text{same problem; but know } x_1 x_2 = \frac{c}{a}$$

$$\text{compute } q = -\frac{1}{2}\left(b + \text{sgn}(b)\sqrt{b^2-4ac}\right);\ x_1 = \frac{q}{a},\ \ x_2 = \frac{c}{q} \quad \text{stable}$$

Illustration of Round-Off Error Problems - 3

Example 6: Implement ∋ round-off errors do not magnify with iteration

- Golden section # φ = .61803398; $\varphi^2=1-\varphi$, $\varphi^3=\varphi-\varphi^2, \ldots$
- In general, $\varphi^{n+1} = \varphi^{n-1}-\varphi^n$ DISATER (Why?)
- If $n = 1$ change Equation to $x^2 + x - 1 = 0 \Rightarrow$

$$x = \frac{\sqrt{5}-1}{2}; \quad -(\frac{\sqrt{5}+1}{2})$$

Try it!!

- Stable $\varphi^{n+1} = \varphi \cdot \varphi^n$
- Suppose interested in only φ^n, $n = 2^L$

$$\gamma \leftarrow \phi$$

DO $i = 1, \ldots, L$

$$\gamma \leftarrow \gamma^2$$

END DO

$\log_2 n$ operations
DOUBLING

- The concept of **doubling** is extremely useful in computing e^{At} and its integrals

Illustration of Round-Off Error Problems - 4

Example 7: Kalman filter covariance equation

Usual → $P_{k+1/k} = \Phi P_{k/k-1}\,\Phi^T + Q_k - \Phi P_{k/k-1} H^T (H P_{k/k-1} H^T + R_k)^{-1} H P_{k/k-1}\,\Phi^T$

Joseph's form → $P_{k+1/k} = \Phi(I-G_kH)P_{k/k-1}(I-G_kH)^T\,\Phi^T + Q_k + G_k R_k G_k^T$

Twice the computation as original

Square root → Achieve stable updates with *approximately* same cost as usual (Lectures 5 and 8).

Example 8: Accumulate all inner products in double precision (e.g., $\underline{x}^T\,\underline{x}$)

```
sum = 0.0D+00
DO i = 1, ..., n
      sum = sum + x(i)*x(i)
END DO
```

Scalar Case:

$$p = \phi^2 p\left(1 - \frac{h^2 p}{h^2 p + r}\right) + q$$

$$= \phi^2 p \frac{r}{h^2 p + r} + q$$

Example 9: Some problems are inherently bad (ill-conditioned)

Roots of a quartic : $x^4 - 4x^3 + 8x^2 - 16x + 15.99999999 = (x-2)^4 - 10^{-8} = 0$

Actual Solution:

$x_1 = 2.01,\; x_2 = 1.99,\; x_3 = 2 + .01i,\; x_4 = 2 - .01i$

Suppose $\varepsilon_m > 10^{-10} \Rightarrow$ computer solves $(x-2)^4 = 0$

Illustration of Round-Off Error Problems - 5

⇒ Small changes in coefficients lead to large changes in solution

⇒ Such problems are termed *ill-conditioned* (not the fault of the algorithm)

Example 10: Two linear equations (intersection of "nearly" parallel lines)

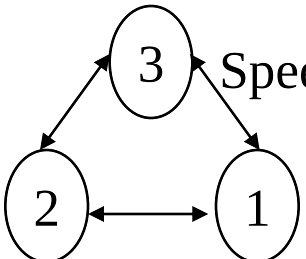
$$\left.\begin{aligned} 0.66x + 3.34y &= 4 \\ 1.99x + 10.01y &= 12 \end{aligned}\right\} \text{solution } = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

- Suppose we change right hand side to (3.96, 11.94),

  the new solution is $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \end{pmatrix}$

⇒ Small changes in coeff. ⇒ Large changes in solution ⇒ ill-conditioned

- For any algorithm want:



Speed < polynomial → GOOD ALGORITHM; non-polynomial

GOOD ALGORITHM + GOOD IMPLEMENTATION > key

Accuracy
finite precision

REALIABLITY
(robustness, stable) → GOOD IMPLEMENTATION; → Good computer (do not have much control)

Background on Matrix Algebra

- Vector – Matrix Notation

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$ a column vector of dimension n

- $x_i \in R$ $\qquad$ $x_i \in [-\infty, \infty]$
- $\underline{x} \in R^n$ $\qquad$ $\underline{x} \in C^n$ for complex numbers
- $A = [a_{ij}]$ $m \times n$ matrix $\in R^{mn}$
- $A^T = [a_{ji}]$ $n \times m$ matrix $\in R^{nm}$
- A square $n \times n$ matrix is symmetric, if $a_{ij} = a_{ji}$ $\begin{bmatrix} 2 & 4 \\ 4 & 11 \end{bmatrix}$ symmetric
- Idempotent if $A^2 = A$ (very useful to validate linear systems software)

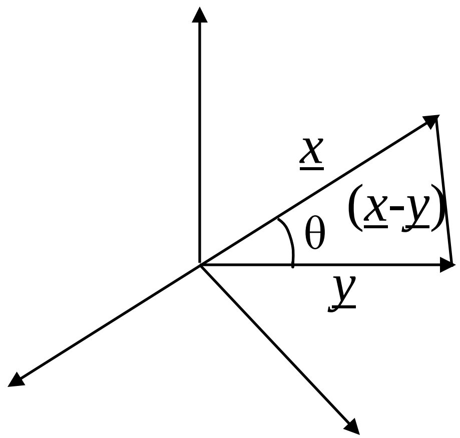
  e.g., $A = \begin{bmatrix} 4/5 & -2/5 \\ -2/5 & 1/5 \end{bmatrix}; e^A = I + \sum_{i=1}^{\infty} \frac{A^i}{i!} = I + (e-1)A = \begin{bmatrix} 2.3746 & -0.6873 \\ -0.6873 & 1.3437 \end{bmatrix}$

- Diagonal matrix: $A = \begin{bmatrix} \mu_1 & & 0 \\ & \mu_2 & \\ 0 & & \mu_n \end{bmatrix} = Diag\,(\mu_1, \mu_2, \ldots, \mu_n) = D\,(\mu_1, \mu_2, \ldots, \mu_n)$
- Identity matrix: $I_n = Diag\,(1\ 1\ \ldots\ 1)$

Matrix-Vector Notation - 1

- A matrix is PD if $\underline{x}^T A \underline{x} > 0 \ \forall \underline{x} \neq \underline{0}$; PSD if $= \underline{x}^T A \underline{x} \geq 0$
- Note: $\underline{x}^T A \underline{x} = \underline{x}^T A^T \underline{x} \Rightarrow \underline{x}^T A \underline{x} = \underline{x}^T [(A+A^T)/2] \underline{x}$

  $\left(\dfrac{A+A^T}{2}\right)$ is called symmetrized part of A
- If A is skew symmetric, $A^T = -A \Rightarrow \underline{x}^T A \underline{x} = 0$ $\quad e.g., A = \begin{bmatrix} 0 & 4 \\ -4 & 0 \end{bmatrix}$
- $A = Diag(\mu_i) \Rightarrow \underline{x}^T A \underline{x} = \sum_{i=1}^{n} \mu_i x_i^2$
- We will study properties of PD matrices later
- Vector $\underline{x}$ is an $n \times 1$ matrix
- $\underline{x}^T \underline{y}$ = inner (dot, scalar) product $= \sum_{i=1}^{n} x_i y_i$ (a scalar)



$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\|\underline{x}-\underline{y}\|^2 = (x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2$$
$$= (x_1^{\,2} + x_2^{\,2} + x_3^{\,2}) + (y_1^{\,2} + y_2^{\,2} + y_3^{\,2})$$
$$-2(x_1 y_1 + x_2 y_2 + x_3 y_3)$$

Also know $(\underline{x}^T \underline{x}) + (\underline{y}^T \underline{y}) - 2\sqrt{(\underline{x}^T \underline{x})(\underline{y}^T \underline{y})}\cos(\theta)$

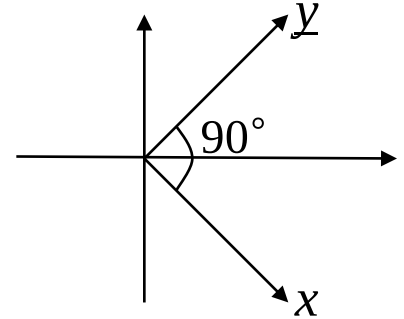
$$\Rightarrow \cos\theta = \frac{\underline{x}^T \underline{y}}{\sqrt{(\underline{x}^T \underline{x})(\underline{y}^T \underline{y})}} = \frac{\underline{x}^T \underline{y}}{\|\underline{x}\|_2 \, \|\underline{y}\|_2}$$

Matrix-Vector Notation - 2

$$\begin{matrix} \underline{x} & \underline{y} \end{matrix}$$
$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \Rightarrow \frac{4}{5} = \cos\theta \Rightarrow \theta = \cos^{-1}(0.8) \cong 36.9^\circ$$

- $\theta = 90^\circ \Rightarrow \underline{x}$ and $\underline{y}$ are perpendicular to each other
  $\Rightarrow$ ORTHOGONAL $\Rightarrow \underline{x}^T\underline{y} = 0$, e.g.,

$$\begin{matrix} \underline{x} & \underline{y} \end{matrix}$$
$$\begin{bmatrix} .8 & .6 \\ -.6 & .8 \end{bmatrix}$$



Matrix-vector product

$$A\underline{x} = \begin{bmatrix} 2 & 4 & 5 \\ 1 & 2 & 6 \\ 3 & 1 & 2 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \\ 4 \end{bmatrix} x_1 + \begin{bmatrix} 4 \\ 2 \\ 1 \\ 5 \end{bmatrix} x_2 + \begin{bmatrix} 5 \\ 6 \\ 2 \\ 6 \end{bmatrix} x_3$$

$\Rightarrow A\underline{x} = \sum \underline{a}_i x_i$ $\quad$ $A\underline{x} \Rightarrow$ linear combinations of columns of A

Linear Independence

- Linearly independent vectors
  - A subspace is what you get by taking **all** linear combinations of n vectors.
  - Suppose have a set of vectors $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_r$
    $\{\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_r\}$ are dependent iff $\exists$ scalars $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_r \ni$

$$\sum_{i=1}^{r} \alpha_i \underline{a}_i = 0 \text{ where at least one } \alpha_i \neq 0$$

  - Independent if $\sum_{i=1}^{r} \alpha_i \underline{a}_i = 0 \Rightarrow \alpha_i = 0 \, \forall i$
    $\Rightarrow$ there does not exist $\alpha_i \neq 0 \ni \sum_{i=1}^{r} \alpha_i \underline{a}_i = 0$

- Rank of an $m \times n$ matrix, A
  - Rank (A) = # of linearly independent columns
    = # of linearly independent rows
    = rank (A^T) = dim [Range (A)] $\leq$ min (m, n)

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & -1 \end{bmatrix} \Rightarrow \text{Independent columns} = 2, \text{ Independent rows} = 2, \text{ Rank} = 2.$$

Linear Spaces Associated with $A\underline{x} = \underline{b}$ - 1

- Linear spaces associated with $A\underline{x}=\underline{b}$
  - Range (A) = $R(A) = \{\underline{y} \in R^m \mid \underline{y} = \sum_{i=1}^{n} \underline{a}_i x_i \text{ for } \forall\ \underline{x} \in R^n\}$

    = column space of (A)
  - dim($R(A)$) = r, rank of (A)
  - The key to answering the question on linear Spaces associated $A\underline{x}=\underline{b}$ is:

    when does $A\underline{x}=\underline{b}$ have a solution?

    that is, when does: $\sum x_i \underline{a}_i = \underline{b}$ *have a solution*?

    $\Rightarrow$ *answer*: *has a solution if $\underline{b}$ can expressed as a linear combination of the columns of A or $\underline{b} \in R(A)$.*
  - In the above example, since in every column $a_{1j} + a_{2j} - a_{3j} = 0\ \forall j$, the right hand side $\underline{b}$ also must have this structure $\Rightarrow$ row 1 + row 2 – row 3 = 0

$$\underline{b} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} \Rightarrow 5\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 3\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \Rightarrow \underline{x} = \begin{bmatrix} 5 \\ 3 \\ 0 \end{bmatrix}$$

  - $b_1+b_2-b_3 = 0 \Rightarrow \underline{y}^T\underline{b} = 0$ where $\underline{y}$=(1 1 -1).
  - So, for a solution to exist, only $\underline{b}$ perpendicular to $\underline{y}$ are allowed. We will see later that $\underline{y}$ is in the so called null space of (A^T)

Linear Spaces Associated with $A\underline{x} = \underline{b}$ 2

$$\underline{b} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \Rightarrow \text{ no solution since not in the column space of } A$$

$$\text{But, } \underline{x} = \begin{bmatrix} 6 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ is also a solution of original problem (why?)}$$

$$A\underline{x} = A\begin{bmatrix} 5 \\ 3 \\ 0 \end{bmatrix} + A\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \underline{b} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} \Rightarrow A\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 0 \Rightarrow \text{infinite \# of solutions } \Rightarrow A\underline{x} = 0 \text{ for } \underline{x} = \alpha\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

- Null space of $A = N(A) = \{\underline{x} \in R^n \mid A\underline{x} = \underline{0}\}$
  .. also called Kernel of A or ker (A)
- Note that $\underline{x} = (0\ 0\ 0)^T$ always satisfies $A\underline{x} = \underline{0}$.
- Key: $\dim(N(A)) = n - r = n - \text{rank}(A)$

❑ To summarize:

- **Existence of a solution for $\underline{b} = \sum \underline{a}_i x_i$ requires that $\underline{b}$ must be in column space or range space, $R(A)$**
- **Uniqueness is related to $N(A) \Rightarrow \dim[N(A)] = 0$**
- **If rank(A) = n then $A\underline{x} = 0 \Rightarrow \underline{x} = 0$;**
  Unique solution to $A\underline{x} = \underline{b}$ if $\underline{b} \in R(A)$ (Note: need $m \geq n$)

Linear Spaces Associated with $A\underline{x} = \underline{b}$ 3

- To complete the characterization of the linear spaces associated with $A\underline{x}=\underline{b}$, we need $R(A^T)$ and $N(A^T)$.
- $R(A^T)=\{\underline{z} \in R^n \mid A^T\underline{y} = \underline{z}\} \Rightarrow$ for solution to exist, $\underline{z}$ should be in the column space of A^T or row space of A
- $N(A^T)=\{\underline{y} \in R^m \mid A^T\underline{y} = \underline{0}\}$ = null space of A^T

	Col. Space	Null space of A
$A\underline{x} = \underline{b}$	$_{(m)}R(A)$	$_{(n)}N(A)$
$A^T\underline{y} = \underline{z}$	$_{(n)}R(A^T)$	$_{(m)}N(A^T)$
	Row space of A	Null space of A^T

❑ KEY: $\dim[R(A^T)]+\dim[N(A)] = r + n - r = n$

$\dim[R(A)]+\dim[N(A^T)] = r + m - r = m$

Rank of A = Rank of $A^T = r$

Linearly ind. col. of (A) = linearly ind. rows of (A)

Example:

$$\overbrace{\begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & -1 \end{bmatrix}}^{A^T}\overbrace{\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}}^{\underline{y}} = \underline{0} \Rightarrow \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \in N(A^T) \quad \Rightarrow \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \text{ are linearly independent}$$

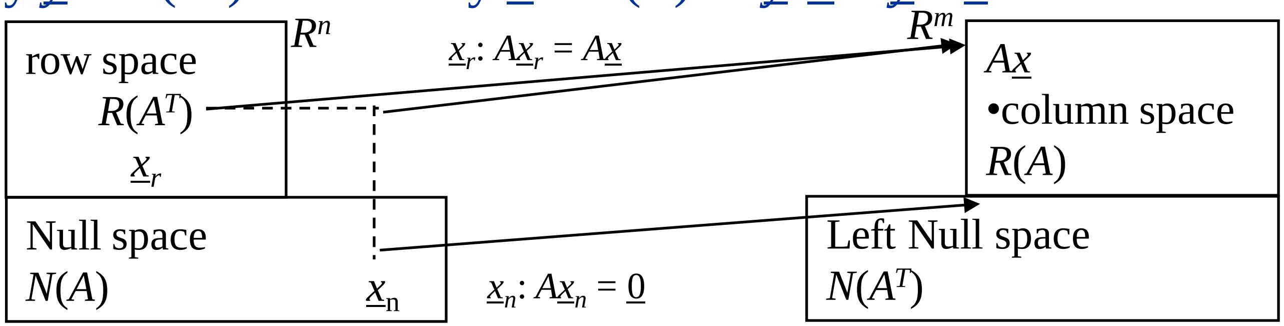
$$\leftarrow R(A) \rightarrow N(A^T)$$

Linear Spaces Associated with $A\underline{x} = \underline{b}$ 4

$$\begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix} \text{ are linearly independent, } \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \in N(A)$$

$$\leftarrow R(A^T) \rightarrow N(A)$$

- Every $\underline{x} \in N(A) \perp^r$ to every $\underline{z} \in R(A^T)$
  $\Rightarrow$ if $A^T\underline{y} = \underline{z}$ and $A\underline{x} = 0 \Rightarrow \underline{x}^T\underline{z} = \underline{x}^T A^T \underline{y} = \underline{0}^T \underline{y} = 0$
- Every $\underline{y} \in N(A^T) \perp^r$ to every $\underline{b} \in R(A) \Rightarrow \underline{y}^T\underline{b} = \underline{y}^T A\underline{x} = 0$



- This characterization of R^n and R^m will be useful in:
  - Least squares (LS) estimation
  - Constructing controllable and uncontrollable subspaces
  - Constructing observable and unobservable subspaces
  - Finding intersection of null spaces
  - Approximating a matrix by another of lower rank (e.g., image compression, data reduction, ….)

Matrix Inverse & Pseudo Inverse

- Matrix inverse and pseudo inverse
  - An $n \times n$ matrix A has rank $(A) = n \Rightarrow A^{-1}$ exists $\Rightarrow |A| \neq 0$
  - $(A^T)^{-1}=(A^{-1})^T$; orthogonal matrix, $Q \Rightarrow Q^{-1} = Q^T$

$$\begin{bmatrix} .8 & -.6 \\ .6 & .8 \end{bmatrix}^{-1} = \begin{bmatrix} .8 & .6 \\ -.6 & .8 \end{bmatrix}$$

  - $(AB)^{-1}=B^{-1}A^{-1}$ if A and B are n by n
  - $|A^{-1}| = 1/|A|$, $|A^T| = |A|$; $|AB| = |A||B|$
  - When rank $(A) < n$ and/or rank $(A) < m$, we define Pseudo inverse or Moore-Penrose Inverse or Generalized Inverse
- Fundamental properties of Pseudo inverse

$AA^{\dagger}A = A$	$(AA^{\dagger})^T = AA^{\dagger}$	(symmetric)
$A^{\dagger}AA^{\dagger} = A^{\dagger}$	$(A^{\dagger}A)^T = (A^{\dagger}A)$	(symmetric)

$$\begin{aligned}&(I_n - A^{\dagger}A)(I_n - A^{\dagger}A)\\ &= (I_n - 2A^{\dagger}A + A^{\dagger}AA^{\dagger}A)\\ &= (I_n - 2A^{\dagger}A + A^{\dagger}A)\\ &= (I_n - A^{\dagger}A)\end{aligned}$$

$\Rightarrow$ Note that both $(I_n - A^{\dagger}A)$ and $(I_m - AA^{\dagger})$ are idempotent

- Concept of pseudo inverse is very useful in least squares, Kalman filtering, and spectral estimation
  - Idea: we find $\underline{x} \ni \|\underline{x}\|_2$ is a minimum.
  - That is, out of the infinite # of possible solutions, we pick one with a minimum norm or smallest "size" … Lecture 6.

Eigen Values & Eigen Vectors - 1

- Eigen values – Eigen vectors
  - Basic property

$$\left\{\begin{matrix}\lambda_i \\ \underline{\xi}_i\end{matrix}\right\} \begin{matrix}\text{eigen values of } A,\ \lambda_i(A) \\ \text{eigen vectors of } A,\ \underline{\xi}_i(A)\end{matrix} \qquad \begin{matrix}\lambda_{\max}(A) = \text{biggest } \lambda_i(A) \\ \lambda_{\min}(A) = \text{smallest } \lambda_i(A)\end{matrix}$$

$$\rho(A) = |\lambda_{\max}(A)| = \text{spectral radius of } A \sim \text{used as measure of size of } A$$

  - Key equation: $A\,\underline{\xi}_i = \lambda_i \underline{\xi}_i$

λ_i solution of $|\lambda I - A| = 0$, characteristic equation of A

$$\lambda^n + a_n \lambda^{n-1} + \ldots + a_2\lambda + a_1 = 0$$

$$A^n + a_n A^{n-1} + \ldots + a_2 A + a_1 I = 0$$ Caley-Hamilton Theorem

  - If $A = A^T \Rightarrow$ symmetric:
    - $\lambda_i(A)$ real; $Q = (\underline{\xi}_1\ \underline{\xi}_2 \ldots \underline{\xi}_n)$ are orthogonal;
    - Q can be made orthonormal $\Rightarrow A = Q\Lambda Q^T$
    - Q is orthonormal $\Rightarrow |\lambda_i(Q)| = 1$

Example: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow \lambda_i = \pm 1; \begin{pmatrix} 0.8 & 0.6 \\ -0.6 & 0.8 \end{pmatrix} \Rightarrow \lambda_i = 0.8 \pm 0.6i \Rightarrow |\lambda_i| = 1$

    - A is PD $\Rightarrow \{\lambda_i(A)\}$ are positive and real

Example: $\begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix} \Rightarrow \lambda_i = 3 \pm \sqrt{5} > 0$ PD

Eigen Values & Eigen Vectors - 2

- A is skew symmetric $\Rightarrow \{\lambda_i(A)\}$ are imaginary

  Example: $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Rightarrow \lambda_i = \pm j$
- In general, $\lambda_i(A)>0 \not\Rightarrow A$ is PD. Example: $\begin{pmatrix} 1 & 100 \\ 0 & 1 \end{pmatrix}$
- But, for symmetric A & $\lambda_i(A)>0 \Rightarrow A$ is PD

  $\Rightarrow$ So, for PD: $\underline{x}^T A \underline{x} > 0 \quad \forall\, \underline{x} \neq \underline{0}$ (A need not be symmetric) or principal minors or eigen values of symmetrized A>0
- Note: A, B are PD $\Rightarrow A + B$, A^2, A^{-1}, and all A^n are PD
- Eigen vectors associated with distinct eigen values are independent.

  Proof: assume dependent $\Rightarrow \alpha_1\xi_1 + \alpha_2\xi_2 + \ldots + \alpha_n\xi_n = 0$

  multiply by A, A^2, …, A^{n-1}

$$\Rightarrow [\alpha_1\underline{\xi}_1 \quad \alpha_2\underline{\xi}_2 \quad \ldots \quad \alpha_n\underline{\xi}_n] \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^{n-1} \\ & \vdots & \vdots & \ldots & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \ldots & \lambda_n^{n-1} \end{bmatrix} = 0$$

Vandermonde matrix

$$\Rightarrow \alpha_i\underline{\xi}_i = 0 \Rightarrow \alpha_i = 0 \quad \text{invertible if } \lambda_i \neq \lambda_j$$

- $tr(A) = \sum_{i=1}^{n} \lambda_i(A), \quad \det(A) = \prod_{i=1}^{n} \lambda_i(A)$

Vector and Matrix Norms - 1

- Similarity Transformations
  - If Q^{-1} exists, $\bar{A} = Q^{-1}AQ \Rightarrow \lambda_i(A) = \lambda_i(\bar{A})$
  - If $\{\lambda_i\}$ are distinct, $Q^{-1}AQ = \Lambda = \text{Diag}(\lambda_i)$
  - In particular, $f(A) = Q\, f(\Lambda)\, Q^{-1}$ e.g., $e^{At} = Q\, e^{\Lambda t}\, Q^{-1}$

    … worst possible way of computing $f(A)$
  - If $\{\lambda_i\}$ are not distinct, need to use Jordan's form….. Messy on computers
  - OK for symmetric matrices $f(A) = Q\, f(\Lambda) Q^T$
- Vector and Matrix Norms:
  - Play an important role in the convergence studies of algorithms.
  - As an example, consider $A\underline{x} = \underline{b}$ problem
  - Simplest and most important problem in Matrix Computations
  - To show its importance, consider a linear system in steady state:

$$\dot{\underline{x}} = \underline{0} = A\underline{x} + B\underline{u} \Rightarrow \underline{x}_{ss} = -A^{-1}B\underline{u} = -A^{-1}\underline{b}$$

$$\Rightarrow \text{ solve } A\underline{x} = -B\underline{u} = \underline{b}$$

  - Mathematically, solution exists *iff* $\underline{b} \in R(A) = \{\underline{x} \in R^m \mid \sum \underline{a}_i\, x_i = \underline{b}\}$

Vector & Matrix Norms - 2

- Unique if $N(A) = \varphi \Rightarrow \sum \underline{a}_i x_i = 0 \Rightarrow x_1 = x_2 = \ldots = x_n = 0$
  $\Rightarrow$ Linearly independent columns of A

Q 1. If A and $\underline{b}$ are perturbed by a small amount δA and $\delta \underline{b}$, how does it affect $\underline{x}$? … the so-called sensitivity (conditioning) problem.

2. What if A is "nearly" singular? what is near singularity?
3. If $\underline{b} \notin R(A)$, then how can we determine $\underline{x} \ni A\underline{x}$ is "close" to $\underline{b}$? $\Rightarrow$ least squares problem
4. How do we measure small Perturbations?
   near singularity?
   distance in vector spaces?

} Norms provide such a language

- Norms generalize the concept of absolute value of a real number to vectors and matrices (measure of "SIZE" of a vector and matrix)
- Vector Norms
  - $\|\underline{x}\|_p = \text{Holder or } p\text{-norm} = [|x_1|^p + |x_2|^p + \ldots + |x_n|^p]^{1/p} = \left[\sum_{i=1}^{n} |x_i|^p\right]^{1/p} \sim \text{"size"}$

most important

$$p = 1 \Rightarrow \|\underline{x}\|_1 = \sum_{i=1}^{n} |x_i| \Rightarrow \text{1-norm or Manhattan Distance}$$

$$p = 2 \Rightarrow \|\underline{x}\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} \text{ (2-norm, root sum square (RSS) or Euclidean norm)}$$

$$p = \infty \Rightarrow \|\underline{x}\|_\infty = \max_i |x_i| \text{ (}\infty\text{-norm)}$$

Properties of Vector Norms

- All norms convey approximately same information.
- Only thing is, some are more convenient to use (e.g., 2-norm).
- However, all satisfy:

(i) $\|\underline{x}+\underline{y}\|_p \leq \|\underline{x}\|_p + \|\underline{y}\|_p$ (Minkowski's inequality)

(ii) $\|\underline{x}+\underline{y}\|_p \geq 0$

(iii) $\|c\underline{x}\|_p = c\|\underline{x}\|_p$

(iv) $\|\underline{x}^T\underline{y}\| = |\underline{x}^T\underline{y}| \leq \|\underline{x}\|_p \|\underline{y}\|_q \; 1/p + 1/q = 1$ (Holder's inequality)

(v) $|\underline{x}^T\underline{y}| \leq \|\underline{x}\|_2 \|\underline{y}\|_2$ (Cauchy-Schwartz-Bunyakovski's inequality)

(vi) $\|Q\underline{x}\|_2^2 = \underline{x}^T Q^T Q \underline{x} = \underline{x}^T\underline{x} = \|\underline{x}\|_2^2$

$\Rightarrow$ **2-norm is invariant under orthogonal transformations … extremely important idea in numerical computations**.

(vii) $\|\underline{x}\|_2 \leq \|\underline{x}\|_1 \leq \sqrt{n}\,\|\underline{x}\|_2$

$\|\underline{x}\|_\infty \leq \|\underline{x}\|_2 \leq \sqrt{n}\,\|\underline{x}\|_\infty$

$\|\underline{x}\|_\infty \leq \|\underline{x}\|_1 \leq n\|\underline{x}\|_\infty$

$\begin{pmatrix}1\\0\end{pmatrix}$ & $\begin{pmatrix}1\\1\end{pmatrix}$ satisfy

LB UB with equality

(viii) $\hat{\underline{x}}$ approx. to $\underline{x} \Rightarrow$ absolute error $\|\underline{x}-\hat{\underline{x}}\|$; relative error $\|\underline{x}-\hat{\underline{x}}\| / \|\underline{x}\|$

∞-norm $\Rightarrow$ # of correct signifiant digits in $\underline{x}$

$10^{-p} \Rightarrow p$ signifiant digits

Properties of Matrix Norms

- Matrix norms: A is m by n

(i) Frobenius norm (or) F-norm = $\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2}$

(ii) p-norm $\|A\|_p = \max_{\underline{x}} \frac{\|A\underline{x}\|_p}{\|\underline{x}\|_p}$ ~ "size" of the matrix

(iii) Note: $\|A\underline{x}\|_p \le \|A\|_p \|\underline{x}\|_p$; $\|AB\|_p \le \|A\|_p \|B\|_p$; $\|A+B\|_p \le \|A\|_p + \|B\|_p$

(iv) $p = 1 \Rightarrow \|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|$ = max. column sum

(v) $p = 2 \Rightarrow \|A\|_2 = \max_{\underline{x}} \frac{\left(\underline{x}^T A^T A \underline{x}\right)^{1/2}}{\left(\underline{x}^T \underline{x}\right)^{1/2}} = \left[\lambda_{\max}(A^T A)\right]^{1/2} = \sigma_{\max}(A)$,

Max. singular value of A

(vi) $p = \infty \Rightarrow \|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|$ = max. row sum

- Use of Matrix Norms:
  - Use $\|A\|$ to approximate $|\lambda_{\max}(A)| = \rho(A)$ spectral radius of A.
  - Can show $\rho(A) \le \|A\|_p \ \forall\, p$.
  - Other properties, see problem set #1

Singular Value Decomposition - 1

- Scale by c = constant > $\rho(A) \Rightarrow A/c$ has all $|\lambda_i(A)| < 1$. Will be useful in evaluating e^{At} and integrals involving e^{At}.
- Matrix norm used in estmaing the convergence rate of algorithms. We see its use in Lecture 4.

❑ Singular Value Decomposition (SVD) :

- Best method for approximating A by $\hat{A}$ & Determining rank(A) is SVD.

❑ What is SVD?

- $A \in R^{mn} = U\Sigma V^T$; U & V orthogonal
- $U = (\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_m) \in R^{mm}$ $\quad V = (\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n) \in R^{nn}$
- $U^TAV = \Sigma = Diag(\sigma_1\ \sigma_2\ \ldots\ \sigma_p)$; $p = \min(m,n)$ $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0$
- $p = m < n \Rightarrow A = U\begin{pmatrix}\Sigma_1 & 0\end{pmatrix}\begin{pmatrix}v_1^T \\ v_2^T\end{pmatrix}$; $\quad p = n < m \Rightarrow \begin{pmatrix}u_1 & u_2\end{pmatrix}\begin{pmatrix}\Sigma_1 \\ 0\end{pmatrix}V^T$
- In general, can have $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ and $\sigma_{r+1} = \ldots = \sigma_p = 0$

$$\Rightarrow \quad A = \begin{pmatrix}U_r & U_2\end{pmatrix}\begin{pmatrix}\Sigma_r & 0 \\ 0 & 0\end{pmatrix}\begin{pmatrix}v_r^T \\ v_2^T\end{pmatrix} = U_r\Sigma_r V_r^T$$

- r = rank(A) $\Rightarrow$ can use SVD to determine rank of A
- $A\,\underline{v}_i = \sigma_i\,\underline{u}_i$, $\underline{u}_i$ is called the left singular vector of A

Singular Value Decomposition - 2

- $A^T \underline{u}_i = \sigma_i \underline{v}_i$, $\underline{v}_i$ is called the right singular vector of A
  $AA^T = U\Sigma^2 U^T \Rightarrow AA^T \underline{u}_i = \sigma^2{}_i \underline{u}_i$
  $\Rightarrow \sigma^2{}_i$ are eigen values of AA^T; $\underline{u}_i$ eigen vectors of AA^T
- Similarly, $A^T A = V\Sigma^2 V^T \Rightarrow A^T A\, \underline{v}_i = \sigma^2{}_i \underline{v}_i$
  $\Rightarrow \sigma^2{}_i$ are eigen values of $A^T A$; $\underline{v}_i$ eigen vectors of $A^T A$
- Symmetric PD $\Rightarrow$ SVD = eigen value analysis (called principal component analysis in statistics)
- Since $\|A\|_2 = [\lambda_{max}(A^T A)]^{1/2} = (\sigma^2{}_{max})^{1/2} = \sigma_{max}$
- Also, $\|A\|_F = (\sigma^2{}_1 + \sigma^2{}_2 + \ldots + \sigma^2{}_p)^{1/2}$
  $\|A^{-1}\|_2 = [\lambda_{max}(A^T A)^{-1}]^{1/2} = 1/\sigma_{min}$

One application of SVD: data compression (reduction)

- Suppose we approximate A by the first r singular vectors. What is the error involved?
- Know

$$A = U\Sigma V^T \qquad \hat{A} = U_r \Sigma_r V_r^T = \sum_{i=1}^{r} \sigma_i \underline{u}_i v_i^T \qquad \left\|A - \hat{A}\right\|_2 = \sigma_{r+1}$$

$$\left\|A - \hat{A}\right\|_F = \left(\sum_{i=r+1}^{n} \sigma_i^2\right)^{1/2}$$

Singular Value Decomposition - 3

- Very useful in spectral estimation and robust control.

❑ Second application: determining the condition number of a matrix A

- Condition number, $k(A) = \|A\|_p \|A^{-1}\|_p$ for $p=2$: $k(A) = \sigma_{max}(A)/\sigma_{min}(A)$
  $\Rightarrow k(A)$ is a measure of non-singularity of A

Example 11: Consider $A = \begin{bmatrix} .66 & 3.34 \\ 1.99 & 10.01 \end{bmatrix}$

$$\sigma_{\max} = 10.7588 \qquad \sigma_{\min} = 0.0037$$

$$k(A) = \sigma_{\max} / \sigma_{\min} = 2894$$

- Large $k(A) > 10^6 \Rightarrow$ bad news in solving $A\underline{x}=\underline{b}$

❑ Third application of SVD: determining the bases of linear spaces associated with $A\underline{x}=\underline{b}$ and $A^T\underline{y}=\underline{c}$ as well as the pseudo inverse

$$\Rightarrow R(A) = (u_1 \ldots u_r); \qquad N(A^T) = (\underline{u}_{r+1} \ldots \underline{u}_m)$$

$$\dim(R(A)) + \dim(N(A^T)) = m$$

$$N(A) = (\underline{v}_{r+1}\underline{v}_{r+2} \ldots \underline{v}_n); \qquad R(A^T) = (\underline{v}_1 \ldots \underline{v}_r)$$

$$\dim(N(A)) + \dim(R(A^T)) = n$$

$$\Rightarrow \text{Pseudo Inverse of } A = A^\dagger = V_r \Sigma_r^\dagger U_r^T = \sum_{i=1}^{r} \frac{1}{\sigma_i} \underline{v}_i \underline{u}_i^T$$