Numerical Linear Algebra (NLA)
MTH 365 / 465

Introduction: NLA is a study of algorithms for performing linear algebra computations.

NLA focuses on 2 questions (roughly):

- Linear systems:
  Q: Given a matrix A & a vector b, what is a vector x that satisfies $Ax = b$?
- Eigen-decompositions:
  Q: Given a matrix A, what is a vector $x \neq 0$ and a scalar λ that satisfy $Ax = \lambda x$?

(Syllabus: list of objectives.)

- NLA used to be a subtopic of Numerical Analysis.
- Applications of NLA: signal processing
  control theory
  heat transfer
  fluid dynamics
  statistics, data mining,
  pattern recognition, etc.
- MATLAB is good for NLA. Other software: Mathematica, LINPACK, LAPACK, etc.

Examples: ("Why shall we study NLA?")

from DEQ's

1. $Ax=b$, A is an $n\times n$ nonsingular matrix, $x=(x_1,\dots,x_n)^T$. If we use Cramer's rule to find x:

$$x_1=\frac{\det(A_1)}{\det(A)},\ x_2=\frac{\det(A_2)}{\det(A)},\dots,\ x_n=\frac{\det(A_n)}{\det(A)}$$

(A_i is a matrix w/ the ith column replaced by b)

Computations: $(n+1)$ determinants with total

$$[n!\,(n-1)](n+1)=(n-1)(n+1)!\ \text{multiplications.}$$

For $n=25$, w/ 10 billion operations/sec, we'll need $\dfrac{24\times 26!}{10^{10}\times 3600\times 24\times 365}\approx 30.6$ billion years!

We'll study Gaussian elimination: $O(n^3)$ ⇒ less than 1 second!

2. Explicitly finding A^{-1} → inaccurate! ($x=A^{-1}b$): don't!
3. Finding eigenvalues computing the zeros of the characteristic polynomials (can be ill-conditioned)

Review of Linear Algebra.

Vectors:

$$\vec{v}=\begin{bmatrix}v_1\\v_2\\\vdots\\v_n\end{bmatrix}=(v_1,v_2,\dots,v_n)^T\in\mathbb{R}^n$$

(or simply, v) (n-vector)

$$\vec{u}+\vec{v}=(u_1+v_1,\dots,u_n+v_n)^T$$

$$c\vec{v}=(cv_1,\dots,cv_n)^T$$

$u \perp v$ if $u \cdot v = 0$
u, v orthogonal

- Inner (dot) product: $\vec{u}^T\vec{v} = \vec{u}\cdot\vec{v} = \langle \vec{u}, \vec{v} \rangle = \sum_{i=1}^{n} u_i v_i$
- Length of v is $\|\vec{v}\| = \|\vec{v}\|_2 = \sqrt{\vec{v}^T\vec{v}} = \sqrt{v_1^2 + \dots + v_n^2}$
  (also, 2-norm or Euclidean norm)
- Lin. idependent set of vectors $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k\}$ in $\mathbb{R}^n$: if $c_1\vec{x}_1 + \dots + c_k\vec{x}_k = 0 \Rightarrow c_1 = c_2 = \dots = c_k = 0$.
  If not all of c_i are zero $\Rightarrow$ set is lin. dependent.

- Span of vectors $\vec{x}_1, \dots, \vec{x}_k \in \mathbb{R}^n$: $\mathrm{Span}(\vec{x}_1, \dots, \vec{x}_k) \overset{\text{or}}{=} \langle \vec{x}_1, \dots, \vec{x}_k \rangle$
  is the set of all linear combinations of $\vec{x}_i$'s.
  $\langle \vec{x}_1, \dots, \vec{x}_k \rangle$ is a subspace of $\mathbb{R}^n$.

V is a <u>vector space</u> if:
$\forall \bar{u}, \bar{v}, \bar{w} \in V$ & $\alpha, \beta \in \mathbb{R}$

- $\bar{u} + \bar{v}$, $\alpha\bar{u} \in V$
- $\bar{u} + \bar{v} = \bar{v} + \bar{u}$, $\bar{u} + (\bar{v} + \bar{w}) = (\bar{u} + \bar{v}) + \bar{w}$
- $\exists \bar{0} \in V$ s.t. $\bar{u} + \bar{0} = \bar{u}$
- $\exists -\bar{u} \in V$ s.t. $\bar{u} + (-\bar{u}) = \bar{0}$
- $\alpha(\bar{u} + \bar{v}) = \alpha\bar{u} + \alpha\bar{v}$
- $(\alpha + \beta)\bar{u} = \alpha\bar{u} + \beta\bar{u}$, $(\alpha\beta)\bar{u} = \alpha(\beta\bar{u})$
- $1 \cdot \bar{u} = \bar{u}$, $0 \cdot \bar{u} = 0$, $(-1)\bar{u} = -\bar{u}$.

<u>Subspace</u>:
$\downarrow$
any subset of V closed under addition & multiplication by a scalar

<u>Basis for V</u>: set $\{\bar{x}_1, \dots, \bar{x}_k\}$, $\bar{x}_i \in V$, s.t.
$V = \mathrm{span}(\bar{x}_1, \dots, \bar{x}_k)$ & $\bar{x}_1, \dots, \bar{x}_k$ are lin. independent
$\dim V = k$
There exist many bases for V, but they all have the same number of elements, k.
$\forall \bar{v} \in V$, $\bar{v} = \sum_{i=1}^{k} c_i \bar{x}_i$ where c_i are unique, they are called coordinates of v w.r.t. the basis.

The Gram-Schmidt Algorithm.

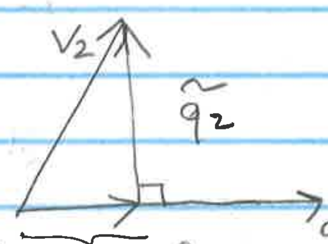
first example of algorithm!

Given linearly independent set of vectors $\vec{v}_1, \dots, \vec{v}_n$ (I will denote them $v_1, \dots, v_k$ for simplicity), one can construct a set of orthonormal vectors $q_1, \dots, q_k$.

$\langle q_i, q_j \rangle = q_i^T q_j = 0$ & $\|q_i\| = 1$
$\forall i,j$ $\qquad \forall i$

Steps: 1) $q_1 = \dfrac{v_1}{\|v_1\|}$ (normalized $v_1 \Rightarrow \|q_1\| = 1$)

2) find the orthogonal projection of v_2 onto q_1, i.e. the closest vector to v_2 from span(q_1)



$\tilde{q}_2 \perp q_1$

orthogonal projection of v_2 onto q_1

$\text{proj}_{q_1}(v_2) = \underbrace{\langle v_2, q_1 \rangle}_{\text{scalar } \alpha} q_1$

why? we need $v_2 - \alpha q_1 \perp q_1$
$\langle v_2 - \alpha q_1, q_1 \rangle = 0$
$\langle v_2, q_1 \rangle = \alpha \underbrace{\langle q_1, q_1 \rangle}_{\overset{\shortparallel}{1}}$

$\tilde{q}_2 = v_2 - \langle v_2, q_1 \rangle q_1 \Rightarrow \langle \tilde{q}_2, q_1 \rangle =$

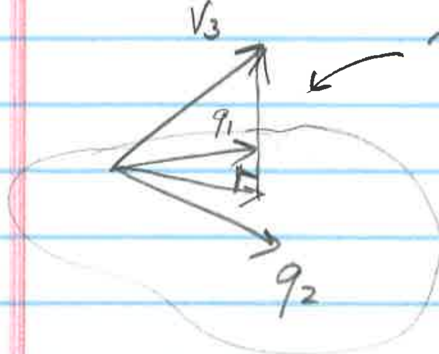
$\langle v_2 - \langle v_2, q_1 \rangle q_1, q_1 \rangle = \langle v_2, q_1 \rangle - \langle v_2, q_1 \rangle \underbrace{\langle q_1, q_1 \rangle}_{\overset{\shortparallel}{1}} = 0$

$\Rightarrow$ indeed, $\tilde{q}_2 \perp q_1$

Now, $q_2 = \dfrac{\tilde{q}_2}{\|\tilde{q}_2\|}$ (normalized $\tilde{q}_2 \Rightarrow \|q_2\| = 1$)

So, we have 2 orthonormal q_1, q_2.

3) Next, take v_3 and find its orthogonal projection onto span (q_1, q_2), i.e. $v_3 \perp q_1$ & $v_3 \perp q_2$



$$\tilde{q}_3 = v_3 - \langle v_3, q_1 \rangle q_1 - \langle v_3, q_2 \rangle q_2$$

$\tilde{q}_3 \perp q_1, q_2$ (check it)

Normalize $\tilde{q}_3$ to get $q_3 = \dfrac{\tilde{q}_3}{\|\tilde{q}_3\|}$.

Repeat the process:
at step k, we have $\tilde{q}_k = v_k - \sum_{i=1}^{k-1} \langle v_k, q_i \rangle q_i$
and $q_k = \dfrac{\tilde{q}_k}{\|\tilde{q}_k\|}$.

G.-S. algorithm is used for QR factorization of a matrix A $(m \times n)$ which is used to solve a linear system $Ax = b$ (§7.6.2)

Matrices & Linear Equations.

- $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n}$
  $= (a_{ij})$

$A + B = (a_{ij} + b_{ij})$

$cA = (ca_{ij})$

$\underset{m\times n}{A}\,\underset{n \times p}{B} = \left(\sum_{k=1}^{n} a_{ik} b_{kj} \right)$

$i = 1, \dots, m$

$j = 1, \dots, p$

($AB \neq BA$! in general)

- Outer product of vectors u & v: $uv^T = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} (v_1 \dots v_m)$

$$= \begin{pmatrix} u_1 v_1 & \dots & u_1 v_m \\ u_2 v_1 & \ddots & u_2 v_m \\ \vdots & & \vdots \\ u_n v_1 & \dots & u_n v_m \end{pmatrix}$$

Example: $u = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $v = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$

Inner product: $u^T v = u \cdot v = (1\,2\,3)\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = 20 \in \mathbb{R}$

Outer product: $u v^T = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}(2\,3\,4) = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 6 & 8 \\ 6 & 9 & 12 \end{pmatrix}$

Transpose $A^T = (a_{ji})$

$(AB)^T = B^T A^T$

$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{pmatrix}$, the identity matrix $n \times n$

A square matrix $A_{n\times n}$ is invertible if $\exists\, B_{n\times n}$ s.t. $BA = AB = I$. The inverse B of A, denoted by A^{-1}, is unique.

Determinant of A:

$$\det A_{2\times 2} = a_{11}a_{22} - a_{12}a_{21}$$

$$\det A_{n\times n} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det \underbrace{(A_{ij})}_{\text{minor}} \overset{\text{or}}{=} \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

column expansion

row expansion

A^{-1} exists $\iff \det A \neq 0$ (A is nonsingular)

(See handout on square matrices.)

Linear Equations

Consider the system

$$\begin{aligned} x+y-2z&=1 \\ 3x-y+z&=0 \\ -x+3y-2z&=4 \end{aligned}$$

$$\underbrace{\begin{pmatrix} 1 & 1 & -2 \\ 3 & -1 & 1 \\ -1 & 3 & -2 \end{pmatrix}}_{\text{matrix } A} \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{\substack{\text{vector} \\ \text{of} \\ \text{unknowns}}} = \underbrace{\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}}_{\text{R.H.S. vector } b}$$

In general: $\underbrace{A}_{m\times n}\underbrace{x}_{n\times 1} = \underbrace{b}_{m\times 1}$

Questions: 1) Under what conditions does it have a solution?

2) When is the solution unique?

Consider the homogeneous system $Ax=0$. It has the trivial solution $x=0$.
If $\exists\, y\neq 0$ s.t. $Ay=0 \Rightarrow \forall \alpha\in\mathbb{R}$, $A(\alpha y) = \alpha(Ay) = 0 \Rightarrow \exists$ infinitely many solutions.
If $b\neq 0$, $Ax=b$ is an inhomogeneous system.

If $Ax = x_1 a_1 + x_2 a_2 + \ldots + x_n a_n = b \Rightarrow b$ is a lin. comb. of the columns

columns of A

$\Rightarrow b\in \mathrm{col}(A)$.

If $Ax=b$ has a solution y & $Ax=0$ has a solution $y_h \neq 0$ $\Rightarrow$

$A(y+\alpha y_h) = Ay + \alpha A y_h = Ay = b$ $\Rightarrow$
$\exists$ infinitely many solutions ($\forall \alpha$)

If $Ax=0$ has the trivial solution only $\Rightarrow$ $Ax=b$ cannot have more than one solution. E.g., if y & z are 2 solutions of $Ax=b$ then
$Ay=b$ & $Az=b \Rightarrow A(y-z) = Ay - Az$
$= b-b=0 \Rightarrow y-z=0 \Rightarrow y=z$.

Thus, $Ax=0$: $x=0$ or infinitely many
If $Ax=b$, $b\neq 0$, then:

1) no solution if $b \notin col(A)$
2) if $b \in col(A)$

$\swarrow$ $Ax=0$ has $x=0$ only

$\Downarrow$

$Ax=b$ has one solution

$\searrow$ $Ax=0$ has sol. $y_h \neq 0$

$\Downarrow$

sol. of $Ax=b$ is in the form

$$\underbrace{y_h}_{\text{sol. to } Ax=0} + \underbrace{y}_{\text{part. sol. of } Ax=b.}$$