A.8 Eigenvalues & Eigenvectors.

For an $n\times n$ matrix A, if $A\vec{v} = \lambda\vec{v}$ for some $\vec{v} \neq 0$ & λ, then λ is said to be an eigenvalue & $\vec{v}$ an eigenvector of A.

$A\vec{v} = \lambda\vec{v} \Rightarrow (A-\lambda I)\vec{v} = \vec{0}$ has a solution $\vec{v} \neq 0$
iff $\underbrace{\det(A-\lambda I)}_{\text{characteristic polynomial } p(\lambda)} = 0$

So, eig(A) are zeros of $p(\lambda)$.

Example: $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \Rightarrow p(\lambda) = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (\lambda-3)(\lambda-1)$

$\Rightarrow \lambda_1 = 3, \ \lambda_2 = 1$

$\vec{v}_1$: $(A-3I)\vec{v} = \vec{0}$

$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$\vec{v}_2$: $(A-I)\vec{v} = \vec{0}$

$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \vec{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

In general, $\lambda \in \mathbb{C}$.

If A has n linearly independent eigenvectors $\vec{v}_1, \dots, \vec{v}_n$ with $A\vec{v}_i = \lambda_i \vec{v}_i$, $i = 1:n$, then A is diagonalizable.
Let $V = [\vec{v}_1 | \dots | \vec{v}_n] \Rightarrow AV = V\Lambda$ with $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, or $V^{-1}AV = \Lambda$

(Note: not all matrices are diagonalizable.)

§7.2.4 Matrices for Which Pivoting is Not Required.

- A is symmetric if $A = A^T$ $(A = A^*)$
- A symmetric matrix is positive definite if $eig(A) > 0$ (or if $\forall \vec{x} \in \mathbb{R}^n, \vec{x} \neq \vec{0}$, $\vec{x}^T A \vec{x} > 0$). We'll write "SPD" for symm., p.def. matrices

Examples of SPD matrices:

1) $A = \begin{bmatrix} 9 & 6 \\ 6 & 5 \end{bmatrix}$

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

Note: $a_{ii} > 0$ for such matrices

Recall: $A = LU$. If A is SPD, we can find a unique L s.t. $A = LL^T$ (L-low-triangular)

Why? Consider a SPD matrix A. Then we can find L & U to get $A = LU$ (from GE)

[Note: pivoting is not required $(a_{ii} > 0)$]

Then $A = L \underbrace{D D^{-1}}_{= I} \overbrace{U}^{M^T}$ with $D = diag\underbrace{(u_{11}, \dots, u_{nn})}_{\text{from } U}$

$A = LDM^T = A^T = (LDM^T) = MDL^T$

$\Rightarrow L = M \Rightarrow A = LDL^T$ ($\underbrace{LDL^T\text{-decomposition}}_{\text{can be found } \forall A}$)

If A is SPD $\Rightarrow d_{ii} = u_{ii} > 0$

$\Rightarrow D^{1/2} = diag(\sqrt{u_{11}}, \dots, \sqrt{u_{nn}})$

$\Rightarrow A = LDL^T = \underbrace{LD^{1/2}}_{\hat{L}} \underbrace{D^{1/2}L^T}_{\hat{L}^T} = \hat{L}\hat{L}^T$ where

$\hat{L} = L \cdot diag(\sqrt{u_{11}}, \dots, \sqrt{u_{nn}})$

See handout w/ an example.

$A = \hat{L}\hat{L}^T$ is called Cholesky factorization (only for SPD matrices).

→ Andre-Louis Cholesky, 1875–1918, French military officer and mathematician

$A =$ 

$\hat{L}$ $\hat{L}^T$

"Cholesky factor"

FLOPS $O\left(\frac{n^3}{3}\right)$

Next: $Ax = b$ ⇒ solve for y: $\hat{L}y = b$ ⇒ y

⇒ solve for x: $\hat{L}^T x = y$ ⇒ x

- Other matrices for which pivoting is not required:

– Strictly (row) diagonally dominant matrices:

ith row: $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$, $i = 1, \dots, n$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \\ a_{n1} & a_{n2} & & a_{nn} \end{pmatrix}$$

→ Ex: $A = \begin{pmatrix} 3 & -2 & 0 \\ 1 & -3 & 1 \\ -1 & 2 & 4 \end{pmatrix}$

(SPD) Banded matrices: type of sparse matrices (or band matrices)

sparse → large fraction of entries are zero

$a_{ij} = 0$ if $|i-j| > m$, where $m \ll n$ ($n = \text{size}(A)$) is the half bandwidth

- Full bandwidth: $2m+1 = 3$ ⇒ tridiagonal matrix

Equivalent (but more general) definition

Examples: Band matrices $a_{ij} = 0$ if $j < i - m_1$ or $j > i + m_2$

(m_1 ↓ lower, m_2 ↓ upper bandwidth)

$m_1 + m_2 + 1 = \text{bandwidth}$

diagonal: $m_1 = m_2 = 0$

$$\begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$$

tridiagonal: $m_1 = m_2 = 1$

$$\begin{pmatrix} a_{11} & a_{12} & 0 & 0 & \cdots & 0 & 0 & 0 \\ a_{12} & a_{22} & a_{23} & 0 & \cdots & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} & & 0 & 0 & 0 \\ & \cdots & & & & \cdots & & \\ 0 & \cdots & & 0 & a_{n-1\,n-2} & a_{n-1\,n-1} & a_{n-1\,n} & \\ 0 & \cdots & & 0 & 0 & a_{n\,n-1} & a_{nn} & \end{pmatrix}$$

$$\begin{pmatrix} \times & \times & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 \\ 0 & \times & \times & \times & 0 \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{pmatrix}$$

Banded structure ⇒ no need to store everything in GE.

Example: $m_1 = m_2 = 2$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & 0 \\ 0 & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ 0 & 0 & a_{53} & a_{54} & a_{55} & a_{56} \\ 0 & 0 & 0 & a_{64} & a_{65} & a_{66} \end{pmatrix} \xrightarrow{\text{reduced cost}} \begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 & 0 \\ 0 & \tilde{a}_{22} & \tilde{a}_{22} & a_{24} & 0 & 0 \\ 0 & \tilde{a}_{32} & \tilde{a}_{33} & a_{34} & a_{35} & 0 \\ 0 & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ 0 & 0 & a_{53} & a_{54} & a_{55} & a_{56} \\ 0 & 0 & 0 & a_{64} & a_{65} & a_{66} \end{pmatrix}$$

already zeros → (the circled zeros a_{41}, a_{51}, a_{61})

Total work in GE for $m = m_1 = m_2$ band matrix is

$$\sum_{j=1}^{n-1} 2 \min\{m, n-j\} \cdot \min\{m+1, n-j+1\} \approx 2m^2 n$$

for $m \ll n \Rightarrow$ large saving over $O(n^3)$

§ 7.2.5 $\rightarrow$ tricks on how to make GE less expensive for large n.

- block algorithms
- parallelism

§ 7.3 Other methods for Solving $Ax=b$.

(1) $Ax = b \Rightarrow x = A^{-1}b$. To find A^{-1}:

$$\underbrace{\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 0 & 0 & 0 & 1 \end{array}\right)}_{A \quad I} \xrightarrow[r_3 - r_1 \times 7]{r_2 - r_1 \times 4} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & -6 & -21 & -7 & 0 & 1 \end{array}\right) \xrightarrow{r_3 - 2r_1} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & 0 & -9 & 1 & -2 & 1 \end{array}\right)$$

(eliminated entries below the diag.)

$$\xrightarrow[r_3 / (-9)]{r_2/(-3)} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 4/3 & -1/3 & 0 \\ 0 & 0 & 1 & -1/9 & 2/9 & -1/9 \end{array}\right) \xrightarrow[r_2 - 2 \cdot r_3]{r_1 - 3 \cdot r_3} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 4/3 & -2/3 & 1/3 \\ 0 & 1 & 0 & 14/9 & -7/9 & 2/9 \\ 0 & 0 & 1 & -1/9 & 2/9 & -1/9 \end{array}\right)$$

scaling diag. entries to 1 — entries above the diagonal

$$\xrightarrow{r_1 - 2 \cdot r_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -16/9 & 8/9 & -1/9 \\ 0 & 1 & 0 & 14/9 & -7/9 & 2/9 \\ 0 & 0 & 1 & -1/9 & 2/9 & -1/9 \end{array}\right)$$

$I \qquad A^{-1}$

We solved 3 linear systems w/ GE! (upper triangular)

In general: n linear systems for $A_{n\times n}$.

The total work is $\underbrace{\frac{2}{3}n^3}_{GE} + \underbrace{n \times 2n^2}_{\text{solving } n \text{ lin. systems}} = \frac{8}{3}n^3 \to$

4 times larger than $\frac{2}{3}n^3$ for GE.

Then $x = A^{-1}b$ requires $2n^2$.

(2) Cramer's Rule:

unappropriate for big systems.

$$Ax = b$$

$$x_i = \frac{\det(A_i)}{\det(A)}, \quad i = 1:n.$$

Work required to compute an $n \times n$ determinant is $> n!$ For $n \geq 20$, with $20! \approx 2.4 \times 10^{18}$, a computer w/ 10^9 FLOPs/sec would work for ≈ 76 years.

Not used on computers!