# Numerical Linear Algebra

# Numerical Linear Algebra

Notation: $A \in \mathbb{R}^{m \times n}$: real matrix, $m$ rows and $n$ columns

$\mathbb{R}^{m \times 1} = \mathbb{R}^{m}$ are (column) vectors

# Common Problems

- Eigenvalue Problem: given $A \in \mathbb{R}^{n\times n}$ find $\lambda \in \mathbb{R}$ or $\mathbb{C}$

  and $x \in \mathbb{R}^n$ or $\mathbb{C}^n$ such that $Ax = \lambda x$

- Linear System: given $A \in \mathbb{R}^{n\times n}, b \in \mathbb{R}^n$ find $x \in \mathbb{R}^n$ s.t. $Ax = b$

- Nonlinear system: almost always reduce to sequence of linear systems eg. Newtons method for $\mathbf{F}(\mathbf{U}) = \mathbf{0}$

  Guess $\mathbf{U}_0$ and for $k = 1, 2, \ldots$
  solve $J\delta = -\mathbf{F}(\mathbf{U}_k)$ and set $\mathbf{U}_{k+1} = \mathbf{U}_k + \delta$

  $$J = \{J_{i,j}\}, J_{i,j} = \frac{\partial F_i}{\partial x_j}(\mathbf{U}_k), \quad \text{Jacobian}$$

- Least Squares (Regression): given $A \in \mathbb{R}^{m\times n}$, $m > n$, $b \in \mathbb{R}^m$ find $x \in \mathbb{R}^m$ such that $||Ax - b||$ is minimum

In different situations $A$ can be:

- Full (dense) - most entries non-zero
- Banded: $\exists b < n$ with $a_{i,j} = 0$ if $|i - j| > b$
  - diagonal: $b = 0$
  - tridiagonal: $b = 1$
- Triangular: upper triangular: $a_{i,j} = 0$ if $i > j$, correspondingly lower
- Hessenberg: upper Hessenberg: $a_{i,j} = 0$ if $i > j + 1$, correspondingly lower
- in Block form: naturally expressed in terms of sub-matrices (matrix blocks)
- Sparse: many zero entries, often very few non-zeros per row

# Block Matrices: If $(m_1, \cdots, m_p)$ and $(n_1, \cdots, n_q)$ are sets of positive integers with $\sum_i m_i = m$, $\sum_j n_j = n$, and $A_{ij} \in \mathbb{R}^{m_i \times n_j}$, $i = 1, \cdots, p$, $j = 1, \cdots, q$ then

$$A = \left[A_{ij}\right] = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & & & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

is a $p \times q$ block matrix.

Names for ordinary matrices apply to block matrices, e.g.

$$T = \begin{bmatrix} A & B \\ O & C \end{bmatrix} \begin{matrix} r \\ s \end{matrix}$$
$$\begin{matrix} r & s \end{matrix}$$

is $2 \times 2$ block upper triangular;
$A \in \mathbb{R}^{r \times r}, C \in \mathbb{R}^{s \times s}, B \in \mathbb{R}^{r \times s}$.

Matrix and vector multiplication work for block matrices: eg. if as above

$$T = \begin{bmatrix} A & B \\ O & C \end{bmatrix} \begin{matrix} r \\ s \end{matrix}, \quad U = \begin{bmatrix} D & E \\ O & F \end{bmatrix} \begin{matrix} r \\ s \end{matrix}, \quad x = \begin{bmatrix} y \\ z \end{bmatrix} \begin{matrix} r \\ s \end{matrix}$$
$$\begin{matrix} r & s \end{matrix} \qquad\qquad \begin{matrix} r & s \end{matrix}$$

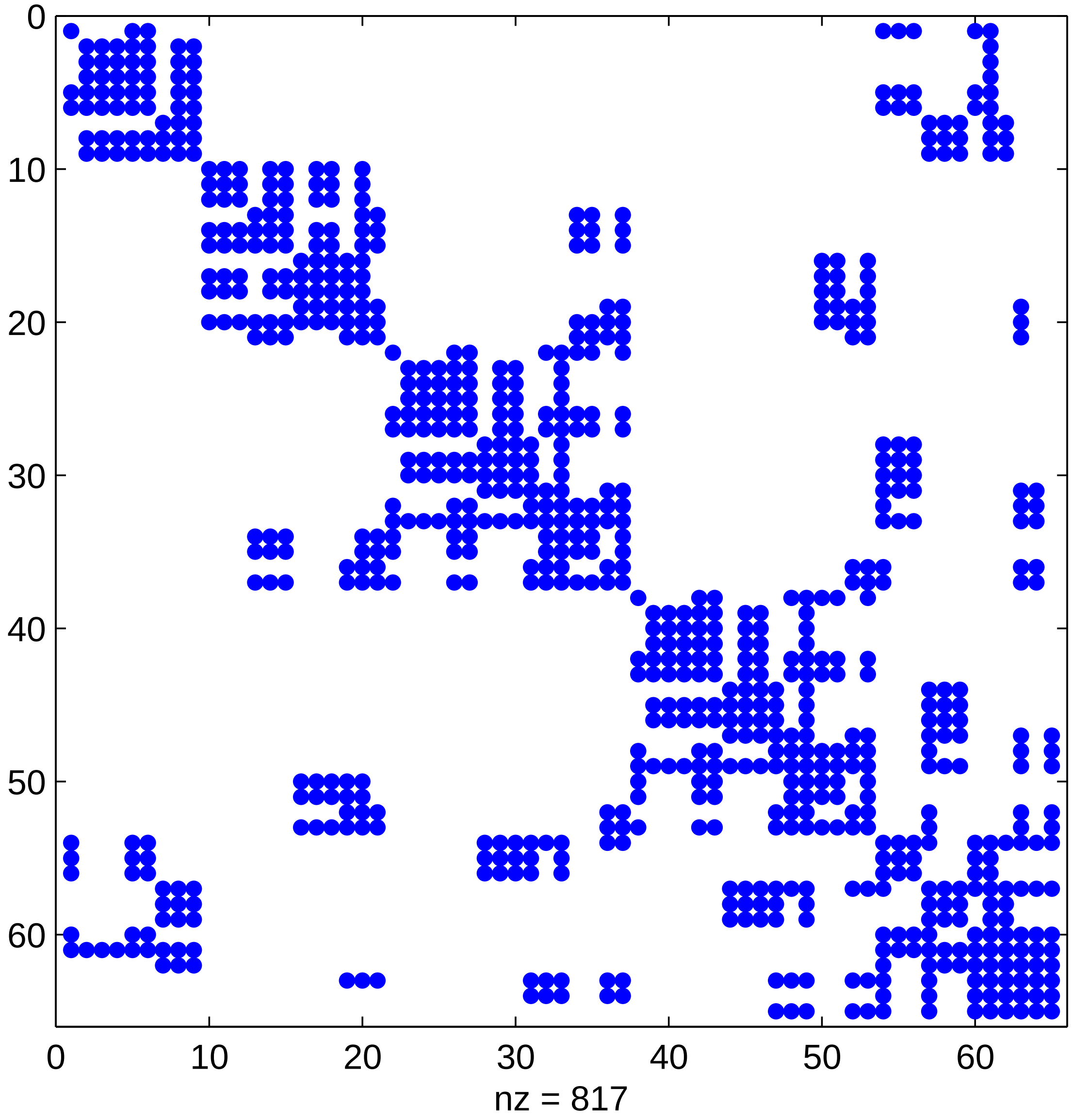
then

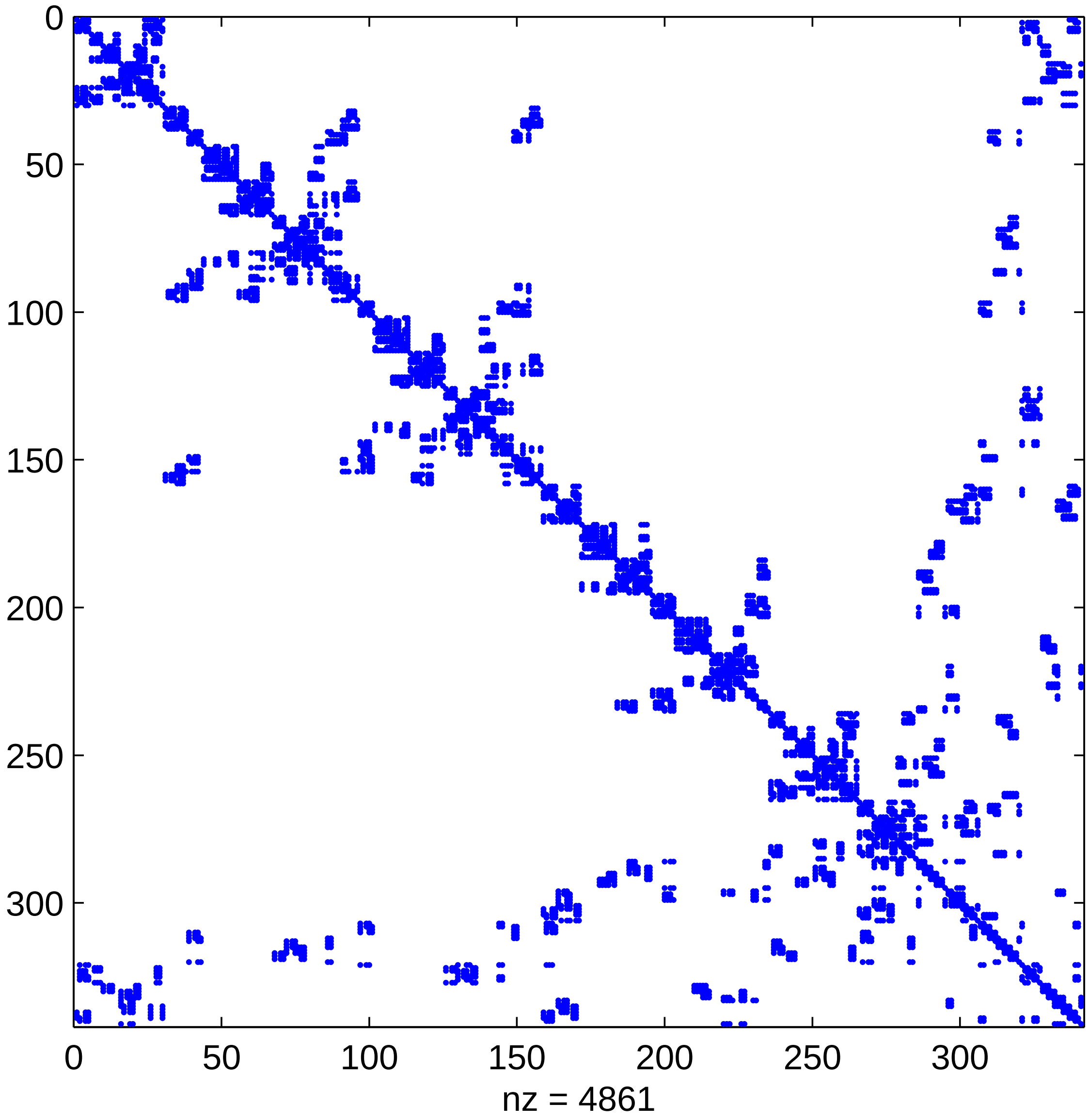
$$TU = \begin{bmatrix} AD & AE + BF \\ O & CF \end{bmatrix}, \quad Tx = \begin{bmatrix} Ay + Bz \\ Cz \end{bmatrix}.$$

Exercise: Deduce that $T^{-1}$ exists if and only if $A^{-1}$ and $C^{-1}$ do, and equals

$$\begin{bmatrix} A^{-1} & -A^{-1}BC^{-1} \\ O & C^{-1} \end{bmatrix}$$

# Sparse Matrices




0
50
100
150
200
250
300
0
50
100
150
200
250
300
nz = 4861

and in any of these structures a square matrix $A \in \mathbb{R}^{n \times n}$ may be

- symmetric: $A^T = A \qquad (a_{i,j} = a_{j,i})$
- skew-symmetric: $A^T = -A$
- positive definite: $x^T A x > 0$ for all non-zero $x \in \mathbb{R}^n$
- positive semi-definite: $x^T A x \geq 0$ for all non-zero $x \in \mathbb{R}^n$
- indefinite: $(x^T A x)(y^T A y) < 0$ for some $x, y \in \mathbb{R}^n$
- $\ldots$

Different structures call for different numerical techniques

Goal of a numerical method is often to convert a given problem to a simpler form (where the solution may be obvious)

# Orthogonal Matrices

$Q \in \mathbb{R}^{n \times n}$ is orthogonal if

**(i)** $Q^T Q = I$ ie.$Q^T = Q^{-1}$

equivalently

**(ii)** $QQ^T = I$

**(iii)** columns of $Q$ are an orthonormal basis for $\mathbb{R}^n$

**(iv)** rows of $Q$ when transposed are an orthonormal basis for $\mathbb{R}^n$

Exercise: **(i)**$\Leftrightarrow$**(ii)**$\Leftrightarrow$**(iii)**$\Leftrightarrow$**(iv)**

2 important orthogonal matrices for computation:

# (a) Householder matrices (elementary reflections)

for $\boldsymbol{w} \in \mathbb{R}^n, \boldsymbol{w} \neq \boldsymbol{0}$ defined by

$$\boldsymbol{H}(\boldsymbol{w}) = \boldsymbol{I} - \boldsymbol{2}\frac{\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}} \quad \in \mathbb{R}^{n \times n}$$

Note

- $\boldsymbol{w}^T\boldsymbol{w} \in \mathbb{R}$: inner product
  $\boldsymbol{w}\boldsymbol{w}^T \in \mathbb{R}^{n \times n}$: outer product
- $\boldsymbol{H}(\boldsymbol{w}) = \boldsymbol{H}(\boldsymbol{w})^T = \boldsymbol{H}(\boldsymbol{w})^{-1}$ Exercise: prove this
- Geometrically $\boldsymbol{H}(\boldsymbol{w})\boldsymbol{x}$ is the reflection of $\boldsymbol{x}$ in the hyperplane $\{\boldsymbol{y} : \boldsymbol{w}^T\boldsymbol{y} = \boldsymbol{0}\}$

# Norms (ways of measuring size)

for vectors:

$$x \in \mathbb{R}^n, \quad \|x\|_2 = \left( \sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} = (x^T x)^{\frac{1}{2}}$$

or more generally $\|x\|_p = \left( \sum_{i=1}^{n} |x_i|^p \right)^{1/p}$

Special/useful cases: $\|x\|_1 = \sum |x_i|, \quad \|x\|_\infty = \max_i |x_i|$

All satisfy the axioms for a norm:

- $\|\alpha x\| = |\alpha| \|x\|, \quad \alpha \in \mathbb{R}, \ x \in \mathbb{R}^n$
- $\|x\| \geq 0$ and $\|x\| = 0 \Leftrightarrow x = 0$
- $\|x + y\| \leq \|x\| + \|y\|$ , triangle inequality

Exercise: check

Norms for matrices: operator norm:

$$\|A\|_p = \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_p}{\|x\|_p} = \sup_{x\in\mathbb{R}^n, \|x\|=1} \|Ax\|_p$$

satisfies matrix norm axioms:

- $\|\alpha A\| = |\alpha|\|A\|, \quad \alpha \in \mathbb{R},\ A \in \mathbb{R}^{m\times n}$
- $\|A\| \geq 0$ and $\|A\| = 0 \Leftrightarrow A = 0$ ($a_{i,j} = 0$ for all $i, j$)
- $\|A + B\| \leq \|A\| + \|B\|$, triangle inequality for matrices

Special cases:

$\|A\|_1 = \max_j \sum_i |a_{i,j}|$ (max absolute column sum)

$\|A\|_\infty = \max_i \sum_j |a_{i,j}|$ (max absolute row sum)

Also: Frobenius norm:

$$\|A\|_F = \left( \sum_i \sum_j a_{i,j}^2 \right)^{\frac{1}{2}}$$

satisfies norm axioms.

Fact: because of compactness (finite dimensionality)

$$\exists x \in \mathbb{R}^n \text{ with } \|x\| = 1 \text{ and } \|A\| = \|Ax\|$$

Orthogonal Matrices and Norms:

If $Q, Z$ orthogonal of appropriate dimensions

$$\|Qx\|_2 = \|x\|_2, \quad \forall x$$

because $\quad \|Qx\|_2^2 = x^T Q^T Q x = x^T x = \|x\|_2^2$

and so also

$$\|QAZ\|_2 = \|A\|_2, \quad \forall A$$

because

$$\frac{\|QAZx\|_2}{\|x\|_2} = \frac{\|AZx\|_2}{\|x\|_2} = \frac{\|A(Zx)\|_2}{\|Zx\|_2} = \frac{\|Ay\|_2}{\|y\|_2}, y = Zx$$

and take supremum over $x$.

Also $\|QAZ\|_F = \|A\|_F$ (see exercises)

# Basic/Important point:

orthogonal matrices do not change 'size' (in $\|\cdot\|_2$ and $\|\cdot\|_F$) so have good numerical properties in finite precision arithmetic