Numerical Methods

§4 - Linear systems

Given $A \in \mathbb{R}^{n \times n}$ and $\underline{b} \in \mathbb{R}^n$, want to find $\underline{x} \in \mathbb{R}^n$ s.t. $\boxed{A\underline{x} = \underline{b}}$.

$\underline{x} = (x_1, \ldots, x_n)^T$
$\underline{b} = (b_1, \ldots, b_n)^T$

With $A = (a_{ij})_{i,j=1}^n$, this is equivalent to the n linear equations $\sum_{j=1}^n a_{ij} x_j = b_i$, $i = 1, \ldots, n$ for n unknowns $x_1, \ldots, x_n$.

e.g. $n=2$:

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

We'll assume a unique solution $\underline{x}$ exists for each $\underline{b}$, equivalently, we assume

- A is invertible (i.e. $\exists A^{-1}$ s.t. $AA^{-1} = A^{-1}A = I$)
- $\det A \neq 0$
- $\operatorname{rank}(A) = n$
- $\nexists$ non-zero $\underline{v}$ s.t. $A\underline{v} = 0$ $\Leftrightarrow$ ($\ker(A) = \{\underline{0}\}$)
- (0 is not an eigenvalue of A)

We know $\underline{x} = A^{-1}\underline{b}$.

But A^{-1} hard to compute in general.

Two numerical methodologies that do work:

- direct methods (e.g. LU factorisation
  i.e. Gaussian elimination)
- iterative methods (what we'll study)

Iterative methods produce a sequence of approximations $\underline{x}_1, \underline{x}_2, \ldots$ to $\underline{x}$

We want $\underline{x}_n \to \underline{x}$ as $n \to \infty$.

And we'd like $\|\underline{x}_n - \underline{x}\|$ to tend to zero rapidly!

Our iterations will involve repeated matrix-vector multiplication.

To study convegence we'll need to consider norms of matrices.

Def'n (matrix norm)

A matrix norm is a norm on the space $\mathbb{R}^{n \times n}$,

i.e. $\|\cdot\| : \mathbb{R}^{n\times n} \to \mathbb{R}_{\geq 0}$ with

- $\|A\| \geq 0 \quad \forall A, \qquad \|A\| = 0 \Leftrightarrow A = 0$ ← 0 matrix.
- $\|\lambda A\| = |\lambda| \, \|A\| \quad \forall A, \; \forall \lambda \in \mathbb{R}$.
- $\|A + B\| \leq \|A\| + \|B\|$.

recall: all norms on $\mathbb{R}^{n\times n}$ are equivalent.

Def'n (convegence): A sequence $(A^{(k)})$ in $\mathbb{R}^{n\times n}$ converges to $A \in \mathbb{R}^{n\times n}$ if $\|A^{(k)} - A\| \to 0$ for some (and hence any) matrix norm $\|\cdot\|$.

Examples:

- "Euclidean norm" in $\mathbb{R}^{n\times n} = \mathbb{R}^{n^2}$ gives the "Frobenius norm":

$$\|A\|_F := \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2}.$$

(not very relevant for us)

- Given a vector norm $\|\cdot\|_V$ on $\mathbb{R}^n$, the induced (or subordinate) matrix norm is given by

V, for "vector"

$$\boxed{\|A\|_M := \sup_{\substack{\underline{x} \in \mathbb{R}^n \\ \underline{x} \neq \underline{0}}} \frac{\|A\underline{x}\|_V}{\|\underline{x}\|_V}}, \quad A \in \mathbb{R}^{n\times n}. \quad (*)$$

"M for "matrix""

(very relevant for us!)

(This is the "operator norm" of the linear operator $\underline{x} \mapsto A\underline{x}$)

Lemma: With $\|\cdot\|_M$ defined as in (✱),

- $\|\cdot\|_M$ is a matrix norm.
- $\|A\underline{x}\|_V \leq \|A\|_M \|\underline{x}\|_V$, $\forall A \in \mathbb{R}^{n \times n}$, $\forall \underline{x} \in \mathbb{R}^n$.

"compatibility" - very important for us.

- $\|I\|_M = 1$

$n \times n$ identity

- $\|AB\|_M \leq \|A\|_M \|B\|_M$ $\forall A, B \in \mathbb{R}^{n \times n}$

Proof: exercise!

Powers of matrices:

We'll use the notation A^k to denote the k-fold product of A with itself.

i.e. $A^0 = I$
$A^1 = A$
$A^2 = A.A$
$A^3 = A.A.A$ etc.

It's easy to show (by induction) that

$$\|A^k\|_M \leq (\|A\|_M)^k \quad \forall k = 0, 1, 2, \ldots$$