Operator (matrix) norms (induced by the corresponding vector norms)

$$\|A\|_{p,q} = \max_{\|x\|_q \le 1} \|Ax\|_p = \max_x \frac{\|Ax\|_p}{\|x\|_q}$$

$$\|A\|_p \equiv \|A\|_{p,p}$$

$$\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}| \quad \text{(max column sum)}$$

$$\|A\|_2 = \sqrt{\max_{1 \le i \le n} \lambda_i (A^T A)} \quad \text{(spectral norm; used most commonly)}$$

$\lambda_i(M)$ – eigenvalue of M

$$\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}| \quad \text{(max row sum)}$$

$$\|A\|_F = \sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2} \quad \text{– (Frobenius "norm"; not really a norm...)}$$

* Ill-Conditioning and Characterization of Errors

Denote

x – the actual (unknown) solution

$\bar{x}$ – the computed solution

$r = b - A\bar{x}$ – the residual

$e = x - \bar{x}$ – the solution error

We have: $Ax = b \Rightarrow x = A^{-1}b$

Using norms: (+)

$$\left.\begin{aligned} \|b\| &= \|Ax\| \le \|A\|\|x\| \\ \|x\| &= \|A^{-1}b\| \le \|A^{-1}\|\|b\| \end{aligned}\right\} \Rightarrow \frac{\|b\|}{\|A\|} \le \|x\| \le \|A^{-1}\|\|b\|$$

Likewise: $r = b - A\tilde{x} = Ax - A\tilde{x} = A(x-\tilde{x}) = Ae$

$$\frac{\|r\|}{\|A\|} \le \|e\| \le \|A^{-1}\|\|r\| \quad (++)$$

Combining (+) and (++)

$$\frac{\|r\|}{\|A\|} \le \|e\| \;/\; \cdot \|x\|$$

$$\left.\begin{aligned} \frac{1}{\|A\|}\frac{\|r\|}{\|x\|} &\le \frac{\|e\|}{\|x\|} \\ \|x\| &\le \|A^{-1}\|\|b\| \end{aligned}\right\} \Rightarrow \frac{1}{\|A\|\|A^{-1}\|}\frac{\|r\|}{\|b\|} \le \frac{\|e\|}{\|x\|}$$

Thus, (by the same argument)

$$\frac{1}{\|A\|\|A^{-1}\|}\frac{\|r\|}{\|b\|} \le \frac{\|e\|}{\|x\|} \le \|A\|\|A^{-1}\|\frac{\|r\|}{\|b\|}$$

condition number $\triangleq \|A\| \cdot \|A^{-1}\| = \mathcal{K}$

So that

$$\frac{1}{\mathcal{K}}\frac{\|r\|}{\|b\|} \le \frac{\|e\|}{\|x\|} \le \mathcal{K}\frac{\|r\|}{\|b\|}$$

Thus, this gives a relative "error bar" on the solution in terms of the condition number $\mathcal{K}$ and the residual r.

$\|r\|$ ~ machine precision ε
$\|b\|$ ~ O(1)

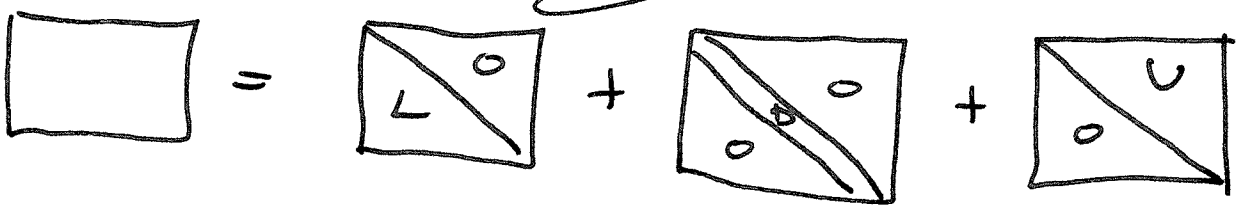
} error e can be quantitatively large if the system matrix A has a large condition number.

ITERATIVE METHODS

Iterative methods start with an initial guess $x^{(0)}$ which is then repeatedly refined. They are a viable alternative for sparse problems.

Jacobi Method

Split the matrix as follows



$$A = L + D + U$$

so that $Ax = (L+D+U)x = b$

Rewrite as

$$Dx^{(n+1)} = -(L+U)x^{(n)} - b$$

$$x^{(n+1)} = -D^{-1}\left[(L+U)x^{(n)} - b\right], \quad n = 1, \ldots, \quad x^{(0)} - \text{given initial guess}$$

$$x_i^{(n+1)} = -\sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{a_{ij}}{a_{ii}} x_j^{(n)} + \frac{b_i}{a_{ii}}, \quad \begin{matrix} i = 1, \ldots, N \\ n = 1, \ldots \end{matrix}$$

This is in fact a fixed-point iteration in $\mathbb{R}^N$ with

$$x^{(n+1)} = G(x^{(n)})$$

Contractivity condition on G necessary for convergence $\Longrightarrow$ diagonal dominance of matrix A

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{N} |a_{ij}| \qquad i = 1, \ldots, N$$

Gauss-Seidel Method

The Jacobi method can be accelerated by using at a given iteration the elements which have already been updated

$$x_i^{(n+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(n+1)} - \sum_{j=i+1}^{N} a_{ij} x_j^{(n)} \right)$$

In the matrix notation

$$(L+D)\, x^{(n+1)} = -U x^{(n)} + b$$

$$x^{(n+1)} = -(L+D)^{-1} U\, x^{(n)} + (L+D)^{-1} b$$

This is also a fixed-point approach. The necessary conditions for convergence are similar to the Jacobi method, but the Gauss-Seidel method is faster.

* Computational cost of iterative methods: $O(mn^2)$
  m - # of iterations

MULTIDIMENSIONAL ROOT-FINDING

(Section 8.2 of Gresselli & Pelinovsky)

Essentially, a generalization of the 1D problem but with important technical modifications.

Problem

Given a continuous function $\bar{F}: \mathbb{R}^n \to \mathbb{R}^n$, find all the solutions of $\bar{F}(\bar{x}) = 0$ for $\bar{x}$ in the region $[a_1, b_1] \times \ldots \times [a_n, b_n]$.