Math 449: Numerical Applied Mathematics
Lecture 19

10/22/2017 Wenzhen

Today's topic: Iterative Methods in Numerical Linear Algebra

Today: Iterative methods numerical lin. alg (continued)

idea: $A\vec{x} = \vec{b}$ construct sequence $x^{(k)} \to$ soln. as $k \to \infty$.

Recall if $\rho(M) < 1$, then $(I-M)^{-1} = I + M + M^2 + M^3 + \cdots$

Can solve $(I-M)\vec{x} = \vec{b}$ using iteration $x^{(0)} = \vec{b}$ $\quad x^{(k+1)} = \vec{b} + Mx^{(k)}$

$$\Rightarrow x^{(k)} = (I + M + \cdots + M^k)\,b$$
$$= (I-M)^{-1} b$$

Generalization: Split

$$A = \underset{\text{nice}}{N} - \underset{\text{permutation}}{P}$$

Iteration:

$$N x^{(0)} = b$$

$$N x^{(k+1)} = b + P x^{(k)}$$

This equiv to
$$(N-P)x = b$$
$$N^{-1}(N-P)x = N^{-1}b$$
$$(I - N^{-1}P)x = N^{-1}b$$
i.e. $M = N^{-1}$

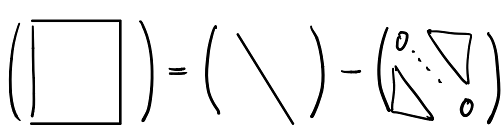
So cvg as $k \to \infty$ when $\rho(N^{-1}P) < 1$

Two main import ways to choose N.

1. Jacobi method.

$$N = \mathrm{diag}(A)$$

$$P = N - A$$

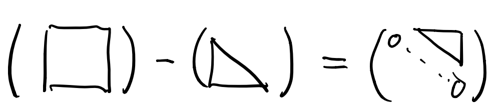


$O(n^2)$

2. Gauss-Seidal methods.

$N = \mathrm{tril}(A)$ (lower triangular part of A)

$$P = N - A$$



Each step $O(n^2)$

When does this converge?

Use the fact that spectral radius is bounded above by any matrix norm.

$\|\cdot\|_1$ and $\|\cdot\|_\infty$ are easy to compute.

Def. $A \in \mathbb{R}^{n\times n}$ is strictly diagonalized dominant if

$$|a_{ii}| > \sum_{j\neq i} |a_{ij}| \quad \text{for } i=1,\dots,n.$$

For Jacobi's method.

$$N^{-1}P = \operatorname{diag}(A)^{-1}[\operatorname{diag}(A) - A] = I.$$

$$= I - \operatorname{diag}(A)^{-1}A$$

$a_{ii} \neq 0.$

Note: If A is s.d.d. then $\operatorname{diag}(A)$ is nonsingular, since $|a_{ii}| > 0$.

Thm: A is strictly ~~di~~ dominant (s.d.d.) $\Leftrightarrow \|I - \operatorname{diag}(A)^{-1}A\|_\infty < 1$

Proof. $\|I - \operatorname{diag}(A)^{-1}A\|_\infty = \max_{i=1,\dots,n} \left[\sum_{j\neq i} \left| \frac{a_{ij}}{a_{ii}} \right| + 1 - 1 \right]$

diag term of $\operatorname{diag}(A^{-1})A$ — diag term of I

$$= \max_{i=1,\dots,n} \frac{1}{|a_{ii}|} \sum_{j\neq i} |a_{ij}| < 1$$

$\Leftrightarrow$ A s.d.d. $\boxtimes$

Cor If A is sdd, then the Jacobi method cvgs.

Proof. $\rho(N^{-1}P) = \rho(I - \operatorname{diag}(A^{-1})A)$

$$\leq \|I - \operatorname{diag}(A^{-1})A\|_\infty$$

$$< 1 \quad \boxtimes$$

This is equivalent to $(N-P)\vec{x} = b$

$$N^{-1}(N-P)\vec{x} = N^{-1}b$$

$$(I - N^{-1}P)\vec{x} = N^{-1}b$$

i.e. $M = N^{-1}P$ so cvgs as $k\to\infty$ when $\rho(N^{-1}P) < 1$

The conjugate gradient method.

Used to solve $A\vec{x} = \vec{b}$ when A is symmetric positive definite (spd.).

Defn. $A \in \mathbb{R}^{n \times n}$ is spd when

- $A^T = A$
- $x^T A x^T > 0$ for $x \neq 0$

Rmle. This is condition for $\langle x, y \rangle = x^T A y$ to be an inner product on $\mathbb{R}^n$

Facts about spd matrices.

Since symmetric, spectral thm says that A is diagonizeble, with real eigs and compute orthonormal basis of eigenvectors.

- since positive definite. $\forall\ \lambda > 0 \quad Av = \lambda v \quad \therefore \vec{v} \neq 0$

$$v^T A v = \lambda \underbrace{(v^T v)}_{>0} > 0.$$

Prop. let $A \in \mathbb{R}^{n \times n}$ be spd. $Ax = b \iff x$ minimizes $f(x) = \frac{1}{2} x^T A x - x^T b$

$\frac{1}{2} A x^2 - bx \quad x(\frac{1}{2} A x - b) = 0.$

Proof. $\nabla f(x) = Ax - b$

$$f(x) = \frac{1}{2} \sum_{i,j} a_{i,j} x_i x_j - \sum_j b_j x_j$$

$$\frac{\partial f}{\partial x_k} = \frac{1}{2} \sum_i a_{ik} x_i + \frac{1}{2} \sum_j a_{kj} x_j - b_k$$

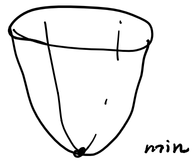
$$= \sum_j \left(\frac{a_{jk} + a_{kj}}{2} \right) x_j - b_k$$

$$= \sum_j a_{kj} x_j - b_k$$

$$= (Ax - b)_k$$

$\boxtimes$

$f(x)$



Given a guess $x^{(k)}$, pick a search direction $s^{(k)}$

and take $$x^{(k+1)} = x^{(k)} + \underbrace{\underset{\text{scalar}}{t^{(k)}}\ \underset{\text{vector}}{s^{(k)}}}_{\Delta x^{(k)}}$$

Move dist $t^{(k)}$ along the line through $x^{(k)}$ in direction $s^{(k)}$

s.t. $f(x^{(k+1)})$ is minimized along the direction.

Expanding $$\frac{d}{dt^{(k)}} f(x^{(k)} + t^{(k)} s^{(k)}) = 0$$

$$\Rightarrow t^{(k)} = -\frac{s^{(k)T}(Ax^{(k)} - b)}{s^{(k)T} A s^{(k)}}$$

Gradient Descent.

Pick $$s^{(k)} = -\nabla f(x^{(k)})$$
$$= -(Ax^{(k)} - b)$$

$$t^{(k)} = \frac{s^{(k)T} s^{(k)}}{s^{(k)T} A s^{(k)}}$$

$$x^{(k+1)} = x^{(k)} + \frac{s^{(k)T} s^{(k)}}{s^{(k)T} A s^{(k)}} s^{(k)}$$

where $s^{(k)} = -(Ax^{(k)} - b)$

To avoid repeated $s^{(k)}$ in same direction, conjugate gradient method ensures $s^{(k)T} A s^{(l)} = 0 \quad \forall l < k$.