Lecture 10

- If $Ax = b$ doesn't satisfy Convergence Criterion for Jacobi, can we scale it to make it Richardson?

  - Can we find "correct" row to update x_i?
    Gauss-Seidel Method!
  - Do we assume $Ax = b$ is spd? ✓
    When it not?

- Convergence Rate for $Ax = b$:

  - Why are λ-vals relevant for Jacobi + Richardson?
  - Why should we read question in Calculus → Unit!
  - Error $\underline{e}^{(k)}$: $\underline{e}^{(k)} = \partial_{\mathcal{F}} g_{\mathcal{A}} \underline{e}^{(k-1)}$ ||
    $\Rightarrow \underline{e}^{(k)} = \frac{\partial_{\mathcal{F}} g_{\mathcal{A}} \cdot A\underline{e}^{(k)}}{\|A\underline{e}^{(k)}\|}$

- Why can't we use Gradient descent to find the Eigenvalues?

  - Common methods for solving Poisson Eq.?
  - Poisson vs. Laplace
    $\nabla u = f$ vs
    $\nabla u = 0$.

Scaling Jacobi/GS?

Note! They are scale invariant!

$Ax = b$ vs. $\alpha Ax = \alpha b$

$x^{(0)}$ $x^{(0)}$
$\vdots$ $\vdots$
$x^{(k)}$ $x^{(k)}$

Show they get the same sequence of iterates.

Gauss-Seidel will always converge of A is sym. pos def.

Set up —

If GS should converge in, say, 1000 steps can I tweak it (gently) that it nearly was converge.

$Ax = \underline{b}$

$\Downarrow$

$\min \|Ax - b\|$

$\Downarrow$

$A^T A x = A^T \underline{b}$

always sym. pos def if A is full-rank.

If Jacobi/GS converges, they give us.

$\underline{e}^{(k)} = (M^{-1}N)^k \underline{e}^{(0)}$

$\underline{e}^{(k)} = \underline{x} - \underline{x}^{(k)} \to$ the error.

So if $\rho(M^{-1}N) < 1$, we know the

$\|\underline{e}^{(k)}\| \to 0$ like $\rho(M^{-1}N)^k$

To get $\|\underline{e}^{(k)}\| \le \varepsilon$, we need $O\left(\frac{\log(\varepsilon)}{\log(\rho(M^{-1}N))}\right)$ iterations.

So with "$O(N \log(\varepsilon))$" time to use Jacobi/GS if they converge!

$\max x^T A x$ s.t. $\|x\|^2 = 1$

$\Downarrow$

$x^{(k+1)} = \frac{x^{(k)} + \gamma_k \cdot Ax^{(k)}}{\|x^{(k)} + \gamma_k \cdot Ax^{(k)}\|}$

as $\gamma \to \infty$ then we get

$\frac{Ax^{(k)}}{\|Ax^{(k)}\|}$

Gradient Ascent on the Lagrangian.

$\mathcal{L}(x, \lambda) = x^T A x - \lambda(x^T x - 1)$

We use stationary pt of $\mathcal{L}$.

$\frac{\partial \mathcal{L}}{\partial x} = 0 = 2Ax - 2\lambda x$ $\quad Ax = \lambda x$

$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 = x^T x - 1.$ $\quad \Leftrightarrow x^T x = 1$

CS 515 - Purdue University - Computer Science - David F. Gleich - 2025