Sparse Systems and Iterative Methods

Paul Heckbert

Computer Science Department
Carnegie Mellon University
PDE’s and Sparse Systems

• A system of equations is *sparse* when there are few nonzero coefficients, e.g. $O(n)$ nonzeros in an $nxn$ matrix.

• Partial Differential Equations generally yield sparse systems of equations.

• Integral Equations generally yield dense (non-sparse) systems of equations.

• Sparse systems come from other sources besides PDE’s.
Example: PDE Yielding Sparse System

- Laplace’s Equation in 2-D: \( \nabla^2 u = u_{xx} + u_{yy} = 0 \)
  - domain is unit square \([0,1]^2\)
  - value of function \(u(x,y)\) specified on boundary
  - solve for \(u(x,y)\) in interior
Sparse Matrix Storage

- Brute force: store nxn array, $O(n^2)$ memory
  - but most of that is zeros – wasted space (and time)!
- Better: use data structure that stores only the nonzeros

\[
\begin{array}{cccccccccc}
\text{col} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
\text{val} & 0 & 1 & 0 & 0 & 0 & 1 & -4 & 1 & 0 & 0 \\
\end{array}
\]

16 bit integer indices: 2, 5, 6, 7, 10
32 bit floats: 1, 1, -4, 1, 1

- Memory requirements, if $kn$ nonzeros:
  - brute force: $4n^2$ bytes, sparse data struc: $6kn$ bytes
An Iterative Method: Gauss-Seidel

• System of equations $Ax=b$
• Solve $i$th equation for $x_i$:
• Pseudocode:
  
  until $x$ stops changing
  for $i = 1$ to $n$
  
  $x[i] \leftarrow (b[i] - \text{sum}\{j \neq i\}{a[i,j] \times x[j]})/a[i,i]$

• modified $x$ values are fed back in immediately
• converges if $A$ is symmetric positive definite
Variations on Gauss-Seidel

• Jacobi’s Method:
  – Like Gauss-Seidel except two copies of $x$ vector are kept, “old” and “new”
  – No feedback until a sweep through $n$ rows is complete
  – Half as fast as Gauss-Seidel, stricter convergence requirements

• Successive Overrelaxation (SOR)
  – extrapolate between old $x$ vector and new Gauss-Seidel $x$ vector, typically by a factor $\omega$ between 1 and 2.
  – Faster than Gauss-Seidel.
Conjugate Gradient Method

• Generally for symmetric positive definite, only.

• Convert linear system $Ax=b$
• into optimization problem: minimize $x^TAx-x^Tb$
  – a parabolic bowl

• Like gradient descent
  – but search in conjugate directions
  – not in gradient direction, to avoid zigzag problem
• Big help when bowl is elongated ($\text{cond}(A)$ large)
Conjugate Directions
Convergence of Conjugate Gradient Method

• If $K = \text{cond}(A) = \frac{\lambda_{\text{max}}(A)}{\lambda_{\text{min}}(A)}$

• then conjugate gradient method converges linearly with coefficient $(\sqrt{K}-1)/(\sqrt{K}+1)$ worst case.

• often does much better: without roundoff error, if $A$ has $m$ distinct eigenvalues, converges in $m$ iterations!