Fundamental Concepts in Computational and Applied Mathematics

Juan Meza
School of Natural Sciences
University of California, Merced

Fall 2014

Consider the following matrix

$$K_n = \begin{bmatrix} 2 & -1 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 \\ \dots & 0 & -1 & 2 & \ddots & 0 \\ 0 & \dots & 0 & \ddots & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 \end{bmatrix}$$

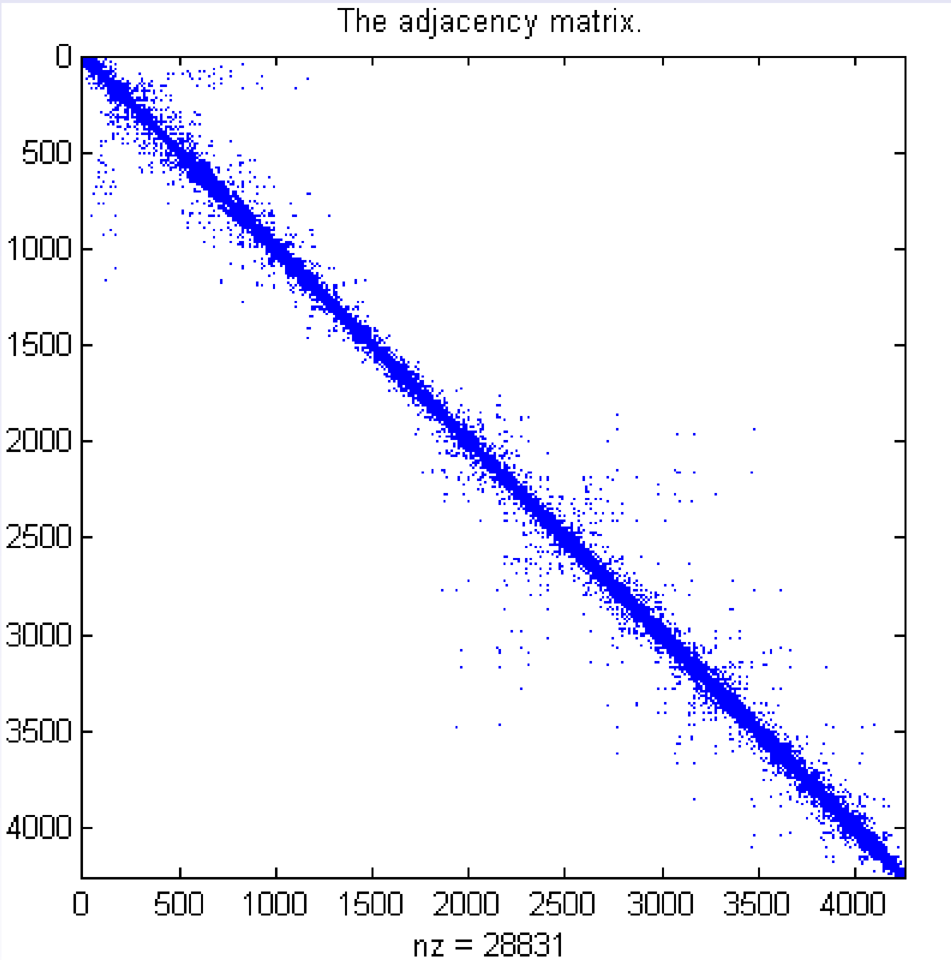
- We will encounter matrices of this form in future talks
- We can take advantage of the special structure to improve our algorithms
- This matrix is known as a Toeplitz matrix (constant diagonal)

Another interesting matrix

$$D = \begin{bmatrix} 0 & 1/2 & 0 & 0 & \dots & -1/2 \\ -1/2 & 0 & 1/2 & 0 & \dots & 0 \\ 0 & -1/2 & 0 & 1/2 & \dots & 0 \\ \dots & 0 & -1/2 & 0 & \ddots & 0 \\ 0 & \dots & 0 & \ddots & 0 & 1/2 \\ 1/2 & 0 & 0 & 0 & -1/2 & 0 \end{bmatrix}$$

- A second order method can be given by the following matrix
- This matrix is Toeplitz, skew-symmetrix, and circulant. All very nice properties.
- Does this structure remind you of anything?

More generally



Definition

Any matrix where the majority of entries are zero is referred to as a *sparse* matrix.

Discussion

So what does *majority* mean?

Sparse methods

- Sparse matrices arise in many science and engineering applications
- Almost all discretization methods you will encounter will have a sparse structure
- Sparse methods take advantage of special structure in matrices to improve computational efficiency

Stationary Iterative Methods

- Based on splitting techniques
- Widely used in engineering, but have been superseded by more modern techniques
- Still useful in certain contexts, e.g. preconditioners (more later)

Basic Stationary Iterative Method (SIM)

Consider $Ax = b$
First we'll "split" the matrix $A = M - N$, where $M = L + D$ and $N = -U$. Now I'll define an iteration for $k = 0, \ldots,$ s.t.

$$
\begin{aligned}
(L+D)x^{k+1} &= b - Ux^k, &(1)\\
Mx^{k+1} &= b + Nx^k, &(2)\\
x^{k+1} &= M^{-1}b + M^{-1}Nx^k, &(3)\\
x^{k+1} &= Gx^k + c, &(4)
\end{aligned}
$$

where $G = M^{-1}N, c = M^{-1}b$.

Note

This is one of several splittings one can use. Can you think of others?

Useful Facts

Fact 1

$e^{k+1} = Ge^k, \quad k = 0, \ldots.$

Fact 2

If we use the SIM in equation (4) $x^k \to \hat{x} \iff \rho(G) < 1.$

Fact 3

Many important matrices resulting from problems arising in CSE satisfy $\rho(G) < 1.$

Discussion: Hestenes and Stiefel

Finite Precision

- In exact arithmetic the CG method will converge to the solution of $Ax = b$ in at most n iterations for A spd.
- In finite precision, roundoff error will contaminate the solution.
- And even in the exact case, what does this mean for a system with dimension 1M or 1B?

CG Convergence Rate

Can show that

$$||x^k - \hat{x}||_A \leq 2 \left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right)^k ||x^0 - \hat{x}||_A \tag{5}$$

Table: Number of iterations to reduce error by 10^{-3}.

$\kappa(A)$	Value	iter
1	0.0	1
10^2	0.82	38
10^4	0.98	376
10^8	0.9998	38,000

Advanced Topic: Krylov subspace methods

CG is an example of what we now call a Krylov space method.

Definition

A *Krylov* space is defined by:

$$\mathcal{K}_k(A, b) = span\{b, Ab, A^2b, A^3b, \ldots, A^{k-1}b\}.$$

Idea

Choose your iterate, x^k, such that it belongs to the Krylov space and it minimizes the distance between x^k and $\hat{x}$.

Advanced Topic – Preconditioning

Fundamental Idea

If you don't like the problem you're given to solve, then change it into something that is easier to work with!

- If the error bound is dependent on the condition number can we do something about it?
- Ideally what would we want the condition number of the problem to be?
- Can we alter the problem to make it better conditioned, without changing the solution of the problem

Advanced Topic – Preconditioning

Consider $Ax = b$, where $\kappa(A)$ is large. Let M be a symmetric positive definite (spd) matrix. Then I can use CG to solve the equivalent problem:

$$M^{-1}Ax^{k+1} = M^{-1}b, \quad k = 0, \ldots \tag{6}$$

and if $\kappa(M^{-1}A) < \kappa(A)$, I should be able to reduce the number of iterations.

Preconditioning

Choose M^{-1} wisely and you will reduce the number of iterations significantly. Choose poorly and all you've done is create more work for yourself!

Exercise

Note that in general $M^{-1}A$ may not be symmetric. How then would you apply CG to this system?

Summary

- Sparse methods can be used to take advantage of special structure.
- Iterative methods are the dominant form of solving large-scale systems of linear equations in modern day applications.
- Most iterative methods must be preconditioned to be effective.

References I

Gene H. Golub and Charles F. Van Loan.
Matrix Computations, 3rd Ed.
Johns Hopkins, 1996.

C. Tim Kelley.
Iterative Methods for Linear and Nonlinear Equations.
SIAM, 1995.

Magnus R. Hestenes and Eduard Stiefel.
Methods of Conjugate Gradients for Solving Linear Systems
J. Res. of NBS, Vol. 49, No. 6, Dec. 1952.

Appendix: Basic CG Method

Consider $Ax = b$, where A is spd. Let $x_0 = 0, r_0 = b, p_0 = r_0$.

for $k = 1, 2, 3, \ldots$

$$\begin{aligned}
\alpha_k &= (r_{k-1}^T r_{k-1})/(p_{k-1}^T A p_{k-1}) \\
x_k &= x_{k-1} + \alpha_k p_{k-1} \\
r_k &= r_{k-1} - \alpha_k A p_{k-1} \\
\beta_k &= (r_k^T r_k)/(r_{k-1}^T r_{k-1}) \\
p_k &= r_k + \beta_k p_{k-1}
\end{aligned}$$