CX 4640 Assignment 6

Wesley Ford

September 29th, 2020

1. Chapter 7, Question 9: Let α be a scalar, and consider the iterative scheme

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha(\boldsymbol{b} - A\boldsymbol{x}_k)$$

This is the gradient descent method with a fixed step size α.

(a) If $A = M - N$ is the splitting associated with this method, state what M and the iteration matrix G are.

We know that $M\boldsymbol{x} = N\boldsymbol{x} + \boldsymbol{b} = (M - A)\boldsymbol{x} + \boldsymbol{b}$. To figure out what M is, we want to manipulate the iterative scheme into this form

Expanding out the iterative scheme and rearranging :

$$\boldsymbol{x_{k+1}} = \boldsymbol{x}_k - \alpha A\boldsymbol{x}_k + \alpha\boldsymbol{b}$$

Dividing by a factor of α and combining common terms

$$\frac{\boldsymbol{x}_{k+1}}{\alpha} = \frac{\boldsymbol{x}_k}{\alpha} - A\boldsymbol{x}_k + \boldsymbol{b}$$
$$\frac{\boldsymbol{x}_{k+1}}{\alpha} = (\frac{I}{\alpha} - A)\boldsymbol{x}_k + \boldsymbol{b}$$

We can see now that $M = \alpha^{-1}I$.

The iteration matrix G is $I - M^{-1}A = I - \alpha A$

(b) Suppose A is SPD and its eigenvalues are $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$.

i. Derive a condition on α that guarantees convergence of the scheme to the solution $\mathbf{x}$ for any initial guess.

For an eigenvalue of A and corresponding eigenvector $\boldsymbol{x}$, the following is true

$$A\boldsymbol{x} = \lambda\boldsymbol{x}$$

Solving for A in terms of G, we get $A = \frac{I-G}{\alpha}$. Plugging this into the above equation

$$(I - G)\boldsymbol{x} = \alpha\lambda\boldsymbol{x}$$
$$G\boldsymbol{x} = (1 - \alpha\lambda)\boldsymbol{x}$$

So if γ is an eigenvalue of G, we can say that $\gamma = 1 - \alpha\lambda$.

Convergence is guaranteed when $\rho(G) = \max_i\{|\gamma_i|\} < 1$. So for all eigenvalues of G, we need

$$|\gamma| = |1 - \alpha\lambda| < 1$$
$$0 < \alpha\lambda < 2$$
$$0 < \alpha < \frac{2}{\lambda}$$

Because this must hold for all λ, the tightest bound is going to be

$$0 < \alpha < \frac{2}{\lambda_1}$$

ii. Show that the best value for the step size in terms of maximizing the speed of convergence is

$$\alpha = \frac{2}{\lambda_1 + \lambda_n}.$$

To find the optimal step size, we want to minimize $\rho(G)$. This means we want to find $\min\{\max\{|\gamma|\}\}$ The maximum values of $|\gamma|$ are either $|1-\alpha\lambda_1|$ or $|1-\alpha\lambda_n|$. To minimize the maximum of these values, we want to find α such that both are equal.

$$|1-\alpha\lambda_1| = |1-\alpha\lambda_n|$$
$$1-\alpha\lambda_1 = \pm(1-\alpha\lambda_n)$$

In the case where $1-\alpha\lambda_1 = 1-\alpha\lambda_n$, $\lambda_n = \lambda_1$, which is not possible with the given eigenvalues, so we know that the only possible case is

$$1-\alpha\lambda_1 = \alpha\lambda_n - 1$$
$$\alpha(\lambda_1+\lambda_2) = 2$$
$$\alpha = \frac{2}{\lambda_1+\lambda_n}$$

So to maximize the speed of convergence, we minimize $\rho(G)$, which is at a minimum when

$$\alpha = \frac{2}{\lambda_1+\lambda_n}$$

Find the spectral radius of the iteration matrix in this case, and express it in terms of the condition number of A.

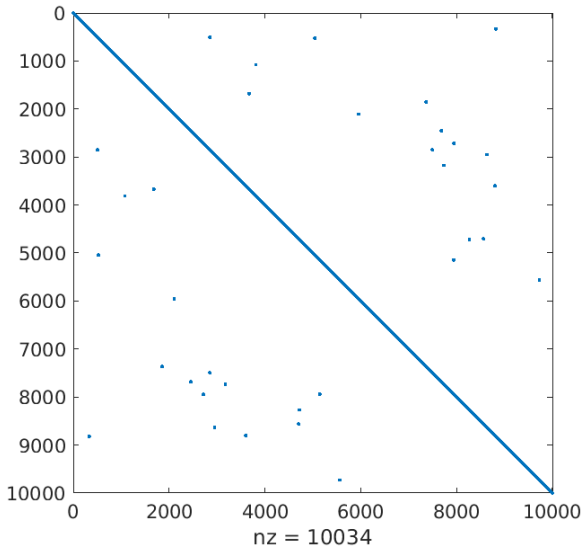
The spectral radius of the iteration matrix $\rho(G)$ will be $|1-\frac{2\lambda_1}{\lambda_1+\lambda_n}| = |\frac{\lambda_1+\lambda_n-2\lambda_1}{\lambda_1+\lambda_n}| = |\frac{\lambda_n-\lambda_1}{\lambda_1+\lambda_n}|$. The condition number of A is norm-2 is $\kappa_2(A) = \frac{\lambda_1}{\lambda_n}$, so $\lambda_1 = \kappa_2(A)\lambda_n$. Plugging these values into our equation for $\rho(G)$, we get $|\frac{\lambda_n-\kappa_2(A)\lambda_n}{\lambda_n\kappa_2(A)+\lambda_n}| = |\frac{1-\kappa_2(A)}{1+\kappa_2(A)}|$. Because $\kappa_2(A)$ will always be greater than 1 ($\lambda_1 > \lambda_n \rightarrow \frac{\lambda_1}{\lambda_n} > 1$), $\frac{1-\kappa_2(A)}{1+\kappa_2(A)} < 0$, so $|\frac{1-\kappa_2(A)}{1+\kappa_2(A)}| = \frac{\kappa_2(A)-1}{\kappa_2(A)+1}$.

(c) Determine whether the following statement is true of false. "If A is strictly diagonally dominant and $\alpha = 1$, then the iterative scheme converges to the solution for any initial guess $\boldsymbol{x}_0$."

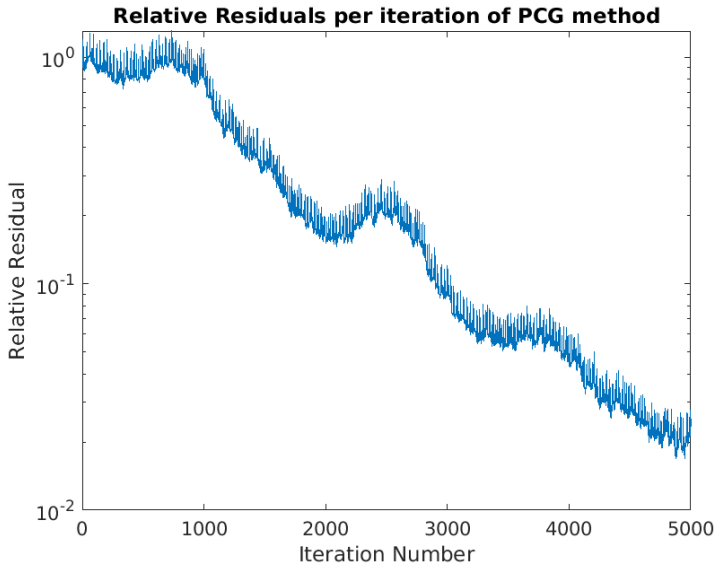
From part (b), we know that $0 < \alpha < \frac{2}{\lambda_1}$. If $\alpha = 1$, then for the iterative scheme to converge, the following inequalities must hold: $0 < \lambda_1 < 2$. If we choose A to be a diagonal matrix with entries greater than 2, clearly $\lambda_1 > 2$, and the iterative method should not converge for any initial guess.

2. Construct a sparse, non-diagonal SPD matrix A of size 10000×10000 such that convergence requires more that 5000 iterations i.e. the relative residual norm is greater than 1e-6 after 5000 iterations. Explain how you construct your matrix A and plot the relative residual norm at each iteration.

   A sparse SPD matrix is created by generating a sparse lower triangular matrix l. Using the Cholsky decomposition, A is generated by $l * l^T$. To ensure that the matrix is SPD, l is generated in such a way that guarantee it to have a fully positive diagonal, which ensures that A has a positive diagonal, a property of SPD matrices. Using Matlab's `spy` and `chol`, we can guarantee that A is indeed non-diagonal, sparse, and SPD. Using `condest`, we find that the condition number of A is on the order of 10^7. A has the following structure



   Plotting the relative residual per iteration yields the following.



   The relative residual does seem start converging over the 5000 iterations, but quite slowly, due to the relatively high condition number of the matrix.

```
%Code for question 2 of Assignemnt 6
clear
n = 10000;
%Create a random diagaonl matrix
dig = sparse(abs(diag(rand(n,1))));
%Ensure that l has a full diagonal
l = sparse(abs(tril(sprand(n, n, 8e-8))) +  dig);
A = l*l';
%Estimate conditon number
condest(A)
%Generate strucure of matrix
spy(A)
%Make sure is SPD
chol(A);
%Get random b
b = rand(n, 1);
%Call PCG
[x, flag, relres, iter, resvec] = pcg(A,b, 1e-6, 5000);
flag
relres
iter
semilogy((0:length(resvec)-1)', resvec/norm(b));
ylabel("Relative Residual")
xlabel("Iteration Number")
title("Relative Residuals per iteration of PCG method")
```