Homework 2

Deadline (for bonus points): 2014-11-28

1. Implement GMRES (Generalized Minumum Residual method) based on the Arnoldi method, with good orthogonalization, that you developed in Homework 1. Consider the linear system $Ax = b$, where A and b are generated by

A typo in the code was corrected on 2014-11-20.

```
alpha=5;  m=100;
rand('state',5);
A = sprand(m,m,0.5);
A = A + alpha*speye(m); A=A/norm(A,1);
b = rand(m,1);
```

(a) Plot the norm of the error as a function of iteration as well as the residual norm (with `semilogy`). You may in this exercise use `A\b` as an exact solution. Generate figures for the values $\alpha = 1, 5, 10, 100$.

(b) Plot the eigenvalues with `plot(eig(full(A)),'*')` for all choices of α in (a). Relate the observed convergence to the convergence theory by plotting the estimated convergence rate (predicted by the eigenvalues) in convergence figures as in (a).

(c) Generate the following table, where `resnorm`=$\|Ax - b\|_2$ and `time` is the CPU-time and n is the number of iterations. In this particular problem it will be sufficient to compute the solution such that it has a residual norm 10^{-5}. Can we beat the the matlab backslash operator \ ? Use $\alpha = 1$ and $\alpha = 100$.

The matlab backslash-command is based on extremely optimized LU-factorizations (or sometimes Cholesky factorizations or Cholesky decompositions).

	$m = 100$		$m = 200$		$m = 500$	
	resnorm	time	resnorm	time	resnorm	time
$n = 5$						
$n = 10$						
$n = 20$						
$n = 50$						
$n = 100$						

2. In this exercise you shall computationally verify the theoretical orthogonality and minimization properties of CG and GMRES. Consider the linear system of equations

$$\begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} x = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

(a) Compute α, β, γ such that $\operatorname{span}(x_1, x_2, x_3, x_4) = \operatorname{span}(c_0, c_1, c_2, c_3)$ where $c_0, \ldots, c_3$ have the structure

Hint: Consider the Krylov matrix K_n and use theory about the span of the iterates in the conjugate gradient method.

$$C = [c_0, c_1, c_2, c_3] = \begin{bmatrix} 1 & \alpha & 0 & \gamma \\ 1 & 0 & \beta & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and $x_0, \ldots, x_3$ are the iterates of the conjugate gradient method for this linear system. This can be solved without a computer.

(b) Implement CG for this problem (TB Algorithm 38.1) and verify the minimization property as follows. In the following code, replace `???` with appropriate formulas such that `x_fmin` and `x_cg` become equal (at least in exact arithmetic and if fminsearch solves the problem exactly).

```
alpha=??; beta=??; gamma=??;
C=[1 alpha 0    gamma
   1 0     beta 0
   0 1     0    0
   0 0     1    7
   0 0     0    1
   0 0     0    0
   0 0     0    0
   0 0     0    0];
opts = optimset('TolFun',1e-10);
z=fminsearch(@(z) ??? ,  [1;1;1;1],opts);
x_fmin=C*z;
x_cg=cg(A,b,3);
```

(c) Make the same experiment as in (b) for GMRES, by changing `???` in `z=fminsearch(@(z) ??? ,  [1;1;1;1],opts)` such that `C*z` becomes the approximation generated by GMRES.

3. Suppose A is a real symmetric matrix with eigenvalues $10, 10.5$ and 100 eigenvalues in the interval $[2, 3]$. How many steps of CG must be carried out in order to reduce the error (measured in $\|Ax_n - b\|_{A^{-1}} = \|x_n - x_*\|_A$) by a factor 10^7. You may assume exact arithmetic and that no premature breakdown occurs.

4. Download `lshape_mod.mat` from the course web page.

The matrix in `lshape_mod.mat` represents a variant of a finite-element discretization of the Laplacian on an L-shaped domain.

(a) Run GMRES for this linear system and plot the residual norm $\|Ax_n - b\|$ in a semilogarithmic plot (semilogy).

(b) Adapt a preconditioner for GMRES as follows. Let D be the diagonal elements of A, extracted as `D=diag(diag(A))`. Multiply the equation $Ax - b = 0$ from the left with D^{-1} and apply GMRES on the new linear system of equations. Plot the residual norm with respect to the original linear system $\|Ax_n - b\|$.

(c) Plot the eigenvalues of A and $D^{-1}A$ and explain the difference in convergence in (a) and (b).

PhD students: see next page.

Only for PhD students taking the course *Numerical linear algebra*:

5. In the lectures we derived convergence bounds of GMRES for diagonalizable matrices. In this exercise you shall show convergence for a class of non-diagonalizable matrices. Suppose $A \in \mathbb{C}^{m\times m}$ is invertible and suppose $\lambda_1 = \lambda_2$ is a double eigenvalue and all other eigenvalues are distinct. Assume that λ_1 has a Jordan block of size two. Moreover, suppose all eigenvalues λ_i, $i = 1,\dots,m$ are contained in an open disk of radius $\rho > 0$ centered at $c \in \mathbb{C}$, such that $\lambda_i \in C(\rho, c)$ for all $i = 1,\dots,m$. Assume $|\rho| < |c|$ and $\lambda_1 \neq c$.

   (a) Let $V\Lambda V^{-1} = A$ be the Jordan canonical form. Prove

   $$\min_{p\in P_n^0} \|p(A)\| \leq \|V\|\|V^{-1}\| \min_{p\in P_n^0} \|p(\Lambda)\|$$

   (b) Prove that for any polynomial $p(z) = a_0 + a_1 z + \cdots + a_n z^n$,

   $$p\left(\begin{pmatrix}\lambda_1 & 1\\ 0 & \lambda_1\end{pmatrix}\right) = \begin{pmatrix}p(\lambda_1) & p'(\lambda_1)\\ 0 & p(\lambda_1)\end{pmatrix}.$$

   (c) Prove

   $$\|p(\Lambda)\| = \max\left(\left\|\begin{pmatrix}p(\lambda_1) & p'(\lambda_1)\\ 0 & p(\lambda_1)\end{pmatrix}\right\|, |p(\lambda_3)|, |p(\lambda_4)|, \dots, |p(\lambda_m)|\right)$$

   (d) Determine α_n and β_n such that

   $$p(z) = (\alpha_n + \beta_n z)\frac{(c-z)^{n-1}}{c^{n-1}}$$

   satisfies $p \in P_n^0$ and $p'(\lambda_1) = 0$ for all $n > 1$.

   (e) Combine (a)-(d) and determine a *bounded* sequence $[\gamma_n]_{n=1}^{\infty}$ such that

   $$\frac{\|Ax_n - b\|}{\|b\|} \leq \|V\|\|V^{-1}\|\gamma_n \frac{\rho^n}{|c|^n}$$

   for all $n > 0$.

   (f) What is the asymptotic penalty to have double eigenvalues in the sense of bounds, i.e., how much worse is the asymptotic convergence predicted by the bound in (e) in comparison to the prediction we derived for diagonalizable matrices when n is large?

Recall from the definition of the Jordan canonical form: If the eigenvalue λ_1 has one Jordan block of size two and all other Jordan blocks are of size one we have the following factorization. There exists an invertible matrix $V \in \mathbb{C}^{m\times m}$ and a matrix

$$\Lambda = \begin{pmatrix}\lambda_1 & 1 & & & \\ & \lambda_1 & 0 & & \\ & & \lambda_3 & \ddots & \\ & & & \ddots & 0\\ & & & & \lambda_m\end{pmatrix}.$$

such that $A = V\Lambda V^{-1}$.

Hint for (c): Show that the 2-norm of a block diagonal matrix is the maximum of the two-norm of the blocks by using the formula for the two-norm in terms of singular values.

From lectures: $P_n^0 = \{p \in P_n : p(0) = 1\}$ where P_n is the set of polynomials of degree less or equal to n.

Connection with current research: Researchers in numerical linear algebra are actively working on gaining further understanding of $\|p(A)\|$. The set $W(A) = \{\frac{x^*Ax}{x^*x} : x \in \mathbb{C}^m\}$ is called the field of values of A. It has been shown that $\|p(A)\| \leq \alpha \cdot \max_{z\in W(A)} |p(z)|$ holds for $\alpha = 11.08$. The open problem called Crouzeix's conjecture states that the bound holds for $\alpha = 2$. What happens if A is a Jordan block of size two with $\lambda_1 = \lambda_2 = 0$ and $p(z) = z$? Can you show that α cannot be smaller than 2? *Spoiler alert* See presentation at an important conference: http://sites.uclouvain.be/HHXIX/Plenaries/Overton.pdf