DELFT UNIVERSITY OF TECHNOLOGY
Faculty of Electrical Engineering, Mathematics and Computer Science

ANSWERS OF THE TEST SCIENTIFIC COMPUTING (wi4201)
Friday January 29 2016, 13:30-16:30

This are short answers, which indicate how the exercises can be answered. In most of the cases more details are needed to give a sufficiently clear answer.

1. (a) True because if symmetric then maximal absolute row sum of A is equal to maximal absolute column sum of A.

   (b) False because the matrices resulting from the discretization of the Poisson equation are counter examples.

   (c) True Since A is an M-matrix, we know that $A^{-1}_{i,j} \geq 0$ for all i and j, and $A_{i,j} \leq 0$ for $i \neq j$. We know

   $$0 < 1 = (AA^{-1})_{i,i} = A_{i,1}A^{-1}_{1,i} + ... + A_{i,n}A^{-1}_{n,i}$$

   Using the sign properties of A and A^{-1} shows that $A_{i,i} > 0$. Since the inverse of $diag(A)$ is given by the diagonal matrix with elements $\frac{1}{A_{i,i}}$, which are all positive, so $diag(A)$ is an M-matrix.

   (d) False because $\|A\|_2 = \sqrt{\lambda_{max}(A^TA)} = \sqrt{16} = 4$.

   (e) False, because the 2-norm is submultiplicative and therefore we have that for any real-valued squared matrix A we have that $1 = \|I\|_2 = \|AA^{-1}\|_2 \leq \|A\|_2\|A^{-1}\|_2 = \text{cond}_2(A)$.

2. (a) The stencil $[A^h]$ of the matrix is given by

   $$[A^h] = \frac{1}{h^2}\begin{bmatrix}-1 & 2 & -1\end{bmatrix} + \begin{bmatrix}0 & 1 & 0\end{bmatrix} = \frac{1}{h^2}\begin{bmatrix}-1 & 2+h^2 & -1\end{bmatrix} \quad (1)$$

   (b) The matrix denoted by A^h is of size 2×2 and given by

   $$A^h = \frac{1}{h^2}\begin{pmatrix}2+h^2 & -1 \\ -1 & 2+h^2\end{pmatrix} \quad (2)$$

   (c) Given that A^h is symmetric all eigenvalues are real-valued. So, it suffices to show that the eigenvalues of A^h are positive. The Gershgorin Theorem provides bounds on the eigenvalues that clearly show that the eigenvalues of A^h are positive. The expression for the eigenvalues derived in the previous exercise shows this as well.

(d) Show that the matrix A^h is an irreducible K-matrix and therefore an M-matrix. As a result of this the inverse of A^h only has positive entries.

(e) The condition number in 2-norm $\mathrm{cond}_2(A^h)$ of A^h is given by

$$\begin{aligned} \mathrm{cond}_2(A^h) &= \frac{\lambda_{max}(A^h)}{\lambda_{min}(A^h)} = \frac{\lambda^{N-1}(A^h)}{\lambda^1(A^h)} \\ &= \frac{2 - 2\cos[(N-1)\,\pi\, h] + h^2}{2 - 2\cos(\pi\, h) + h^2} \end{aligned} \tag{3}$$

(f) We have that

$$B^h_{JAC(w)} = I - w(D^h)^{-1}A^h\,. \tag{4}$$

(g) We have that

$$\begin{aligned} \rho(B^h_{JAC}) &= \rho[I - (D^h)^{-1}A^h] = 1 - \frac{h^2}{2+h^2}\lambda^1(A^h) \\ &= 1 - \frac{2 - 2\cos(\pi\, h) + h^2}{2+h^2} \\ &= \frac{2\cos(\pi\, h)}{2+h^2}\,. \end{aligned} \tag{5}$$

3. (a) We have that

$$\begin{aligned} \mathbf{u}^{k+1} &= \mathbf{u}^k + \mathbf{e}^k \\ &= \mathbf{u}^k + \widehat{A}^{-1}\mathbf{r}^k \\ &= \mathbf{u}^k + I^h_H[A^H]^{-1}I^H_h\mathbf{r}^k \end{aligned} \tag{6}$$

(b) We have that

$$A^h = \begin{pmatrix} A_{RR} & A_{RB} \\ A_{BR} & A_{BB} \end{pmatrix} \tag{7}$$

where A_{RR} and A_{BB} are diagonal matrices and where A_{RB} and A_{BR} are matrices with two diagonals below and above the main diagonal.

(c) The Gauss-Seidel iteration is such that during the k-th iteration the components of the residual vector $\mathbf{r}^k$ are consecutively made zero. In a red-black ordering this means that the residual vector in first the red nodes and then in the black nodes is made equal zero. The decoupling between the black nodes is such that making the residual equal zero in the second black node does not affect the residual in the first black node. After one red-black Gauss-Seidel sweep the residual in all the black nodes is equal to zero.

(d) We have that

$$\mathbf{u}^{k+1/2} = \mathbf{u}^k + M^{-1}_{GS}\mathbf{r}^k \text{ [Gauss-Seidel pre-smoothing step]} \tag{8}$$

and

$$\mathbf{u}^{k+1} = \mathbf{u}^{k+1/2} + I^h_H[A^H]^{-1}I^H_h\mathbf{r}^{k+1/2} \text{ [defect-correction step]} \tag{9}$$

4. (a) Note that $\|b - Au_1\|_2$ is minimal if $\|b - Au_1\|_2^2$ is minimal. Expanding $\|b - Au_1\|_2^2$ leads to:

$$\|b - Au_1\|_2^2 = (b - \alpha Ab)^T(b - \alpha Ab) = b^T b - 2\alpha b^T Ab + \alpha^2 (Ab)^T(Ab)$$

Take the derivative to α equal to zero:

$$2\alpha(Ab)^T(Ab) = 2b^T Ab$$

so the norm is minimized for $\alpha = \frac{b^T Ab}{(Ab)^T(Ab)}$.

(b) The Krylov-space of dimension K is $K^k(A; b) := \text{span}\left\{b, Ab, \ldots, A^{k-1}b\right\}$

(c) The minimization property of GMRES is:

$$\|r_k\|_2 = \|b - Au_k\|_2 = \min_{z \in K^k(A; r_0)} \|r_0 - Az\|_2. \tag{10}$$

After n iterations we know that the Krylov subspace is equal to $\mathbb{R}^n$ and thus the exact solution is an element of $K^n(A; r_0)$. Due to the minimization property this implies that $u_n = u$.

(d) In the algorithm the following vectors are stored: b, u, r, s_k, and v_k. This implies that $(2i + 3)n$ memory positions are needed. Per iteration one needs to do one matrix vector multiplication, $i + 1$ inner products, $2i$ vector updates, and 2 scalings. In total one matrix vector multiplication and $(6i + 4)n$ flops for the remaining parts.

5. (a) In order to solve the linear system $Au = b$ with LU-decomposition without pivotting, we do the following steps:

   - Find a lower triangular matrix L and an upper triangular matrix U, such that $LU = A$ and the diagonal elements of L are equal to 1.
   - Solve y from $Ly = b$.
   - Solve u from $Uu = y$.

   Since L and U are triangular matrices, this solution process is easy to implement. For the derivation of the costs see the lecture notes. The answer for a full matrix is for the decomposition the cost is $\frac{2}{3}n^3$ and for both solution steps together $2n^2$.

(b) If we do the multiplication:

$$(I - \alpha^{(k)}\mathbf{e}_k^T)(I + \alpha^{(k)}\mathbf{e}_k^T)$$

we obtain the following:

$$I - \alpha^{(k)}\mathbf{e}_k^T + \alpha^{(k)}\mathbf{e}_k^T + \alpha^{(k)}\mathbf{e}_k^T\alpha^{(k)}\mathbf{e}_k^T =$$

$$I + \alpha^{(k)}\mathbf{e}_k^T\alpha^{(k)}\mathbf{e}_k^T$$

Due to the zero structure of $\mathbf{e}_k$ and $\alpha^{(k)}$ the product $\mathbf{e}_k^T\alpha^{(k)}$ is equal to zero, so the last term is equal to zero, so

$$(I - \alpha^{(k)}\mathbf{e}_k^T)(I + \alpha^{(k)}\mathbf{e}_k^T) = I$$

which proves the claim that $M_k^{-1} = I + \alpha^{(k)}\mathbf{e}_k^T$.

(c) For the perturbation analysis see Section 4.3.1 of the lecture notes.

(d) From the construction of L and U it follows that there are only zeroes outside the band with bandwidth m. Within the band, elements which are zero in A become in general non-zero in L and U due to fill in.