Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4205 Numerical Linear Algebra Fall 2015

Solutions to exercise set 6

1 **a)** Given the data, the "naïve" basis for Krylov subspace is orthonormal; in fact it is the canonical basis in $\mathbb{R}^5$: $A^k r_0 = A^k b = A^k e_1 = e_{k+1}$, $k = 0,\dots,4$. Therefore $V_m = [e_1,\dots,e_m]$ after m steps of Arnoldi.
A direct computation shows that

$$\bar{H}_5 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

from which the submatrices $\bar{H}_k$, $k = 1,\dots,4$ are deducible. Note that $h_{6,5} = 0$, that is, Arnoldi process (understandably) "breaks" at this iteration, as $K_5(A, r_0)$ spans the whole space.

b) At iteration 1 we need to find a Givens rotation that "eliminates" the second component of the vector $(0,1)^{\mathrm{T}}$. Using e.g. formulas at the top of p. 168 in [Saad], we find

$$\Omega_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \text{and} \quad \bar{R}_1 = \Omega_1 \bar{H}_1 = \Omega_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{1}$$

The same story holds at iterations $k = 2,3,4$ with $\Omega_k = \Omega_1$ and $R_k = I_k$, $k \times k$ identity matrix.
$\bar{R}_5$ can be computed knowing the QR factorization of $\bar{H}_4$ as follows. First, we compute

$$\begin{pmatrix} Q_4 & 0 \\ 0 & 1 \end{pmatrix} \bar{H}_5 = \begin{pmatrix} Q_4 \bar{H}_4 & Q_4 e_1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \bar{R}_4 & e_5 \\ 0 & 0 \end{pmatrix},$$

where Q_4 being the result of successive application of $\Omega_1,\dots,\Omega_4$ to rows $(1,2),\dots,(4,5)$, e_i is the i-th canonical basis vector in $\mathbb{R}^5$, and we used the easily verifiable formula $Q_k e_1 = (-1)^k e_{k+1}$, $k = 1,\dots,4$. Since the resulting matrix is already upper triangular, we can simply take $\Omega_5 = I_{2\times 2}$, the 2×2 identity matrix. In particular, $R_5 = I_{5\times 5}$.

c) To solve the least squares problems appearing in GMRES we need to find $\bar{g}_k = Q_k(\beta e_1)$, where $\beta = \|r_0\|_2 = 1$. For $k = 1,\dots,4$ we get $\bar{g}_k = (-1)^k e_{k+1}$, and fpr $k = 5$ we get $\bar{g}_5 = e_5 \in \mathbb{R}^6$. Knowing that $R_k = I_{k\times k}$, $k = 1,\dots,5$ we compute $y_k = R_k^{-1} g_k$, which results in $y_k = 0$, $k = 1,\dots,4$ and $y_5 = e_5$. Finally we get $x_k = x_0 + V_k y_k$, resulting in $x_k = x_0$, $k = 1,\dots,4$, and $x_5 = e_5$.

2 Let H_{m-1}, $\bar{H}_{m-1}$, $V_{m-1} = [v_1,\dots,v_{m-1}]$, $V_m = [v_1,\dots,v_m]$ be the usual matrices obtained after $m-1$ steps of Arnoldi process starting from $v_1 = Av_0/\|Av_0\|_2$, where $v_0 = r_0$. We will also write $\bar{V}_m = [v_0, V_{m-1}]$.

Since V_{m-1} contains the orthonormal basis for $\mathscr{K}_{m-1}(A, Ar_0)$, the columns of $\bar{V}_m$ form a basis for $\mathscr{K}_m(A, r_0)$.

a) As in any residual-projection algorithm, in GMRES we are looking for an approximate solution in the form $x_m = x_0 + \bar{V}_m y_m$, where the unknowns $y_m \in \mathbb{R}^m$ are chosen in such a way as to minimize the 2-norm of the residual

$$\begin{aligned} r_m = r_0 - A\bar{V}_m y_m = r_0 - [Av_0, AV_{m-1}]y_m = r_0 - [\|Av_0\|_2 v_1, V_m \bar{H}_{m-1}]y_m \\ = r_0 - V_m[\|Av_0\|_2 e_1, \bar{H}_{m-1}]y_m. \end{aligned}$$

As a result

$$\|r_m\|_2 = \|(I - V_m V_m^{\mathrm{T}})r_0 + V_m(V_m^{\mathrm{T}} r_0 - [\|Av_0\|_2 e_1, \bar{H}_{m-1}]y_m)\|_2$$

Note that $V_m V_m^{\mathrm{T}}$ *an orthogonal projector* onto the subspace spanned by $[v_1, \dots, v_m]$ (see Section 1.12.3 in [Saad]) and $I - V_m V_m^{\mathrm{T}}$ is an orthogonal projector onto the orthogonal complement of the subspace spanned by $[v_1, \dots, v_m]$. Therefore, $(I - V_m V_m^{\mathrm{T}})r_0 \perp V_m(V_m^{\mathrm{T}} r_0 - [\|Av_0\|_2 e_1, \bar{H}_{m-1}]y_m)$ and

$$\begin{aligned} \|r_m\|_2^2 &= \|(I - V_m V_m^{\mathrm{T}})r_0\|_2^2 + \|V_m(V_m^{\mathrm{T}} r_0 - [\|Av_0\|_2 e_1, \bar{H}_{m-1}]y_m)\|_2^2 \\ &= \|(I - V_m V_m^{\mathrm{T}})r_0\|_2^2 + \|V_m^{\mathrm{T}} r_0 - [\|Av_0\|_2 e_1, \bar{H}_{m-1}]y_m\|_2^2. \end{aligned}$$

The first term is independent from y_m, whereas the second can be eliminated by solving a linear algebraic system with a triangular matrix $[\|Av_0\|_2 e_1, \bar{H}_{m-1}]y_m = V_m^{\mathrm{T}} r_0$ (recall: $\bar{H}_{m-1}$ is upper Hessenberg).

b) From the computations above we get that $r_m = (I - V_m V_m^{\mathrm{T}})r_0 \perp \operatorname{span}(v_1, \dots, v_m)$.

c) Owing to the same argument: $\|r_m\|_2 = \|(I - V_m V_m^{\mathrm{T}})r_0\|_2$, which is computable without the knowledge of y_m or x_m.

$\boxed{3}$ After applying m steps of Arnoldi process to a matrix A we obtain the matrix V_m containing the orthonormal basis for $\mathscr{K}_m(A, v_1)$ and an upper Hessenberg matrix H_m satisfying the equality $H_m = V_m^{\mathrm{T}} A V_m$. Assuming that $A^{\mathrm{T}} = -A$, the matrix on the right hand side of the equality sign is anti-symmetric. Therefore H_m is also antisymmetric and thus has only two non-zero diagonals:

$$H_m = \begin{pmatrix} 0 & -h_{2,1} & \dots & 0 \\ h_{2,1} & 0 & \ddots & 0 \\ 0 & \ddots & \ddots & -h_{m,m-1} \\ 0 & \dots & h_{m,m-1} & 0 \end{pmatrix} \tag{2}$$

As a result, Arnoldi process simplifies to (note: only the sub-diagonal of H is computed):

$v_1 := v/\|v\|_2$, $v_0 := 0$, $h_{1,0} := 0$
for j=1,...,m **do**
  $w_j := Av_j + h_{j,j-1} v_{j-1}$
  $h_{j+1,j} := \|w_j\|_2$
  if $h_{j+1,j} = 0$ **then stop**
  end if
  $v_{j+1} := w_j / h_{j+1,j}$

end for

4 The short story is: all residuals need to be scaled by δ, but the search space and the constraint space remain the same. Indeed, $\mathrm{span}\langle r_0, Ar_0, \dots, A^{m-1} r_0 \rangle = \mathrm{span}\langle \delta r_0, \delta^2 Ar_0, \dots, \delta^m A^{m-1} r_0 \rangle$. Furthermore, $r_m \perp \mathscr{L}$ if and only if $\delta r_m \perp \mathscr{L}$, since $\mathscr{L}$ is a linear space.