Numerical Linear Algebra: Solutions for Sheet 6

1. We just substitute $x = \cos\theta$ and obtain

$$2^{2-n-m}\langle T_n, T_m\rangle = 2^{2-n-m}\int_0^\pi \frac{\cos m\theta\cos n\theta}{\sin\theta}\sin\theta d\theta = 0, \qquad n \neq m.$$

2. If $h_{l+1,l} = 0 \Rightarrow Av_l - \sum_{j=1}^{l}(v_j^T A v_l)v_j = 0$, i.e.

$$A^l r_0 \in span\{v_1, \ldots, v_l\} = \mathcal{K}_l(A, r_0)$$

i.e.

$$A^l r_0 = \alpha_0 I r_0 + \alpha_1 A r_0 + \ldots + \alpha_{l-1}A^{l-1}r_0$$

$\Rightarrow$

$$A^{l-1}r_0 - \alpha_{l-1}A^{l-2}r_0 - \ldots - \alpha_1 I r_0 = \alpha_0 A^{-1}r_0.$$

Now $\alpha_0 \neq 0$ else we would previously have had $h_{l,l-1} = 0$ since if $\alpha_0 = 0$ we have

$$A^{l-1}r_0 = \alpha_1 I r_0 + \ldots + \alpha_{l-1}A^{l-2}r_0$$

so

$$A^{-1}r_0 = A^{-1}(b - Ax_0) = x - x_0 = \frac{1}{\alpha_0}(A^{l-1}r_0 - \alpha_{l-1}A^{l-2}r_0 - \ldots - \alpha_1 I r_0)$$

i.e.

$$x \in x_0 + \mathcal{K}_l(A, r_0)$$

hence $x = x_l$ as $|r_l|$ is minimal for $x_l \in x_0 + \mathcal{K}_l(A, r_0)$.

3. Consider

$$J_k^T J_{k-1}^T \cdots J_1^T \hat{H}_k = \begin{bmatrix} m_{11} & \cdots & \cdots & m_{1k} \\ 0 & m_{22} & & \vdots \\ & \ddots & \ddots & \vdots \\ & & 0 & m_{kk} \\ & & & 0 \end{bmatrix} = R_k$$

so for the next Hessenberg matrix in the Arnoldi process we get

$$J_k^T J_{k-1}^T \cdots J_1^T \hat{H}_{k+1} = \begin{bmatrix} m_{11} & \cdots & \cdots & m_{1k} & m_{1k+1} \\ 0 & m_{22} & & \vdots & m_{2k+1} \\ & \ddots & \ddots & \vdots & \vdots \\ & & 0 & m_{kk} & m_{kk+1} \\ & & & 0 & \hat{m}_{k+1k+1} \\ & & & & h_{k+2k+1} \end{bmatrix}$$

where

$$\begin{bmatrix} m_{1k+1} \\ m_{2k+1} \\ \vdots \\ m_{kk+1} \\ \hat{m}_{k+1k+1} \end{bmatrix} = J_k^T J_{k-1}^T \cdots J_1^T \begin{bmatrix} h_{1k+1} \\ h_{2k+1} \\ \vdots \\ h_{kk+1} \\ h_{k+1k+1} \end{bmatrix}$$

and we see that we only require

$$J_{k+1}^T = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & c & -s \\ & & & s & c \end{bmatrix}$$

with

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} \hat{m}_{k+1k+1} \\ h_{k+2k+1} \end{bmatrix} = \begin{bmatrix} m_{k+1k+1} \\ 0 \end{bmatrix}.$$

We now get that $\hat{H}_{k+1} = Q_{k+1}R_{k+1}$ with $Q_{k+1} = J_1J_2 \dots J_kJ_{k+1}$ and

$$R_{k+1} = \begin{bmatrix} m_{11} & \cdots & \cdots & m_{1k+1} \\ 0 & m_{22} & & \vdots \\ & \ddots & \ddots & \vdots \\ & & 0 & m_{k+1k+1} \\ & & & 0 \end{bmatrix}.$$

4. The linear least squares problem

$$y = argmin \left\| \|r_0\|e_1 - \hat{H}_k y \right\|_2 = argmin \left\| Q_k^T \|r_0\|e_1 - R_k y \right\|_2$$

has least squares error equal to the $(k+1)$ component of $Q_k^T\|r_0\|e_1$ since y is calculated to solve the leading $k \times k$ triangular system. For $k = 1$ we have

$$R_1 = \begin{bmatrix} * \\ 0 \end{bmatrix}$$

and

$$Q_1^T\|r_0\|e_1 = \|r_0\| \begin{bmatrix} c_1 & -s_1 \\ s_1 & c_1 \end{bmatrix} e_1 = \|r_0\| \begin{bmatrix} * \\ s_1 \end{bmatrix}$$

so the least squares error is $\|r_1\| = |s_1|\|r_0\| =$. For $k = 2$ we get

$$\begin{aligned} Q_2^T\|r_0\|e_1 = \|r_0\|J_2^TJ_1^Te_1 &= \|r_0\|J_2^T \begin{bmatrix} * \\ * \\ 0 \end{bmatrix} \\ &= \|r_0\| \begin{bmatrix} 1 & & \\ & c_2 & -s_2 \\ & s_2 & c_2 \end{bmatrix} \begin{bmatrix} * \\ s_1 \\ 0 \end{bmatrix} \\ &= \|r_0\| \begin{bmatrix} * \\ * \\ s_2s_1 \end{bmatrix} \end{aligned}$$

which gives $\|r_2\| = |s_2|\|r_1\|$. This continues inductively to $\|r_k\| = |s_k|\|r_{k-1}\|$.
Non-monotonicity: Consider $|s_k| = 1 \Rightarrow c_k = 0$ (as sin and cos of the same angle) and this happens when

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{m}_{k+1,k+1} \\ h_{j+2,k+1} \end{bmatrix} = \begin{bmatrix} m_{k+1,k+1} \\ 0 \end{bmatrix}$$

i.e. when $\hat{m}_{k+1,k+1} = 0$, i.e. if when we apply Q_k^T to the new column

$$\begin{bmatrix} h_{1,k+1} \\ h_{2,k+1} \\ \cdots \\ h_{k+2,k+1} \end{bmatrix}$$

of $\hat{H}_{k+1}$ we obtain a zero in the position of $h_{k+1,k+1}$. This can certainly happen: see Question 5.

5. Taking $x_0 = 0$ means we have:

$$r_0 = b = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad Ar_0 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad A^2 r_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad A^3 r_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad A^4 r_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = x$$

So we have for some α_i, β_i, $x_0 = 0$,

$$x_1 \in span\{r_0\} \Rightarrow x_1 = \begin{bmatrix} \alpha_1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \Rightarrow r_1 = b - Ax_1 = \begin{bmatrix} 1 \\ -\alpha_1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

but the GMRES algorithm minimises $\|r_1\|$ so we must have $\alpha_1 = 0$ and $x_1 = x_0$, $r_1 = r_0$. We can continue and apply this to x_2,

$$x_2 \in span\{r_0, Ar_0\} \Rightarrow x_2 = \begin{bmatrix} \alpha_2 \\ \beta_2 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \Rightarrow r_2 = b - Ax_2 = \begin{bmatrix} 1 \\ -\alpha_2 \\ -\beta_2 \\ 0 \\ 0 \end{bmatrix}$$

and again, since the GMRES algorithm minimises $\|r_2\|$ we must have $\alpha_2 = \beta_2 = 0$ and $x_2 = x_1$, $r_2 = r_1$. This process can be continued to show that $x_4 = x_3 = x_2 = x_1 = x_0$ and hence $r_4 = r_3 = r_2 = r_1 = r_0$. Finally we find that $r_5 = 0$ and $x_5 = x$. We therefore have a convergence graph as in figure 1.

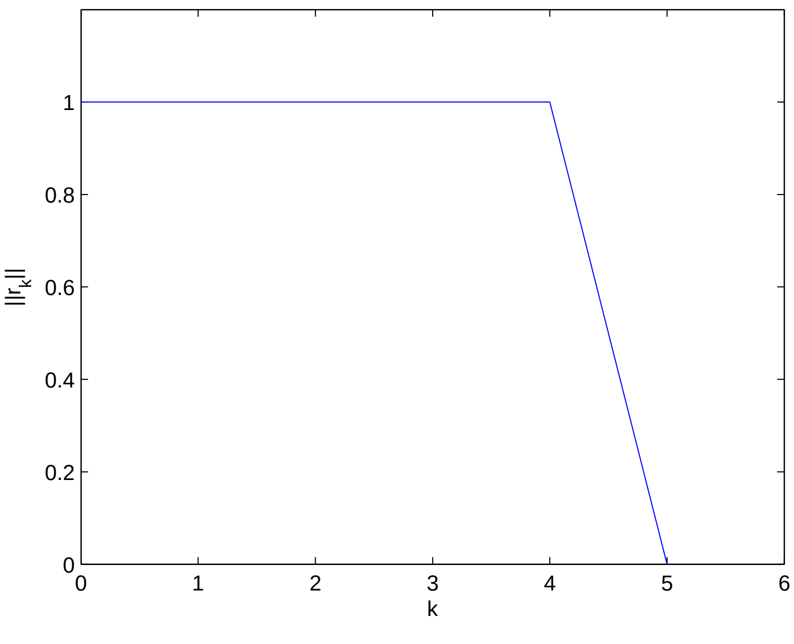

Figure 1: GMRES convergence

6. We have

$$A = \begin{bmatrix} I & B_2 & & & & \\ & I & B_3 & & & \\ & & \ddots & \ddots & & \\ & & & I & B_{k-1} & \\ & & & & I & B_k \\ & & & & & I \end{bmatrix}, \quad \text{so } (I - A) = \begin{bmatrix} 0 & -B_2 & & & & \\ & 0 & -B_3 & & & \\ & & \ddots & \ddots & & \\ & & & 0 & -B_{k-1} & \\ & & & & 0 & -B_k \\ & & & & & 0 \end{bmatrix}$$

so

$$(I - A)^2 = \begin{bmatrix} 0 & 0 & C_2 & & & \\ & 0 & 0 & C_3 & & \\ & & 0 & \ddots & \ddots & \\ & & & 0 & 0 & C_{k-1} \\ & & & & 0 & 0 \\ & & & & & 0 \end{bmatrix}, \quad \text{and } (I - A)^3 = \begin{bmatrix} 0 & 0 & 0 & D_2 & & & \\ & 0 & 0 & 0 & D_3 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & 0 & 0 & 0 & D_{k-2} \\ & & & & 0 & 0 & 0 \\ & & & & & 0 & 0 \\ & & & & & & 0 \end{bmatrix}$$

where $C_j = B_j B_{j+1}$ and $D_j = C_j C_{j+1}$. Inductively we have that $(I - A)^k = 0$.

If we now assume that l is the lowest power such that $(I - A)^l = 0$ then we have

$$\sum_{j=0}^{l} (-1)^j \begin{pmatrix} l \\ j \end{pmatrix} A^j = 0$$

and therefore

$$r_0 = -\sum_{j=1}^{l} (-1)^j \begin{pmatrix} l \\ j \end{pmatrix} A^j r_0$$

and since A^{-1} exists ($\det(A) = 1$) we have that

$$x - x_0 = A^{-1} r_0 = -\sum_{j=1}^{l} (-1)^j \begin{pmatrix} l \\ j \end{pmatrix} A^{j-1} r_0$$

ie $x - x_0 = q_{l-1}(A) r_0$ with $q_{l-1} \in \Pi_{l-1}$ and since the GMRES algorithm minimises the residual at the i^{th} stage amongst all $x_i \in x_0 + K_i(A, r_0)$, we must have $x_l = x$. In other words, (full) GMRES terminates in at most k iterations.

7. We use Matlab to investigate the GMRES algorithm using various different test matrices.

```
close all

for i=1:5
    switch i
    case 1
        A = rand(47);
        b = ones(47,1);
    case 2
        A = sprandn(100,100,0.1);
        b = ones(100,1);
    case 3
        A = sprandn(100,100,0.1) + 2*eye(100,100);
        b = ones(100,1);
    case 4
        A = sprandn(100,100,0.1) + 4*eye(100,100);
        b = ones(100,1);
    case 5
        X = randn(9,9); A = X*diag([1,1,-4,3,3,-4,-4,-4,3])/X;
        b = ones(9,1);
    end

    [x, flag, relres, iter, resvec] = gmres(A,b,[],1.e-6, size(A,1))

    subplot(3,2,i);
    plot(resvec)
    xlabel('Iteration Number')
    ylabel('Norm of Residual')
    title(['GMRES convergence - Q7 part ' int2str(i)]);
end
```

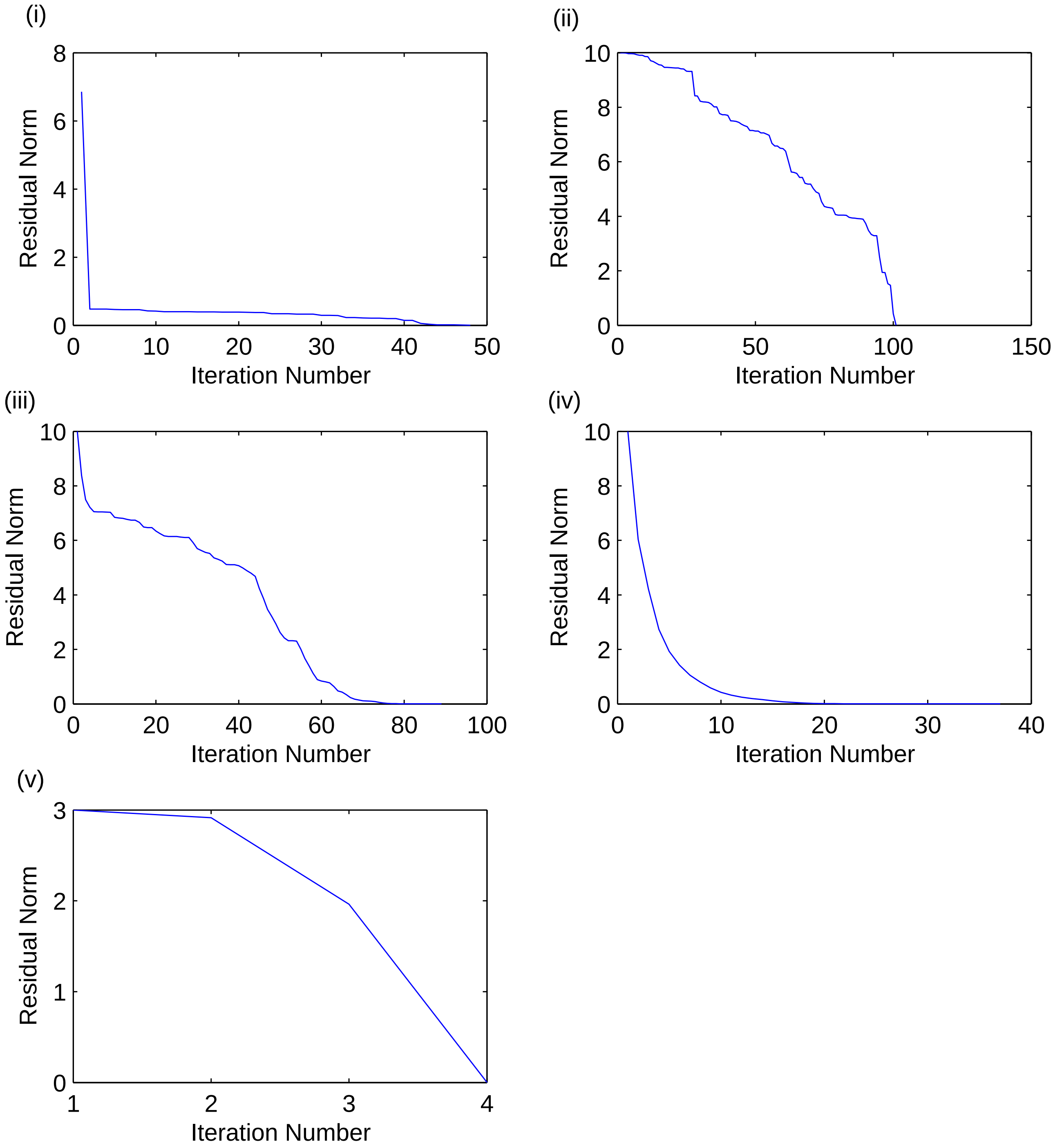

Figure 2: GMRES convergence for various matrices - question 7