Norwegian University of Science and Technology
Department of Mathematical Sciences

Solutions to exercise set 5

1 One can insert scaled variables into all parts of the algorithm and observe that the output at each step is the same. However, the short story is: all residuals need to be scaled by δ, but the search space and the constraint space remain the same. Indeed, $\text{span}\langle r_0, Ar_0, \ldots, A^{m-1}r_0\rangle = \text{span}\langle \delta r_0, \delta^2 Ar_0, \ldots, \delta^m A^{m-1}r_0\rangle$. Furthermore, $r_m \perp \mathscr{L}$ if and only if $\delta r_m \perp \mathscr{L}$, since $\mathscr{L}$ is a linear space.

2 Let H_{m-1}, $\bar{H}_{m-1}$, $V_{m-1} = [v_1, \ldots, v_{m-1}]$, $V_m = [v_1, \ldots, v_m]$ 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 $[v_1, \ldots, v_m]$ (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 $[v_1, \ldots, v_m]$. 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 $r_m = (I - V_m V_m^{\mathrm{T}})r_0 \perp \text{span}(v_1, \ldots, v_m)$.

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.