30 Convergence of Orthomin(1)

30.1 Numerical range

(See [Gre97, §1.3.6]).

Definition 30.1 (Numerical range and radius)**.** The numerical range (a.k.a., *field of values*) of a matrix $A \in \mathbb{C}^{m\times m}$ is the set of complex numbers

$$\mathcal{F}(A) = \left\{ \frac{y^\star A y}{y^\star y} \ : \ y \in \mathbb{C}^m, y \neq \mathbf{0} \right\}.$$

The *numerical radius* $\eta(A)$, is defined as

$$\eta(A) := \max\{|z| \ : \ z \in \mathcal{F}(A)\}.$$

Properties of the numerical range include that

$$\mathcal{F}(A + \alpha I) = \mathcal{F}(A) + \alpha, \qquad \text{and} \qquad \mathcal{F}(\alpha A) = \alpha \mathcal{F}(A).$$

30.2 Orthomin(1), again

Recall that the Orthomin(1) method (29.3)+(29.4) comes from setting

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + a_k(\mathbf{b} - A\mathbf{x}^{(k)}), \tag{30.1a}$$

where

$$a_k = \frac{(\mathbf{r}^{(k)}, A\mathbf{r}^{(k)})}{(A\mathbf{r}^{(k)}, A\mathbf{r}^{(k)})}. \tag{30.1b}$$

So (see [Gre97, §1.3.2]), since

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - a_k A \mathbf{r}^{(k)},$$

we can think of the method as setting $\mathbf{r}^{(k+1)}$ as $\mathbf{r}^{(k)}$ minus its projection on to $A\mathbf{r}^{(k)}$.

Theorem 30.2 (Thm. 2.2.1 of [Gre97])**.** *If Orthomin(1) method* (30.1) *generates the sequence* $\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \dots\}$*, and* $\mathbf{r}^{(0)}$*,* $\mathbf{r}^{(1)}$*, ... are the associated residuals, then* $\|\mathbf{r}^{(k+1)}\|_2 < \|\mathbf{r}^{(k)}\|_2$ *for any* $\mathbf{x}^{(0)}$*, if and only if* $0 \notin \mathcal{F}(A)$*.*

However, in practice, it is not enough to know that the residual decreases from step k to step $k+1$, we need to know by how much it decreases. This is given by:

Theorem 30.3 (Thm. 2.2.2 of [Gre97])**.** *The iteration* (30.1) *converges to* $A^{-1}b$ *for any initial guess* $\mathbf{x}^{(0)}$ *if* $0 \notin \mathcal{F}(A)$*, and*

$$\|\mathbf{r}^{(k+1)}\|_2 \leqslant \|\mathbf{r}^{(k)}\| \sqrt{1 - d^2/\|A\|_2^2},$$

where d *is the distance from the origin to* $\mathcal{F}(A)$*.*

Note: as observed during class (thanks, Olga!) the proof discussed was not complete: we also need that $d \leqslant \|A\|$*. This was addressed at the start of Lecture 31.*

30.3 Exercises

Exercise 30.1. Recall Definition 30.1.

(a) Show that if $A \in \mathbb{R}^{m\times m}$, then $\mathcal{F}(A) \subset \mathbb{R}$.

(b) Show that if $A \in \mathbb{C}^{m\times m}$, and $A^\star = A$, then $\mathcal{F}(A) \subset \mathbb{R}$.

Exercise 30.2. (a) Show that if λ is an eigenvalue of A, then $\lambda \in \mathcal{F}(A)$.

(b) Suppose that Q is unitary. Show that $\mathcal{F}(Q^\star A Q) = \mathcal{F}(A)$.