

Chapter 4

The Richardson Iteration

Remark 4.1. Motivation. The Richardson iteration by itself is not of that much interest in practice. However, it provides the idea for a tool that is then used for the construction of advanced iterative methods, namely the consideration of Krylov subspaces. \square

Definition 4.2. Richardson iteration. Let $\underline{x}^{(0)} \in \mathbb{R}^n$ be a given initial iterate. The Richardson¹ iteration for computing a sequence of vectors $\underline{x}^{(k)} \in \mathbb{R}^n, k = 0, 1, 2, \dots$, has the form

$$\underline{r}^{(k)} = \underline{b} - A\underline{x}^{(k)}, \quad \underline{x}^{(k+1)} = \underline{x}^{(k)} + \alpha_k \underline{r}^{(k)} \quad (4.1)$$

with appropriately chosen numbers $\alpha_k \in \mathbb{R}$. The vector $\underline{r}^{(k)}$ is called residual. A straightforward calculation shows that this method can be written as fixed-point iteration in the following form

$$\underline{x}^{(k+1)} = (I - \alpha_k A) \underline{x}^{(k)} + \alpha_k \underline{b}.$$

Notice that the iteration (4.1) is of the same principal type as the classical iteration schemes, e.g., compare (3.6) and (3.7), where the fixed matrices multiplying the residual vector in the classical iteration schemes are replaced with a scalar α_k that might change from step to step. \square

Definition 4.3. Co-domain of a matrix. The set

$$\mathcal{R}(A) = \left\{ \frac{\underline{y}^* A \underline{y}}{\underline{y}^* \underline{y}} : \underline{y} \in \mathbb{C}^n, \underline{y} \neq \underline{0} \right\} \subset \mathbb{C}$$

is called co-domain of A . \square

Remark 4.4. On the co-domain of a matrix. The co-domain of A can be defined by using only the vectors from the unit sphere of \mathbb{C}^n , since

$$\frac{\underline{y}^* A \underline{y}}{\underline{y}^* \underline{y}} = \frac{\underline{y}^* A \underline{y}}{\|\underline{y}^*\|_2 \|\underline{y}\|_2} = \underbrace{\frac{\underline{y}^*}{\|\underline{y}^*\|_2}}_{\|\cdot\|_2=1} A \underbrace{\frac{\underline{y}}{\|\underline{y}\|_2}}_{\|\cdot\|_2=1}.$$

The unit sphere is a compact set (bounded and closed) and the mapping $\underline{y} \mapsto \underline{y}^* A \underline{y} / \underline{y}^* \underline{y}$ is continuous. It follows that $\mathcal{R}(A)$ is also a compact set, see literature. \square

Lemma 4.5. Co-domain of the inverse matrix. Let $A \in \mathbb{R}^{n \times n}$ with $\mathcal{R}(A) \subset \{\lambda \in \mathbb{C} : \operatorname{Re}(\lambda) > 0\}$, i.e., the co-domain of A is a subset of the right half of the complex plane. Then

¹ Lewis Fry Richardson (1881 – 1953)

$$\mathcal{R}(A^{-1}) \subset \{\lambda \in \mathbb{C} : \operatorname{Re}(\lambda) > 0\}.$$

Proof. From the assumption, it follows that A is non-singular. Otherwise, there would be a vector $\underline{z} \in \ker(A)$, $\underline{z} \neq \underline{0}$, and

$$\operatorname{Re}\left(\frac{\overbrace{\underline{z}^* A \underline{z}}^{=0}}{\underline{z}^* \underline{z}}\right) = \operatorname{Re}(0) = 0.$$

This statement contradicts the assumption on $\mathcal{R}(A)$.

Let $\underline{y} \in \mathbb{C}^n$, $\underline{y} \neq \underline{0}$, be arbitrary and $\underline{z} = A^{-1}\underline{y} \neq \underline{0}$. Hence, \underline{z} is also an arbitrary vector. Using the definition of \underline{y} , that the real parts of a complex number and of its conjugate are the same, and $\|A\underline{z}\|_2 \leq \|A\|_2 \|\underline{z}\|_2$ yields

$$\begin{aligned} \operatorname{Re}\left(\frac{\underbrace{\underline{y}^* A^{-1} \underline{y}}_{\in \mathbb{R}}}{\|\underline{y}\|_2^2}\right) &= \frac{1}{\|\underline{y}\|_2^2} \operatorname{Re}(\underline{y}^* A^{-1} \underline{y}) = \frac{1}{\|A\underline{z}\|_2^2} \operatorname{Re}(\underbrace{(A\underline{z})^* A^{-1} A \underline{z}}_{=I}) \\ &= \frac{1}{\|A\underline{z}\|_2^2} \operatorname{Re}(\underbrace{\underline{z}^* A^* \underline{z}}_{\in \mathbb{C}}) = \frac{1}{\|A\underline{z}\|_2^2} \operatorname{Re}((\underline{z}^* A^* \underline{z})^*) = \frac{1}{\|A\underline{z}\|_2^2} \operatorname{Re}(\underline{z}^* A \underline{z}) \\ &= \frac{\|\underline{z}\|_2^2}{\|A\underline{z}\|_2^2} \operatorname{Re}\left(\frac{\underline{z}^* A \underline{z}}{\underline{z}^* \underline{z}}\right) \geq \frac{1}{\|A\|_2^2} \operatorname{Re}\left(\frac{\underline{z}^* A \underline{z}}{\underline{z}^* \underline{z}}\right) > 0, \end{aligned}$$

where the last statement follows from the assumption on $\mathcal{R}(A)$. ■

Theorem 4.6. Convergence of the Richardson iteration. *Let $A \in \mathbb{R}^{n \times n}$ with $\mathcal{R}(A) \subset \{\lambda \in \mathbb{C} : \operatorname{Re}(\lambda) > 0\}$. Then the Richardson iteration (4.1) converges to the solution of the linear system $A\underline{x} = \underline{b}$ for every initial iterate if $\alpha_k = \alpha$, $k = 0, 1, 2, \dots$, with*

$$0 < \alpha < \min \left\{ \beta = \operatorname{Re}(\lambda), \lambda \in \mathcal{R}(A^{-1}) \right\}. \quad (4.2)$$

Proof. Note that $\mathcal{R}(A^{-1})$ is a compact set such that the minimum in (4.2) exists. By Lemma 4.5, the minimum is positive such that a positive value α as given in (4.2) exists.

Let \underline{x} be the solution of (1.1). It will be shown that the error $\|\underline{x} - \underline{x}^{(k)}\|_2$ decreases strongly monotonically and the rate of decrease is strictly lower (uniformly with respect to k) than one. Using (4.1) and $\underline{b} = A\underline{x}$, one has the recursion

$$\begin{aligned} \underline{x} - \underline{x}^{(k+1)} &= \underline{x} - \underline{x}^{(k)} - \alpha \underline{r}^{(k)} = \underline{x} - \underline{x}^{(k)} - \alpha (\underline{b} - A\underline{x}^{(k)}) \\ &= \underline{x} - \underline{x}^{(k)} - \alpha A (\underline{x} - \underline{x}^{(k)}). \end{aligned}$$

Hence, it is

$$\begin{aligned} \|\underline{x} - \underline{x}^{(k+1)}\|_2^2 &= \left(\underline{x} - \underline{x}^{(k)} - \alpha A (\underline{x} - \underline{x}^{(k)}), \underline{x} - \underline{x}^{(k)} - \alpha A (\underline{x} - \underline{x}^{(k)}) \right) \\ &= \|\underline{x} - \underline{x}^{(k)}\|_2^2 - 2\alpha (\underline{x} - \underline{x}^{(k)})^T A (\underline{x} - \underline{x}^{(k)}) + \alpha^2 \|A (\underline{x} - \underline{x}^{(k)})\|_2^2. \end{aligned} \quad (4.3)$$

Denoting $\underline{y} = A (\underline{x} - \underline{x}^{(k)})$, using that the transposed of a real number is the same number and (4.2), one obtains

$$\begin{aligned} \frac{(\underline{x} - \underline{x}^{(k)})^T A (\underline{x} - \underline{x}^{(k)})}{\|A (\underline{x} - \underline{x}^{(k)})\|_2^2} &= \frac{(\underline{x} - \underline{x}^{(k)})^T A^T A^{-T} A (\underline{x} - \underline{x}^{(k)})}{\|A (\underline{x} - \underline{x}^{(k)})\|_2^2} = \frac{\overbrace{\underline{y}^T A^{-T} \underline{y}}^{\in \mathbb{R}}}{\underline{y}^T \underline{y}} \\ &= \frac{\underline{y}^T A^{-1} \underline{y}}{\underline{y}^T \underline{y}} \geq \min \left\{ \operatorname{Re}(\lambda) : \lambda \in \mathcal{R}(A^{-1}) \right\} > \alpha, \\ &\iff \\ \alpha^2 \|A (\underline{x} - \underline{x}^{(k)})\|_2^2 &< \alpha (\underline{x} - \underline{x}^{(k)})^T A (\underline{x} - \underline{x}^{(k)}). \end{aligned}$$

Applying this estimate to the last term of (4.3) yields

$$\begin{aligned} \|\underline{x} - \underline{x}^{(k+1)}\|_2^2 &< \|\underline{x} - \underline{x}^{(k)}\|_2^2 - \alpha (\underline{x} - \underline{x}^{(k)})^T A (\underline{x} - \underline{x}^{(k)}) \\ &= \|\underline{x} - \underline{x}^{(k)}\|_2^2 \left(1 - \alpha \frac{(\underline{x} - \underline{x}^{(k)})^T A (\underline{x} - \underline{x}^{(k)})}{\|\underline{x} - \underline{x}^{(k)}\|_2^2} \right). \end{aligned} \quad (4.4)$$

Since $\mathcal{R}(A)$ is compact, there is a $\varepsilon > 0$ such that $\operatorname{Re}(\lambda) \geq \varepsilon$ for all $\lambda \in \mathcal{R}(A)$ (there is no sequence that can converge to the imaginary axis). Hence, it holds that

$$\frac{(\underline{x} - \underline{x}^{(k)})^T A (\underline{x} - \underline{x}^{(k)})}{\|\underline{x} - \underline{x}^{(k)}\|_2^2} \geq \varepsilon.$$

Choosing ε such that $\alpha\varepsilon < 1$, then it follows from (4.4) that

$$\|\underline{x} - \underline{x}^{(k+1)}\|_2^2 < \|\underline{x} - \underline{x}^{(k)}\|_2^2 (1 - \alpha\varepsilon) =: q \|\underline{x} - \underline{x}^{(k)}\|_2^2$$

with q independent of k and $q \in (0, 1)$. One obtains by induction

$$\|\underline{x} - \underline{x}^{(k)}\|_2 \leq q^{k/2} \|\underline{x} - \underline{x}^{(0)}\|_2$$

such that $\underline{x}^{(k)} \rightarrow \underline{x}$ as $k \rightarrow \infty$. ■

Remark 4.7. Choice of α for s.p.d. matrices. Let A be symmetric and positive definite. Using the Rayleigh quotient (2.5) yields for an arbitrary vector $\underline{y} \in \mathbb{C}^n$

$$\begin{aligned} \frac{\operatorname{Re}(\underline{y}^* A^{-1} \underline{y})}{\|\underline{y}\|_2^2} &= \frac{1}{\|\underline{y}\|_2^2} \left((\operatorname{Re}(\underline{y}))^T A^{-1} \operatorname{Re}(\underline{y}) + (\operatorname{Im}(\underline{y}))^T A^{-1} \operatorname{Im}(\underline{y}) \right) \\ &= \frac{1}{\|\underline{y}\|_2^2} \left(\|\operatorname{Re}(\underline{y})\|_2^2 \frac{(\operatorname{Re}(\underline{y}))^T A^{-1} \operatorname{Re}(\underline{y})}{\|\operatorname{Re}(\underline{y})\|_2^2} \right. \\ &\quad \left. + \|\operatorname{Im}(\underline{y})\|_2^2 \frac{(\operatorname{Im}(\underline{y}))^T A^{-1} \operatorname{Im}(\underline{y})}{\|\operatorname{Im}(\underline{y})\|_2^2} \right) \\ &\geq \frac{1}{\|\underline{y}\|_2^2} \left(\|\operatorname{Re}(\underline{y})\|_2^2 \lambda_{\min}(A^{-1}) + \|\operatorname{Im}(\underline{y})\|_2^2 \lambda_{\min}(A^{-1}) \right) \\ &= \lambda_{\min}(A^{-1}) = \frac{1}{\lambda_{\max}(A)} = \frac{1}{\rho(A)}. \end{aligned}$$

That means, the choice $\alpha < 1/\rho(A)$ guarantees the convergence of the Richardson method. □

Remark 4.8. Residual minimization for choosing α_k . One possibility to choose α_k in practice consists in the minimization of the norm of the residual

$$\begin{aligned} \|\underline{r}^{(k+1)}\|_2^2 &= \|\underline{b} - A\underline{x}^{(k+1)}\|_2^2 = \|\underline{b} - A\underline{x}^{(k)} - \alpha_k A\underline{r}^{(k)}\|_2^2 = \|\underline{r}^{(k)} - \alpha_k A\underline{r}^{(k)}\|_2^2 \\ &= \|\underline{r}^{(k)}\|_2^2 - \alpha_k (\underline{r}^{(k)})^T A \underline{r}^{(k)} - \alpha_k \underbrace{(A\underline{r}^{(k)})^T \underline{r}^{(k)}}_{\in \mathbb{R}} + \alpha_k^2 \|A\underline{r}^{(k)}\|_2^2 \\ &= \|\underline{r}^{(k)}\|_2^2 - 2\alpha_k (\underline{r}^{(k)})^T A \underline{r}^{(k)} + \alpha_k^2 \|A\underline{r}^{(k)}\|_2^2. \end{aligned}$$

The necessary condition for a minimum

$$\frac{d}{d\alpha_k} \left\| \underline{r}^{(k+1)} \right\|_2^2 = -2 \left(\underline{r}^{(k)} \right)^T A \underline{r}^{(k)} + 2\alpha_k \left\| A \underline{r}^{(k)} \right\|_2^2 = 0$$

gives

$$\alpha_k = \frac{\left(\underline{r}^{(k)} \right)^T A \underline{r}^{(k)}}{\left\| A \underline{r}^{(k)} \right\|_2^2}. \quad (4.5)$$

Since

$$\frac{d^2}{d\alpha_k^2} \left\| \underline{r}^{(k+1)} \right\|_2^2 = 2 \left\| A \underline{r}^{(k)} \right\|_2^2 > 0,$$

if $\underline{r}^{(k)} \neq \underline{0}$, one obtains in fact a minimum. □

Remark 4.9. Spaces spanned by the iterates. It is by (4.1)

$$\begin{aligned} \underline{x}^{(1)} &\in \underline{x}^{(0)} + \text{span} \left\{ \underline{r}^{(0)} \right\}, \\ \underline{x}^{(2)} &\in \underline{x}^{(1)} + \text{span} \left\{ \underline{r}^{(1)} \right\} = \underline{x}^{(0)} + \text{span} \left\{ \underline{r}^{(0)}, \underline{r}^{(1)} \right\}. \end{aligned}$$

It holds

$$\underline{r}^{(1)} = \underline{b} - A \underline{x}^{(1)} = \underline{b} - A \underline{x}^{(0)} - \alpha_0 A \underline{r}^{(0)} = \underline{r}^{(0)} - \alpha_0 A \underline{r}^{(0)}$$

and consequently

$$\underline{x}^{(2)} \in \underline{x}^{(0)} + \text{span} \left\{ \underline{r}^{(0)}, A \underline{r}^{(0)} \right\}.$$

One obtains by induction

$$\underline{x}^{(k)} \in \underline{x}^{(0)} + \text{span} \left\{ \underline{r}^{(0)}, A \underline{r}^{(0)}, \dots, A^{k-1} \underline{r}^{(0)} \right\}.$$

□

Definition 4.10. Krylov subspace. Let $\underline{q} \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$. Then, the space

$$K_m(\underline{q}, A) := \text{span} \left\{ \underline{q}, A \underline{q}, \dots, A^{m-1} \underline{q} \right\}$$

is called the Krylov² subspace of order m that is spanned by \underline{q} and A . □

Remark 4.11. Next goal. It holds that $\underline{x}^{(k)} \in \underline{x}^{(0)} + K_k \left(\underline{r}^{(0)}, A \right)$. In the following, Richardson's method will be improved by constructing the iterates $\underline{x}^{(k)}$ in this manifold with respect to certain optimality criteria. □

² Aleksey Nikolaevich Krylov (1863 – 1945)