Chapter 4

The Richardson Iteration

Definition 4.1 Richardson iteration. Let $\mathbf{x}^{(0)} \in \mathbb{R}^n$ be a given initial iterate. The Richardson[1] iteration for computing a sequence of vectors $\mathbf{x}^{(k)} \in \mathbb{R}^n, k = 0, 1, 2, \ldots$, has the form

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

with appropriately chosen numbers $\alpha_k \in \mathbb{R}$. □

Definition 4.2 Co-domain of a matrix. The set

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

is called co-domain of A. □

Remark 4.3 *On the co-domain of a matrix.* The co-domain of A is the co-domain of the unit sphere of $\mathbb{C}^n$ since

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

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

Lemma 4.4 Co-domain of the inverse matrix. *Let $A \in \mathbb{R}^{n \times n}$ with $\mathcal{R}(A) \subset \{\lambda \in \mathbb{C} \ : \ Re(\lambda) > 0\}$, i.e., the co-domain of A is a subset of the right half of the complex plane. Then*

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

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

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

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

[1] Lewis Fry Richardson (1881 – 1953)

Let $\mathbf{y} \in \mathbb{C}^n$, $\mathbf{y} \neq \mathbf{0}$, be arbitrary and $\mathbf{z} = A^{-1}\mathbf{y} \neq \mathbf{0}$. Hence, $\mathbf{z}$ is also an arbitrary vector. One has

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

where $\|A\mathbf{z}\|_2 \leq \|A\|_2\|\mathbf{z}\|_2$ has been used. ■

Theorem 4.5 Convergence of the Richardson iteration. *Let $A \in \mathbb{R}^{n\times n}$ with $\mathcal{R}(A) \subset \{\lambda \in \mathbb{C} : Re(\lambda) > 0\}$. Then the Richardson iteration (4.1) converges to the solution of the linear system $A\mathbf{x} = \mathbf{b}$ for every initial iterate if $\alpha_k = \alpha$, $k = 0, 1, 2, \ldots$, with*

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

Proof: Note that $\mathcal{R}\left(A^{-1}\right)$ is a compact set such that the minimum exists. Let $\mathbf{x}$ be the solution of (1.1). It will be shown that the error $\left\|\mathbf{x} - \mathbf{x}^{(k)}\right\|_2$ decreases strongly monotonically and the rate of decrease is strictly lower than one. Using $\mathbf{b} = A\mathbf{x}$ and (4.1), one has the recursion

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

Hence,

$$\begin{aligned}
\left\|\mathbf{x} - \mathbf{x}^{(k+1)}\right\|_2^2 &= \left(\mathbf{x} - \mathbf{x}^{(k)} - \alpha A\left(\mathbf{x} - \mathbf{x}^{(k)}\right), \mathbf{x} - \mathbf{x}^{(k)} - \alpha A\left(\mathbf{x} - \mathbf{x}^{(k)}\right)\right) \qquad (4.2)\\
&= \left\|\mathbf{x} - \mathbf{x}^{(k)}\right\|_2^2 - 2\alpha\left(\mathbf{x} - \mathbf{x}^{(k)}\right)^T A\left(\mathbf{x} - \mathbf{x}^{(k)}\right) + \alpha^2\left\|A\left(\mathbf{x} - \mathbf{x}^{(k)}\right)\right\|_2^2.
\end{aligned}$$

Denoting $\mathbf{y} = A\left(\mathbf{x} - \mathbf{x}^{(k)}\right)$, one obtains

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

Applying this estimate to the last term of (4.2) gives

$$\begin{aligned}
\left\|\mathbf{x} - \mathbf{x}^{(k+1)}\right\|_2^2 &\leq \left\|\mathbf{x} - \mathbf{x}^{(k)}\right\|_2^2 - \alpha\left(\mathbf{x} - \mathbf{x}^{(k)}\right)^T A\left(\mathbf{x} - \mathbf{x}^{(k)}\right) \\
&= \left\|\mathbf{x} - \mathbf{x}^{(k)}\right\|_2^2\left(1 - \alpha\frac{\left(\mathbf{x} - \mathbf{x}^{(k)}\right)^T A\left(\mathbf{x} - \mathbf{x}^{(k)}\right)}{\|\mathbf{x} - \mathbf{x}^{(k)}\|_2^2}\right). \qquad (4.3)
\end{aligned}$$

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

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

Choose ε such that $\alpha\varepsilon \le 1$, then it follows from (4.3) that

$$\left\|\mathbf{x} - \mathbf{x}^{(k+1)}\right\|_2^2 \le \left\|\mathbf{x} - \mathbf{x}^{(k)}\right\|_2^2 (1 - \alpha\varepsilon) =: q^2 \left\|\mathbf{x} - \mathbf{x}^{(k)}\right\|_2^2$$

with $0 < q < 1$ independent of k. One obtains by induction

$$\left\|\mathbf{x} - \mathbf{x}^{(k)}\right\|_2 \le q^k \left\|\mathbf{x} - \mathbf{x}^{(0)}\right\|_2$$

such that $\mathbf{x}^{(k)} \to \mathbf{x}$ as $k \to \infty$. ■

Remark 4.6 *Choice of α for s.p.d. matrices.* Let A be symmetric and positive definite. Using Rayleigh's coefficient (2.3), one gets

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

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

Remark 4.7 *Residual minimzation for choosing α_k.* One possibility to choose α_k in practice consists in the minimization of the norm of the residual

$$\begin{aligned}
\left\|\mathbf{r}^{(k+1)}\right\|_2^2 &= \left\|\mathbf{b} - A\mathbf{x}^{(k+1)}\right\|_2^2 = \left\|\mathbf{b} - A\mathbf{x}^{(k)} - \alpha_k A\mathbf{r}^{(k)}\right\|_2^2 = \left\|\mathbf{r}^{(k)} - \alpha_k A\mathbf{r}^{(k)}\right\|_2^2 \\
&= \left\|\mathbf{r}^{(k)}\right\|_2^2 - 2\alpha_k \mathbf{r}^{(k)T} A r^{(k)} + \alpha_k^2 \left\|A\mathbf{r}^{(k)}\right\|_2^2.
\end{aligned}$$

The necessary condition for a minimum

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

gives

$$\alpha_k = \frac{\mathbf{r}^{(k)T} A\mathbf{r}^{(k)}}{\left\|A\mathbf{r}^{(k)}\right\|_2^2}. \tag{4.4}$$

Since

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

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

Remark 4.8 *Spaces spanned by the iterates.* It is by (4.1)

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

It holds

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

and consequently

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

One obtains by induction

$$\mathbf{x}^{(k)} \in \mathbf{x}^{(0)} + \operatorname{span}\left\{\mathbf{r}^{(0)}, A\mathbf{r}^{(0)}, \ldots, A^{k-1}\mathbf{r}^{(0)}\right\}.$$

□

Definition 4.9 Krylov subspace. Let $\mathbf{q} \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$. Then, the space

$$K_m(\mathbf{q}, A) := \operatorname{span}\left\{\mathbf{q}, A\mathbf{q}, \ldots, A^{k-1}\mathbf{q}\right\}$$

is called the Krylov[2] subspace of order m which is spanned by $\mathbf{q}$ and A. □

Remark 4.10 *Next goal.* It holds $\mathbf{x}^{(k)} \in \mathbf{x}^{(0)} + K_k\left(\mathbf{r}^{(0)}, A\right)$. In the following, Richardson's method will be generalized by constructing the iterates $\mathbf{x}^{(k)}$ in this manifold with respect to certain optimality criteria. □

[2] Krylov