Chapter 3
Classical Iterative Schemes

3.1 General Theory

Remark 3.1. Basic idea, transform to a fixed-point equation. The construction of a classical iterative scheme for solving (1.1) starts with the decomposition

$$A = M - N, \quad M, N \in \mathbb{R}^{n\times n}, \quad M \text{ is non-singular},$$

of the system matrix A. Using this decomposition, (1.1) can be transformed into the fixed-point equation

$$M\underline{x} = \underline{b} + N\underline{x} \quad \Longleftrightarrow \quad \underline{x} = M^{-1}\left(\underline{b} + N\underline{x}\right). \tag{3.1}$$

Given an initial iterate $\underline{x}^{(0)} \in \mathbb{R}^n$, one can try to solve (3.1) with the fixed-point iteration

$$\underline{x}^{(k+1)} = M^{-1}\left(\underline{b} + N\underline{x}^{(k)}\right), \quad k = 0, 1, 2, \ldots. \tag{3.2}$$

Banach's[1] fixed-point theorem gives information on the convergence of this iteration. □

Theorem 3.2. Banach's fixed-point theorem. *Let $(\mathcal{X}, d)$ be a complete metric space and let $f : \mathcal{X} \to \mathcal{X}$ be a contraction (f is Lipschitz[2] continuous with the Lipschitz constant $L < 1$). Then, the equation $x = f(x)$ possesses a unique solution $\hat{x} \in \mathcal{X}$ (a fixed point). The iterative scheme*

$$x^{(k+1)} = f\left(x^{(k)}\right), \quad k = 0, 1, 2, \ldots$$

converges to $\hat{x}$ for any initial iterate $x^{(0)} \in \mathcal{X}$.

Proof. Basic course on calculus. ■

Theorem 3.3. Condition on the iteration matrix of (3.2) for convergence. *The iterative scheme (3.2) converges to the solution $\underline{x}$ of (1.1) for any initial iterate $\underline{x}^{(0)}$ if and only if the spectral radius of the iteration matrix $G = M^{-1}N$ is smaller than one: $\rho(G) < 1$.*

Proof. i) Lipschitz continuity. The iteration (3.2) is a fixed-point iteration with

$$f : \mathbb{R}^n \to \mathbb{R}^n, \quad \underline{x} \mapsto M^{-1}N\underline{x} + M^{-1}\underline{b}.$$

Let $\mathbb{R}^n$ be equipped with the vector norm $\|\cdot\|_*$ defined in (2.2) with sufficiently small δ. The operator $G = M^{-1}N$ is linear and bounded since $\|G\|_*$ is finite. Hence, G is continuous and even Lipschitz continuous. Since f is continuously differentiable, the Lipschitz constant is given by

---
[1] Stefan Banach (1892 – 1945)
[2] Rudolf Lipschitz (1832 – 1903)

$$L_* = \sup_{\underline{x}\in\mathbb{R}^n} \|J(f(\underline{x}))\|_* = \sup_{\underline{x}\in\mathbb{R}^n} \|G\|_* = \|G\|_*,$$

where $J(f(\underline{x}))$ is the (constant) Jacobian of $f(\underline{x})$.

ii) Convergence for $\rho(G) < 1$. Let $\rho(G) < 1$. Then, it is possible to find a matrix norm $\|\cdot\|_*$ such that, according to Lemma 2.8, $\|G\|_* \le \rho(G) + \varepsilon < 1$ with $\varepsilon > 0$. Hence, $L_* < 1$ and $f(\underline{x})$ is a contraction. Now, the statement follows with Theorem 3.2.

iii) $\rho(G) \ge 1$. Let $\rho(G) \ge 1$. An initial guess will be constructed for which the fixed-point iteration does not converge. Without loss of generality, consider the case $\underline{b} = \underline{0}$ such that the solution of (1.1) is $\underline{x} = \underline{0}$.

Since $\rho(G) \ge 1$, there is an eigenvalue $\lambda \in \mathbb{C}$ of G with $|\lambda| \ge 1$. This eigenvalue can be written in the form

$$\lambda = |\lambda|(\cos(\varphi) + i\sin(\varphi)), \tag{3.3}$$

where φ is the argument of λ. Let $\underline{z} \in \mathbb{C}^n, \underline{z} \ne \underline{0}$, be a corresponding eigenvector:

$$G\underline{z} = \lambda\underline{z}. \tag{3.4}$$

From the conjugate of this equation $\overline{G\underline{z}} = \overline{\lambda\underline{z}}$, it follows that $G\overline{\underline{z}} = \overline{\lambda}\overline{\underline{z}}$ since G is a real matrix.

Choose the initial iterate $\underline{x}^{(0)} = \underline{z} + \overline{\underline{z}} \in \mathbb{R}^n$.

One has to exclude that $\underline{x}^{(0)} = \underline{0}$. If $\underline{x}^{(0)} = \underline{0}$, then $\underline{z} = i\underline{v}$ with $\underline{v} \in \mathbb{R}^n$. One obtains from the eigenvalue equation that $iG\underline{v} = i\lambda\underline{v}$ which is equivalent to $G\underline{v} = \lambda\underline{v}$. On the left-hand side of this equation, there is a real vector. Since $\underline{v}$ is a real vector, it follows that λ must be real, too. But in this case, the corresponding eigenvector $\underline{z}$ is also real and it cannot be of the form $\underline{z} = i\underline{v}$. Hence, an eigenvector of form $\underline{z} = i\underline{v}$ cannot exist and $\underline{x}^{(0)} \ne \underline{0}$.

Using the definition of the initial iterate, the eigenvalue problem (3.4), and some basic properties of eigenvalues and complex numbers, it follows that

$$\underline{x}^{(k)} = \underbrace{G\left(G\left(\cdots G\underline{x}^{(0)}\right)\right)}_{k \text{ times}} = G^k\underline{x}^{(0)} = G^k\underline{z} + G^k\overline{\underline{z}} = \lambda^k\underline{z} + \overline{\lambda}^k\overline{\underline{z}} = 2\mathrm{Re}\left(\lambda^k\underline{z}\right), \ k = 0, 1, \ldots.$$

The iteration converges to the solution $\underline{x} = \underline{0}$ if

$$\begin{aligned}\underline{0} &= \lim_{k\to\infty} 2\mathrm{Re}\left(\lambda^k\underline{z}\right) = \lim_{k\to\infty} 2|\lambda|^k \,\mathrm{Re}\Big((\cos(k\varphi) + i\sin(k\varphi))\,\underline{z}\Big) \\ &= \lim_{k\to\infty} 2|\lambda|^k (\cos(k\varphi)\,\mathrm{Re}(\underline{z}) - \sin(k\varphi)\,\mathrm{Im}(\underline{z})),\end{aligned}$$

where (3.3) and basic properties of the real part of complex numbers were used. The factor $|\lambda|^k$ is always larger or equal to 1, since $|\lambda| \ge 1$. Hence, the second factor has to converge to zero if the iteration should converge to $\underline{x} = \underline{0}$. Note that the second factor is a vector. It converges to zero if and only if each of its components converges to zero. There is at least one component z_l with $z_l \ne 0$, since $\underline{z}$ is an eigenvector. Let ζ be the argument of z_l, i.e., the angle in the complex plane. Using the definition of the real and imaginary part of z_l and the trigonometric identity for $\cos(\alpha + \beta)$ yields

$$\begin{aligned}\cos(k\varphi)\,\mathrm{Re}(z_l) - \sin(k\varphi)\,\mathrm{Im}(z_l) &= |z_l|(\cos(k\varphi)\cos(\zeta) - \sin(k\varphi)\sin(\zeta)) \\ &= |z_l|\cos(k\varphi + \zeta).\end{aligned}$$

The only possibility to obtain convergence to zero for $k \to \infty$ for the periodic cosine function and for fixed increment φ is the case $\zeta = \pm\pi/2$ and $\varphi \in \{0, \pi\}$, because then $k\varphi + \zeta$ is $\pi/2$ plus an integer multiple of π, so that the term with the cosine vanishes. Since $\zeta = \pm\pi/2$, it follows that $\mathrm{Re}(z_l) = 0$ and $\lambda \in \mathbb{R}$. However, if $\lambda \in \mathbb{R}$, then the eigenvector $\underline{z}$ is real, too, which is a contradiction to $\mathrm{Re}(z_l) = 0$.

In summary, the iteration (3.2) does not converge for the initial iterate $\underline{x}^{(0)} = \underline{z} + \overline{\underline{z}}$. That means, if the iteration (3.2) converges for all initial iterates, then $\rho(G) \ge 1$ cannot hold. ∎

Remark 3.4. Complex-valued systems. The last part of the proof simplifies much if one considers complex-valued systems of linear equations. Then, one can take the initial iterate $\underline{x}^{(0)} = \underline{z} \ne \underline{0}$, finds that $\underline{x}^{(k)} = \lambda^k\underline{z}$, and concludes that $\left\|\underline{x}^{(k)}\right\|_2 = |\lambda|^k \|\underline{z}\|_2 \not\to 0$, since $\|\underline{z}\|_2 \ne 0$ and $|\lambda|^k \ge 1$.

However, if a real-valued system is given, computations in practice are performed usually only with real-valued vectors. Thus, the goal of the proof was to show that there is a real-valued initial iterate for which the iteration does not converge. Showing that there is a complex-valued initial iteration for which the iteration does not converge does not exclude that it converges for all real-valued initial iterates. In particular, the real-valued situation is not a special case of the complex-valued case. □