Chapter 1

Introduction

Remark 1.1 *Contents of the lecture notes.* These lecture notes consider solvers for a linear system of equation

$$A\mathbf{x} = \mathbf{b}, \quad A \in \mathbb{R}^{n\times n}, \mathbf{x}, \mathbf{b} \in \mathbb{R}^n \tag{1.1}$$

with non-singular matrix A. The solution of such systems is the core of many algorithms.

In particular, systems with the following features will be considered in these notes:

- the dimension n of the systems is very large,
- the system matrix A is sparse, i.e. the number of non-zero entries in A is only a small percentage, usually $\mathcal{O}(n)$, of the total number of entries that is n^2.

Systems with these features arise, e.g., in the discretization of partial differential equations.

Throughout the lecture notes, vectors are denoted by small bold-faced letters, components of vectors by small letters, matrices by capital letters, scalars by Greek letters, and indices by the letters i, j, l, m. The iteration index in iterative scheme is denoted by k.

Main parts of these lecture notes follow Starke (2001). □

Chapter 2

Some Basics on Vectors and Matrices

Remark 2.1 *Contents.* This chapter gives an overview on vector and matrix properties which will be used in these lecture notes. □

Remark 2.2 *Norms of vectors.* Let $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ be a vector. The l^p-norm is defined by

$$\begin{aligned} \|\mathbf{x}\|_p &:= \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}, \quad p \in [1,\infty), \\ \|\mathbf{x}\|_\infty &:= \max_{i=1,\ldots,n} |x_i|. \end{aligned}$$

If $p = 1$, the norm is called sum norm, in the case $p = 2$ one speaks of the Euclidean norm and for $p = \infty$ of the maximum norm. □

Remark 2.3 *Norms of matrices.* Let

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

The induced matrix p-norm is defined by

$$\|A\|_p := \max_{\mathbf{x}\in\mathbb{R}^n, \mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \max_{\mathbf{x}\in\mathbb{R}^n, \|\mathbf{x}\|_p \le 1} \|A\mathbf{x}\|_p = \max_{\mathbf{x}\in\mathbb{R}^n, \|\mathbf{x}\|_p = 1} \|A\mathbf{x}\|_p. \tag{2.1}$$

Special cases are

$$\begin{aligned} \|A\|_1 &= \max_{j=1,\ldots,n} \sum_{i=1}^n |a_{ij}|, \quad \text{maximum absolute column sum norm,} \\ \|A\|_2 &= \left(\lambda_{\max}\left(A^T A\right)\right)^{1/2}, \quad \text{Euclidean norm, spectral norm,} \\ \|A\|_\infty &= \max_{i=1,\ldots,n} \sum_{j=1}^n |a_{ij}|, \quad \text{maximum absolute row sum norm.} \end{aligned}$$

Another norm is the Frobenius[1] norm given by

$$\|A\|_F = \left(\sum_{i,j=1}^n |a_{ij}|^2\right)^{1/2}.$$

[1] Ferdinand Georg Frobenius (1849 - 1917)

□

Remark 2.4 *Properties of matrix norms.* From (2.1) it follows immediately for all $\mathbf{x} \in \mathbb{R}^n$ that

$$\|A\|_p \geq \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} \quad \Longleftrightarrow \quad \|A\mathbf{x}\|_p \leq \|A\|_p \|\mathbf{x}\|_p .$$

It holds also $\|A\mathbf{x}\|_2 \leq \|A\|_F \|\mathbf{x}\|_2$. By induction it follows for $B \in \mathbb{R}^{n\times n}$ that

$$\begin{aligned} \|AB\mathbf{x}\|_p &\leq \|A\|_p \|B\mathbf{x}\|_p \leq \|A\|_p \|B\|_p \|\mathbf{x}\|_p \quad \Longleftrightarrow \\ \|AB\|_p &= \max_{\mathbf{x}\in\mathbb{R}^n, \mathbf{x}\neq 0} \frac{\|AB\mathbf{x}\|_p}{\|\mathbf{x}\|_p} \leq \|A\|_p \|B\|_p . \end{aligned}$$

□

Lemma 2.5 Properties of non-singular quadratic matrices. *Let* $A \in \mathbb{R}^{n\times n}$. *The following properties are equivalent:*
- *A is non-singular.*
- *The inverse A^{-1} of A exists.*
- *The linear system* (1.1) *possesses for each right hand side* **b** *a unique solution.*
- *The determinant of A does not vanish:* $\det(A) \neq 0$.
- *All eigenvalues of A are different from zero.*

Proof: This lemma was proved in the course on basic linear algebra. ■

Definition 2.6 Eigenvalues, eigenvectors, spectral radius. A complex number $\lambda \in \mathbb{C}$ is called eigenvalue of $A \in \mathbb{C}^{n\times n}$ if there is a vector $\mathbf{x} \in \mathbb{C}^n, \mathbf{x} \neq \mathbf{0}$, such that

$$A\mathbf{x} = \lambda\mathbf{x}.$$

The vector $\mathbf{x}$ is called eigenvector. Note that all real (complex) eigenvalues will be associated to real (complex) eigenvectors.

The spectral radius of a matrix A is defined by

$$\rho(A) = \max\{|\lambda| \ : \ \lambda \text{ is eigenvalue of A}\}.$$

□

Remark 2.7 *On eigenvalues.* For every eigenvalue $\lambda_j \in \mathbb{C}$ of A it holds $|\lambda_j| \leq \|A\|$ for any matrix norm which is given in Remark 2.3. It follows that $\rho(A) \leq \|A\|$. □

Lemma 2.8 Existence of a matrix norm that is arbitrarily close to the spectral radius. *Let* $A \in \mathbb{R}^{n\times n}$ *and* $\varepsilon > 0$ *be given. Then, there is a vector norm* $\|\cdot\|_*$ *such that for the induced matrix norm it holds*

$$\rho(A) \leq \|A\|_* \leq \rho(A) + \varepsilon.$$

Proof: The proof uses Schur's[2] triangulation theorem: Every matrix $A \in \mathbb{R}^{n\times n}$ can be factored in the form $A = U^*TU$, where $U \in \mathbb{C}^{n\times n}$ is a unitary matrix, $U^* = U^{-1}$ (the adjoint matrix is the inverse matrix), and T an upper triangular matrix of the form

$$T = \begin{pmatrix} \lambda_1 & t_{12} & \cdots & t_{1n} \\ 0 & \lambda_2 & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

[2] Issai Schur (1875 - 1941)

with the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A, e.g. see (Marcus and Minc, 1992, p. 67). The vector norm $\|\cdot\|_*$ is defined with the diagonal matrix $D_\delta = \operatorname{diag}\left(1, \delta, \ldots, \delta^{n-1}\right)$, $\delta > 0$:

$$\|\mathbf{x}\|_* := \left\|D_\delta^{-1} U \mathbf{x}\right\|_\infty .$$

For the induced matrix norm it follows, using the Schur triangulation of A, that

$$\|A\|_* := \max_{\mathbf{x}\in\mathbb{R}^n, \mathbf{x}\neq\mathbf{0}} \frac{\left\|D_\delta^{-1} U A \mathbf{x}\right\|_\infty}{\left\|D_\delta^{-1} U \mathbf{x}\right\|_\infty} = \max_{\mathbf{x}\in\mathbb{R}^n, \mathbf{x}\neq\mathbf{0}} \frac{\left\|D_\delta^{-1} U U^* T U \mathbf{x}\right\|_\infty}{\left\|D_\delta^{-1} U \mathbf{x}\right\|_\infty}.$$

Setting $\mathbf{y} = D_\delta^{-1} U \mathbf{x}$ it follows that $\mathbf{x} = U^* D_\delta \mathbf{y}$ since the matrices U and D_δ are non-singular and $D_\delta^{-1} U$ is a bijection from $\mathbb{R}^n$ to $\mathbb{R}^n$. Inserting this expression gives

$$\|A\|_* = \max_{\mathbf{y}\in\mathbb{R}^n, \mathbf{y}\neq\mathbf{0}} \frac{\left\|D_\delta^{-1} U U^* T U U^* D_\delta \mathbf{y}\right\|_\infty}{\|\mathbf{y}\|_\infty} = \left\|D_\delta^{-1} T D_\delta\right\|_\infty .$$

The diagonal matrix D_δ^{-1} scales just the rows of T and the matrix D_δ just the columns of T. Thus, the product is again an upper triangular matrix and a straigthforward calculation shows that

$$D_\delta^{-1} T D_\delta = \begin{pmatrix} \lambda_1 & \delta t_{12} & \cdots & \delta^{n-1} t_{1n} \\ 0 & \lambda_2 & \cdots & \delta^{n-2} t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \delta t_{n-1,n} \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

and hence $\left\|D_\delta^{-1} T D_\delta\right\|_\infty \le \rho(A) + \varepsilon$ if δ is chosen sufficiently small. ■

Definition 2.9 Spectral condition number. The spectral condition number $\kappa_2(A)$ of a non-singular matrix $A \in \mathbb{R}^{n\times n}$ is defined by

$$\kappa_2(A) = \|A\|_2 \left\|A^{-1}\right\|_2 .$$

□

Definition 2.10 Definiteness. The matrix $A \in \mathbb{R}^{n\times n}$ is called positive definite if

$$\mathbf{x}^T A \mathbf{x} > 0 \quad \forall\, \mathbf{x} \in \mathbb{R}^n \setminus \{0\}. \tag{2.2}$$

If the equal sign can occur, A is called positv semi-definite. □

Remark 2.11 *On definiteness.* Applying the standard basis vectors

$$\mathbf{e}^{(i)} = (0, \ldots, 0, \underbrace{1}_{i}, 0, \ldots, 0)^T, \quad i = 1, \ldots, n,$$

in (2.2) shows that if A is positive (semi-)definite then also the diagonal matrix $\operatorname{diag}(a_{ii})$ is positive (semi-)definite. □

Remark 2.12 *On symmetric matrices.* A matrix $A \in \mathbb{R}^{n\times n}$ is called symmetric if $A = A^T$. It is called skew-symmetric if $A^T = -A$.

One of the most important properties of symmetric matrices is that all eigenvalues are real numbers. It holds, see e.g. Saad (2003),

$$\lambda_{\max}(A) = \max_{\mathbf{x}\in\mathbb{R}^n, \mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}, \quad \lambda_{\min}(A) = \min_{\mathbf{x}\in\mathbb{R}^n, \mathbf{x}\neq 0} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}. \tag{2.3}$$

The quotient on the right hand side is called Rayleigh[3] quotient. A symmetric matrix is positive definite (s.p.d.) if and only if all of its eigenvalues are positive. It is positive semi-definite if and only if all of its eigenvalues are non-negative.

[3] John William Strutt (Lord Rayleigh) (1842 - 1919)

In the case of $A \in \mathbb{R}^{n\times n}$ being symmetric and positive definite, one obtains for the spectral condition number of A

$$\|A\|_2 = \left(\lambda_{\max}\left(A^TA\right)\right)^{1/2} = \left(\lambda_{\max}\left(A^2\right)\right)^{1/2} = \left(\left(\lambda_{\max}\left(A\right)\right)^2\right)^{1/2} = \lambda_{\max}\left(A\right).$$

Since $\lambda_{\max}\left(A^{-1}\right) = \left(\lambda_{\min}\left(A\right)\right)^{-1}$, one has $\left\|A^{-1}\right\|_2 = \left(\lambda_{\min}\left(A\right)\right)^{-1}$ and

$$\kappa_2\left(A\right) = \frac{\lambda_{\max}\left(A\right)}{\lambda_{\min}\left(A\right)} \geq 1.$$

□

Definition 2.13 Diagonal dominance. A matrix A is called diagonally dominant if

$$|a_{ii}| \geq \sum_{j=1, j\neq i}^{n} |a_{ij}| \quad \text{for all } i = 1, \ldots, n.$$

If for all i the larger sign holds, then A is called stronly diagonally dominant. □

Definition 2.14 Normal matrix. The matrix $A \in \mathbb{R}^{n\times n}$ is called normal, if $A^TA = AA^T$. □

Remark 2.15 *On normal matrices.* It is known that A is normal if and only if it is unitary similar to a diagonal matrix, i.e., there is a unitary matrix $Q \in \mathbb{R}^{n\times n}$ such that

$$A = Q^*\Lambda Q, \quad \Lambda = \operatorname{diag}\left(\lambda_1, \ldots, \lambda_n\right),$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. Obviously, symmetric matrices are normal. □