# Advanced Numerical Analysis

Vladimir Kazeev

March 18, 2025

# 1 Linear algebra

## 1.1 Vectors and matrices

In this section the field $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$. $m$ and $n$ always denote natural numbers.

**Definition 2.1.** Let $V$ be a vector space over $\mathbb{F}$. A function $\|\cdot\| : V \to \mathbb{R}$ is called a norm on V if for all $v, w \in V$ and $\alpha \in \mathbb{F}$ the following properties hold:

1. $\|v\| \geq 0$
2. $\|v\| \neq 0 \quad \forall v \neq 0$
3. $\|\alpha v\| = |\alpha| \|v\|$
4. $\|v + w\| \leq \|v\| + \|w\|$

**Example 2.2.** Let $V = \mathbb{F}^n$

- $\|\cdot\|_\infty : V \to \mathbb{R} : \|v\|_\infty = \max_{i=1}^n |v_i| \quad \forall v \in V$
- $\|\cdot\|_p : V \to \mathbb{R} : \|v\|_p = \sqrt[p]{\sum_{i=1}^n |v_i|^p} \quad \forall v \in V$ and $p \in [1, \infty)$

Also $\lim_{p\to\infty} \|v\|_p = \|v\|_\infty$

**Example 2.3.** $V = \mathbb{F}^{m\times n}$. Then we define $\|\cdot\|_{\max}, \|\cdot\|_{\mathrm{F}} : \mathbb{F}^{m\times n} \to \mathbb{R}$ as follows:

- $\|A\|_{\max} = \max_{i,j} |a_{ij}|$ (maximum absolute value norm / Chebyshev norm)
- $\|A\|_{\mathrm{F}} = \sqrt{\sum_{i,j} |a_{ij}|^2}$ (Frobenius norm)

**Proposition 2.4.** *Let $V, U$ be $\mathbb{F}$-vector spaces. $\mathcal{L}$ denotes the space of continuous (w.r.t. $\|\cdot\|_V$, $\|\cdot\|_U$) linear mappings from $V$ to $U$. Then $\|\cdot\|: \mathcal{L} \to \mathbb{R}$ given by*

$$\|\varphi\| = \sup_{\substack{v\in V \\ \|v\|_V=1}} \|\varphi(v)\|_U \quad \forall \varphi \in \mathcal{L}$$

*is a norm.*

**Definition 2.5.** The norm given in Proposition 2.4 is called the *operator norm* on $\mathcal{L}$ induced by the norms $\|\cdot\|_V$ and $\|\cdot\|_U$.

**Definition 2.6.** $V = \mathbb{F}^n$, $U = \mathbb{F}^m$. $\mathcal{L}$ is identified with $W = \mathbb{F}^{m\times n}$ using the standard basis.

$$\varphi \in \mathcal{L} \quad \longleftrightarrow \quad A = \mathrm{Mat}(\varphi) \in W$$
$$\varphi(v) = Av$$

Let $\|\cdot\|$ be the operator norm on $\mathcal{L}$ induced by $\|\cdot\|_V$ and $\|\cdot\|_U$. Then $\|\cdot\| \cdot \mathrm{Mat}^{-1}$ : $\mathbb{F}^{m\times n} \to \mathbb{R}$ is called the *matrix operator norm* induced by $\|\cdot\|_V$ and $\|\cdot\|_U$.

**Example 2.7.** For $p, q \in [1, \infty]$, $W = \mathbb{F}^{m\times n}$.

$$\|\cdot\|_{p,q} : \; W \to \mathbb{R} \text{ given by } \|A\|_{p,q} = \max_{\substack{v\in\mathbb{F}^n \\ \|v\|_q=1}} \|Av\|_p \quad \forall A \in W$$

is an (matrix) operator norm induced by $\|\cdot\|_p$ and $\|\cdot\|_q$.

**Definition 2.8.** For $p = q \in [1, \infty]$ we write $\|\cdot\|_{p,q} = \|\cdot\|_p$ and $\|\cdot\|_p$ is called the matrix p-norm on $\mathbb{F}^{m\times n}$.

**Proposition 2.9.** $\mathbb{F}^{n\times 1} \simeq \mathbb{F}^n$. *The matrix p-norm on $\mathbb{F}^{n\times 1}$ coincides with the vector p-norm on $\mathbb{F}^n$.*

**Proposition 2.10.** *For $A \in \mathbb{F}^{m\times n}$ the following holds:*

$$\begin{aligned}
\|A\|_1 &= \max_{j=1,\dots,n} \sum_{i=1}^{m} |a_{ij}| && \textit{(column sum norm)} \\
\|A\|_\infty &= \max_{i=1,\dots,m} \sum_{j=1}^{n} |a_{ij}| && \textit{(row sum norm)} \\
\|A\|_2 &= \sqrt{\lambda_{\max}(A^*A)} = \sigma_{\max}(A) && \textit{(spectral norm)} \\
&= \max_{\substack{u\in\mathbb{F}^m \\ v\in\mathbb{F}^n \\ \|u\|_2=\|v\|_2=1}} u^*Av
\end{aligned}$$

*where $\lambda_{\max}$ is the largest eigenvalue and $\sigma_{\max}$ is the largest singular value of $A$.*

**Definition 2.11.** $U = \mathbb{F}^{k\times m}, V = \mathbb{F}^{m\times n}, W = \mathbb{F}^{k\times n}$. Let $\|\cdot\|_U, \|\cdot\|_V, \|\cdot\|_W$ be norms on $U, V, W$ respectively. These norms are called *consistent* (or *submultiplicative*) if

$$\|AB\|_W \leq \|A\|_U \|B\|_V \quad \forall A \in U, B \in V$$

For $U = V = W$ and $\|\cdot\|_U = \|\cdot\|_V = \|\cdot\|_W$ this reduces to

$$\|AB\|_W \leq \|A\|_W \|B\|_W \quad \forall A, B \in W.$$

**Proposition 2.12.**

- *$p$-norm on $\mathbb{F}^{n\times n}$ is consistent for $p \in \{1, 2, \infty\}$*
- *Frobenius norm on $\mathbb{F}^{n\times n}$ is consistent*
- *Chebyshev norm on $\mathbb{F}^{n\times n}$ is* not *consistent*

  *e.g.* $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}: \ \|A \cdot A\|_{\max} = \|\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}\|_{\max} = 2 \not\leq 1 = \|A\|_{\max}\|A\|_{\max}$

**Proposition 2.13.** *$U \in \mathbb{F}^{n\times n}$ invertible and $\|\cdot\|$ a norm on $\mathbb{F}^{n\times n}$. Consider $\|\cdot\|_*, \|\cdot\|_{**}, \|\cdot\|_{***} : \mathbb{F}^{n\times n} \to \mathbb{R}$ given by $\|A\|_* = \|UA\|,\ \|A\|_{**} = \|AU\|,\ \|A\|_{***} = \|U^{-1}AU\|$. These 3 functions are norms on $\mathbb{F}^{n\times n}$ and they are consistent if $\|\cdot\|$ is consistent.*

## 1.2 Eigenvalues of matrices

**Definition 2.14.** $A \in F^{n\times n}, \lambda \in \mathbb{F}$. If $\ker(A - \lambda I) \neq \{0\}$ then $\lambda$ is called an eigenvalue of $A$ and every non-zero vector from $\ker(A - \lambda I)$ is called an eigenvector of $A$ associated with the eigenvalue $\lambda$.

**Definition 2.15.** $A \in \mathbb{F}^{n\times n}$. $\chi_A : \ \mathbb{F} \to \mathbb{F}$ given by $\chi_A(\lambda) = \det(A - \lambda I)\ \forall \lambda \in \mathbb{F}$ is called *characteristic polynomial.*

**Proposition 2.16.** *$A \in \mathbb{F}^{n\times n}$. $\chi_A$ is an algebraic polynomial of degree $n$ with leading coefficient $(-1)^n$. For any $\lambda \in \mathbb{F}$, $\lambda$ is an eigenvalue of $A$ if and only if $\chi_A(\lambda) = 0$.*

**Definition 2.17.** $A \in \mathbb{F}^{n\times n}, \lambda \in \mathbb{F}$ eigenvalue of $A$. The *algebraic multiplicity* of $\lambda$ is the multiplicity of $\lambda$ as a root of $\chi_A$.

**Definition 2.18.** The *geometric multiplicity* of $\lambda$ is the dimension of $\ker(A - \lambda I)$. $\lambda$ is called *defective* if its geometric multiplicity is less than its algebraic multiplicity. If the geometric multiplicity of $\lambda$ is equal to its algebraic multiplicity then $\lambda$ is called *non-defective* eigenvalue of $A$.

**Example.** $A = I \in \mathbb{F}^{n\times n}$. $\chi_A(\lambda) = \det(I - \lambda I) = (1-\lambda)^n$. So $\lambda = 1$ is the only eigenvalue of $I$ with algebraic multiplicity $n$. We have that $\dim(\ker(A - I)) = \dim(\ker(0)) = n$.

If $A \in \mathbb{F}^{n\times n}$ is a Jordan block of size $n \geq 2$, then there is only one eigenvalue, $\lambda = 1$, with algebraic multiplicity $n$ and geometric multiplicity $\dim(\ker(A - I)) = 1 < n$. So $\lambda = 1$ is a defective eigenvalue of $A$.

## 1.3 Schur canonical form

**Definition 2.19.** $A \in \mathbb{C}^{n\times n}$. Assume that $Q \in \mathbb{C}^{n\times n}$ is unitary and that $T = Q^*AQ$ (which is equivalent to $A = QTQ^*$) is upper triangular. Then the factorization $A = QTQ^*$ is called a Schur decomposition of $A$ and $T$ is called a *Schur canonical form.*

**Proposition 2.20.** *In the context of the previous definition, the diagonal entries of $T$ are the eigenvalues of $A$ repeated according to their algebraic multiplicities.*

**Theorem 2.21.** *Let $A \in \mathbb{C}^{n\times n}$, $\lambda_1, \dots, \lambda_n$ be the eigenvalues of $A$ repeated according to their algebraic multiplicities. Then there exists a unitary matrix $Q \in \mathbb{C}^{n\times n}$ such that $T = Q^*AQ$ is upper triangular with diagonal entries $\lambda_1, \dots, \lambda_n$.*

*Proof.* Let $x_1$ be a normalized eigenvector of $A$ associated with $\lambda_1$. Consider a matrix $X = \left[\begin{array}{c|c} x_1 & X_1 \end{array}\right] \in \mathbb{C}^{n\times n}$ unitary (with $X_1 \in \mathbb{C}^{n\times(n-1)}$). Then

$$\begin{aligned} X^*AX &= \left[\begin{array}{c} x_1^* \\ \hline X_1 \end{array}\right] A \left[\begin{array}{c|c} x_1 & X_1 \end{array}\right] = \left[\begin{array}{c|c} x_1^*Ax_1 & x_1^*AX_1 \\ \hline X_1^*Ax_1 & X_1^*AX_1 \end{array}\right] \\ &= \left[\begin{array}{c|c} X^*Ax_1 & \begin{array}{c} x_1^*AX_1 \\ X_1^*AX_1 \end{array} \end{array}\right] = \left[\begin{array}{c|c} \lambda_1 & x_1^*AX_1 \\ \hline 0 & X_1^*AX_1 \end{array}\right] = \left[\begin{array}{c|c} \lambda_1 & t_1^* \\ \hline 0 & A_1 \end{array}\right] \end{aligned}$$

where $t_1 = X_1^*A^*x_1 \in \mathbb{C}^{n-1}$ and $A_1 = X_1^*AX_1 \in \mathbb{C}^{(n-1)\times(n-1)}$. For any $\lambda \in \mathbb{C}$, we have that $\det(A - \lambda I) = \det(X^*AX - \lambda I) = (\lambda_1 - \lambda)\det(A_1 - \lambda I)$. The following are equivalent for all $\lambda \in \mathbb{C}, m \in \mathbb{N}$:

(i) $\lambda$ is a root of $\chi_A$ of multiplicity $m$

(ii) $\lambda$ is a root of $\chi_{A_1}$ of multiplicity $\begin{cases} m & \text{if } \lambda \neq \lambda_1 \\ m-1 & \text{if } \lambda = \lambda_1 \end{cases}$

Then by Proposition 2.16 and Definition 2.17 $\lambda_1, \dots, \lambda_n$ are the eigenvalues of $A$ repeated according to their algebraic multiplicities. Assume that there exists a unitary matrix $Q_1 \in \mathbb{C}^{(n-1)\times(n-1)}$ such that $T_1 = Q_1^* A_1 Q_1$ is upper triangular with diagonal entries $\lambda_2, \ldots, \lambda_n$. Then $Q = X \left[\begin{array}{c|c} 1 & \\ \hline & Q_1 \end{array}\right]$ and $T = \left[\begin{array}{c|c} \lambda_1 & t_1^* Q_1 \\ \hline & T_1 \end{array}\right]$ fullfill the claim for $A$:

$$Q^*AQ = \left[\begin{array}{c|c} 1 & \\ \hline & Q_1^* \end{array}\right] X^*AX \left[\begin{array}{c|c} 1 & \\ \hline & Q_1 \end{array}\right] = \left[\begin{array}{c|c} 1 & \\ \hline & Q_1^* \end{array}\right] \left[\begin{array}{c|c} \lambda_1 & t_1^* \\ \hline & A_1 \end{array}\right] \left[\begin{array}{c|c} 1 & \\ \hline & Q_1 \end{array}\right] = \left[\begin{array}{c|c} \lambda_1 & t_1^* Q_1 \\ \hline & T_1 \end{array}\right]$$

For $n = 1$: $Q = [1], T = A$ fullfill the claim. By induction the claim holds (for any $n \in \mathbb{N}$). $\square$

*Remark.* For square matrices $A \in \mathbb{C}^{n\times n}$ and $S \in \mathbb{C}^{n\times n}$ invertible, the following holds:

$$\begin{aligned} \det(S^{-1}AS) &= \det(S^{-1})\det(A)\det(S) = \det(A) \\ \chi_{S^{-1}AS}(\lambda) &= \det(S^{-1}AS - \lambda I) = \det(S^{-1}(A - \lambda I)S) \\ &= \det(S^{-1})\det(A - \lambda I)\det(S) = \det(A - \lambda I) = \chi_A(\lambda) \end{aligned}$$

That is, the eigenvalues of $A$ and $S^{-1}AS$ coincide.

**Theorem** (Spectral theorem)**.** *Let $n \in N$. $A \in \mathbb{F}^{n\times n}$ is diagonalizable by …*

*$\mathbb{F} = \mathbb{C}$: … an unitary similarity transformation $\Leftrightarrow$ $A$ is normal*

*$\mathbb{F} = \mathbb{R}$: … an orthogonal similarity transformation $\Leftrightarrow$ $A$ is symmetric*

*Remark.* There can be non-hermitian normal matrices, e.g. $A = \begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix}$

*Remark* 3.1. For $A \in \mathbb{C}^{n\times n}$. By theorem 2.21 $A$ has a Schur form $T$. It is easy to check that $A$ is normal if and only if $T$ is normal.

$$(A = QTQ^*, \text{ so } A^*A - AA^* = Q(T^*T - TT^*)Q^* = 0 \Leftrightarrow T^*T = TT^*)$$

It is left as an exercise to show that

$$T \text{ is normal} \Leftrightarrow T \text{ is diagonal.}$$

So the Schur form is a generalization of diagonalization by unitary similarity transformations for normal matrices to abitrary matrices.

## 1.4 Spectral radius of a matrix: The behavior of matrix powers

**Definition 3.2.** For $A \in \mathbb{F}^{n\times n}$ the set of eigenvalues of $A$ is called the *spectrum* of $A$. We will denote the spectrum of $A$ by $\boldsymbol{\lambda}(A)$. (i.e., $\boldsymbol{\lambda}(A)$ is the zero set of $\chi_A$).

**Definition 3.3.** For $A \in \mathbb{F}^{n\times n}$, $\rho(A) = \max_{\lambda\in\boldsymbol{\lambda}(A)} |\lambda|$ is called the *spectral radius* of $A$. Any $\lambda \in \boldsymbol{\lambda}(A)$ with $|\lambda| = \rho(A)$ is called a *dominant eigenvalue* of $A$.

**Lemma 3.4.** *Let $\|\cdot\|$ be a consistent norm on $\mathbb{F}^{n\times n}$. Then $\rho(A) \le \|A\|$ for all $A \in \mathbb{F}^{n\times n}$.*

*Proof.* (Auxiliary result: Consider $y \in \mathbb{C}^n$ non-zero. Let $\|\cdot\|_* : \mathbb{F}^n \to \mathbb{R}$ be given by $\|x\|_* = \|xy^*\| \quad \forall x \in \mathbb{F}^n$. Then $\|Ax\|_* = \|(Ax)y^*\| = \|A(xy^*)\| \le \|A\|\|xy^*\| = \|A\|\|x\|_*$. So the norms $\|\cdot\|_*, \|\cdot\|, \|\cdot\|_*$ are consistent norms.)

Let $\|\cdot\|_*$ be a norm on $\mathbb{F}^{n\times n}$ with which $\|\cdot\|$ is consistent (i.e. $\|\cdot\|_*, \|\cdot\|, \|\cdot\|_*$ are consistent). Let $\lambda \in \mathbb{F}$ be an eigenvalue of $A$ and $x \in \mathbb{F}^n$ an associated eigenvector of unit length w.r.t. $\|\cdot\|_*$. Then $\|A\| = \|A\|\|x\|_* \ge \|Ax\|_* = \|\lambda x\|_* = |\lambda|\|x\|_* = |\lambda|$. $\square$

**Lemma 3.5.** *Let $A \in \mathbb{F}^{n\times n}$ and $\varepsilon > 0$. Then there exists a consistent norm $\|\cdot\|_{A,\varepsilon}$ on $\mathbb{F}^{n\times n}$ such that $\|A\|_{A,\varepsilon} \le \rho(A) + \varepsilon$.*

*If the dominant eigenvalues of $A$ are non-defective, then there exists a consistent norm $\|\cdot\|_A$ on $\mathbb{F}^{n\times n}$ such that $\|A\|_A = \rho(A)$.*

*Proof.* By theorem 2.21, $A$ has a Schur decomposition $A = QTQ^*$ with $Q \in \mathbb{C}^{n\times n}$ unitary and $T \in \mathbb{C}^{n\times n}$ upper triangular. Let $\Lambda = \operatorname{diag}(T)$ and $U = T - \Lambda = \operatorname{offdiag}(T)$ (so $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ in the context of theorem 2.21, $U$ is strictly upper triangular).

For $\eta > 0$ consider $D_\eta = \operatorname{diag}(\eta^0, \eta^1, \ldots, \eta^{n-1}) \in \mathbb{C}^{n\times n}$ then, for all $i, j \in \{1, \ldots, n\}$, we have

$$(D_\eta^{-1} U D_\eta)_{ij} = \eta^{1-i} U_{ij} \eta^{j-1} = \begin{cases} 0 & \text{if } i \ge j \\ \eta^{j-i} U_{ij} & \text{if } i < j \end{cases}$$

So there exists $\eta_* > 0$ such that $\|D_{\eta_*}^{-1} U D_{\eta_*}\|_\infty < \varepsilon$. Let $D = D_{\eta_*}$. Then

$$\begin{aligned} \|D^{-1}Q^*AQD\|_\infty &= \|D^{-1}\Lambda D + D^{-1}UD\|_\infty = \|\Lambda + D^{-1}UD\|_\infty \\ &\le \|\Lambda\|_\infty + \|D^{-1}UD\|_\infty < \rho(A) + \varepsilon \end{aligned}$$

Let us define $\|\cdot\|_{A,\varepsilon} : \mathbb{C}^{n\times n} \to \mathbb{R}$ by $\|B\|_{A,\varepsilon} = \|D^{-1}Q^*BQD\|_\infty$. By proposition 2.13 $\|\cdot\|_{A,\varepsilon}$ is a consistent norm on $\mathbb{C}^{n\times n}$. On the other hand, $\|\cdot\|_A < \rho(A) + \varepsilon$.

For the second claim, let us assume that $\lambda_1, \ldots, \lambda_k$ with $k \in \{1, \ldots, n\}$ are the dominant eigenvalues of $A$ (i.e. $|\lambda_1| = \ldots = |\lambda_k| = \rho(A) > |\lambda_{k+1}|, \ldots, |\lambda_n|$) and that they are non-defective.

If $\rho(A) = 0$, then $\lambda_1 = \ldots = \lambda_n = 0$, so 0 is a non-defective eigenvalue of $A$ with algebraic multiplicity = geometric multiplicity = $n$, so $A = 0$. Then any consistent norm fullfills the claim.

If $k = n$, all eigenvalues are non-defective. Then $A$ is diagonalizable, i.e. there is $S \in \mathbb{C}^{n\times n}$ invertible such that $A = S\Lambda S^{-1}$ with $\Lambda = S^{-1}AS$ diagonal. Let $\|\cdot\|_A : \mathbb{C}^{n\times n} \to \mathbb{R}$ be given by $\|B\|_A = \|S^{-1}BS\|_\infty$ for all $B \in \mathbb{C}^{n\times n}$. As discussed earlier, $\|\cdot\|_A$ is a consistent norm and $\|A\|_A = \|\Lambda\|_\infty = \rho(A)$.

For the remainder of the proof, assume that $k < n$. Let $\Lambda_1 = \operatorname{diag}(\lambda_1, \ldots, \lambda_k) \in \mathbb{C}^{k\times k}$ and $\Lambda_2 = \operatorname{diag}(\lambda_{k+1}, \ldots, \lambda_n) \in \mathbb{C}^{(n-k)\times(n-k)}$. Then $\Lambda = \left[\begin{array}{c|c} \Lambda_1 & \\ \hline & \Lambda_2 \end{array}\right] \in \mathbb{C}^{n\times n}$. We consider a Schur decomposition $A = QTQ^*$ with $Q$ unitary and $T$ upper triangular with $\operatorname{diag}(T) = \Lambda$. Partition T as $T = \left[\begin{array}{c|c} T_1 & T_{12} \\ \hline & T_2 \end{array}\right]$ with $T_{11} \in \mathbb{C}^{k\times k}$. We have:

(i) Every dominant eigenvalue of $A$ is not an eigenvalue of $T_2$ but is an eigenvalue of $T_1$.

(ii) For all $\lambda \in \{\lambda_1, \ldots, \lambda_k\}$, $T_2 - \lambda I$ is invertible, so we have $\dim(\ker(A - \lambda I)) = \dim(\ker(T - \lambda I)) = \dim(\ker(T_1 - \lambda I))$.

So $T_1$ is diagonalizable: $\exists S_1 \in \mathbb{C}^{k\times k}$ invertible such that $S_1^{-1}T_1S_1 = \Lambda_1$. Let us consider the matrix $S = \left[\begin{array}{c|c} S_1 & \\ \hline & I_2 \end{array}\right]$ with $I_2 \in \mathbb{C}^{(n-k)\times(n-k)}$ the identity matrix. We have

$$S^{-1}Q^*AQS = S^{-1}TS = \left[\begin{array}{c|c} \Lambda_1 & \\ \hline & \Lambda_2 \end{array}\right] + \left[\begin{array}{c|c} 0 & T_{12} \\ \hline 0 & U_2 \end{array}\right]$$

where $U_2 = T_2 - \Lambda_2 = \operatorname{offdiag}(T_2)$ is strictly upper triangular.

Consider $\eta > 0, D = \operatorname{diag}(\eta^0, \ldots, \eta^{n-1}) = \left[\begin{array}{c|c} D_1 & \\ \hline & D_2 \end{array}\right]$ with $D_1 \in \mathbb{C}^{k\times k}$.

$$D^{-1}S^{-1}Q^*AQSD = \left[\begin{array}{c|c} \Lambda_1 & \\ \hline & \Lambda_2 \end{array}\right] + \left[\begin{array}{c|c} 0 & Z_{12} \\ \hline 0 & Z_2 \end{array}\right]$$

where $Z_{12} = D_1^{-1}T_{12}D_2$ and $Z_2 = D_2^{-1}U_2D_2$. We will consider $\|\cdot\|_A : \mathbb{C}^{n\times n} \to \mathbb{R}$ given by $\|B\|_A = \|D^{-1}S^{-1}Q^*BQSD\|_1$ for all $B \in \mathbb{C}^{n\times n}$. Again, $\|\cdot\|_A$ is a consistent norm.

Block $Z_2$: $U_2$ is strictly upper triangular, so $Z_2$ is strictly upper triangular. For all $i, j \in \{1, \ldots, n-k\}$ such that $i < j$ we have

$$(Z_2)_{ij} = \eta^{j-i}(U_2)_{ij} \xrightarrow{\eta \to 0} 0$$

Block $Z_{12}$: For all $i \in \{1, \ldots, k\}$ and $j \in \{1, \ldots, n-k\}$ we have

$$(Z_{12})_{ij} = \frac{\eta^{k+j}}{\eta^i}(T_{12})_{ij} = \eta^{k-i}\eta^j(T_{12})_{ij} \xrightarrow{\eta \to 0} 0$$

So $\left\| \left[ \begin{array}{c} Z_{12} \\ \hline Z_2 \end{array} \right] \right\|_1 \xrightarrow{\eta \to 0} 0$. So there exists $\eta_* > 0$ such that $\left\| \left[ \begin{array}{c} Z_{12} \\ \hline Z_2 \end{array} \right] \right\|_1 < \frac{1}{2}(\|\Lambda_1\|_1 - \|\Lambda_2\|_1)$. For $D$ defined with $\eta = \eta_*$ we have

$$\|A\|_A = \|D^{-1}S^{-1}Q^*AQSD\|_1 = \left\| \left[ \begin{array}{c|c} \Lambda_1 & \\ \hline & \Lambda_2 \end{array} \right] + \left[ \begin{array}{c|c} 0 & Z_{12} \\ \hline 0 & Z_2 \end{array} \right] \right\|_1$$

$$= \max\left\{ \|\Lambda_1\|_1, \left\| \left[ \begin{array}{c} Z_{12} \\ \hline \Lambda_2 + Z_2 \end{array} \right] \right\|_1 \right\} = \|\Lambda_1\|_1 = \rho(A)$$

where we used

$$\left\| \left[ \begin{array}{c} Z_{12} \\ \hline \Lambda_2 + Z_2 \end{array} \right] \right\|_1 \le \|\Lambda_2\|_1 + \left\| \left[ \begin{array}{c} Z_{12} \\ \hline Z_2 \end{array} \right] \right\|_1 < \|\Lambda_2\|_1 + \frac{1}{2}(\|\Lambda_1\|_1 - \|\Lambda_2\|_1)$$

$$= \frac{1}{2}(\|\Lambda_1\|_1 + \|\Lambda_2\|_1) < \|\Lambda_1\|_1 = \|\Lambda\|_1 = \rho(A)$$

□

**Lemma 3.6.** *Let $A \in \mathbb{C}^{n \times n}$ and $\|\cdot\|$ be a norm on $\mathbb{C}^{n \times n}$, $\varepsilon > 0$, then there exists $c > 0$ (depending on $n$, $\|\cdot\|$, but not on $A$ or $\varepsilon$) and $C > 0$ (depending on $n$, $\|\cdot\|$, $A$ and $\varepsilon$) such that*

$$c\rho^k \le \|A^k\| \le C(\rho + \varepsilon)^k \quad \forall k \in \mathbb{N} \quad \textit{where } \rho = \rho(A).$$

*If the dominant eigenvalues of $A$ are non-defective, the same holds with $\varepsilon = 0$.*

*Proof.* (i) For the lowerbound, consider a dominant eigenvalue $\lambda \in \mathbb{C}$ of $A$ and a corresponding eigenvector, s.t. $\|x\|_2 = 1$. Then $\|A^k x\|_2 = \|\lambda^k x\|_2 = |\lambda|^k \|x\|_2 = \rho^k$. By the equivalence of norms, there exists $c > 0$ such that $c\|B\|_2 \le \|B\|$ for all $B \in \mathbb{C}^{n \times n}$. So $\|A^k\| \ge c\|A^k x\|_2 = c\rho^k$ for all $k \in \mathbb{N}$.

(ii) For the upperbound: Let $\|\cdot\|_{A,\varepsilon}$ be a consistent norm on $\mathbb{C}^{n\times n}$ such that $\|A\|_{A,\varepsilon} \le \rho(A)+\varepsilon$ (Lemma 3.5). Consistency yields $\|A^k\|_{A,\varepsilon} \le \|A\|_{A,\varepsilon}^k$, so $\|A^k\|_{A,\varepsilon} \le (\rho+\varepsilon)^k$ for all $k \in \mathbb{N}$.

By the equivalence of norms, there exists $C > 0$ (depending on $n$, $\|\cdot\|$, $A$ and $\varepsilon$) such that $\|B\| \le C\|B\|_{A,\varepsilon}$ for all $B \in \mathbb{C}^{n\times n}$. So $\|A^k\| \le C\|A^k\|_{A,\varepsilon} \le C(\rho+\varepsilon)^k$ for all $k \in \mathbb{N}$.

(iii) If the dominant eigenvalues of $A$ are non-defective, the same holds with $\varepsilon = 0$.

□

*Remark.* Taking k-th root and limit $k \to \infty$ in the previous lemma yields

$$\rho(A) \le \lim_{k\to\infty} \|A^k\|^{1/k} \le \rho(A) + \varepsilon$$

if the middle limit exists.

**Definition 3.7.** $A \in \mathbb{C}^{n\times n}$ is called

- row-wise (non-)strictly diagonally dominant if

$$|a_{ii}| > (\ge) \sum_{j\in\{1,\dots,n\}\setminus\{i\}} |a_{ij}| \quad \forall i \in \{1,\dots,n\}$$

- column-wise (non-)strictly diagonally dominant if

$$|a_{jj}| > (\ge) \sum_{i\in\{1,\dots,n\}\setminus\{j\}} |a_{ij}| \quad \forall j \in \{1,\dots,n\}$$

**Theorem 3.8** (Levy-Desplanques)**.** *For any $A \in \mathbb{C}^{n\times n}$, if $A$ is row-wise or column-wise strictly diagonally dominant, then it is invertible.*

**Definition 3.9** (Gerschgorin dishes)**.** Let $A \in \mathbb{C}^{n\times n}$. For each $k \in \{1,\dots,n\}$

$$\mathcal{R}_k = \left\{ z \in \mathbb{C} : \; |z - a_{kk}| \le \sum_{j\in\{1,\dots,n\}\setminus\{k\}} |a_{kj}| \subset \mathbb{C} \right\}$$

$$\mathcal{C}_k = \left\{ z \in \mathbb{C} : \; |z - a_{kk}| \le \sum_{i\in\{1,\dots,n\}\setminus\{k\}} |a_{ik}| \subset \mathbb{C} \right\}$$

are called the $k$-th row-wise and column-wise *Gerschgorin dishes* of $A$.

**Theorem 3.10** (1st Gerchgorin theorem)**.** *Let $A \in \mathbb{C}^{n\times n}$. Then*

$$\boldsymbol{\lambda}(A) \subseteq \bigcup_{k=1}^{n} \mathcal{R}_k, \;\; \boldsymbol{\lambda}(A) \subseteq \bigcup_{k=1}^{n} \mathcal{C}_k.$$