Appendix A Basics in linear algebra

A.1 Norms

Definition A.1 (ℓ_p-norms)**.** For $x \in \mathbb{C}^n$ and $p \in [1, \infty)$, the ℓ_p-norm is

$$\|x\|_p = \big(\sum_{i=1}^n |x_i|^p\big)^{1/p},$$

and ($p = \infty$)

$$\|x\|_\infty = \max_{i=1}^n |x_i|.$$

In particular

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \quad \|x\|_2 = \big(\sum_{i=1}^n |x_i|^2\big)^{1/2}.$$

If $C \in \mathbb{C}^{n\times n}$ is non-singular and $\|\cdot\|$ is a norm, then

$$\|x\|_C := \|Cx\|$$

is a norm, too.

Definition A.2 (operator norm)**.** Let $\|\cdot\|$ be a norm on $\mathbb{C}^n$ and let $A \in \mathbb{C}^{n\times n}$. Then

$$\|A\| := \max\{\frac{\|Ax\|}{\|x\|} : x \in \mathbb{C}^n, x \neq 0\} = \max\{\|Ax\| : x \in \mathbb{C}^n, \|x\| = 1\}$$

is a norm on $\mathbb{C}^{n\times n}$, the *operator norm* induced by $\|\cdot\|$ on $\mathbb{C}^n$.

For operator norms we have

$$\begin{aligned} \|Ax\| &\leq \|A\| \cdot \|x\| \text{ for all } x \\ \|AB\| &\leq \|A\| \cdot \|B\| \quad \text{(“sub-multiplicativity”)} \\ \|I\| &= 1 \end{aligned}$$

Theorem A.3 (explicit expressions for some operator norms). *For $A \in \mathbb{C}^{n\times n}$ we have*

(i) $\|A\|_1 = \max_{j=1} \sum_{i=1}^n |a_{ij}|$ *(column sum norm)*

(ii) $\|A\|_2 = \sigma_{\max}(A) = \sqrt{\lambda_{\max}(A^H A)}$ *(maximum singular value)*

(iii) $\|A\|_\infty = \max_{i=1} \sum_{j=1}^n |a_{ij}|$ *(row sum norm)*

Remark A.4. If $A = A^H$, then $\|A\| = \max\{|\lambda : \lambda \in \text{spec}(A)\}$. If p is a polynomial and $A = A^H$, then $p(A)$ is not necessarily Hermitian if p has complex coefficients, but still $\|p(A)\|_2 = \max |p(\lambda) : \lambda \in \text{spec}(A)|$. The latter can be seen by using the polynomial $q = \bar{p}p$ ($\bar{p}$ has the complex conjugate coefficients) with $p(A)^H p(A) = q(A)$, and $\text{spec}(q(A)) = \{q(\lambda) : \lambda \in \text{spec}(A)\}$ with $q(\lambda) = |p(\lambda)|^2$.

The *Neumann series* $\sum_{k=0}^\infty A^k$ is a helpful tool in many places.

Theorem A.5. *Let $A \in \mathbb{C}^{n\times n}$ have $\|A\| < 1$ for some operator norm $\|\cdot\|$. Then $(I - A)^{-1}$ exists and*

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k. \tag{A.1}$$

Moreover,

$$\|(I - A)^{-1} \le \frac{1}{1 - \|A\|}$$

Proof. Dor any $m \in \mathbb{N}$ we have

$$(I - A) \sum_{k=0}^{m-1} = I - A^m.$$

Here, $\|A^m\| \le \|A\|^m \le$ and thus $\lim_{m\to\infty} A^m = 0$. This gives

$$(I - A) \sum_{k=0}^{\infty} = I,$$

which implies that $(I - A)$ is non-singular and that (A.1) holds. The "moreover" part follows using the sub-multiplicativity of the norm and the formula for the value of the geometric series. □

A.2 Spectra

Definition A.6 (eigenpair, -value, -vector, spectrum). (λ, v) is an *eigenpair* of $A \in \mathbb{C}^{n\times n}$ if $v \neq 0$ and $Ax = \lambda x$. Here, λ is the *eigenvalue* and v is the *eigenvector*. The set of all eigenvalues is the *spectrum* $\operatorname{spec}(A)$.

Theorem A.7. *$A \in \mathbb{C}^{n\times n}$ has at most n different eigenvalues. Eigenvectors to different eigenvalues are linearly independent. If the sum of the dimensions of all eigenspaces to the different eigenvalues is n, then $\mathbb{C}^n$ admits a basis of eigenvectors of A. This need not necessarily be the case.*

If A is diagonizable and $v_1, \dots, v_n$ is a basis of eigenvectors of A, then $Av_i = \lambda_i v_i$ can be summarized as the matrix equation

$$AV = V\Lambda, \text{ where} V = [v_1 \mid \cdots \mid v_n], \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n). \tag{A.2}$$

Note that the same eigenvalue may be counted several times in this representation. A consequence of (A.2) is

$$V^{-1}AV = \Lambda. \tag{A.3}$$

This is termed the *eigen decomposition* of A.

Definition A.8 (spectral radius). The *spectral radius* of $A \in \mathbb{C}^{n\times n}$ is

$$\rho(A) := \max\{|\lambda| : \lambda \in \operatorname{spec}(A)\}.$$

Lemma A.9. *Let $A \in \mathbb{C}^{n\times n}$. For any induced operator norm we have*

$$\rho(A) = \limsup_{k\to\infty} \|A^k\|^{1/k}.$$

In particular, $\rho(A) \leq \|A\|$.

Proof. "$\leq$" For any eigenpair (λ, v) with $\|v\| = 1$ of A we have $A^k v = \lambda^k v$ and thus $|\lambda^k| = \|A^k v\| \leq \|A^k\|$.

"$\geq$" Let σ be any number with $\sigma > \limsup_{k\to\infty} \|A^k\|^{1/k}$. We show that an eigenvalue $\lambda \in \operatorname{spec}(A)$ satisfies $|\lambda| < \sigma$ from which we then get the assertion. Taking the limit $k \to \infty$ in

$$(I - \frac{1}{\sigma}A)\sum_{i=0}^{k}(\frac{1}{\sigma}A)^k = I - (\frac{1}{\sigma}A)^{k+1}$$

with $\lim_{k\to\infty} \|(\frac{1}{\sigma}A)^{k+1}\| = 0$ shows that the series $\sum_{i=0}^{\infty}(\frac{1}{\sigma}A)^i$ converges and that

$$\sum_{i=0}^{\infty}(\frac{1}{\sigma}A)^i = (I - \frac{1}{\sigma}A)^{-1}.$$

Multiplying the above equality with an eigenvector v of A with eigenvalue λ gives

$$\sum_{i=0}^{\infty}(\frac{\lambda}{\sigma})^i v = (1 - \frac{\lambda}{\sigma})^{-1} v.$$

and thus

$$\sum_{i=0}^{\infty}(\frac{\lambda}{\sigma})^i = (1 - \frac{\lambda}{\sigma})^{-1}.$$

So the sum to the left converges which implies $|\frac{\lambda}{\sigma}| < 1$, i.e. if $|\lambda| < \sigma$. The "in particular" statement is a direct consequence using $\|A^k\| \leq \|A\|^k$ □

Lemma A.10 (invariance of the spectrum under transformations). *Let $A, B, C \in \mathbb{C}^{n\times n}$, C non-singular. Then*

(i) $\operatorname{spec}(A) = \operatorname{spec}(CAC^{1-})$,

(ii) $\operatorname{spec}(AB) = \operatorname{spec}(BA)$.

Proof. If $Av = \lambda v$ with $v \neq 0$, then $CAv = \lambda Cv \Leftrightarrow CAC^{-1}(Cv) = \lambda(Cv)$. This proves (i) and at the same time shows that the eigenvectors v transform as $x \to Cx$. Of course, (ii) is a special case of (i) of A or B are non-singular. If both are singular, $0 \in \operatorname{spec}(AB)$and $0 \in \operatorname{spec}(BA)$. For any eigenpair $\lambda, v)$ of AB with $\lambda \neq 0$ we have $w = Bv \neq 0$ and $BAw = BABv = \lambda BAw$, so $\lambda \in \operatorname{spec}(BA)$. This gives $\operatorname{spec}(AB) \subseteq \operatorname{spec}(BA)$. The same reasoning also gives the inclusion in the other direction and thus (ii). □

A.3 Inner products

Definition A.11 (inner product)**.** An inner product $(,\cdot,\cdot) : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ satisfies

$$\begin{aligned} (\alpha x + y, z) &= \alpha(x,z) + (y,z) \text{ for all } x,y,z \in \mathbb{C}^n, \alpha \in \mathbb{C}, \\ (x,y) &= \overline{(y,x)} \text{ for all } x,y \in \mathbb{C}^n, \\ (x,x) &> 0 \text{ if } x \neq 0. \end{aligned}$$

As a consequence, $(x,x)^{1/2}$ is a norm.

Definition A.12 (standard inner product)**.** The standard inner product on $\mathbb{C}^n$ is

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \bar{y}_i$$

Definition A.13 (adjoint operator, self-adjoint and anti-selfadjoint)**.** For $A \in \mathbb{C}^{n\times n}$ and an inner product $(\cdot,\cdot)$, the adjoint A^* is defined via

$$(Ax, y) = (x, A^* y) \text{ for all } x, y.$$

A is called *self-adjoint* if $A^* = A$ and *anti-self-adjoint* if $A^* = -A$. For the standard inner product we write A^H instead of A^*, which coincides with the interpretation of ^{H} as "conjugate transpose".

Definition A.14 (Hermitian, skew-Hermitian)**.** A is called *Hermitian* if $A^H = A$ and *skew-Hermitian* if $A^H = -A$.

Definition A.15 (positive (semi-)definite)**.** $A \in \mathbb{C}^{n\times n}$ is called *positive (semi-)definite* (wrt. $(\cdot,\cdot)$) if $\Re(Ax,x) > (\geq 0)$ for all $x \neq 0$. If $(\cdot,\cdot)$ is the standard inner product and A is, in addition, Hermitian, we say that A is hpd (hspd) (Hermitian positive (semi-)definite).

Definition A.16. On the set of all self-adjoint matrices we define the *Loewner ordering* $\prec$ via

$$A \prec B \Leftrightarrow B - A \text{ is positive semidefinite}$$

Remark A.17. So the Loewner ordering depends on the inner product. It is a partial ordering, meaning that it is transitive and anti-symmetric, but that any two self-adjoint matrices are not necessarily comparable in that ordering. We also write $B \succeq A$ if $A \preceq B$, and we denote $\prec B$ if $B - A$ is positive definite. The matrix A is positive definite if $A \succ 0$. If we don't explicitly specify the inner product, then the Loewner ordering refers to the standard inner product.

Definition A.18 (unitary matrices). $U \in \mathbb{C}^{n\times n}$ is called *unitary* (wrt. the inner product $(\cdot,\cdot)$), if $U^*U = I$.

Theorem A.19 (diagonalization of self-adjoint matrices). *Let A be self-adjoint wrt. the inner product $(\cdot,\cdot)$. Then $\mathbb{C}^n$ admits a basis of orthonormal (wrt. $(\cdot,\cdot)$) eigenvectors, i.e. there exists $Q \in \mathbb{C}^{n\times n}$ such that*

$$U^*AU = \operatorname{diag}(\lambda_1, \cdots, \lambda_n).$$

Moreover, all eigenvalues λ_i ar real. If A is, in addition, positive (semi-)definite, then all eigenvalues are positive (non-negative).

Theorem A.20 (diagonalization of anti-self-adjoint matrices). *Let A be anti-self-adjoint wrt. the inner product $(\cdot,\cdot)$. Then $\mathbb{C}^n$ admits a basis of orthonormal (wrt. $(\cdot,\cdot)$) eigenvectors, i.e. there exists $U \in \mathbb{C}^{n\times n}$ such that*

$$U^*AU = \operatorname{diag}(\lambda_1, \cdots, \lambda_n).$$

In addition, all eigenvalues λ_i are purely imaginary, $\frac{1}{i}\lambda_j \in \mathbb{R}$.

Proof. Use Theorem A.19 on iA. □

Definition A.21 (energy inner product). For $Q \in \mathbb{C}^{n\times n}$, the *$Q$-inner product* or *energy inner product* is defined as

$$\langle x, y\rangle_Q = \langle Qx, y\rangle.$$

One easily checks that the energy inner product is indeed an inner product.

Theorem A.22 (representation of inner products). *If $(\cdot,\cdot)$ is an inner product, then there exists Q hpd such that*

$$(\cdot,\cdot) = \langle\cdot,\cdot\rangle_Q$$

Proof. Let e_i denote the ith unit vector and take $b_{ij} = (e_i, e_j)$. Then for $x = \sum_{i=1}^n \xi_i e_i, y = \sum_{j=1}^n \eta_j e_j$

$$(x, y) = \sum_{i,j=1}^{n} \xi\bar{\eta}_j(e_i e_j) = \sum_{i,j=1}^{n} \xi\bar{\eta}_j b_{ij} = \langle Qx, y\rangle.$$

□

So all inner products are Q-inner products. We can characterize the adjoint for the Q inner product using the standard inner product.

Lemma A.23 (representation of adjoints). *Let A^* be the adjoint of A wrt. the Q-inner product. Then*

$$A^* = Q^{-1}A^H Q$$

In particular, A is self-adjoint w.r.t. the Q-inner product if

$$QA = A^H Q,$$

and skew-self adjoint if

$$QA = -A^H Q.$$

Proof. This follows from

$$\begin{aligned} \langle Ax, y\rangle_Q &= \langle QAx, y\rangle \\ &= \langle x, A^H Q y\rangle \\ &= \langle Qx, Q^{-1}A^H by\rangle \\ &= \langle x, Q^{-1}A^H Q y\rangle. \end{aligned}$$

□

Definition A.24 (normal matrices). Let $Q \in \mathbb{C}^{n\times n} \succ 0$. Then $A \in \mathbb{C}^{n\times n}$ is termed *Q-normal* if it commutes with its adjoint wrt. $\langle \cdot, \cdot \rangle_Q$, i.e.

$$AQ^{-1}A^H Q = Q^{-1}A^H QA.$$

Theorem A.25 (eigendecomposition of normal matrices). . *Let $A \in \mathbb{C}^{n\times n}$ be Q-normal. Then $\mathbb{C}^n$ admits a basis of orthonormal (wrt. $\langle \cdot, \cdot \rangle_Q$) eigenvectors, i.e. there exists $U \in \mathbb{C}^{n\times n}$ such that*

$$U^* AU = \operatorname{diag}(\lambda_1, \cdots, \lambda_n).$$

Note that as opposed to the self-adjoint case, the spectrum need not be real. Anti self-adjoint matrices are normal, for example.

A.4 Chebyshev polynomials

Let Π denote the space of all polynomials of arbitrary degree, $\Pi = \cup_{m=0}^{\infty} \Pi_m$.

Definition A.26. On Π we define the inner product

$$[p, q] = \int_{-1}^{1} \frac{1}{\sqrt{1-t^2}} p(t)\bar{q}(t) dt.$$

The sequence T_m of orthogonal polynomials of degree m wrt. $[\cdot, \cdot]$ are the *Chebyshev* polynomials.

German inspired notation: Chebyshev is spelled **T**schebyscheff in German.

The T_m can be obtained by successively orthonormalizing tT_m against $T_0, \dots, T_m$. This gives

$$\begin{aligned} T_0 &= 1 \\ W_{m+1} &= tT_m - \sum_{i=0}^{m} [tT_m, T_i] T_i, m = 0, 1, \dots \\ T_{m+1} &= \frac{1}{\sqrt{[W_{m+1}, W_{m+1}]}} \cdot W_{m+1}. \end{aligned}$$

And since $[tT_m, T_i] = [T_m, tT_i]$ with $tT_i \in \Pi_{i+1}$, we have $[tT_m, T_i] = 0$ if $i < m - 1$. So we actually have a three term recurrence for W_{m+1}, and one obtains the coefficients by evaluating the integrals using the substitution $t = \cos(z)$ according to the following lemma.

Lemma A.27 (Recursion for Chebyshev-polynomials). *We have*

$$\begin{aligned} T_0 &= 1, \\ T1(t) &= t, \\ T_{m+1}(t) &= 2tT_m(t) - T_{m-1}, m = 1, 2, \dots \end{aligned} \tag{A.4}$$

Proof. For $\cos(z)$ one has the equalities

$$\cos((m+1)z) = 2\cos(z)\cos(mz)) - \cos((m-1)z), m = 1, 2,$$

which follows from the elementary trigonometric equality $\cos(\alpha \pm \beta) = \cos(\alpha)\cos(\beta) \mp \sin(\alpha)\sin(\beta)$. So for $t \in [-1 - 1]$, the polynomials T_m from the recurrence (A.4) the lemma may be written as $T_m(t) = \cos(mz)$ with $t = \cos(z), z \in [-pi, \pi]$. Use the substitution $t = \cos(z)$ in the integrals, trigonometric identities and your knowledge from elementary calculus to see

$$\int_{-1}^{1} \frac{1}{\sqrt{1-t^2}} T_m(t)^2 dt \;=\; \int_{-\pi}^{\pi} \cos(mz)^2 dz \;=\; 1$$

$$\begin{aligned}
&\int_{-1}^{1} \frac{1}{\sqrt{1-t^2}} (2tT_m - T_{m-1}(t))T_m(t)(t)dt \\
&= \int_{-\pi}^{\pi} 2\cos(z)\cos^2(mz) - \cos(mz)\cos((m-1)z)dz \;=\; 0, \\
&\int_{-1}^{1} \frac{1}{\sqrt{1-t^2}} (2tT_m - T_{m-1}(t))T_{m-1}(t)(t)dt \\
&= \int_{-\pi}^{\pi} 2\cos(z)\cos(mz)(\cos((m-1)z) - \cos^2((m-1)z)dz \;=\; 0,
\end{aligned}$$

□

Lemma A.28 (Expressions for the Chebyshev polynomials). . *We have*

(i) for $t \in [-1, 1]$*:* $T_m(t) = \cos(m\,\mathrm{Arcosh}(t))$

(ii) for $t \geq 1$*:* $T_m(t) = \cosh(m\,\mathrm{Arcosh}(t))$

(ii)i for $t < -1$*:* $T_m(t) = (-1)^m \cosh(m\,\mathrm{Arcosh}(-t))$

Proof. This follows using standard identities for the trigonometric and hyperbolic functions. □