Chapter 12
Distances on Numbers, Polynomials, and Matrices

12.1 Metrics on Numbers

Here we consider the most important metrics on the classical number systems: the semiring $\mathbb{N}$ of natural numbers, the ring $\mathbb{Z}$ of integers, and the fields $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ of rational, real, complex numbers, respectively. We consider also the algebra $\mathcal{Q}$ of quaternions.

- **Metrics on natural numbers**

  There are several well-known metrics on the set $\mathbb{N}$ of natural numbers:

  1. $|n - m|$; the restriction of the **natural metric** (from $\mathbb{R}$) on $\mathbb{N}$;
  2. $p^{-\alpha}$, where α is the highest power of a given prime number p dividing $m - n$, for $m \neq n$ (and equal to 0 for $m = n$); the restriction of the **p-adic metric** (from $\mathbb{Q}$) on $\mathbb{N}$;
  3. $\ln \frac{lcm(m,n)}{gcd(m,n)}$; an example of the **lattice valuation metric**;
  4. $w_r(n - m)$, where $w_r(n)$ is the *arithmetic r-weight* of n; the restriction of the **arithmetic r-norm metric** (from $\mathbb{Z}$) on $\mathbb{N}$;
  5. $\frac{|n-m|}{mn}$ (cf. **M-relative metric** in Chap. 5);
  6. $1 + \frac{1}{m+n}$ for $m \neq n$ (and equal to 0 for $m = n$); the **Sierpinski metric**.

  Most of these metrics on $\mathbb{N}$ can be extended on $\mathbb{Z}$. Moreover, any one of the above metrics can be used in the case of an arbitrary countable set X. For example, the **Sierpinski metric** is defined, in general, on a countable set $X = \{x_n : n \in \mathbb{N}\}$ by $1 + \frac{1}{m+n}$ for all $x_m, x_n \in X$ with $m \neq n$ (and is equal to 0, otherwise).

- **Arithmetic r-norm metric**

  Let $r \in \mathbb{N}$, $r \geq 2$. The *modified r-ary form* of an integer x is a representation

  $$x = e_n r^n + \cdots + e_1 r + e_0,$$

  where $e_i \in \mathbb{Z}$, and $|e_i| < r$ for all $i = 0, \ldots, n$.

M.M. Deza, E. Deza, *Encyclopedia of Distances*,
DOI 10.1007/978-3-662-52844-0_12

An r-ary form is called *minimal* if the number of nonzero coefficients is minimal. The minimal form is not unique, in general. But if the coefficients e_i, $0 \le i \le n-1$, satisfy the conditions $|e_i + e_{i+1}| < r$, and $|e_i| < |e_{i+1}|$ if $e_i e_{i+1} < 0$, then the above form is unique and minimal; it is called the *generalized nonadjacent form*.

The *arithmetic r-weight* $w_r(x)$ of an integer x is the number of nonzero coefficients in a *minimal r-ary form* of x, in particular, in the generalized nonadjacent form. The **arithmetic r-norm metric** on $\mathbb{Z}$ (see, for example, [Ernv85]) is defined by

$$w_r(x-y).$$

- **Distance between consecutive primes**

  The **distance between consecutive primes** (or *prime gap*, *prime difference function*) is the difference $g_n = p_{n+1} - p_n$ between two successive prime numbers.

  It holds $g_n \le p_n$, $\overline{\lim}_{n\to\infty} g_n = \infty$ and (Zhang, 2013) $\underline{\lim}_{n\to\infty} g_n < 7 \times 10^7$, improved to ≤ 246 (conjecturally, to ≤ 6) by Polymath8, 2014. There is no $\lim_{n\to\infty} g_n$ but $g_n \approx \ln p_n$ for the average g_n.

  Open *Polignac's conjecture*: for any $k \ge 1$, there are infinitely many n with $g_n = 2k$; the case $k = 1$ (i.e., that $\underline{\lim}_{n\to\infty} g_n = 2$ holds) is the *twin prime conjecture*.
- **Distance Fibonacci numbers**

  Fibonacci numbers are defined by the recurrence $F_n = F_{n-1} + F_{n-2}$ for $n \ge 2$ with initial terms $F_0 = 0$ and $F_1 = 1$. **Distance Fibonacci numbers** are three following generalizations of them in the distance sense, considered by Wloch et al..

  Kwaśnik–Wloch, 2000: $F(k,n) = F(k,n-1) + F(k,n-k)$ for $n > k$ and $F(k,n) = n+1$ for $n \le k$.

  Bednarz et al., 2012: $Fd(k,n) = Fd(k,n-k+1) + Fd(k,n-k)$ for $n \ge k > 1$ and $Fd(k,n) = 1$ for $0 \le n < k$.

  Wloch et al., 2013: $F_2(k,n) = F_2(k,n-2) + F_2(k,n-k)$ for $n \ge k \ge 1$ and $F_2(k,n) = 1$ for $0 \le n < k$.
- **p-adic metric**

  Let p be a prime number. Any nonzero rational number x can be represented as $x = p^{\alpha}\frac{c}{d}$, where c and d are integers not divisible by p, and α is a unique integer. The *p-adic norm* of x is defined by $|x|_p = p^{-\alpha}$. Moreover, $|0|_p = 0$ is defined.

  The **p-adic metric** is a **norm metric** on the set $\mathbb{Q}$ of rational numbers defined by

  $$|x-y|_p.$$

  This metric forms the basis for the algebra of p-adic numbers. The **Cauchy completions** of the metric spaces $(\mathbb{Q}, |x-y|_p)$ and $(\mathbb{Q}, |x-y|)$ with the **natural metric** $|x-y|$ give the fields $\mathbb{Q}_p$ of *p-adic numbers* and $\mathbb{R}$ of real numbers, respectively.

The **Gajić metric** is an **ultrametric** on the set $\mathbb{Q}$ of rational numbers defined, for $x \neq y$ (via the integer part $\lfloor z \rfloor$ of a real number z), by

$$\inf\{2^{-n} : n \in \mathbb{Z},\ \lfloor 2^n(x-e) \rfloor = \lfloor 2^n(y-e) \rfloor\},$$

where e is any fixed irrational number. This metric is **equivalent** to the **natural metric** $|x-y|$ on $\mathbb{Q}$.

- **Continued fraction metric on irrationals**

  The **continued fraction metric on irrationals** is a complete metric on the set *Irr* of irrational numbers defined, for $x \neq y$, by

$$\frac{1}{n},$$

  where n is the first index for which the continued fraction expansions of x and y differ. This metric is **equivalent** to the **natural metric** $|x-y|$ on *Irr* which is noncomplete and disconnected. Also, the *Baire* 0-*dimensional space* $B(\aleph_0)$ (cf. **Baire metric** in Chap. 11) is homeomorphic to *Irr* endowed with this metric.

- **Natural metric**

  The **natural metric** (or **absolute value metric**, **line metric**, *the distance between numbers*) is a metric on $\mathbb{R}$ defined by

$$|x-y| = \begin{cases} y-x, \text{ if } x-y<0, \\ x-y, \text{ if } x-y \geq 0. \end{cases}$$

  On $\mathbb{R}$ all l_p**-metrics** coincide with the natural metric. The metric space $(\mathbb{R}, |x-y|)$ is called the *real line* (or *Euclidean line*).

  There exist many other metrics on $\mathbb{R}$ coming from $|x-y|$ by some **metric transform** (Chap. 4). For example: $\min\{1, |x-y|\}$, $\frac{|x-y|}{1+|x-y|}$, $|x|+|x-y|+|y|$ (for $x \neq y$) and, for a given $0<\alpha<1$, the **generalized absolute value metric** $|x-y|^{\alpha}$.

  Some authors use $|x-y|$ as the *Polish notation* (parentheses-free and computer-friendly) of the distance function in any metric space.

- **Zero bias metric**

  The **zero bias metric** is a metric on $\mathbb{R}$ defined by

$$1+|x-y|$$

  if one and only one of x and y is strictly positive, and by

$$|x-y|,$$

  otherwise, where $|x-y|$ is the **natural metric** (see, for example, [Gile87]).

- **Sorgenfrey quasi-metric**
  The **Sorgenfrey quasi-metric** is a quasi-metric d on $\mathbb{R}$ defined by

$$y - x$$

if $y \geq x$, and equal to 1, otherwise. Some similar quasi-metrics on $\mathbb{R}$ are:

1. $d_1(x, y) = \max\{y - x, 0\}$ (in general, $\max\{f(y) - f(x), 0\}$ is a quasi-metric on a set X if $f : X \to \mathbb{R}_{\geq 0}$ is an injective function);
2. $d_2(x, y) = \min\{y - x, 1\}$ if $y \geq x$, and equal to 1, otherwise;
3. $d_3(x, y) = y - x$ if $y \geq x$, and equal to $a(x - y)$ (for fixed $a > 0$), otherwise;
4. $d_4(x, y) = e^y - e^x$ if $y \geq x$, and equal to $e^{-y} - e^{-x}$ otherwise.

- **Real half-line quasi-semimetric**
  The **real half-line quasi-semimetric** is defined on the half-line $\mathbb{R}_{>0}$ by

$$\max\{0, \ln \frac{y}{x}\}.$$

- **Janous–Hametner metric**
  The **Janous–Hametner metric** is defined on the half-line $\mathbb{R}_{>0}$ by

$$\frac{|x - y|}{(x + y)^t},$$

where $t = -1$ or $0 \leq t \leq 1$, and $|x - y|$ is the **natural metric**.

- **Extended real line metric**
  An **extended real line metric** is a metric on $\mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$. The main example (see, for example, [Cops68]) of such metric is given by

$$|f(x) - f(y)|,$$

where $f(x) = \frac{x}{1+|x|}$ for $x \in \mathbb{R}$, $f(+\infty) = 1$, and $f(-\infty) = -1$.
  Another metric, commonly used on $\mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$, is defined by

$$|\arctan x - \arctan y|,$$

where $-\frac{1}{2}\pi < \arctan x < \frac{1}{2}\pi$ for $-\infty < x < \infty$, and $\arctan(\pm\infty) = \pm\frac{1}{2}\pi$.

- **Complex modulus metric**
  The **complex modulus metric** on the set $\mathbb{C}$ of complex numbers is defined by

$$|z - u|,$$

where, for any $z = z_1 + z_2 i \in \mathbb{C}$, the number $|z| = \sqrt{z\overline{z}} = \sqrt{z_1^2 + z_2^2}$ is the *complex modulus*. The *complex argument* θ is defined by $z = |z|(\cos(\theta) + i\sin(\theta))$.

The metric space $(\mathbb{C}, |z-u|)$ is called the *complex* (or *Wessel–Argand*) *plane*. It is isometric to the Euclidean plane $(\mathbb{R}^2, ||x-y||_2)$. So, the metrics on $\mathbb{R}^2$, given in Chaps. 19 and 5, can be seen as metrics on $\mathbb{C}$. For example, the **British Rail metric** on $\mathbb{C}$ is $|z| + |u|$ for $z \neq u$. The p**-relative** (if $1 \leq p < \infty$) and **relative metric** (if $p = \infty$) on $\mathbb{C}$ are defined for $|z| + |u| \neq 0$ respectively, by

$$\frac{|z-u|}{\sqrt[p]{|z|^p + |u|^p}} \text{ and } \frac{|z-u|}{\max\{|z|, |u|\}}.$$

- $\mathbb{Z}(\eta_m)$**-related norm metrics**

A *Kummer* (or *cyclotomic*) *ring* $\mathbb{Z}(\eta_m)$ is a subring of the ring $\mathbb{C}$ (and an extension of the ring $\mathbb{Z}$), such that each of its elements has the form $\sum_{j=0}^{m-1} a_j \eta_m^j$, where η_m is a primitive m-th root $\exp(\frac{2\pi i}{m})$ of unity, and all a_j are integers.

The *complex modulus* $|z|$ of $z = a + b\eta_m \in \mathbb{C}$ is defined by

$$|z|^2 = z\bar{z} = a^2 + (\eta_m + \overline{\eta_m})ab + b^2 = a^2 + 2ab\cos(\frac{2\pi i}{m}) + b^2.$$

Then $(a+b)^2 = q^2$ for $m = 2$ (or 1), $a^2 + b^2$ for $m = 4$, and $a^2 + ab + b^2$ for $m = 6$ (or 3), i.e., for the ring $\mathbb{Z}$ of usual integers, $\mathbb{Z}(i)$ of *Gaussian integers* and $\mathbb{Z}(\rho)$ of *Eisenstein–Jacobi* (or *EJ*) *integers*.

The set of units of $\mathbb{Z}(\eta_m)$ contain $\eta_m^j, 0 \leq j \leq m-1$; for $m = 5$ and $m \geq 6$, units of infinite order appear also, since $\cos(\frac{2\pi i}{m})$ is irrational. For $m = 2, 4, 6$, the set of units is $\{\pm 1\}$, $\{\pm 1, \pm i\}$, $\{\pm 1, \pm \rho, \pm \rho^2\}$, where $i = \eta_4$ and $\rho = \eta_6 = \frac{1+i\sqrt{3}}{2}$.

The norms $|z| = \sqrt{a^2 + b^2}$ and $||z||_i = |a| + |b|$ for $z = a + bi \in \mathbb{C}$ give rise to the **complex modulus** and i**-Manhattan** metrics on $\mathbb{C}$. They coincide with the Euclidean (l_2-) and Manhattan (l_1-) metrics, respectively, on $\mathbb{R}^2$ seen as the complex plane. The restriction of the i-Manhattan metric on $\mathbb{Z}(i)$ is the path metric of the square grid $\mathbb{Z}^2$ of $\mathbb{R}^2$; cf. **grid metric** in Chap. 19.

The ρ**-Manhattan metric** on $\mathbb{C}$ is defined by the norm $||z||_\rho$, i.e.,

$$\min\{|a|+|b|+|c| : z = a+b\rho+c\rho^2\} = \min\{|a|+|b|, |a+b|+|b|, |a+b|+|a| : z = a+b\rho\}.$$

The restriction of the ρ-Manhattan metric on $\mathbb{Z}(\rho)$ is the path metric of the triangular grid of $\mathbb{R}^2$ (seen as the *hexagonal lattice* $A_2 = \{(a,b,c) \in \mathbb{Z}^3 : a+b+c = 0\}$), i.e., the **hexagonal metric** (Chap. 19).

Let f denote either i or $\rho = \frac{1+i\sqrt{3}}{2}$. Given a $\pi \in \mathbb{Z}(f) \setminus \{0\}$ and $z, z' \in \mathbb{Z}(f)$, we write $z \equiv z' \pmod{\pi}$ if $z - z' = \delta\pi$ for some $\delta \in \mathbb{Z}(f)$. For the quotient ring $\mathbb{Z}_\pi(f) = \{z \pmod{\pi} : z \in \mathbb{Z}(f)\}$, it holds $|\mathbb{Z}_\pi(f)| = ||\pi||_f^2$.

Call two congruence classes $z \pmod{\pi}$ and $z' \pmod{\pi}$ *adjacent* if $z - z' \equiv f^j \pmod{\pi}$ for some j. The resulting graph on $\mathbb{Z}_\pi(f)$ called a *Gaussian network* or *EJ network* if, respectively, $f = i$ or $f = \rho$. The path metrics of these networks coincide with their norm metrics, defined (Fan–Gao, 2004) for $z\,(\mathrm{mod}\,\pi)$ and $z'\,(\mathrm{mod}\,\pi)$, by

$$\min ||u||_f : u \in z - z'\,(\mathrm{mod}\,\pi).$$

These metrics are different from the previously defined ([Hube94a, Hube94b]) distance on $\mathbb{Z}_\pi(f)$: $||v||_f$, where $v \in z - z'\,(\mathrm{mod}\,\pi)$ is selected by minimizing the complex modulus. For $f = i$, this is the **Mannheim distance** (Chap. 16), which is not a metric.

- **Chordal metric**

  The **chordal metric** d_χ is a metric on the set $\overline{\mathbb{C}}$=$\mathbb{C} \cup \{\infty\}$ defined by

$$d_\chi(z,u) = \frac{2|z-u|}{\sqrt{1+|z|^2}\sqrt{1+|u|^2}} \text{ and } d_\chi(z,\infty) = \frac{2}{\sqrt{1+|z|^2}}$$

for all $u, z \in \mathbb{C}$ (cf. ***M*-relative metric** in Chap. 5).

The metric space $(\overline{\mathbb{C}}, d_\chi)$ is called the *extended complex plane*. It is homeomorphic and conformally equivalent to the *Riemann sphere*, i.e., the *unit sphere* $S^2 = \{(x_1,x_2,x_3) \in \mathbb{E}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$ (considered as a metric subspace of $\mathbb{E}^3$), onto which $(\overline{\mathbb{C}}, d_\chi)$ is one-to-one mapped under stereographic projection.

The plane $\overline{\mathbb{C}}$ can be identified with the plane $x_3 = 0$ such that the and imaginary axes coincide with the x_1 and x_2 axes. Under stereographic projection, each point $z \in \mathbb{C}$ corresponds to the point $(x_1,x_2,x_3) \in S^2$, where the ray drawn from the “north pole” $(0,0,1)$ to the point z meets the sphere S^2; the “north pole” corresponds to the point at ∞. The chordal (spherical) metric between two points $p, q \in S^2$ is taken to be the distance between their preimages $z, u \in \overline{\mathbb{C}}$.

The chordal metric is defined equivalently on $\overline{\mathbb{R}}^n = \mathbb{R}^n \cup \{\infty\}$:

$$d_\chi(x,y) = \frac{2||x-y||_2}{\sqrt{1+||x||_2^2}\sqrt{1+||y||_2^2}} \text{ and } d_\chi(x,\infty) = \frac{2}{\sqrt{1+||x||_2^2}}.$$

The restriction of the metric d_χ on $\mathbb{R}^n$ is a **Ptolemaic metric**; cf. Chap. 1.

Given $\alpha > 0$, $\beta \geq 0$, $p \geq 1$, the **generalized chordal metric** is a metric on $\mathbb{C}$ (in general, on $(\mathbb{R}^n, ||.||_2)$ and even on any *Ptolemaic space* $(V, ||.||)$), defined by

$$\frac{|z-u|}{\sqrt[p]{\alpha + \beta|z|^p} \cdot \sqrt[p]{\alpha + \beta|u|^p}}.$$

- **Metrics on quaternions**

  Quaternions are members of a noncommutative division algebra $\mathcal{Q}$ over the field $\mathbb{R}$, geometrically realizable in $\mathbb{R}^4$ ([Hami66]). Formally,

$$\mathcal{Q} = \{q = q_1 + q_2 i + q_3 j + q_4 k : q_i \in \mathbb{R}\},$$

where the *basic units* $1, i, j, k \in \mathcal{Q}$ satisfy $i^2 = j^2 = k^2 = -1$ and $ij = -ji = k$.

The *quaternion norm* is defined by $||q|| = \sqrt{q\overline{q}} = \sqrt{q_1^2 + q_2^2 + q_3^2 + q_4^2}$, where $\overline{q} = q_1 - q_2 i - q_3 j - q_4 k$. The **quaternion metric** is the norm metric $||q - q'||$ on $\mathcal{Q}$.

The set of all *Lipschitz integers* and *Hurwitz integers* are defined, respectively, by

$$L = \{q_1 + q_2 i + q_3 j + q_4 k : q_i \in \mathbb{Z}\} \text{ and}$$

$$H = \{q_1 + q_2 i + q_3 j + q_4 k : \text{ all } q_i \in \mathbb{Z} \text{ or all } q_i + \frac{1}{2} \in \mathbb{Z}\}.$$

A quaternion $q \in L$ is *irreducible* (i.e., $q = q'q''$ implies $\{q', q''\} \cap \{\pm 1, \pm i, \pm j, \pm k\} \neq \emptyset$) if and only if $||q||$ is a prime. Given an irreducible $\pi \in L$ and $q, q' \in H$, we write $q \equiv q' \pmod{\pi}$ if $q - q' = \delta\pi$ for some $\delta \in L$.

For the rings $L_\pi = \{q \pmod{\pi} : q \in L\}$ and $H_\pi = \{q \pmod{\pi} : q \in H\}$ it holds $|L_\pi| = ||\pi||^2$ and $|H_\pi| = 2||\pi||^2 - 1$.

The **quaternion Lipschitz metric** on L_π is defined (Martinez et al., 2009) by

$$d_L(\alpha, \beta) = \min \sum_{1 \le s \le 4} |q_s| : \alpha - \beta \equiv q_1 + q_2 i + q_3 j + q_4 k \pmod{\pi}.$$

The ring H is additively generated by its subring L and $w = \frac{1}{2}(1 + i + j + k)$. The **Hurwitz metric** on the ring H_π is defined (Guzëltepe, 2013) by

$$d_H(\alpha, \beta) = \min \sum_{1 \le s \le 5} |q_s| : \alpha - \beta \equiv q_1 + q_2 i + q_3 j + q_4 k + q_5 w \pmod{\pi}.$$

Cf. the **hyper-Kähler** and **Gibbons–Manton** metrics in Sect. 7.3 and the **unit quaternions** and **joint angle** metrics in Sect. 18.3.

12.2 Metrics on Polynomials

A *polynomial* is a sum of powers in one or more variables multiplied by coefficients. A *polynomial in one variable* (or *monic polynomial*) with constant real (complex) coefficients is given by $P = P(z) = \sum_{k=0}^{n} a_k z^k$, $a_k \in \mathbb{R}$ ($a_k \in \mathbb{C}$). The set $\mathcal{P}$ of all real (complex) polynomials forms a ring $(\mathcal{P}, +, \cdot, 0)$. It is also a vector space over $\mathbb{R}$ (over $\mathbb{C}$).

- **Polynomial norm metric**

  A **polynomial norm metric** is a **norm metric** on the vector space $\mathcal{P}$ of all real (complex) polynomials defined by

$$||P - Q||,$$

where $||.||$ is a *polynomial norm*, i.e., a function $||.|| : \mathcal{P} \rightarrow \mathbb{R}$ such that, for all $P, Q \in \mathcal{P}$ and for any scalar k, we have the following properties:

1. $||P|| \geq 0$, with $||P|| = 0$ if and only if $P \equiv 0$;
2. $||kP|| = |k|||P||$;
3. $||P + Q|| \leq ||P|| + ||Q||$ (triangle inequality).

The l_p-*norm* and L_p-*norm* of a polynomial $P(z) = \sum_{k=0}^{n} a_k z^k$ are defined by

$$||P||_p = (\sum_{k=0}^{n} |a_k|^p)^{1/p} \text{ and } ||P||_{L_p} = (\int_0^{2\pi} |P(e^{i\theta})|^p \frac{d\theta}{2\pi})^{\frac{1}{p}} \text{ for } 1 \leq p < \infty,$$

$$||P||_{\infty} = \max_{0 \leq k \leq n} |a_k| \text{ and } ||P||_{L_\infty} = \sup_{|z|=1} |P(z)| \text{ for } p = \infty.$$

The values $||P||_1$ and $||P||_{\infty}$ are called the *length* and *height* of polynomial P.

- **Distance from irreducible polynomials**

  For any field $\mathbb{F}$, a polynomial with coefficients in $\mathbb{F}$ is said to be *irreducible over* $\mathbb{F}$ if it cannot be factored into the product of two nonconstant polynomials with coefficients in $\mathbb{F}$. Given a metric d on the polynomials over $\mathbb{F}$, the **distance** (of a given polynomial $P(z)$) **from irreducible polynomials** is $d_{ir}(P) = \inf d(P, Q)$, where $Q(z)$ is any irreducible polynomial of the same degree over $\mathbb{F}$.

  Polynomial conjecture of Turán, 1967, is that there exists a constant C with $d_{ir}(P) \leq C$ for every polynomial P over $\mathbb{Z}$, where $d(P, Q)$ is the *length* $||P - Q||_1$ of $P - Q$.

  Lee–Ruskey–Williams, 2007, conjectured that there exists a constant C with $d_{ir}(P) \leq C$ for every polynomial P over the Galois field $\mathbb{F}_2$, where $d(P, Q)$ is the **Hamming distance** between the $(0, 1)$-sequences of coefficients of P and Q.

- **Bombieri metric**

  The **Bombieri metric** (or **polynomial bracket metric**) is a **polynomial norm metric** on the set $\mathcal{P}$ of all real (complex) polynomials defined by

$$[P - Q]_p,$$

  where $[.]_p$, $0 \leq p \leq \infty$, is the *Bombieri p-norm*.

  For a polynomial $P(z) = \sum_{k=0}^{n} a_k z^k$ it is defined by

$$[P]_p = (\sum_{k=0}^{n} \binom{n}{k}^{1-p} |a_k|^p)^{\frac{1}{p}}.$$

- **Metric space of roots**

  The **metric space of roots** is (Ćurgus–Mascioni, 2006) the space (X, d) where X is the family of all multisets of complex numbers with n elements and the distance between multisets $U = \{u_1, \ldots, u_n\}$ and $V = \{v_1, \ldots, v_n\}$ is defined by the following analog of the **Fréchet metric**:

$$\min_{\tau \in Sym_n} \max_{1 \leq j \leq n} |u_j - v_{\tau(j)}|,$$

where τ is any permutation of $\{1, \dots, n\}$. Here the set of roots of some monic complex polynomial of degree n is considered as a multiset with n elements. Cf. **metrics between multisets** in Chap. 1.

The function assigning to each polynomial the multiset of its roots is a **homeomorphism** between the metric space of all monic complex polynomials of degree n with the **polynomial norm metric** l_∞ and the metric space of roots.

12.3 Metrics on Matrices

An $m \times n$ matrix $A = ((a_{ij}))$ over a field $\mathbb{F}$ is a table consisting of m rows and n columns with the entries a_{ij} from $\mathbb{F}$. The set of all $m \times n$ matrices with real (complex) entries is denoted by $M_{m,n}$ or $\mathbb{R}^{m \times n}$ ($\mathbb{C}^{m \times n}$). It forms a *group* $(M_{m,n}, +, 0_{m,n})$, where $((a_{ij})) + ((b_{ij})) = ((a_{ij} + b_{ij}))$, and the matrix $0_{m,n} \equiv 0$. It is also an mn-dimensional vector space over $\mathbb{R}$ ($\mathbb{C}$).

The *transpose* of a matrix $A = ((a_{ij})) \in M_{m,n}$ is the matrix $A^T = ((a_{ji})) \in M_{n,m}$. A $m \times n$ matrix A is called a *square matrix* if $m = n$, and a *symmetric matrix* if $A = A^T$. The *conjugate transpose* (or *adjoint*) of a matrix $A = ((a_{ij})) \in M_{m,n}$ is the matrix $A^* = ((\overline{a}_{ji})) \in M_{n,m}$. An *Hermitian matrix* is a complex square matrix A with $A = A^*$.

The set of all square $n \times n$ matrices with real (complex) entries is denoted by M_n. It forms a *ring* $(M_n, +, \cdot, 0_n)$, where $+$ and 0_n are defined as above, and $((a_{ij})) \cdot ((b_{ij})) = ((\sum_{k=1}^n a_{ik} b_{kj}))$. It is also an n^2-dimensional vector space over $\mathbb{R}$ (over $\mathbb{C}$). The *trace* of a square $n \times n$ matrix $A = ((a_{ij}))$ is defined by $\mathrm{Tr}(A) = \sum_{i=1}^n a_{ii}$.

The *identity matrix* is $1_n = ((c_{ij}))$ with $c_{ii} = 1$, and $c_{ij} = 0$, $i \neq j$. An *unitary matrix* $U = ((u_{ij}))$ is a square matrix defined by $U^{-1} = U^*$, where U^{-1} is the *inverse matrix* of U, i.e., $UU^{-1} = 1_n$. A matrix $A \in M_{m,n}$ is *orthonormal* if $A^*A = 1_n$. A matrix $A \in \mathbb{R}^{n \times n}$ is *orthogonal* if $A^T = A^{-1}$, *normal* if $A^TA = AA^T$ and *singular* if its determinant is 0.

If for a matrix $A \in M_n$ there is a vector x such that $Ax = \lambda x$ for some scalar λ, then λ is called an *eigenvalue* of A with corresponding *eigenvector* x. Given a matrix $A \in \mathbb{C}^{m \times n}$, its *singular values* $s_i(A)$ are defined as $\sqrt{\lambda(A^*A)}$. A real matrix A is *positive-definite* if $v^T A v > 0$ for all nonzero real vectors v; it holds if and only if all eigenvalues of $A_H = \frac{1}{2}(A + A^T)$ are positive. An Hermitian matrix A is *positive-definite* if $v^* A v > 0$ for all nonzero complex vectors v; it holds if and only if all $\lambda(A)$ are positive.

The *mixed states* of a n-dimensional *quantum system* are described by their *density matrices*, i.e., positive-semidefinite Hermitian $n \times n$ matrices of trace 1. The set of such matrices is convex, and its extremal points describe the *pure states*. Cf. **monotone metrics** in Chap. 7 and **distances between quantum states** in Chap. 24.

- **Matrix norm metric**

  A **matrix norm metric** is a **norm metric** on the set $M_{m,n}$ of all real (complex) $m \times n$ matrices defined by

$$||A - B||,$$

where $||.||$ is a *matrix norm*, i.e., a function $||.|| : M_{m,n} \to \mathbb{R}$ such that, for all $A, B \in M_{m,n}$, and for any scalar k, we have the following properties:

1. $||A|| \geq 0$, with $||A|| = 0$ if and only if $A = 0_{m,n}$;
2. $||kA|| = |k|||A||$;
3. $||A + B|| \leq ||A|| + ||B||$ (triangle inequality).
4. $||AB|| \leq ||A|| \cdot ||B||$ (*submultiplicativity*).

All matrix norm metrics on $M_{m,n}$ are equivalent. The simplest example of such metric is the **Hamming metric** on $M_{m,n}$ (in general, on the set $M_{m,n}(\mathbb{F})$ of all $m \times n$ matrices with entries from a field $\mathbb{F}$) defined by $||A - B||_H$, where $||A||_H$ is the *Hamming norm* of $A \in M_{m,n}$, i.e., the number of nonzero entries in A. Example of a *generalized* (i.e., not submultiplicative one) *matrix norm* is the *max element norm* $||A = ((a_{ij}))||\max = \max_{i,j} |a_{ij}|$; but $\sqrt{mn}||A||_{\max}$ is a matrix norm.

- **Natural norm metric**

  A **natural** (or *operator*, *induced*) **norm metric** is a **matrix norm metric** on the set M_n defined by

$$||A - B||_{\text{nat}},$$

where $||.||_{\text{nat}}$ is a *natural* (or *operator*, *induced*) *norm* on M_n, induced by the vector norm $||x||$, $x \in \mathbb{R}^n$ ($x \in \mathbb{C}^n$), is a matrix norm defined by

$$||A||_{\text{nat}} = \sup_{||x|| \neq 0} \frac{||Ax||}{||x||} = \sup_{||x||=1} ||Ax|| = \sup_{||x|| \leq 1} ||Ax||.$$

  The natural norm metric can be defined in similar way on the set $M_{m,n}$ of all $m \times n$ real (complex) matrices: given vector norms $||.||_{\mathbb{R}^m}$ on $\mathbb{R}^m$ and $||.||_{\mathbb{R}^n}$ on $\mathbb{R}^n$, the *natural norm* $||A||_{\text{nat}}$ of a matrix $A \in M_{m,n}$, induced by $||.||_{\mathbb{R}^n}$ and $||.||_{\mathbb{R}^m}$, is a matrix norm defined by $||A||_{\text{nat}} = \sup_{||x||_{\mathbb{R}^n}=1} ||Ax||_{\mathbb{R}^m}$.

- **Matrix p-norm metric**

  A **matrix p-norm metric** is a **natural norm metric** on M_n defined by

$$||A - B||^p_{\text{nat}},$$

where $||.||^p_{\text{nat}}$ is the *matrix* (or *operator*) *p-norm*, i.e., a *natural norm*, induced by the vector l_p*-norm*, $1 \leq p \leq \infty$:

$$||A||^p_{\text{nat}} = \max_{||x||_p=1} ||Ax||_p, \quad \text{where} \quad ||x||_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p}.$$

The **maximum absolute column** and **maximum absolute row metric** are the **matrix** 1**-norm** and **matrix** ∞**-norm metric** on M_n. For a matrix $A = ((a_{ij})) \in M_n$, the *maximum absolute column* and *maximum absolute row sum norm* are

$$||A||^1_{\text{nat}} = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}| \text{ and } |A||^{\infty}_{\text{nat}} = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|.$$

The **spectral norm metric** is the **matrix** 2**-norm metric** $||A-B||^2_{\text{nat}}$ on M_n. The matrix 2-norm $||.||^2_{\text{nat}}$, induced by the vector l_2*-norm*, is also called the *spectral norm* and denoted by $||.||_{sp}$. For a symmetric matrix $A = ((a_{ij})) \in M_n$, it is

$$||A||_{sp} = s_{\max}(A) = \sqrt{\lambda_{\max}(A^*A)},$$

where $A^* = ((\overline{a}_{ji}))$, while $s_{\max}$ and $\lambda_{\max}$ are largest singular value and eigenvalue.

- **Frobenius norm metric**

The **Frobenius norm metric** is a **matrix norm metric** on $M_{m,n}$ defined by

$$||A-B||_{Fr},$$

where $||.||_{Fr}$ is the *Frobenius* (or *Hilbert–Schmidt*) *norm*. For $A = ((a_{ij}))$, it is

$$||A||_{Fr} = \sqrt{\sum_{i,j} |a_{ij}|^2} = \sqrt{\text{Tr}(A^*A)} = \sqrt{\sum_{1\le i\le \text{rank}(A)} \lambda_i} = \sqrt{\sum_{1\le i\le \text{rank}(A)} s_i^2},$$

where λ_i, s_i are the eigenvalues and singular values of A.

This norm is strictly convex, is a differentiable function of its elements a_{ij} and is the only unitarily invariant norm among $||A||_p = (\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^p)^{\frac{1}{p}}, p \ge 1$.

The **trace norm metric** is a matrix norm metric on $M_{m,n}$ defined by

$$||A-B||_{tr},$$

where $||.||_{tr}$ is the *trace norm* (or *nuclear norm*) on $M_{m,n}$ defined by

$$||A||_{tr} = \sum_{i=1}^{\min\{m,n\}} s_i(A) = \text{Tr}(\sqrt{A^*A}).$$

- **Schatten norm metric**

Given $1 \le p < \infty$, the **Schatten norm metric** is a **matrix norm metric** on $M_{m,n}$ defined by

$$||A-B||^p_{Sch},$$

where $||.||^p_{Sch}$ is the *Schatten p-norm* on $M_{m,n}$. For a matrix $A \in M_{m,n}$, it is defined as the p-th root of the sum of the p-th powers of all its *singular values*:

$$||A||^p_{Sch} = (\sum_{i=1}^{\min\{m,n\}} s_i^p(A))^{\frac{1}{p}}.$$

For $p = \infty$, 2 and 1, one obtains the **spectral norm metric**, **Frobenius norm metric** and **trace norm metric**, respectively.

- **(c, p)-norm metric**

  Let $k \in \mathbb{N}$, $k \leq \min\{m, n\}$, $c \in \mathbb{R}^k$, $c_1 \geq c_2 \geq \cdots \geq c_k > 0$, and $1 \leq p < \infty$. The (c, p)**-norm metric** is a **matrix norm metric** on $M_{m,n}$ defined by

$$||A - B||^k_{(c,p)},$$

where $||.||^k_{(c,p)}$ is the (c, p)*-norm* on $M_{m,n}$. For a matrix $A \in M_{m,n}$, it is defined by

$$||A||^k_{(c,p)} = (\sum_{i=1}^{k} c_i s_i^p(A))^{\frac{1}{p}},$$

where $s_1(A) \geq s_2(A) \geq \cdots \geq s_k(A)$ are the first k *singular values* of A.

  If $p = 1$, it is the *c-norm*. If, moreover, $c_1 = \cdots = c_k = 1$, it is the *Ky Fan k-norm*.

- **Ky Fan k-norm metric**

  Given $k \in \mathbb{N}$, $k \leq \min\{m, n\}$, the **Ky Fan k-norm metric** is a **matrix norm metric** on $M_{m,n}$ defined by

$$||A - B||^k_{KF},$$

where $||.||^k_{KF}$ is the *Ky Fan k-norm* on $M_{m,n}$. For a matrix $A \in M_{m,n}$, it is defined as the sum of its first k *singular values*:

$$||A||^k_{KF} = \sum_{i=1}^{k} s_i(A).$$

For $k = 1$ and $k = \min\{m, n\}$, one obtains the **spectral** and **trace** norm metrics.

- **Cut norm metric**

  The **cut norm metric** is a **matrix norm metric** on $M_{m,n}$ defined by

$$||A - B||_{cut},$$

where $||.||_{cut}$ is the *cut norm* on $M_{m,n}$ defined, for a matrix $A = ((a_{ij})) \in M_{m,n}$, as:

$$||A||_{cut} = \max_{I\subset\{1,\dots,m\},J\subset\{1,\dots,n\}} |\sum_{i\in I,j\in J} a_{ij}|.$$

Cf. in Chap. 15 the **rectangle distance on weighted graphs** and the **cut semimetric**, but the **weighted cut metric** in Chap. 19 is not related.

- **Matrix nearness problems**

A norm $||.||$ is *unitarily invariant* on $M_{m,n}$ if $||B|| = ||UBV||$ for all $B \in M_{m,n}$ and all unitary matrices U, V. All *Schatten p-norms* are unitarily invariant.

Given a unitarily invariant norm $||.||$ on $M_{m,n}$, a matrix property $\mathcal{P}$ defining a subspace or compact subset of $M_{m,n}$ (so that $d_{||.||}(A, \mathcal{P})$ below is well defined) and a matrix $A \in M_{m,n}$, then the *distance to* $\mathcal{P}$ is the **point-set distance** on $M_{m,n}$

$$d(A) = d_{||.||}(A, \mathcal{P}) = \min\{||E|| : A + E \text{ has property } \mathcal{P}\}.$$

A **matrix nearness problem** is ([High89]) to find an explicit formula for $d(A)$, the $\mathcal{P}$*-closest matrix* (or matrices) $X_{||.||}(A) = A + E$, satisfying the above minimum, and efficient algorithms for computing $d(A)$ and $X_{||.||}(A)$. The *componentwise nearness problem* is to find $d'(A) = \min\{\epsilon : |E| \leq \epsilon|A|, A + E$ has property $\mathcal{P}\}$, where $|B| = ((|b_{ij}|))$ and the matrix inequality is interpreted componentwise.

The most used norms for $B = ((b_{ij}))$ are the *Schatten* 2- and ∞-*norms* (cf. **Schatten norm metric**): the *Frobenius norm* $||B||_{Fr} = \sqrt{\text{Tr}(B^*B)} = \sqrt{\sum_{1\leq i\leq \text{rank}(B)} s_i^2}$ and the *spectral norm* $||B||_{sp} = \sqrt{\lambda_{\max}(B^*B)} = s_1(B)$.

Examples of closest matrices $X = X_{||.||}(A, \mathcal{P})$ follow.

Let $A \in \mathbb{C}^{n\times n}$. Then $A = A_H + A_S$, where $A_H = \frac{1}{2}(A + A^*)$ is Hermitian and $A_H = \frac{1}{2}(A - A^*)$ is *skew-Hermitian* (i.e., $A_H^* = -A_H$). Let $A = U\Sigma V^*$ be a *singular value decomposition* (SVD) of A, i.e., $U \in M_m$ and $V^* \in M_n$ are unitary, while $\Sigma = \text{diag}(s_1, s_2, \dots, s_{\min\{m,n\}})$ is an $m \times n$ diagonal matrix with $s_1 \geq s_2 \geq \dots \geq s_{\text{rank}(A)} > 0 = \dots = 0$. Fan and Hoffman, 1955, showed that, for any unitarily invariant norm, A_H, A_S, UV^* are closest Hermitian (symmetric), skew-Hermitian (skew-symmetric) and unitary (orthogonal) matrices, respectively. Such matrix $X_{Fr}(A)$ is a unique minimizer in all three cases.

Let $A \in \mathbb{R}^{n\times n}$. Gabriel, 1979, found the closest normal matrix $X_{Fr}(A)$. Higham found in 1988 a unique closest symmetric positive-semidefinite matrix $X_{Fr}(A)$ and, in 2001, the closest matrix of this type with unit diagonal (i.e., ab correlation matrix).

Given a SVD $A = U\Sigma V^*$ of A, let A_k denote $U\Sigma_k V^*$, where Σ_k is a diagonal matrix $\text{diag}(s_1, s_2, \dots, s_k, 0, \dots, 0)$ containing the largest k singular values of A. Then (Mirsky, 1960) A_k achieves $\min_{\text{rank}(A+E)\leq k} ||E||$ for any unitarily invariant norm. So, $||A - A_k||_{Fr} = \sqrt{\sum_{i=k+1}^{\text{rank}(A)} s_i^2}$ (Eckart–Young, 1936) and $||A - A_k||_{sp} = s_{max}(A - A_k) = s_{k+1}(A)$. A_k is a unique minimizer $X_{Fr}(A)$ if $s_k > s_{k+1}$.

Let $A \in \mathbb{R}^{n \times n}$ be nonsingular. Then its **distance to singularity** $d(A, Sing) = \min\{||E|| : A + E \text{ is singular}\}$ is, for both above norms, $s_n(A) = \frac{1}{s_1(A^{-1})} = \frac{1}{||A^{-1}||_{sp}} = \sup\{\delta : \delta\mathbb{B}_{\mathbb{R}^n} \subseteq A\mathbb{B}_{\mathbb{R}^n}\}$; here $\mathbb{B}_{\mathbb{R}^n} = \{x \in \mathbb{R}^n : ||x|| \leq 1\}$.

Given a closed convex cone $C \subseteq \mathbb{R}^n$, call a matrix $A \in \mathbb{R}^{m \times n}$ *feasible* if $\{Ax : x \in C\} = \mathbb{R}^m$; so, for $m = n$ and $C = \mathbb{R}^n$, feasibly means nonsingularity. Renegar, 1995, showed that, for feasible matrix A, its **distance to infeasibility** $\min\{||E||_{\text{nat}} : A + E \text{ is not feasible}\}$ is $\sup\{\delta : \delta\mathbb{B}_{\mathbb{R}^m} \subseteq A(\mathbb{B}_{\mathbb{R}^n} \cap C)\}$.

Lewis, 2003, generalized this by showing that, given two real normed spaces X, Y and a surjective *convex process* (or *set valued sublinear mapping*) F from X to Y, i.e., a multifunction for which $\{(x, y) : y \in F(x)\}$ is a closed convex cone, it holds

$$\min\{||E||_{\text{nat}} : E \text{ is any linear map } X \to Y, F + E \text{ is not surjective}\} = \frac{1}{||F^{-1}||_{\text{nat}}}.$$

Donchev et al. 2002, extended this, computing **distance to irregularity**; cf. **metric regularity** (Chap. 1). Cf. the above four *distances to ill-posedness* with **distance to uncontrollability** (Chap. 18) and **distances from symmetry** (Chap. 21).

- **$Sym(n, \mathbb{R})^+$ and $Her(n, \mathbb{C})^+$ metrics**

  Let $Sym(n, \mathbb{R})^+$ and $Her(n, \mathbb{C})^+$ be the cones of $n \times n$ symmetric real and Hermitian complex positive-definite $n \times n$ matrices. The $Sym(n, \mathbb{R})^+$ **metric** is defined, for any $A, B \in Sym(n, \mathbb{R})^+$, as

$$\left(\sum_{i=1}^{n} \log^2 \lambda_i\right)^{\frac{1}{2}},$$

where λ_1, c, λ_n are the *eigenvalues* of the matrix $A^{-1}B$ (the same as those of $A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$). It is the **Riemannian distance**, arising from the Riemannian metric $ds^2 = \text{Tr}((A^{-1}(dA))^2)$. This metric was rediscovered in Förstner–Moonen, 1999, and Pennec et al., 2004, via *generalized eigenvalue problem*: $det(\lambda A - B) = 0$.

The $Her(n, \mathbb{C})^+$ **metric** is defined, for any $A, B \in Her(n, \mathbb{C})^+$, by

$$d_R(A, B) = ||\log(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})||_{Fr},$$

where $||H||_{Fr} = (\sum_{i,j} |h_{ij}|^2)^{\frac{1}{2}}$ is the *Frobenius norm* of the matrix $H = ((h_{ij}))$. It is the **Riemannian distance** arising from the Riemannian metric of nonpositive curvature, defined locally (at H) by $ds = ||H^{-\frac{1}{2}}\, dH\, H^{-\frac{1}{2}}||_{Fr}$. In other words, this distance is the **geodesic distance**

$$\inf\{L(\gamma) : \gamma \text{ is a (differentiable) path from A to B}\},$$

where $L(\gamma) = \int_A^B ||\gamma^{-\frac{1}{2}}(t)\gamma'(t)\gamma^{-\frac{1}{2}}(t)||_{Fr}dt$ and the geodesic $[A, B]$ is parametrized by $\gamma(t) = A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^t A^{\frac{1}{2}}$ in the sense that $d_R(A, \gamma(t)) = td_R(A, B)$ for each $t \in [0, 1]$. In particular, the geodesic midpoint $\gamma(\frac{1}{2})$ of $[A, B]$ can be seen as the *geometric mean* of two positive-definite matrices A and B.

The space $(Her(n, \mathbb{C})^+, d_R))$ is an **Hadamard** (i.e., complete and CAT(0)) **space**, cf. Chap. 6. But $Her(n, \mathbb{C})^+$ is not complete with respect to matrix norms; it has a boundary consisting of the singular positive-semidefinite matrices.

Above $Sym(n, \mathbb{R})^+$ and $Her(n, \mathbb{C})^+$ metrics are the special cases of the distance $d_R(x, y)$ among **invariant distances on symmetric cones** in Chap. 9.

Cf. also, in Chap. 24, the **trace distance** on all Hermitian of trace 1 positive-definite $n \times n$ matrices and in Chap. 7, the **Wigner–Yanase–Dyson metrics** on all complex positive-definite $n \times n$ matrices.

The **Bartlett distance** between two matrices $A, B \in Her(n, \mathbb{C})^+$, is defined (Conradsen et al., 2003, for radar applications) by

$$\ln\left(\frac{(det(A + B))^2}{4det(A)det(B)}\right).$$

- **Siegel distance**

The *Siegel half-plane* is the set SH_n of $n \times n$ matrices $Z = X + iY$, where X, Y are symmetric or Hermitian and Y is positive-definite. The **Siegel–Hua metric** (Siegel, 1943, and independently, Hua, 1944) on SH_n is defined by

$$ds^2 = \text{Tr}(Y^{-1}(dZ)Y^{-1}(d\overline{Z})).$$

It is unique metric preserved by any automorphism of SH_n. The Siegel–Hua metric on the *Siegel disk* $SD_n = \{W = (Z - iI)(Z + iI)^{-1} : Z \in SH_n\}$ is defined by

$$ds^2 = \text{Tr}((I - WW^*)^{-1}dW(I - W^*W)^{-1}dW^*).$$

For n=1, the Siegel–Hua metric is the **Poincaré metric** (cf. Chap. 6) on the *Poincaré half-plane* SH_1 and the *Poincaré disk* SD_1, respectively.

Let $A_n = \{Z = iY : Y > 0\}$ be the imaginary axe on the Siegel half-plane. The Siegel–Hua metric on A_n is (cf. [Barb12]) the Riemannian **trace metric** $ds^2 = \text{Tr}((P^1 dP)^2)$. The corresponding distances are $Sym(n, \mathbb{R})^+$ **metric** or $Her(n, \mathbb{C})^+$ **metric**. The **Siegel distance** on $SH_n \setminus A_n$ is defined by

$$d^2_{Siegel}(Z_1, Z_2) = \sum_{i=1}^{n} \log^2(\frac{1 + \sqrt{\lambda_i}}{1 - \sqrt{\lambda_i}});$$

$\lambda_1, \ldots, \lambda_n$ are the *eigenvalues* of the matrix $(Z_1 - Z_2)(Z_1 - \overline{Z_2}){-}1(\overline{Z_1} - \overline{Z_2})(\overline{Z_1} - Z_2)^{-1}$.

- **Barbaresco metrics**

  Let $z(k)$ be a complex temporal (discrete time) *stationary* signal, i.e., its mean value is constant and its *covariance function* $\mathbb{E}[z(k_1)z^*(k_2)]$ is only a function of $k_1 - k_2$. Such signal can be represented by its covariance $n \times n$ matrix $R = ((r_{ij}))$, where $r_{ij} = \mathbb{E}[z(i), z*(j)] = \mathbb{E}[z(n)z*(n-i+j)]$. It is a positive-definite *Toeplitz* (i.e. diagonal-constant) Hermitian matrix. In radar applications, such matrices represent the Doppler spectra of the signal. Matrices R admit a parametrization (complex ARM, i.e., m-th order autoregressive model) by *partial autocorrelation coefficients* defined recursively as the complex correlation between the forward and backward prediction errors of the $(m-1)$-th order complex ARM.

  Barbaresco ([Barb12]) defined, via this parametrization, a **Bergman metric** (Chap. 7) on the bounded domain $\mathbb{R} + xD_n \subset \mathbb{C}^n$ of above matrices R; here D is a *Poincaré disk*. He also defined a related **Kähler metric** on $M \times S_n$, where M is the set of positive-definite Hermitian matrices and SD_n is the *Siegel disk* (cf. **Siegel distance**). Such matrices represent spatiotemporal stationary signals, i.e., in radar applications, the Doppler spectra and spatial directions of the signal.

  Ben Jeuris, 2015, extended above metrics on *block Toeplitz matrices*, i.e., those having blocks that are repeated (as elements of a Toeplitz matrix) down the diagonals of the matrix.

  Cf. **Ruppeiner metric** (Chap. 7) and **Martin cepstrum distance** (Chap. 21).

- **Distances between graphs of matrices**

  The *graph $G(A)$ of a complex $m \times n$ matrix A* is the *range* (i.e., the span of columns) of the matrix $R(A) = ([IA^T])^T$. So, $G(A)$ is a subspace of $\mathbb{C}^{m+n}$ of all vectors v, for which the equation $R(A)x = v$ has a solution.

  A **distance between graphs of matrices** A and B is a distance between the subspaces $G(A)$ and $G(B)$. It can be an **angle distance between subspaces** or, for example, the following distance (cf. also the **Kadets distance** in Chap. 1 and the **gap metric** in Chap. 18).

  The **spherical gap distance** between subspaces A and B is defined by

  $$\max\{\max_{x \in S(A)} d_E(x, S(B)), \max_{y \in S(B)} d_E(y, S(A))\},$$

  where $S(A), S(B)$ are the unit spheres of the subspaces A, B, $d(z, C)$ is the **point-set distance** $\inf_{y \in C} d(z, y)$ and $d_E(z, y)$ is the Euclidean distance.

- **Angle distances between subspaces**

  Consider the *Grassmannian space $G(m, n)$* of all n-dimensional subspaces of Euclidean space $\mathbb{E}^m$; it is a compact *Riemannian manifold* of dimension $n(m-n)$.

  Given two subspaces $A, B \in G(m, n)$, the *principal angles* $\frac{\pi}{2} \geq \theta_1 \geq \cdots \geq \theta_n \geq 0$ between them are defined, for $k = 1, \dots, n$, inductively by

  $$\cos \theta_k = \max_{x \in A} \max_{y \in B} x^T y = (x^k)^T y^k$$

  subject to the conditions $||x||_2 = ||y||_2 = 1, x^T x^i = 0, y^T y^i = 0$, for $1 \leq i \leq k-1$, where $||.||_2$ is the Euclidean norm.

The principal angles can also be defined in terms of orthonormal matrices Q_A and Q_B spanning subspaces A and B, respectively: in fact, n ordered singular values of the matrix $Q_A Q_B \in M_n$ can be expressed as cosines $\cos\theta_1, \dots, \cos\theta_n$.

The **Grassmann distance** between subspaces A and B of the same dimension is their geodesic distance defined by

$$\sqrt{\sum_{i=1}^{n} \theta_i^2}.$$

The **Martin distance** between subspaces A and B is defined by

$$\sqrt{\ln \prod_{i=1}^{n} \frac{1}{\cos^2\theta_i}}.$$

In the case when the subspaces represent ARMs (*autoregressive models*), the Martin distance can be expressed in terms of the *cepstrum* of the autocorrelation functions of the models. Cf. the **Martin cepstrum distance** in Chap. 21.

The **Asimov distance** between subspaces A and B is defined by θ_1. The **spectral distance** (or *chordal* 2-*norm distance*) is defined by $2\sin(\frac{\theta_1}{2})$.

The **containment gap distance** (or *projection distance*) is $\sin\theta_1$. It is the l_2-*norm* of the difference of the *orthogonal projectors* onto A and B. Many versions of this distance are used in Control Theory (cf. **gap metric** in Chap. 18).

The **Frobenius distance** and **chordal distance** between subspaces A and B are

$$\sqrt{2\sum_{i=1}^{n} \sin^2\theta_i} \quad \text{and} \quad \sqrt{\sum_{i=1}^{n} \sin^2\theta_i}, \quad \text{respectively.}$$

It is the *Frobenius norm* of the difference of above projectors onto A and B.

Similar distances $\sqrt{1 - \prod_{i=1}^{n} \cos^2\theta_i}$ and $\arccos(\prod_{i=1}^{n} \cos\theta_i)$ are called the **Binet–Cauchy distance** and (cf. Chap. 7) **Fubini–Study distance**, respectively.

- **Larsson–Villani metric**

Let A and B be two arbitrary orthonormal $m \times n$ matrices of full rank, and let θ_{ij} be the angle between the i-th column of A and the j-th column of B.

We call **Larsson–Villani metric** the distance between A and B (used by Larsson and Villani, 2000, for multivariate models) the square of which is defined by

$$n - \sum_{i=1}^{n} \sum_{j=1}^{n} \cos^2\theta_{ij}.$$

The square of usual Euclidean distance between A and B is $2(1-\sum_{i=1}^{n}\cos\theta_{ii})$. For $n=1$, above two distances are $\sin\theta$ and $\sqrt{2(1-\cos\theta)}$, respectively.

- **Lerman metric**

Given a finite set X and real symmetric $|X|\times|X|$ matrices $((d_1(x,y)))$, $((d_2(x,y)))$ with $x,y\in X$, their **Lerman semimetric** (cf. **Kendall τ distance** on permutations in Chap. 11) is defined by

$$|\{(\{x,y\},\{u,v\}):(d_1(x,y)-d_1(u,v))(d_2(x,y)-d_2(u,v))<0\}|\binom{|X|+1}{2}^{-2},$$

where $(\{x,y\},\{u,v\})$ is any pair of unordered pairs of elements x,y,u,v from X.

Similar **Kaufman semimetric** between $((d_1(x,y)))$ and $((d_2(x,y)))$ is

$$\frac{|\{(\{x,y\},\{u,v\}):(d_1(x,y)-d_1(u,v))(d_2(x,y)-d_2(u,v))<0\}|}{|\{(\{x,y\},\{u,v\}):(d_1(x,y)-d_1(u,v))(d_2(x,y)-d_2(u,v))\neq 0\}|}.$$