Mathematical Methods for Quantum Information Theory

Part I: Matrix Analysis

Koenraad Audenaert (RHUL, UK)

September 14, 2008

Preface

- Books on Matrix Analysis:
  - R. Bhatia, Matrix Analysis, Springer, 1997.
  - R.A. Horn and C.R. Johnson, Matrix Analysis, Cambridge, 1985.
  - R.A. Horn and C.R. Johnson, Topics in Matrix Analysis, Cambridge, 1991.
  - X. Zhan, Matrix Inequalities, Springer, 1999.
- Journals: Linear Algebra and its Applications, and others
- Purpose of Lecture: introduction to those aspects of matrix analysis that are/may be/have been useful in QIT
- No Proofs

Contents (roughly)

- Classes of matrices
- Operations and functions on matrices
- Matrix Decompositions
- Matrix Norms
- Applications in QIT: Schmidt decomposition, distance measures

- For this Lecture, I've set 10 questions, but don't do all of them.
- Marks are between square brackets [ ]. Hard questions receive more marks. Aim at a total of about 10 marks.

1. The very basics

Matrices

- If you don't know what a matrix is, there are 3 good movies about them.
- Matrix product is non-commutative: $AB \neq BA$
- In QIT, matrices occur whenever systems are finite-dimensional (in one way or another): density matrices, observables, Hamiltonians, POVMs, maps, channels.

Basic Matrix Operations

- Inverse: A^{-1}, satisfies $AA^{-1} = \mathbf{I}$; need not always exist
- Transpose: A^T, $(A^T)_{i,j} = A_{j,i}$
- Complex Conjugate: $\overline{A}$, $\overline{A}_{i,j} = \overline{A_{i,j}}$
- Hermitian Conjugate: $A^* = \overline{A^T}$
  - Note: in physics: $A^\dagger$, in engineering: A^H
  - Example: if $A = |\psi\rangle$, then $A^* = \langle\psi|$
- $(AB)^T = B^T A^T$
- $\overline{AB} = \overline{A}\ \overline{B}$
- $(AB)^* = B^* A^*$
- Trace: for square matrices $\mathrm{Tr}(A) = \sum_i A_{i,i}$
  - Linear: $\mathrm{Tr}(aA + bB) = a\,\mathrm{Tr}\,A + b\,\mathrm{Tr}\,B$
  - Cyclicity property: $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, $\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA)$,...

Matrix Classes

- Diagonal matrix: square matrix with non-zero elements on diagonal only: $A_{i,j} = a_i \delta_{i,j}$ or $A = \mathrm{Diag}(a_1, a_2, \ldots)$
- Identity matrix $\mathbf{I}$: diagonal matrix with all 1's on the diagonal: $\mathbf{I}_{i,j} = \delta_{i,j}$
- Scalar matrix: $A = a\mathbf{I}$
- Hermitian matrix: $A = A^*$
- Positive semi-definite (PSD) matrix: a matrix that has square root(s)

$$A \geq 0 \iff \exists B : A = B^* B$$

- Unitary matrix: square matrix U with $U^* U = \mathbf{I}$
- Projector: a Hermitian matrix equal to its own square: $P = P^2$.

Characterisations

- Examples of Hermitian matrices: observables, Hamiltonians
- Example of PSD matrices: density matrices; e.g. $A = |\psi\rangle\langle\psi|$: $B^* = |\psi\rangle$
- Examples of unitary matrices: any evolution operator, Pauli matrices, CNOT
- A matrix A is Hermitian iff all its expectation values are real:
  $\forall\psi : \langle\psi|A|\psi\rangle \in \mathbb{R}$
- A matrix A is PSD iff all its expectation values are real and non-negative:
  $\forall\psi : \langle\psi|A|\psi\rangle \geq 0.$
- **Exercise 1 [3]**: prove this last statement from the definition of PSD.
- A matrix is unitary iff its column vectors form an orthonormal basis
- For square U, $U^*U = \mathbf{I}$ implies $UU^* = \mathbf{I}$

2. The Density Matrix Formalism

Dealing with Statistical Uncertainty

- State vectors are used mainly in undergraduate QM courses and in quantum field theory.
- In real experiments, we have to deal with many uncertainties and uncontrollable factors.
- E.g. preparation of a particle in some state is never perfect. What we get is $\psi = (\cos\alpha, \sin\alpha)^T$, with some α close to the desired value, but with errors.
- How can we efficiently deal with those and other errors in QM?
- Naïve method: specify distribution of parameters (α) or of state itself.
- That's both complicated and unnecessary. What we can measure are only expectation values, like $\langle\psi|\hat{O}|\psi\rangle$.

Dealing with Statistical Uncertainty

- Because of statistical uncertainty on ψ, expressed by the probability density $p(\psi)d\psi$, we measure $\int d\psi\, p(\psi)\langle\psi|\hat{O}|\psi\rangle$.
- Rewrite this as $\mathrm{Tr}[(\int d\psi\, p(\psi)|\psi\rangle\langle\psi|)\hat{O}]$.
- We can calculate all expectation values, once we know the matrix

$$\int d\psi\, p(\psi)|\psi\rangle\langle\psi|.$$

- Hence, this is "the" state! We call it the *density matrix* (cf. probability density). Usual symbol ρ.
- **Exercise 2 [2]**: Prove that a density matrix is PSD and has trace 1.

Dealing with Statistical Uncertainty

- A set of state vectors ψ_i with given probabilities p_i is called an *ensemble*.
- A density matrix is the barycenter of the ensemble.
- Different ensembles may yield the same density matrix:

$$\{p_1 = 1/2, \psi_1 = (1,0) \ \ , \ p_2 = 1/2, \psi_2 = (0,1)\}$$
$$\{p_1 = 1/2, \psi_1 = (1,1)/\sqrt{2} \ \ , \ p_2 = 1/2, \psi_2 = (1,-1)/\sqrt{2}\}$$

  both yield the density matrix $\rho = \mathbf{I}/2$, the *maximally mixed state*.
- We can never figure out which ensemble a density matrix originated from!
- A state with density matrix of the form $\rho = \psi\psi^* = |\psi\rangle\langle\psi|$ is a *pure state* and corresponds to a state vector ψ.
- Otherwise, we call the state a *mixed state* (cf. statistical mixing).

3. Tensor Products, Partial Traces and Partial Transposes

Tensor Product of Vectors

- Suppose I have 2 independent particles. Particle 1 is in state ϕ, and particle 2 in state θ.
- The particles taken together are then in the state ψ, which is the tensor product of ϕ and θ.
- Notation $|\psi\rangle = |\phi \otimes \theta\rangle = |\phi\rangle|\theta\rangle$.
- E.g. $(1, 2) \otimes (3, 4) = (3, 4, 6, 8)$.
- Note the order! "The indices of particle 2 change fastest"

$$\psi = (\phi_\uparrow, \phi_\downarrow) \otimes (\theta_\uparrow, \theta_\downarrow) = (\psi_{\uparrow\uparrow}, \psi_{\uparrow\downarrow}, \psi_{\downarrow\uparrow}, \psi_{\downarrow\downarrow}), \text{ with } \psi_{ij} = \phi_i \theta_j.$$

- To do the same for matrices, it is beneficial to use block matrix notation.

Block matrices

- A block matrix can be seen as being a matrix whose elements are matrices themselves (of equal size).
- Example: $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$.
- Indexing is more complicated. We need row and column indexes to single out a block, and row and column indexes to single out an element within that block. Hence the need to use *composite indices*.
- Let i, k be (row/col) indexes pointing to a block, and j, l indexes pointing within the block. Then (i, j) denotes a *composite* row index, and (k, l) a composite column index.
- The elements of a block matrix can then be denoted by $A_{(i,j),(k,l)}$, and

$$A = \sum_{i,j,k,l} A_{(i,j),(k,l)} \ |i\rangle|j\rangle \ \langle k|\langle l| = \sum_{i,j,k,l} A_{(i,j),(k,l)} \ |i\rangle\langle k| \otimes |j\rangle\langle l|.$$

Tensor Product of Matrices

- The tensor product, a.k.a. *Kronecker Product*, of matrices A and B, $A \otimes B$, can be represented by a block matrix with elements

$$(A \otimes B)_{(i,j),(k,l)} = A_{i,k} B_{j,l}$$

- E.g. when A is 2×2

$$A \otimes B = \begin{pmatrix} A_{11}B & A_{12}B \\ A_{21}B & A_{22}B \end{pmatrix}$$

- Trace rule: $\mathrm{Tr}(A \otimes B) = \mathrm{Tr}(A)\,\mathrm{Tr}(B)$

Partial Trace

- To "ignore" a particle in a group of particles in a given state, "trace out" that particle.
- The *partial trace* of the ith factor in a tensor product is obtained by replacing the ith factor with its trace:

$$\begin{aligned}\mathrm{Tr}_1(A \otimes B) &= \mathrm{Tr}(A) \otimes B = \mathrm{Tr}(A)B \\ \mathrm{Tr}_2(A \otimes B) &= A \otimes \mathrm{Tr}(B) = \mathrm{Tr}(B)A\end{aligned}$$

- In block matrix form:

$$\begin{aligned}\mathrm{Tr}_1(A \otimes B) &= \mathrm{Tr}_1 \begin{pmatrix} A_{11}B & A_{12}B \\ A_{21}B & A_{22}B \end{pmatrix} = A_{11}B + A_{22}B \\ \mathrm{Tr}_2(A \otimes B) &= \mathrm{Tr}_2 \begin{pmatrix} A_{11}B & A_{12}B \\ A_{21}B & A_{22}B \end{pmatrix} = \begin{pmatrix} A_{11}\,\mathrm{Tr}\,B & A_{12}\,\mathrm{Tr}\,B \\ A_{21}\,\mathrm{Tr}\,B & A_{22}\,\mathrm{Tr}\,B \end{pmatrix}\end{aligned}$$

Partial Trace

- By linearity of the trace, this extends to all block matrices:

$$\mathrm{Tr}_1 \begin{pmatrix} A & B \\ C & D \end{pmatrix} = A + D$$

$$\mathrm{Tr}_2 \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \mathrm{Tr}\, A & \mathrm{Tr}\, B \\ \mathrm{Tr}\, C & \mathrm{Tr}\, D \end{pmatrix}$$

- Equivalent definition:

$$\begin{aligned} \mathrm{Tr}((\mathbf{I} \otimes X)A) &= \mathrm{Tr}(X \,\mathrm{Tr}_1 A), \forall X \\ \mathrm{Tr}((X \otimes \mathbf{I})A) &= \mathrm{Tr}(X \,\mathrm{Tr}_2 A), \forall X \end{aligned}$$

- **Exercise [1000]**: Relate the eigenvalues of A to those of $\mathrm{Tr}_1 A$ and $\mathrm{Tr}_2 A$.

Partial Transpose

- Another "partial" operation on block matrices is the partial transpose.
- Take again a 2-qubit state with density matrix ρ written as a block matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$.
- The partial transpose w.r.t. the first particle: $\rho^{\Gamma_1} = \begin{pmatrix} A & C \\ B & D \end{pmatrix}$.
- The partial transpose w.r.t. the second particle: $\rho^{\Gamma_2} = \begin{pmatrix} A^T & B^T \\ C^T & D^T \end{pmatrix}$.
- The partial transpose of a state need no longer be a state; it is if ρ is separable.

4. Completely Positive (CP) maps

Operations on States

- There exist various ways of operating on states:
  - Unitary evolution: $|\psi\rangle \to U|\psi\rangle$
  - Adding particles (in a determined state): $|\psi\rangle \to |\psi\rangle \otimes |0\rangle$
  - Removing/ignoring particles: $|\psi\rangle\langle\psi| \to \mathrm{Tr}_1\,|\psi\rangle\langle\psi|$
  - Measurements: $|\psi\rangle \to \langle\psi|E_i|\psi\rangle$
  - Combinations thereof
  - Measurement outcomes may even determine the choice of subsequent operations
- Absolutely astonishing fact about QM #31: all of this can be combined into one simple formula!

Operations on States

- Every quantum operation, composed of the above basic operations, can be written as a *completely positive, trace preserving, linear map* or *CPT map* Φ acting on the density matrix: $\rho \mapsto \Phi(\rho)$
- Completely positive = positivity preserving when acting on any subset of the state's particles: because a state should remain a state.
- Non-example: The matrix transpose is a positive, trace preserving linear map, but not a completely positive one: when it acts on 1 particle of an EPR state, one gets a non-positive matrix.

Characterisation of CP(T) maps

- By dropping the trace-preservation requirement, we get a CP map.
- Any linear map can be represented using its *Choi-matrix* $\mathbf{\Phi}$:
  - A block matrix with $d_{in} \times d_{in}$ blocks of size $d_{out} \times d_{out}$
  - Block i, j of $\mathbf{\Phi}$ is given by $\Phi(|i\rangle\langle j|)$
  - $\Phi(\rho) = \sum_{i,j} \rho_{ij} \Phi(|i\rangle\langle j|) = \mathrm{Tr}_1[\mathbf{\Phi}.(\rho^T \otimes \mathbf{I})]$.
- A map Φ is CP if and only if its Choi-matrix $\mathbf{\Phi}$ is PSD [Choi].
- **Exercise 3 [8]**: Prove this. Hint: operate the map on one particle of the EPR state $\psi = \sum_{i=1}^{d_{in}} |i\rangle|i\rangle$.
- **Exercise 4 [5]**: Find the Choi matrix of matrix transposition (for qubit states) and use it to show why transposition is not a CP map.

Characterisation of CP(T) maps

- Since the Choi-matrix is a block matrix, we can define its partial traces: $\mathrm{Tr}_1 = \mathrm{Tr}_{in}$ and $\mathrm{Tr}_2 = \mathrm{Tr}_{out}$
- **Exercise 5 [4]**: Show that a CP map is trace preserving if and only $\mathrm{Tr}_{out}\, \mathbf{\Phi} = \mathbf{I}$.

5. Matrix Decompositions

Matrix Functions

- Problem: to calculate von Neumann entropy $S(\rho) = -\operatorname{Tr}[\rho \log \rho]$, we need to calculate functions of matrices.
- Analytic functions can be represented by (formal) power series $f(z) = \sum_{k=0} a_k z^k$.
- Since we know how to multiply matrices we can calculate $\sum_{k=0} a_k A^k$
- This (formally) defines a matrix function $f(A)$
- Example: $\exp(A) = \sum_{k=0} A^k/k!$
- Series are not the most convenient way to work with matrix functions

Eigenvalues

- Many of the presented concepts get "easier" descriptions when the matrix has an eigenvalue decomposition.
- Eigenvalue/eigenvector: $Ax = \lambda x$, $\det(A - \lambda \mathbf{I}) = 0$.
- Stack $x^{(i)}$ columnwise in matrix S, and λ_i in diagonal matrix Λ:
  $AS = S\Lambda$
- If S is invertible, we get $A = S\Lambda S^{-1}$
- A matrix is *diagonalisable* if there exists an invertible S such that $S^{-1}AS$ is diagonal.
- A matrix is *unitarily diagonalisable* if there exists a unitary U such that $U^{-1}AU = U^*AU$ is diagonal; then $A = U\Lambda U^*$.

Eigenvalues

- Theorem: A matrix A is unitarily diagonalisable (UD) iff the matrix is normal ($AA^* = A^*A$)
- The eigenvalue decomposition (EVD) of a normal matrix A is $A = U\Lambda U^*$
- A Hermitian matrix is UD, with real eigenvalues
- A PSD matrix is UD, with non-negative eigenvalues
- A Projector ($P = P^2$) has eigenvalues ...

Matrix Functions

- Matrix functions of Hermitian (or PSD) matrices: $f(A) = U f(\Lambda) U^*$, where f operates entrywise on the diagonal elements (eigenvalues)
- Example: for PSD A, with $A = U \Lambda U^*$, THE square root is

$$A^{1/2} = U \,\mathrm{Diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots) U^*$$

- Matrix absolute value (or modulus): $|A| = (A^* A)^{1/2}$
- **Exercise 6 [3]**: For Hermitian H, express $\mathrm{Tr}(H)$, $|H|$ and $\mathrm{Tr}\,|H|$ in terms of its EVD. For $A \geq 0$, what is $|A|$?

Singular values

- Not all square matrices are diagonalisable, and none of the non-square matrices are.
- All matrices, even the non-square ones, have a singular value decomposition (SVD), and it is essentially unique: $A = U\Sigma V^*$, where U and V are unitary and Σ is "diagonal".
- One can find U and V s.t. the diagonal elements of Σ are non-negative reals and sorted in non-ascending fashion; then the diagonal elements of Σ, $\sigma_i(A)$, are the singular values of A.
- Use a computer with (Matlab, Maple, Mathematica)
- **Exercise 7 [3]**: show that for $A \geq 0$, $\sigma_i(A) = \lambda_i(A)$.

Singular values and Rank

- One of the ways to check invertibility of a square matrix is to inspect its singular values: A is invertible iff all $\sigma_i(A) > 0$, strictly.
- The number of non-zero singular values of A equals the *rank* of $A =$ the number of independent column (or row) vectors of A.
- The density matrix of a pure state has rank 1.

Schmidt decomposition

- Tilde notation: converts a pure bipartite state vector ψ to a matrix, denoted $\tilde{\psi}$.
- For $\psi = \sum_{i=1..d_1, j=1..d_2} x_{i,j}|i\rangle|j\rangle$, $\tilde{\psi}$ is the matrix $x_{i,j}$.
- Example: 2-qubit states $\psi = (\psi_1, \psi_2, \psi_3, \psi_4)$

$$\begin{aligned} |1\rangle|1\rangle &= (1,0,0,0) \qquad |1\rangle|2\rangle = (0,1,0,0) \\ |2\rangle|1\rangle &= (0,0,1,0) \qquad |2\rangle|2\rangle = (0,0,0,1) \end{aligned}$$

- Thus $\tilde{\psi} = \begin{pmatrix} \psi_1 & \psi_2 \\ \psi_3 & \psi_4 \end{pmatrix}$.
- In general, for vectors u and v, $\widetilde{u \otimes v} = uv^T$.
- In bra-ket notation: $\widetilde{|u\rangle|v\rangle} = |u\rangle\langle\overline{v}|$.

Schmidt decomposition

- Schmidt decomposition: one can find orthonormal bases $\{u_k\}$ and $\{v_k\}$ of the two systems, respectively, and non-negative Schmidt coefficients σ_k, such that

$$\psi = \sum_k \sigma_k |u_k\rangle |v_k\rangle$$

- Proof: The Schmidt decomposition of ψ is the SVD of the matrix $\tilde{\psi}$.
- Write $\tilde{\psi} = U \Sigma V^*$, and denote the k-th column vector of U by $|u_k\rangle$, of V by $|v_k\rangle$. Denote the k-th diagonal element of Σ by σ_k.
- Because U is unitary, the $\{|u_k\rangle\}$ are orthonormal. Same for V.

Schmidt decomposition

- Then

$$\tilde{\psi} = \sum_k \sigma_k |u_k\rangle\langle v_k|,$$

so that

$$\psi = \sum_k \sigma_k |u_k\rangle|\overline{v_k}\rangle.$$

- Product state: if $\psi = \phi \otimes \theta$, then $\tilde{\psi} = \phi\theta^T$, which has rank 1. I.e. only 1 non-zero Schmidt coefficient.
- A pure states ψ is entangled iff $\tilde{\psi}$ has rank > 1, i.e. has more than 1 non-zero Schmidt coefficients.

6. Matrix Norms

Matrix Norms

- A matrix norm $|||A|||$ is a mapping from the space of matrices to $\mathbb{R}_+$ obeying:
  - $|||A||| = 0$ iff $A = 0$
  - Homogeneous: $|||zA||| = |z|\,|||A|||$
  - Triangle inequality: $|||A + B||| \leq |||A||| + |||B|||$
  - Submultiplicative: $|||AB||| \leq |||A|||\,|||B|||$
- Of particular interest are the unitarily invariant (UI) matrix norms: $|||UAV||| = |||A|||$, i.e. they depend only on $\sigma(A)$

UI Matrix Norms

- Operator norm: $||A|| = ||A||_\infty = \sigma_1(A)$, largest singular value
- Trace norm: $||A||_{\mathrm{Tr}} = ||A||_1 = \sum_{i=1}^n \sigma_i(A) = \mathrm{Tr}\,|A|$
- Frobenius or Hilbert-Schmidt norm:
  $||A||_2 = (\sum_{i=1}^n \sigma_i^2(A))^{1/2} = (\mathrm{Tr}\,|A|^2)^{1/2} = (\sum_{i,j} |A_{i,j}|^2)^{1/2}$
- Schatten q-norms: $||A||_q = (\sum_{i=1}^n \sigma_i^q(A))^{1/q} = (\mathrm{Tr}\,|A|^q)^{1/q}$

Matrix norms in QIT

- Matrix norms are important in QIT for many reasons
- A Schatten norm of a state is a measure of its purity: it is 1 for pure states, and strictly less than 1 for mixed states. Minimal for maximally mixed state.
- **Exercise 8 [2]**: calculate the Schatten q-norm of the d-dimensional maximally mixed state $\mathbf{I}_d/d$.
- The entanglement measure "negativity" is defined as the trace norm of the partial transpose of a bipartite state: $N = ||\rho^\Gamma||_1 = \mathrm{Tr}\,|\rho^\Gamma|$.
- **Exercise 9 [3]**: Show that N is equal to 1 minus 2 times the sum of the negative eigenvalues of ρ^Γ.
- Matrix norms of $\rho - \sigma$ can be used as a distance measure between states: the states are equal iff their difference has zero norm.

Matrix norms in QIT

- The von Neumann entropy $S(\rho)$ is closely related to the Schatten q norms.
- Note the following:

$$\frac{d}{dq}x^q = x^q \log x$$

- Thus

$$x \log x = \frac{d}{dq}|_{q=1} x^q$$

and

$$S(\rho) = -\operatorname{Tr} \rho \log \rho = -\frac{d}{dq}|_{q=1} \operatorname{Tr} \rho^q$$

- Note that $\operatorname{Tr} \rho^q = (||\rho||_q)^q$.

Matrix norms in QIT

- Alternative relation:

$$-x \log x = \lim_{q \to 1} \frac{x - x^q}{q - 1}$$

thus

$$S(\rho) = \lim_{q \to 1} \frac{1 - \operatorname{Tr} \rho^q}{q - 1}$$

- One more tool for proving things about entropy

7. Distance measures between states

Need for State Distance Measures

- Example 1. Given an initial state ρ, and a class of maps, find the map Φ such that $\Phi(\rho)$ comes as close to a desired σ as possible.
- Example 2. L. Hardy's "Crazy Qubits": find the "best" physical (CPT) approximation of non-physical (non-CP, non-TP, non-linear) maps
- We could ask for the approximating map's output states to be as close to the hypothetical map's output states as possible.
- $\rightarrow$ Thus the need for distance measures between states.

Linear Fidelity

- Pure states: *Overlap* = *Linear Fidelity*:

$$F_L(\psi, \phi) = |\langle\psi|\phi\rangle|^2 = \mathrm{Tr}[|\psi\rangle\langle\psi|\ |\phi\rangle\langle\phi|].$$

  This is 1 iff $|\psi\rangle\langle\psi| = |\phi\rangle\langle\phi|$, and less than 1 otherwise. It is 0 for orthogonal states.

- For mixed states, the linear fidelity $F_L(\rho, \sigma) = \mathrm{Tr}[\rho\sigma]$ is not very useful.

- **Exercise 10 [2]**: why not?

Uhlmann Fidelity

- For mixed states, we can use the Uhlmann fidelity.

$$F_U(\rho, \sigma) = \mathrm{Tr}\,\sqrt{\rho^{1/2}\sigma\rho^{1/2}}.$$

- Characterisation: $(F_U)^2$ is the linear fidelity between *purifications*

$$\begin{aligned} F_U(\rho, \sigma) \;=\; \max_{\psi,\phi}\{|\langle\psi|\phi\rangle| :& \\ &\mathrm{Tr}_2(|\psi\rangle\langle\psi|) = \rho, \\ &\mathrm{Tr}_2(|\phi\rangle\langle\phi|) = \sigma\}. \end{aligned}$$

- F_U coincides with $\sqrt{F_L}$ when ρ or σ is pure.
- *Bures Distance*: $D_B(\rho, \sigma) = 2\sqrt{1 - F_U(\rho, \sigma)}$.

Trace Distance

- *Trace Distance*: $T(\rho, \sigma) = ||\rho - \sigma||_1/2$. Between 0 and 1.
- Obeys triangle inequality.
- Easier than Bures distance.
- Statistical interpretation: error probability of optimal POVM for distinguishing between ρ and σ is

$$P_e = (1 - T(\rho, \sigma))/2.$$

- Behaves “erratically” under tensor powers. One can find states such that

$$\begin{aligned} T(\rho, \sigma) &< T(\tau, \upsilon) \\ \text{but } T(\rho \otimes \rho, \sigma \otimes \sigma) &> T(\tau \otimes \tau, \upsilon \otimes \upsilon) \end{aligned}$$

Relative Entropy

- *Relative Entropy*: $S(\rho||\sigma) = \mathrm{Tr}\,\rho(\log\rho - \log\sigma)$.
- Statistical interpretation: error exponent of optimal asymmetric hypothesis test.
- Behaves nicely under tensor powers:

$$S(\rho^{\otimes n}||\sigma^{\otimes n}) = n\,S(\rho||\sigma).$$

- Relative entropy does not obey triangle inequality.
- Asymmetric in its arguments.
- For pure states, either 0 (same states) or infinite (different states).
- *Quantum Chernoff Distance* combines the best of T and S. It is the regularisation of T w.r.t. taking tensor powers.

The Quantum Chernoff distance

- Recall: error probability of optimal POVM for distinguishing between ρ and σ is

$$P_e = (1 - T(\rho, \sigma))/2.$$

- Now do the same for n copies of the two states: error probability of the optimal POVM is

$$P_e = (1 - T(\rho^{\otimes n}, \sigma^{\otimes n}))/2.$$

- This P_e goes down exponentially with n at a rate

$$\lim_{n \to \infty} -\frac{1}{n} \log(1 - T(\rho^{\otimes n}, \sigma^{\otimes n})).$$

The Quantum Chernoff distance

- A closed-form expression of the rate is given by the quantum Chernoff distance:

$$-\log Q, \quad \text{with } Q = \min_{0 \le s \le 1} \mathrm{Tr}[\rho^s \sigma^{1-s}].$$

- Just like the relative entropy, $-\log Q$ is multiplicative:

$$-\log Q(\rho^{\otimes n}, \sigma^{\otimes n}) = -n \log Q(\rho, \sigma).$$

- For pure states, $-\log Q$ attains all values between 0 and ∞. For equal states 0, for orthogonal states ∞.
- Does not obey triangle inequality.