Mathematical preliminaries

I. LINEAR ALGEBRA

A. Definition of matrix

Definition 1 ***Complex Euclidean spaces***.

Let Σ be an alphabet, then the set of all functions from Σ to the complex numbers $\mathbb{C}$, denoted $\mathbb{C}^{\Sigma}$, forms a vector space of dimension $|\Sigma|$ with

1. *Addition:* $(u+v)(a) = u(a) + v(a)\ \forall a \in \Sigma$.
2. *Scalar multiplication:* $(\alpha u)(a) = \alpha u(a)\ \forall \alpha \in \mathbb{C}\ \forall a \in \Sigma$.

- *Inner product is defined as*

$$\langle u, v\rangle = \sum_{a\in\Sigma} \overline{u(a)} v(a) \tag{1}$$

- *Euclidean norm of vectors is defined as*

$$\|u\| = \sqrt{\langle u, u\rangle} = \left(\sum_{a\in\Sigma} |u(a)|^2\right)^{\frac{1}{2}} \tag{2}$$

- *p-norm of vectors is defined as*

$$\|u\|_p = \left(\sum_{a\in\Sigma} |u(a)|^p\right)^{\frac{1}{p}} \tag{3}$$

- *Direct sum of complex Euclidean spaces*

$$\mathcal{X}_1 \oplus \cdots \oplus \mathcal{X}_n = \mathbb{C}^{\Sigma_1 \cup \cdots \cup \Sigma_n} \tag{4}$$

- *Tensor product of complex Euclidean spaces*

$$\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n = \mathbb{C}^{\Sigma_1 \times \cdots \times \Sigma_n} \tag{5}$$

Theorem 1 *Let $\mathcal{X}$ and $\mathcal{Y}$ be linear spaces. Then*

$$\dim(\mathcal{X}) + \dim(\mathcal{Y}) = \dim(\mathcal{X}+\mathcal{Y}) + \dim(\mathcal{X}\cap\mathcal{Y}) \tag{6}$$

Definition 2 *Linear transformations and matrices.*

- *Linear operators are linear maps $A : \mathcal{X} \to \mathcal{Y}$ betweem two vector spaces such that*

$$\forall u, v \in \mathcal{X} \quad \forall \lambda, \mu \in \mathbb{C} \quad A(\lambda u + \mu v) = \lambda A u + \mu A v \tag{7}$$

  The collection of all linear mappings of the above form is denoted $\mathsf{Lin}(\mathcal{X}, \mathcal{Y})$.

- *Matrices*

$$M : \Gamma \times \Sigma \to \mathbb{C} \tag{8}$$

- *Let $M : \Gamma \times \Lambda \to \mathbb{C}$ and $N : \Lambda \times \Sigma \to \mathbb{C}$. $MN : \Gamma \times \Sigma$ is defined as*

$$(MN)(a, b) = \sum_{c\in\Lambda} M(a, c) N(c, b) \tag{9}$$

Each operator $A : \mathbb{C}^\Sigma \to \mathbb{C}^\Gamma$ is associated with a matrix $M : \Gamma \times \Sigma \to \mathbb{C}$

$$M(a,b) = \langle e_a, Ae_b \rangle \tag{10}$$

And conversely, a matrix $M : \Gamma \times \Sigma \to \mathbb{C}$ represents an operator $A : \mathbb{C}^\Sigma \to \mathbb{C}^\Gamma$

$$(Au)(a) = \sum_{b\in\Sigma} M(a,b)u(b) \qquad \forall u \in \mathbb{C}^\Sigma \ \forall a \in \Gamma \tag{11}$$

Let $A : \mathbb{C}^\Lambda \to \mathbb{C}^\Gamma$ and $B : \mathbb{C}^\Sigma \to \mathbb{C}^\Lambda$. Let $M : \Gamma \times \Lambda \to \mathbb{C}$ and $N : \Lambda \times \Sigma \to \mathbb{C}$ be the corresponding matrices. Then $\forall a \in \Gamma$

$$\begin{aligned}
(ABu)(a) &= \sum_{c\in\Lambda} M(a,c)(Bu)(c) && (12)\\
&= \sum_{c\in\Lambda} M(a,c) \sum_{b\in\Sigma} N(c,b)u(b) && (13)\\
&= \sum_{b\in\Sigma} \left(\sum_{c\in\Lambda} M(a,c)N(c,b) \right) u(b) && (14)
\end{aligned}$$

Thus MN corresponds to the composite operator AB.

The standard basis $\{E_{a,b} : a \in \Gamma,\ b \in \Sigma\}$ of a space of operators

$$E_{a,b}(c,d) = \begin{cases} 1 & (c,d) = (a,b) \\ 0 & \text{otherewise} \end{cases} \tag{15}$$

The number of elements in this basis is, of course, consistent with the fact that the dimension of $\mathsf{Lin}(\mathcal{X}, \mathcal{Y})$ is given by $\dim(\mathsf{Lin}(\mathcal{X}, \mathcal{Y})) = \dim(\mathcal{X}) \dim(\mathcal{Y})$.

Remark 1 *Generally, matrices and linear maps can be viewed as the same thing.*

B. General properties of matrices

Definition 3 *The entry-wise conjugate, transpose, and adjoint of* $A : \Gamma \times \Lambda \to \mathbb{C}$

$$\begin{aligned}
\overline{A}(a,b) &= \overline{A(a,b)} && (16)\\
A^T(b,a) &= A(a,b) && (17)\\
\langle v, Au \rangle &= \langle A^* v, u \rangle && (18)
\end{aligned}$$

Property 1 $A^*(a,b) = \overline{A(b,a)}$

$$\begin{aligned}
\langle v, Au \rangle = \sum_{a\in\Gamma} \overline{v(a)}(Au)(a) &= \sum_{b\in\Sigma} \overline{(A^*v)(b)}u(b) = \langle A^* v, u \rangle \\
\Longleftrightarrow \sum_{a\in\Gamma} \overline{v(a)} \sum_{b\in\Sigma} A(a,b)u(b) &= \sum_{b\in\Sigma} \sum_{a\in\Gamma} \overline{A^*(b,a)v(a)}u(b) \\
\Longleftrightarrow A^*(b,a) = \overline{A(a,b)}
\end{aligned}$$

Definition 4 *Kernel, image and rank of* $A \in \mathsf{Lin}(\mathcal{X}, \mathcal{Y})$

$$\begin{aligned}
\mathsf{ker}(A) &= \{u \in \mathcal{X} : Au = 0\} && (19)\\
\mathsf{im}(A) &= \{Au : u \in \mathcal{X}\} && (20)\\
\mathsf{rank}(A) &= \dim(\mathsf{im}(A)) && (21)
\end{aligned}$$

Theorem 2 *Useful formulas*

1. *Let* $A \in \mathsf{Lin}(\mathcal{X}, \mathcal{Y})$*, then*

$$\dim(\mathsf{ker}(A)) + \mathsf{rank}(A) = \dim(\mathcal{X}) \tag{22}$$

   Proof. *Let* $\{v_1, \cdots, v_k, v_{k+1}, \cdots, v_n\}$ *be a basis of* $\mathcal{X}$ *such that*

$$\mathsf{Span}\{v_1, \cdots, v_k\} = \mathsf{ker}(A) \qquad \mathsf{Span}\{Av_{k+1}, \cdots, Av_n\} = \mathsf{im}(A) \tag{23}$$

$$\begin{aligned} & c_{k+1}Av_{k+1} + \cdots + c_n Av_n = 0 && (24)\\ \implies & A(c_{k+1}v_{k+1} + \cdots + c_n v_n) = 0 && (25)\\ \implies & c_{k+1}v_{k+1} + \cdots + c_n v_n = c_1 v_1 + \cdots + c_k v_k && (26)\\ \implies & c_1 = \cdots = c_n = 0 && (27) \end{aligned}$$

   Thus $\{Av_{k+1}, \cdots, Av_n\}$ *is a linear-independent set. Then*

$$\mathsf{rank}(A) = \dim(\mathsf{im}(A)) = \dim(\mathcal{X}) - \dim(\mathsf{ker}(A)) \tag{28}$$

2. *Let* $A \in \mathsf{Lin}(\mathcal{Y}, \mathcal{Z})$ *and* $B \in \mathsf{Lin}(\mathcal{X}, \mathcal{Y})$*. Then*

$$\mathsf{rank}(A) + \mathsf{rank}(B) - \dim(\mathcal{Y}) \le \mathsf{rank}(AB) \le \min(\mathsf{rank}(A), \mathsf{rank}(B)) \tag{29}$$

3. *Let* $A, B \in \mathsf{Lin}(\mathcal{X}, \mathcal{Y})$*, then*

$$\mathsf{rank}(A + B) \le \mathsf{rank}(A) + \mathsf{rank}(B) \tag{30}$$

4. $\mathsf{ker}(A) = \mathsf{ker}(A^*A)$
   Proof.

$$A^*Au = 0 \implies u^*A^*Au = 0 \implies |Au|^2 = 0 \implies Au = 0 \tag{31}$$

$$Au = 0 \implies A^*Au = A^*0 = 0 \tag{32}$$

5. $\mathsf{rank}(A) = \mathsf{rank}(AA^*) = \mathsf{rank}(A^*A)$
   Proof.

$$\mathsf{ker}(A) = \mathsf{ker}(A^*A) \implies \mathsf{rank}(A^*A) = \mathsf{rank}(A) \tag{33}$$

$$\mathsf{ker}(A^*) = \mathsf{ker}(AA^*) \implies \mathsf{rank}(AA^*) = \mathsf{rank}(A^*) = \mathsf{rank}(A) \tag{34}$$

6. $\mathsf{im}(A) = \mathsf{im}(AA^*)$
   Proof.

$$\begin{cases} \mathsf{im}(AA^*) \subset \mathsf{im}(A) \\ \mathsf{rank}(AA^*) = \mathsf{rank}(A) \end{cases} \implies \mathsf{im}(AA^*) = \mathsf{im}(A) \tag{35}$$

Definition 5 *Identity operator, inverse operator, trace, inner product, determinant.*

- *Identity operator is defined as*

$$\mathbb{1}(a, b) = \begin{cases} 1 & a = b \\ 0 & a \neq b \end{cases} \tag{36}$$

- *An operator* $A \in \mathsf{Lin}(\mathcal{X})$ *is invertible iff* $\exists B \in \mathsf{Lin}(\mathcal{X}),\ BA = I$*. It can be shown that* $BA = I \implies AB = I$*. Thus* $B = A^{-1}$ *is unique.*

- *Trace of* $X \in \mathsf{Lin}(\mathbb{C}^\Sigma)$ *is defined as*

$$\mathrm{Tr}(X) = \sum_{a \in \Sigma} X(a,a) \tag{37}$$

- *For an arbitrary operator* $A \in \mathsf{Lin}(\mathcal{X}, \mathcal{Y})$, *a* ***pseudoinverse*** *of* A *is defined as a matrix* $A^+ \in \mathsf{Lin}(\mathcal{Y}, \mathcal{X})$ *satisfying all of the following four criteria, known as the Moore-Penrose conditions:*
  1. $AA^+A = A$
  2. $A^+AA^+ = A^+$
  3. $(AA^+)^* = AA^+$
  4. $(A^+A)^* = A^+A$

  - *The pseudoinverse exists and is unique.*
  - *If* A *has real entries, then so does* A^+.
  - *If* A *is invertible,* $A^+ = A^{-1}$.
  - $O^+ = O^T$
  - $(A^+)^+ = A$
  - $(A^T)^+ = (A^+)^T, (\overline{A})^+ = \overline{A^+}, (A^*)^+ = (A^+)^*$
  - $P = AA^+$ *and* $Q = A^+A$ *are orthogonal projection operators.* $PA = AQ = A$.
  - P *is the orthogonal projector onto the range of* A.
  - Q *is the orthogonal projector onto the range of* A^*.
  - $\mathsf{ker}(A^+) = \mathsf{ker}(A^*) = \mathsf{ker}(AA^*)$
  - $\mathsf{im}(A^+) = \mathsf{im}(A^*) = \mathsf{im}(A^*A)$

- *Inner product on the space* $\mathsf{Lin}(\mathcal{X}, \mathcal{Y})$ *is defined as*

$$\langle X, Y \rangle = \mathrm{Tr}(X^*Y) \tag{38}$$

- *The determinant of a square operator* $X \in \mathsf{Lin}(\mathbb{C}^\Sigma)$ *is defined by the equation*

$$\det(X) = \sum_{\pi \in \mathrm{Sym}(\Sigma)} \mathrm{sign}(\pi) \prod_{a \in \Sigma} X(a, \pi(a)) \tag{39}$$

Property 2 *Some properties about trace*

1. *Let* $\{u_a : a \in \Gamma\} \subset \mathbb{C}^\Sigma$ *be an orthnormal basis. Then*

$$\sum_{a \in \Gamma} u_a u_a^* = \mathbb{1} \tag{40}$$

2. *Cyclic property of the trace*

$$\mathrm{Tr}(XY) = \mathrm{Tr}(YX) \tag{41}$$

3. *Trace is the sum of expectations of the operator on an ONB.*

$$\mathrm{Tr}(X) = \mathrm{Tr}(X \sum_{a \in \Gamma} u_a u_a^*) = \sum_{a \in \Gamma} u_a^* X u_a \tag{42}$$

Definition 6 *Eigenvectors and eigenvalues*

$$Xu = \lambda u \tag{43}$$

Characteristic polynomial of X

$$P_X(\alpha) = \det(\alpha \mathbb{1} - X) \tag{44}$$

The spectrum of A, *denoted* $\mathrm{spec}(A)$, *is the* ***multiset*** *containing the roots of the polynomial* P_A , *where each root appears a number of times equal to its multiplicity. Each element* $\lambda \in \mathrm{spec}(A)$ *is necessarily an eigenvalue of* A, *and every eigenvalue of* A *is contained in* $\mathrm{spec}(A)$.

Property 3

$$\mathrm{Tr}(X) = \sum_{\lambda \in \mathrm{spec}(X)} \lambda \tag{45}$$

$$\det(X) = \prod_{\lambda \in \mathrm{spec}(X)} \lambda \tag{46}$$

$$\mathrm{spec}(XY) = \mathrm{spec}(YX) \qquad \forall X \in \mathsf{Lin}(\mathcal{X}, \mathcal{Y}) \; \forall Y \in \mathsf{Lin}(\mathcal{Y}, \mathcal{X}) \tag{47}$$

Hint: $\det(I_{\mathcal{Y}} - AB) = \det(I_{\mathcal{X}} - BA)$

Definition 7 *A set $\mathcal{A} \subset \mathsf{Lin}(\mathcal{X})$ is a **subalgebra** of $\mathsf{Lin}(\mathcal{X})$ if it is closed under addition, scalar multiplication, and operator composition.*

- *Self-adjoint: $X \in \mathcal{A} \implies X^* \in \mathcal{A}$.*
- *Unital: $\mathbb{1} \in \mathcal{A}$.*

Lie bracket $[X, Y] \in \mathsf{Lin}(\mathcal{X})$is defined as

$$[X, Y] = XY - YX \tag{48}$$

Commutant of $\mathcal{A}$

$$\mathrm{comm}(\mathcal{A}) = \{Y \in \mathsf{Lin}(\mathcal{X}) : \forall X \in \mathcal{A} \; [X, Y] = 0\} \tag{49}$$

The commutant of every subset of $\mathsf{Lin}(\mathcal{X})$ is a unital subalgebra of $\mathsf{Lin}(\mathcal{X})$.

C. Important classes of operators

Definition 8 *Important classes of operators*

1. *Normal: $XX^* = X^*X$*
2. *Hermitian: $H^* = H$*
3. *Positive semidefinite: $P = B^*B \quad B \in \mathsf{Lin}(\mathcal{X}, \mathcal{Y})$*
4. *Positive definite: $X \in \mathsf{Pos}(\mathcal{X}), \quad \det(X) \neq 0$.*
5. *Density: $\rho \in \mathsf{Pos}(\mathcal{X}), \quad \mathrm{Tr}(\rho) = 1$*
6. *Projection: $\Pi \in \mathsf{Pos}(\mathcal{X}), \quad \Pi = \Pi^2$*
7. *Isometries: $V^*V = \mathbb{1}_{\mathcal{X}} \quad V \in \mathsf{Lin}(\mathcal{X}, \mathcal{Y})$*
8. *Unitary: $UU^* = U^*U = \mathbb{1}$*
9. *Diagnal: $X(a, b) = 0 \quad a \neq b$*

Remark 2 *The sum of two Hermitian operators is Hermitian, as is a real scalar multiple of a Hermitian operator. The inner product of two Hermitian operators is real as well. For every choice of a complex Euclidean space $\mathcal{X}$, the space $\mathsf{Herm}(\mathcal{X})$ therefore forms a vector space over the real numbers on which an inner product is defined.*

Theorem 3 $\mathsf{Herm}(\mathbb{C}^{\Sigma})$ *and the real Euclidean space $\mathbb{R}^{\Sigma \times \Sigma}$ are isometrically isomorphic.*

Proof. One way to define a mapping ϕ as above is as follows. First, assume that a total ordering of Σ has been fixed, and define an ONB

$$H_{a,b} = \begin{cases} E_{a,a} & a = b \\ \frac{1}{\sqrt{2}}(E_{a,b} + E_{b,a}) & a < b \\ \frac{1}{\sqrt{2}}(iE_{a,b} - iE_{b,a}) & a > b \end{cases} \tag{50}$$

The mapping ϕ is defined by the equation

$$\phi(e_{(a,b)}) = H_{a,b} \tag{51}$$

Let $v = \sum_{a,b\in\Sigma} \alpha_{ab} e_{(a,b)}$. Notice that

$$\|v\| = \sum_{a,b\in\Sigma} \alpha_{ab}^2$$

$$\|\phi(v)\| = \langle \phi(v), \phi(v) \rangle = \sum_{a,b\in\Sigma} \alpha_{ab}^2$$

■

Definition 9 $H \in \mathsf{Herm}(\mathcal{X})$. *Define a vector*

$$\lambda(\mathrm{H}) = (\lambda_1(H), \cdots, \lambda_n(H)) \in \mathbb{R}^n \tag{52}$$

such that

$$\mathrm{spec}(H) = \{\lambda_1(H), \cdots, \lambda_n(H)\} \tag{53}$$

$$\lambda_1(H) \geq \cdots \geq \lambda_n(H) \tag{54}$$

Theorem 4 *(Courant-Fischer theorem) Let $\mathcal{X}$ be a complex Euclidean space of dimension n and let $H \in \mathsf{Herm}(\mathcal{X})$ be a Hermitian operator. For every $k \in \{1, \cdots, n\}$ it holds that*

$$\begin{aligned}\lambda_k(H) &= \max_{u_1,\cdots,u_{n-k}\in \mathrm{S}(\mathcal{X})} \min_{v\in \mathsf{Span}\{u_1,\cdots,u_{n-k}\}^\perp} v^* H v \\ &= \min_{u_1,\cdots,u_{k-1}\in \mathrm{S}(\mathcal{X})} \max_{v\in \mathsf{Span}\{u_1,\cdots,u_{k-1}\}^\perp} v^* H v\end{aligned}$$

Proof. Suppose H has the following spectrum decomposition

$$H = \sum_{k=1}^{n} \lambda_k(H) v_k v_k^* \tag{55}$$

Let $\mathcal{Y}$ be a k-dimensional subspace and $\mathcal{Z} = \mathsf{Span}\{v_k, \cdots, v_n\}$.

$$\dim(\mathcal{Y} \cap \mathcal{Z}) \geq 1 \tag{56}$$

Choose $x \in \mathrm{S}(\mathcal{Y} \cap \mathcal{Z})$, then

$$x^* H x = \sum_{i=k}^{n} \lambda_k |\langle x, v_i \rangle|^2 \leq \lambda_k(H) \tag{57}$$

Thus

$$\forall \mathcal{Y} \inf_{x\in \mathrm{S}(\mathcal{Y})} x^* H x \leq \lambda_k(H) \tag{58}$$

Therefore

$$\sup_{\mathcal{Y}} \inf_{x\in \mathrm{S}(\mathcal{Y})} x^* H x \leq \lambda_k(H) \tag{59}$$

Choose $x = v_k$ to reach the inf and sup. ■

Theorem 5 *Properties of positive operators. The following statements are equivalent.*

1. $P \in \mathsf{Pos}(\mathcal{X})$
2. $P = A^* A$, $A \in \mathsf{Lin}(\mathcal{X}, \mathcal{Y})$
3. $P \in \mathsf{Herm}(\mathcal{X})$ *and eigenvalues are nonnegative.*
4. $\forall u \in \mathcal{X}, x^* P x \geq 0$
5. $\forall Q \in \mathsf{Pos}(\mathcal{X}), \langle Q, P \rangle \geq 0$
6. $P(a,b) = \langle u_a, u_b \rangle$ *where* $\{u_a : a \in \Sigma\} \subset \mathcal{Y}$.

D. Linear maps on operators

Definition 10 *Linear maps on square operators*

- *The set of linear maps on square operators* $\mathsf{Lin}(\mathcal{X}) \to \mathsf{Lin}(\mathcal{Y})$ *are denoted* $\mathsf{T}(\mathcal{X}, \mathcal{Y})$*.*
- *For a given map* $\Phi \in \mathsf{T}(\mathcal{X}, \mathcal{Y})$*, the adjoint of* Φ *is defined to be the unique map* $\Phi^* \in \mathsf{T}(\mathcal{Y}, \mathcal{X})$ *that satisfies*

$$\langle \Phi^*(Y), X \rangle = \langle Y, \Phi(X) \rangle \quad \forall X \in \mathsf{Lin}(\mathcal{X}) \; \forall Y \in \mathsf{Lin}(\mathcal{Y}) \tag{60}$$

- *The identity map is defined as*

$$\mathbb{1}_{\mathsf{Lin}(\mathcal{X})}(X) = X \quad \forall X \in \mathsf{Lin}(\mathcal{X}) \tag{61}$$

- *The trace function can be viewed as*

$$\mathrm{Tr} \in \mathsf{T}(X, \mathbb{C}) \tag{62}$$

- *Partial trace*

$$\mathrm{Tr}_{\mathcal{X}} = \mathrm{Tr} \otimes \mathbb{1}_{\mathsf{Lin}(\mathcal{Y})} \in \mathsf{T}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Y}) \tag{63}$$

$$\mathrm{Tr}_{\mathcal{Y}} = \mathbb{1}_{\mathsf{Lin}(\mathcal{X})} \otimes \mathrm{Tr} \in \mathsf{T}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{X}) \tag{64}$$

- *The following classes of maps are among those that are discussed in greater detail later:*

  1. *Hermitian-preserving maps.*

  $$\forall H \in \mathsf{Herm}(\mathcal{X}) \quad \Phi(H) \in \mathsf{Herm}(\mathcal{Y})$$

  2. *Positive maps.*

  $$\forall P \in \mathsf{Pos}(\mathcal{X}) \quad \Phi(P) \in \mathsf{Pos}(\mathcal{Y})$$

  3. *Completely positive maps. For arbitrary Euclidean space* $\mathcal{Z}$

  $$\Phi \otimes \mathbb{1}_{\mathsf{Lin}(\mathcal{Z})}$$

  is a positive map.

  4. *Trace-preserving maps.*

  $$\forall X \in \mathsf{Lin}(\mathcal{X}) \quad \mathrm{Tr}(\Phi(X)) = \mathrm{Tr}(X)$$

  5. *Unital maps.*

  $$\Phi(\mathbb{1}_{\mathcal{X}}) = \mathbb{1}_{\mathcal{Y}}$$

Theorem 6 *Suppose* $\Phi \in \mathsf{T}(\mathcal{X}, \mathcal{Y})$ *has the following decomposition*

$$\Phi(X) = \sum_{a \in \Sigma} A_a X B_a^* \quad \forall X \in \mathsf{Lin}(\mathcal{X}) \tag{65}$$

Then

$$\Phi^*(Y) = \sum_{a \in \Sigma} A_a^* Y B_a \quad \forall Y \in \mathsf{Lin}(\mathcal{Y}) \tag{66}$$

Proof.

$$\langle \Phi(X), Y \rangle = \mathrm{Tr}((\sum_{a \in \Sigma} B_a X^* A_a^*) Y) \tag{67}$$

$$= \mathrm{Tr}(X^* \sum_{a \in \Sigma} A_a^* Y B_a) \tag{68}$$

$$= \langle X, \sum_{a \in \Sigma} A_a^* Y B_a \rangle \tag{69}$$

Definition 11 *The operator-vector correspondence. There is a correspondence between the spaces* $\mathsf{Lin}(\mathcal{Y}, \mathcal{X})$ *and* $\mathcal{X} \otimes \mathcal{Y}$*, for any choice of complex Euclidean spaces* $X = \mathbb{C}^{\Sigma}$ *and* $Y = \mathbb{C}^{\Gamma}$*.*

$$\mathrm{vec} : \mathsf{Lin}(\mathcal{Y}, \mathcal{X}) \to \mathcal{X} \otimes \mathcal{Y} \tag{70}$$

defined by the action

$$\mathrm{vec}(E_{a,b}) = e_a \otimes e_b \tag{71}$$

Property 4 $u, v \in \mathcal{X} \otimes \mathcal{Y},\ A, B \in \mathsf{Lin}(\mathcal{Y}, \mathcal{X})$

1. $\mathrm{vec}(uv^*) = u \otimes \overline{v}$
2. $\langle A, B\rangle = \langle \mathrm{vec}(A), \mathrm{vec}(B)\rangle$
3. $(A_0 \otimes A_1)\,\mathrm{vec}(B) = \mathrm{vec}(A_0 B A_1^T) \quad \forall A_0 \in \mathsf{Lin}(\mathcal{X}_0, Y_0),\ A_1 \in \mathsf{Lin}(\mathcal{X}_1, Y_1),\ B \in \mathsf{Lin}(\mathcal{X}_1, \mathcal{X}_0)$
4. $\mathrm{Tr}_{\mathcal{Y}}(\mathrm{vec}(A)\,\mathrm{vec}(B)^*) = AB^*$
5. $\mathrm{Tr}_{\mathcal{X}}(\mathrm{vec}(A)\,\mathrm{vec}(B)^*) = A^T\overline{B}$

Theorem 7 *Some important theorems*

1. *Spectrmum theorem*
2. *Jordan-Hahn decompositions*

$$H \in \mathsf{Herm}(\mathcal{X}) \implies \exists P, Q \in \mathsf{Pos}(\mathcal{X}) \begin{cases} H = P - Q \\ PQ = 0 \end{cases} \tag{72}$$

3. *Singular value theorem*
4. *Polar decompositions*
5. *Schmidt decompositions*

Definition 12 *Norms of operators*

1. ***Schatten p-norms***

   This family includes the three most commonly used norms in quantum information theory: the spectral norm, the Frobenius norm, and the trace norm.

$$\|A\|_p = \left\{ \mathrm{Tr}\left[(A^*A)^{p/2} \right] \right\}^{1/p} \tag{73}$$

   The Schatten p-norm of an operator A coincides with the ordinary vector p-norm of the vector of singular values of A:

$$\|A\|_p = \|s(A)\|_p \tag{74}$$

2. ***The spectrum norm***

$$\|A\|_\infty = \max\{\|Au\| : u \in X,\ \|u\| \le 1\} \tag{75}$$

   The spectral norm is the most important norm we use.

3. ***The Frobenius norm***

$$\|A\|_2 = (\mathrm{Tr}(A^*A))^{1/2} = \sqrt{\langle A^*A\rangle} = \|\mathrm{vec}(A)\| = \sqrt{\sum_{a,b} |A(a,b)|^2} \tag{76}$$

4. ***The trace norm***

$$\|A\|_1 = \mathrm{Tr}(\sqrt{A^*A}) = \max\{|\langle U, A\rangle| : U \in \mathsf{U}(\mathcal{X})\} \tag{77}$$

Property 5 *1. Schatten p-norms are non-increasing in p*

$$1 \le p \le q \le \infty \implies \|A\|_p \ge \|A\|_q \tag{78}$$

2. For every nonzero operator A and for $1 \le p \le q \le \infty$, it holds that

$$\|A\|_p \le \mathsf{rank}(A)^{\frac{1}{p}-\frac{1}{q}} \|A\|_q \tag{79}$$

3. For every $p \in [1,\infty]$, the Schatten p-norm is isometrically invariant (and therefore unitarily invariant): for every $A \in \mathsf{Lin}(\mathcal{X},\mathcal{Y})$, $U \in \mathsf{U}(\mathcal{Y},\mathcal{Z})$, and $V \in \mathsf{U}(\mathcal{X},\mathcal{W})$ it holds that

$$\|A\|_p = \|UAV^*\|_p \tag{80}$$

4. For every operator $A \in \mathsf{Lin}(\mathcal{X},\mathcal{Y})$, it holds that the Schatten p-norm and p-norm are dual

$$\|A\|_p = \max\{|\langle B, A\rangle| : B \in \mathsf{Lin}(\mathcal{X},\mathcal{Y}), \|B\|_{p^*} \le 1\} \tag{81}$$

One consequence is the inequality

$$|\langle B, A\rangle| \le \|A\|_p \|B\|_{p^*} \tag{82}$$

where $1/p + 1/p^* = 1$.

5. $A \in \mathsf{Lin}(\mathcal{Z},\mathcal{W})$, $B \in \mathsf{Lin}(\mathcal{Y},\mathcal{Z})$ and $C \in \mathsf{Lin}(\mathcal{X},\mathcal{Y})$

$$\|ABC\|_p \le \|A\|_\infty \|B\|_p \|C\|_\infty \tag{83}$$

It follows that the Schatten $p-norm$ is submultiplicative:

$$\|AB\|_p \le \|A\|_p \|B\|_p \tag{84}$$

6. It holds that

$$\|A\|_p = \|A^*\|_p = \|A^T\|_p = \|\overline{A}\|_p \tag{85}$$

7. The spectrum norm is induced by the Euclidean norm. It has the following property

$$\|A^*A\|_\infty = \|AA^*\|_\infty = \|A\|_\infty^2 \tag{86}$$

8. Let $A \in \mathsf{Lin}(\mathcal{X} \otimes \mathcal{Y})$, then

$$\begin{aligned}
\| \mathrm{Tr}_{\mathcal{Y}}(A)\|_1 &= \|(I_{\mathcal{X}} \otimes \mathrm{Tr})A\|_1 \\
&= \max\{|\langle U, (I_{\mathcal{X}} \otimes \mathrm{Tr})A\rangle| : U \in \mathsf{U}(\mathcal{X})\} \\
&= \max\{\mathrm{Tr}[(I_{\mathcal{X}} \otimes \mathrm{Tr})(U^* \otimes I_{\mathcal{Y}})A] : U \in \mathsf{U}(\mathcal{X})\} \\
&= \max\{\mathrm{Tr}[(U^* \otimes I_{\mathcal{Y}})A] : U \in \mathsf{U}(\mathcal{X})\} \\
&= \max\{\langle U \otimes I_{\mathcal{Y}}, A\rangle : U \in \mathsf{U}(\mathcal{X})\} \\
&\le \max\{|\langle V, A\rangle| : V \in \mathsf{U}(\mathcal{X} \otimes \mathcal{Y})\} \\
&= \|A\|_1
\end{aligned}$$

9. Let $\alpha, \beta \ge 0$ and $u, v \in \mathcal{X}$

$$\|\alpha uu^* - \beta vv^*\|_1 = \sqrt{(\alpha+\beta)^2 - 4\alpha\beta|\langle u, v\rangle|^2} \tag{87}$$

Remark 3 ***Spectrum norm is so important*** *that when we always omit the ∞ symbol when we write it. The spectrum norm of a vector (viewed as an operator on $\mathsf{Lin}(\mathbb{C},\mathcal{X})$) coincides with its Euclidean norm.*

II. ANALYSIS

Definition 13 *Nets*

Let V be a vector space. Let $W \subset V$. A set of vectors N is an ϵ-net if

$$\forall u \in W \ \exists v \in N \ \|u - v\| \leq \epsilon$$

Theorem 8 *Let X be a complex Euclidean space of dimension n and let $\epsilon > 0$ be a positive real number. With respect to the Euclidean norm on X, there exists an ϵ-net $N \subset B(\mathcal{X})$ for the unit ball $B(\mathcal{X})$ such that*

$$|N| \leq \left(1 + \frac{2}{\epsilon}\right)^{2n} \tag{88}$$

Definition 14 ***Borel sets and functions***

$\mathcal{A} \subset \mathcal{V}$ and $\mathcal{B} \subset \mathcal{W}$ denote fixed subsets of finite-dimensional real or complex vector spaces $\mathcal{V}$ and $\mathcal{W}$.

A set $\mathcal{C} \subset \mathcal{A}$ is said to be a Borel subset of $\mathcal{A}$ if one or more of the following inductively defined properties holds:

1. *$\mathcal{C}$ is an open set relative to $\mathcal{A}$*
2. *$\mathcal{C}$ is the complement of a Borel subset of $\mathcal{A}$*
3. *For $\{\mathcal{C}_1, \mathcal{C}_2, \cdots\}$ being a countable collection of Borel subsets of $\mathcal{A}$, it holds that $\mathcal{C}$ is equal to the union*

$$\mathcal{C} = \bigcup_{k=1}^{\infty} \mathcal{C}_k \tag{89}$$

A function $f : \mathcal{A} \to \mathcal{B}$ is a Borel function if $f^{-1}(\mathcal{C}) \in \mathrm{Borel}(\mathcal{A})$ for all $\mathcal{C} \in \mathrm{Borel}(\mathcal{B})$.

Continous function and the characteristic function of a Borel subset are both Borel functions.

1. If B is a vector space, αf and $f + g$ are also Borel functions.
2. If B is a subalgebra of $\mathsf{Lin}(Z)$, for Z being a real or complex Euclidean space, and $f, g : A \to B$ are Borel functions, then the function $h : A \to B$ defined by

$$h(u) = f(u)g(u)$$

is also a Borel function.

Definition 15 *Measures on Borel sets*

A Borel measure (or simply a measure) defined on $\mathrm{Borel}(\mathcal{A})$ is a function

$$\mu : \mathrm{Borel}(\mathcal{A}) \to [0, \infty] \tag{90}$$

that possesses two properties:

1. *$\mu(\emptyset) = 0$*
2. *For any countable collection $\mathcal{C}_1, \mathcal{C}_2, \cdots \subset \mathrm{Borel}(\mathcal{A})$ of pairwise disjoint Borel subsets of $\mathcal{A}$, it holds that*

$$\mu(\bigcup_{k=1}^{\infty} \mathcal{C}_k) = \sum_{k=1}^{\infty} \mu(\mathcal{C}_k) \tag{91}$$

Definition 16 *Let V be a vector space over the real or complex numbers. A subset C of V is convex if*

$$\lambda u + (1 - \lambda)v \in C \qquad \forall u, v \in C \ \forall \lambda \in [0, 1] \tag{92}$$

A set $K \subset V$ is a cone if

$$\lambda u \in K \qquad \forall u \in K \ \forall \lambda \geq 0 \tag{93}$$

The cone generated by a set $A \subset V$ is defined as

$$\mathrm{cone}(A) = \{\lambda u : u \in A, \ \lambda \geq 0\} \tag{94}$$

A convex cone is simply a cone that is also convex.

Theorem 9 *If A is a compact set that does not include 0, then* cone(A) *is necessarily a closed set.*

Note that If A contains 0, cone(A) may not be closed. For instance, $A = \{(x, y) : (x-1)^2 + y^2 \leq 1,\ x, y \in \mathbb{R}\}$ is closed, but cone$(A) = \{(x, y) : x > 0,\ y \in \mathbb{R}\}$ is not closed.

Theorem 10 *A cone K is convex if and only if it is closed under addition*

$$u + v \in K \quad \forall u, v \in K \tag{95}$$

Proof. If

$$u, v \in K\ \lambda \in [0,1] \implies \lambda u, (1-\lambda)v \in K \implies \lambda u + (1-\lambda)v \in K \tag{96}$$

Only if

$$u, v \in K\ \lambda \in (0,1) \implies \frac{u}{\lambda}, \frac{v}{1-\lambda} \in K \implies u + v = \lambda \cdot \frac{u}{\lambda} + (1-\lambda) \cdot \frac{v}{1-\lambda} \in K \tag{97}$$

Definition 17 *$C \subset V$ is convex. Convex function $f : C \to \mathbb{R}$ if*

$$f(\lambda u + (1-\lambda)v) \leq \lambda f(u) + (1-\lambda) f(v) \tag{98}$$

Theorem 11 *A convex function f of one real variable defined on some open interval C is continuous on C and Lipschitz continuous on any closed subinterval.*

Theorem 12 *$C \subset V$ is a convex set.*

$$\begin{cases} f(\frac{u+v}{2}) \leq \frac{f(u)+f(v)}{2} & \forall u, v \in C \\ f \text{ is continuous} \end{cases} \implies f \text{ is convex} \tag{99}$$

Definition 18 *The convex hull of a set $A \subset V$ is defined as*

$$\mathrm{conv}(A) = \{\sum_{a \in \Sigma} p(a) u_a : p \in \mathcal{P}(\Sigma),\ \{u_a : a \in \Sigma\} \subset A\} \tag{100}$$

Theorem 13 *The convex hull* conv(A) *of a closed set A need not itself be closed. However, if A is compact, then so too is* conv(A).

Theorem 14 *Let V be a real vector space and let $A \subset V$. A is contained in an affine subspace of V having dimension n. For every vector $v \in$ conv(A) in the convex hull of A, there exist $m \leq n+1$ vectors $u_1, \cdots, u_m \in A$ such that $v \in$ conv$(\{u_1, \cdots, u_m\})$.*

Definition 19 *A point $w \in C$ in a convex set C is said to be an extreme point of C if*

$$\forall u, v \in C\ \forall \lambda \in (0,1)\ w = \lambda u + (1-\lambda)v \implies u = v = w \tag{101}$$

Theorem 15 *Let V be a finite-dimensional vector space over the real or complex numbers, let $C \subset V$ be a compact and convex set, and let $A \subset C$ be the set of extreme points of C. It holds that $C =$ conv(A).*

1. The spectral norm unit ball. For any complex Euclidean space X, the set

$$\{X \in \mathsf{Lin}(\mathcal{X}) : \|X\|_\infty \leq 1\} \tag{102}$$

is a convex and compact set. The extreme points of this set are the unitary operators $\mathsf{U}(\mathcal{X})$.

2. The trace norm unit ball. For any complex Euclidean space X, the set

$$\{X \in \mathsf{Lin}(\mathcal{X}) : \|X\|_1 \leq 1\} \tag{103}$$

is a convex and compact set. The extreme points of this set are those operators of the form uv for $u, v \in S(\mathcal{X})$ unit vectors.

3. Density operators. For any complex Euclidean space $\mathcal{X}$, the set $\mathsf{D}(\mathcal{X})$ of density operators acting on $\mathcal{X}$ is convex and compact. The extreme points of $\mathsf{D}(\mathcal{X})$ coincide with the rank-one projection operators. These are the operators of the form uu for $u \in S(\mathcal{X})$ being a unit vector.

4. Probability vectors. For any alphabet Σ, the set of probability vectors $\mathcal{P}(\Sigma)$ is convex and compact. The extreme points of this set are the elements of the standard basis$\{e_a : a \in \Sigma\}$ of $\mathbb{R}^{\Sigma}$.

Convex sets in real Euclidean spaces possess a fundamentally important property: every vector lying outside of a given convex set in a real Euclidean space can be separated from that convex set by a hyperplane. That is, if the underlying real Euclidean space has dimension n, then there exists an affine subspace of that space having dimension $n1$ that divides the entire space into two half-spaces: one contains the convex set and the other contains the chosen point lying outside of the convex set. The following theorem represents one specific formulation of this fact.

Theorem 16 *Let V be a real Euclidean space, let $C \subset V$ be a closed, convex subset of V, and let $u \in V$ be a vector with $u \notin C$. There exists a vector $v \in V$ and a scalar $\alpha \in \mathbb{R}$ such that*

$$\langle v, u\rangle < \alpha \leq \langle v, w\rangle \qquad \forall w \in C \tag{104}$$

If C is a cone, then v may be chosen so that it holds for $\alpha = 0$.

Consider the 2-dimensional case. A line in $\mathbb{R}^2$ is

$$Ax + By + C = 0 \qquad (A, B) \neq (0, 0) \tag{105}$$

Then the plane is divided into 2 parts:

$$\begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq -C \qquad \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} > -C$$

If C is a cone, then there exists a line passing $(0, 0)$.

Theorem 17 *Let X and Y be real or complex Euclidean spaces, let $A \subset X$ and $B \subset Y$ be convex sets with B compact, and let $f : A \times B \to \mathbb{R}$ be a continuous function such that*

1. $u \mapsto f(u, v)$ is a convex function on A for all $v \in B$.

2. $v \mapsto f(u, v)$ is a concave function on B for all $u \in A$.

It holds that

$$\inf_{u\in A} \max_{v\in B} f(u, v) = \max_{v\in B} \inf_{u\in A} f(u, v) \tag{106}$$

III. PROBABILITY THEORY

Definition 20 *Suppose A is a subset of a finite-dimensional real or complex vector space V and*

$$\mu : \mathrm{Borel}(A) \to [0, 1] \tag{107}$$

is a probability measure (by which it is meant that μ is a normalized Borel measure). A random variable X distributed with respect to μ is a real-valued, integrable Borel function of the form

$$X : A \to \mathbb{R} \tag{108}$$

For every Borel subset $B \subset R$ of the real numbers, the probability that X takes a value in B is defined as

$$\Pr(X \in B) = \mu(\{u \in A : X(u) \in B\}) \tag{109}$$

The expected value (or mean value) of a random variable X, distributed with respect to a probability measure $\mu : \mathrm{Borel}(A) \to [0, 1]$, is defined as

$$E(X) = \int X(u)\mathrm{d}\mu(u) \tag{110}$$

If X is a random variable taking nonnegative real values, then it holds that

$$E(X) = \int_0^\infty \lambda \Pr(\lambda \leq X < \lambda + \mathrm{d}\lambda) = \int_0^\infty \Pr(X \geq \lambda)\mathrm{d}\lambda \tag{111}$$

Definition 21 *Random variables for discrete distributions. Consider the set* $\{1, \cdots, n\} \subset \mathbb{R}$ *for some choice of a positive integer* n*, and observe that every subset of* $\{1, \cdots, n\}$ *is a Borel subset of this set. The Borel probability measures*

$$\mu : \mathrm{Borel}(\{1, \cdots, n\}) \to [0, 1] \tag{112}$$

coincide precisely with the set of all probability vectors $p \in P(\{1, \cdots, n\})$*. through the equations*

$$\mu(B) = \sum_{b \in B} p(b) \qquad p(a) = \mu(\{a\}) \tag{113}$$

for every $B \subset \{1, \cdots, n\}$ *and* $a \in \{1, \cdots, n\}$*.*

Definition 22 *Independent:*

$$\Pr((X, Y) \in A \times B) = \Pr(X \in A) \Pr(Y \in B) \qquad \forall A, B \subset \mathbb{R} \tag{114}$$

Identically distributed:

$$\Pr(X \in A) = \Pr(Y \in A) \qquad \forall A \subset \mathbb{R} \tag{115}$$

Theorem 18 *A few fundamental theorems*

Markovs inequality: *Let* X *be a random variable taking nonnegative real values, and let* $\epsilon > 0$ *be a positive real number. It holds that*

$$\Pr(X \geq \epsilon) \leq \frac{E(X)}{\epsilon} \tag{116}$$

Jensens inequality: *Suppose that* X *is a random variable and* $f : \mathbb{R} \to \mathbb{R}$ *is a convex function. It holds that*

$$f(E(X)) \leq E(f(X)) \tag{117}$$

Weak law of large numbers: *Let* X *be a random variable and let* $\alpha = E(X)$*. Assume, moreover, for every positive integer* n*, that* $\mathcal{X}_1, \cdots, \mathcal{X}_n$ *are independent random variables identically distributed to* X*. For every positive real number* $\epsilon > 0$*, it holds that*

$$\lim_{n \to \infty} \Pr\left(\left| \frac{X_1 + \cdots + X_n}{n} - \alpha \right| \geq 0 \right) = 0 \tag{118}$$

Hoeffdings inequality: *Let* $\mathcal{X}_1, \cdots, \mathcal{X}_n$ *be independent and identically distributed random variables taking values in the interval* $[0, 1]$ *and having mean value* α*. For every positive real number* $\epsilon > 0$ *it holds that*

$$\Pr\left(\left| \frac{X_1 + \cdots + X_n}{n} - \alpha \right| \geq \epsilon \right) \leq 2e^{-2n\epsilon^2} \tag{119}$$

Definition 23 *Gaussian measure and normally distributed random variables. The standard Gaussian measure on* R *is the Borel probability measure*

$$\gamma : \mathrm{Borel}(\mathbb{R}) \to [0, 1] \tag{120}$$

defined as

$$\gamma(A) = \frac{1}{\sqrt{2\pi}} \int_A \exp(-\frac{\alpha^2}{2}) \mathrm{d}\alpha \tag{121}$$

A random variable X is a standard normal random variable if it holds that $\Pr(X \in A) = \gamma(A)$ for every $A \in \mathrm{Borel}(\mathbb{R})$. This is equivalent to saying that X is identically distributed to a random variable $Y(\alpha) = \alpha$ distributed with respect to the standard Gaussian measure γ on $\mathbb{R}$.

The following integrals are among many integrals of a similar sort that are useful when reasoning about standard normal random variables:

1. For every positive real number $\lambda > 0$ and every real number $\beta \in \mathbb{R}$ it holds that

$$\int \exp(-\lambda\alpha^2 + \beta\alpha)\mathrm{d}\alpha = \sqrt{\frac{\pi}{\lambda}} \exp\left(\frac{\beta^2}{4\lambda}\right) \tag{122}$$

2. For every positive integer n, it holds that

$$\int_0^\infty \alpha^n \mathrm{d}\gamma(\alpha) = \frac{2^{\frac{n}{2}}\Gamma(\frac{n+1}{2})}{2\sqrt{\pi}} \tag{123}$$

   where the Γ-function may be defined at positive half-integer points as follows:

$$\Gamma\left(\frac{m+1}{2}\right) = \begin{cases} \sqrt{\pi} & m = 0 \\ 1 & m = 1 \\ \frac{m-1}{2}\Gamma(\frac{m-1}{2}) & m \geq 2 \end{cases} \tag{124}$$

3. For every positive real number $\lambda > 0$ and every pair of real numbers ${}_{0}, {}_{1} \in \mathbb{R}$ with $\beta_0 \leq \beta_1$ it holds that

$$\int_{\beta_0}^{\beta_1} \alpha \exp(-\lambda\alpha^2)\mathrm{d}\alpha = \frac{1}{2\lambda}\exp(-\lambda\beta_0^2) - \frac{1}{2\lambda}\exp(-\lambda\beta_1^2) \tag{125}$$

Definition 24 *For every positive integer n, the standard Gaussian measure on $\mathbb{R}_n$ is the Borel probability measure*

$$\gamma_n : \mathrm{Borel}(\mathbb{R}^n) \to [0,1] \tag{126}$$

obtained by taking the n-fold product measure of γ with itself. Equivalently

$$\gamma_n(A) = (2\pi)^{-\frac{n}{2}} \int_A \exp\left(-\frac{\|u\|^2}{2}\right) \mathrm{d}\nu_n(u) \tag{127}$$

where ν_n denotes the n-fold product measure of the standard Borel measure ν with itself and the norm is the Euclidean norm.

Definition 25 *Let X and Y be complex Euclidean spaces, let $\Phi \in T(X,Y)$ be a Hermitian-preserving map, and let $A \in \mathsf{Herm}(\mathcal{X})$ and $B \in \mathsf{Herm}(\mathcal{X})$. A semidefinite program is a triple (Φ, A, B), with which the following pair of optimization problems is associated:*

$$\begin{array}{llll} \textit{Primary problem} & & \textit{Dual problem} & \\ \textit{maximize:} & \langle A, X\rangle & \textit{minimize:} & \langle B, Y\rangle \\ \textit{subject to:} & \Phi(X) = B & \textit{subject to:} & \Phi^*(Y) \geq A \\ & X \in \mathsf{Pos}(\mathcal{X}) & & Y \in \mathsf{Herm}(\mathcal{Y}) \end{array} \tag{128}$$