Chapter 2
Pure State Extensions in Linear Algebra

In this chapter we introduce the concept of a pure state extension by means of a concrete example: we consider the matrix algebra

$$M := M_n(\mathbb{C}),$$

for some fixed $n \in \mathbb{N}$. We often denote an element $a \in M$ by

$$a = \sum_{i,j} a_{ij} \left|e_i\right\rangle \left\langle e_j\right| ,$$

where $\{e_i\}$ is the standard basis of $\mathbb{C}^n$ and we use the shorthand notation $|x\rangle \langle y|$ for the operator which satisfies $|x\rangle \langle y| (z) = \langle y, z\rangle x$. This means that a_{ij} is the element in the i-th row and j-th column of the matrix a. Furthermore, we consider the diagonal matrices

$$D := \{a \in M | a_{ij} = 0 \text{ if } i \neq j\},$$

which form a unital subalgebra of M.

The algebra M also has a *-operation that is an involution, defined by:

$$a^* = \sum_{i,j} \overline{a_{ji}} \left|e_i\right\rangle \left\langle e_j\right| .$$

We call a^* the **adjoint** of a. Note that D is also closed under this operation.

M. Stevens, *The Kadison-Singer Property*,
SpringerBriefs in Mathematical Physics 14, DOI 10.1007/978-3-319-47702-2_2

2.1 Density Operators and Pure States

M is not merely an algebraic object; it also has its defining action on $\mathbb{C}^n$, which is a vector space with a natural inner product $\langle x, y\rangle = \sum_i \overline{x_i} y_i$ (i.e. we take the standard inner product that is linear in the second argument). Using this, we can define a special class of matrices.

Definition 2.1 $a \in M$ is called **positive** if for each $x \in \mathbb{C}^n$ we have $\langle x, ax\rangle \geq 0$. We write this condition as $a \geq 0$.

Now we can define our main object of study: states.

Definition 2.2 A **state** on M is a linear map $f : M \to \mathbb{C}$ that is positive, meaning that $f(a) \geq 0$ for all $a \geq 0$, and unital, i.e. $f(1) = 1$. The set of all states on M is denoted by $S(M)$, which we call the **state space** of M.

In turns out that all states on M are of a specific form. To make this more precise, we need two more definitions.

Definition 2.3 The **trace** of a matrix $a \in M$ is defined as $\mathrm{Tr}(a) = \sum_i a_{ii}$.

Lemma 2.4 *1.* $\mathrm{Tr}(ab) = \mathrm{Tr}(ba)$ *for all* $a, b \in M$
2. For any basis $\{v_i\}$ *of* $\mathbb{C}^n$*, we have* $\mathrm{Tr}(a) = \sum_i \langle v_i, av_i\rangle$

Proof 1. $\mathrm{Tr}(ab) = \sum_i (ab)_{ii} = \sum_i \sum_k a_{ik} b_{ki} = \sum_k \sum_i b_{ki} a_{ik} = \sum_k (ba)_{kk} = \mathrm{Tr}(ba)$.

2. Note that by definition, $\mathrm{Tr}(a) = \sum_i \langle e_i, ae_i\rangle$. For another basis $\{v_i\}$ there is a unitary $u \in M$, i.e. $uu^* = u^*u = 1$, such that $ue_i = v_i$ for all i. Then:

$$\sum_i \langle v_i, av_i\rangle = \sum_i \langle ue_i, aue_i\rangle = \sum_i \langle e_i, u^*aue_i\rangle = \mathrm{Tr}(u^*au) = \mathrm{Tr}(auu^*) = \mathrm{Tr}(a),$$

where we used part 1 of this lemma. □

There is a connection between states on M and so-called *density operators*.

Definition 2.5 A **density operator** $\rho \in M$ is a positive operator that satisfies $\mathrm{Tr}(\rho) = 1$. We write $\mathscr{D}(M)$ for the set of all density operators in M.

Theorem 2.6 *There is a bijective correspondence between states* f *on* M *and density operators* $\rho \in M$*, given by* $f(a) = \mathrm{Tr}(\rho a)$ *for all* $a \in M$*.*

Proof We prove that $S(M) \cong \mathscr{D}(M)$ as sets. We construct $\Phi : S(M) \to \mathscr{D}(M)$ via

$$\Phi(f) = \sum_{i,j} \rho_{ij} \left|e_i\right\rangle \left\langle e_j\right|,$$

where $\rho_{ij} = f(\left|e_j\right\rangle \langle e_i|)$.

To see that Φ is well defined, note that

$$\mathrm{Tr}(\Phi(f)) = \sum_i f(|e_i\rangle\langle e_i|) = f(\sum_i |e_i\rangle\langle e_i|) = f(1) = 1$$

and for $x \in \mathbb{C}^n$, say $x = \sum_i c_i e_i$,

$$\langle x, \Phi(f)x\rangle = \sum_{i,j} \overline{c_i} c_j \langle e_i, \Phi(f)e_j\rangle = \sum_{i,j} \overline{c_i} c_j f(\left|e_j\right\rangle\langle e_i|) = f(|x\rangle\langle x|) \geq 0,$$

which means that $\Phi(f)$ is indeed a density operator.

Next, define $\Psi : \mathscr{D}(M) \to S(M)$ by

$$\Psi(\rho)(a) = \mathrm{Tr}(\rho a)$$

for all $a \in M$.

To see that Ψ is well defined, first note that $\Psi(\rho)(1) = \mathrm{Tr}(\rho) = 1$. Next, let $\rho \in \mathscr{D}(M)$ and $a \in M$ positive. Then ρ has a spectral decomposition

$$\rho = \sum_i p_i |v_i\rangle\langle v_i|,$$

for some orthonormal basis (v_i), where all $p_i \geq 0$. Since a is positive,

$$a = \sum_{i,j} \lambda_{ij} |v_i\rangle\left\langle v_j\right|,$$

with all $\lambda_{ii} \geq 0$. Then $\rho a = \sum_{i,j} p_i \lambda_{ij} |v_i\rangle\left\langle v_j\right|$, so

$$\Psi(\rho)(a) = \mathrm{Tr}(\rho a) = \sum_i p_i \lambda_{ii} \geq 0,$$

so $\Psi(\rho)$ is positive, and hence a state. Now, note that

$$\begin{aligned}\Psi(\Phi(f))(a) &= \mathrm{Tr}(\Phi(f)a) = \mathrm{Tr}((\sum_{i,j} \rho_{ij} |e_i\rangle\left\langle e_j\right|)(\sum_{l,k} a_{lk} |e_l\rangle\langle e_k|)) \\ &= \sum_{i,j} \rho_{ij} a_{ji} = \sum_{i,j} a_{ji} f(\left|e_j\right\rangle\langle e_i|) = f(\sum_{i,j} a_{ji} \left|e_j\right\rangle\langle e_i|) \\ &= f(a),\end{aligned}$$

meaning that $\Psi \circ \Phi = \mathrm{Id}$.

Next,

$$\Phi(\Psi(\rho))_{ij} = \Psi(\rho)(\left|e_j\right\rangle\langle e_i|) = \mathrm{Tr}(\rho\left|e_j\right\rangle\langle e_i|) = \langle e_i, \rho e_j\rangle = \rho_{ij},$$

meaning that $\Phi \circ \Psi = \mathrm{Id}$. Hence, $\mathscr{D}(M) \cong S(M)$ as sets, and writing $\Psi(\rho) = f$ the given formula $f(a) = \mathrm{Tr}(\rho a)$ holds. □

Note that $S(M)$ and $\mathscr{D}(M)$ have more structure than that of a set, since they are also convex. A function $f : A \to B$ between two convex sets is called **affine** if it preserves the convex structure, i.e. if $f(tx + (1-t)y) = tf(x) + (1-t)f(y)$ for all $t \in [0, 1]$ and $x, y \in A$. Note that the bijection in Theorem 2.6 is an affine function.

For a convex set C, a point $c \in C$ is called **extreme** if for any $c_1, c_2 \in C$ and $t \in (0, 1)$ such that $c = tc_1 + (1-t)c_2$ we have $c_1 = c_2 = c$. The set of extreme points of a convex set C is called the **extreme boundary** of C, often denoted as $\partial_e C$.

Since $S(M)$ is a convex set, we can consider its boundary, which plays a crucial role in our discussion. For the elements in this boundary, i.e. the extreme points of $S(M)$, we have a special name.

Definition 2.7 A state $f \in S(M)$ is called a **pure state** if it is an extreme point of $S(M)$.

To determine the pure states on M, we use the following lemma.

Lemma 2.8 *Suppose that C and D are convex sets and that there is an affine isomorphism between them. Then $\partial_e C$ is isomorphic to $\partial_e D$.*

Proof Suppose that the map $\phi : C \to D$ is an affine isomorphism. First of all, we claim that $\phi(\partial_e C) \subseteq \partial_e D$.

To see this, first note that ϕ^{-1} is an affine isomorphism as well. Now suppose $x \in \partial_e C$ and $t \in [0, 1]$, $a, b \in D$ such that $\phi(x) = ta + (1-t)b$. Then

$$x = \phi^{-1}(ta + (1-t)b) = t\phi^{-1}(a) + (1-t)\phi^{-1}(b).$$

Then, since $x \in \partial_e C$, $x = \phi^{-1}(a) = \phi^{-1}(b)$, but then also $\phi(x) = a = b$, so we have that $\phi(x) \in \partial_e D$.

Hence $\phi(\partial_e C) \subseteq \partial_e D$, so by the same token $\phi^{-1}(\partial_e D) \subseteq \partial_e C$, whence ϕ maps $\partial_e C$ bijectively to $\partial_e D$. Therefore $\partial_e C$ and $\partial_e D$ are isomorphic. □

We can now give an explicit description of the pure states on M.

Corollary 2.9 *There is a bijective correspondence between pure states f on M and one-dimensional projections $|\psi\rangle\langle\psi|$, such that $f(a) = \langle\psi, a\psi\rangle$ for all $a \in M$.*

Proof By Theorem 2.6 we know that $S(M)$ corresponds bijectively to $\mathscr{D}(M)$ via the formula $f(a) = \mathrm{Tr}(\rho a)$. Since this formula is affine and the pure states on M are exactly $\partial_e S(M)$, we only need to determine $\partial_e \mathscr{D}(M)$, by Lemma 2.8.

Suppose that $\rho \in \partial_e \mathscr{D}(M)$ and let $\rho = \sum_i p_i |v_i\rangle\langle v_i|$ be its spectral decomposition. Then since ρ is positive and has unit trace, we know that the $\{v_i\}$ are orthonormal, all $p_i \geq 0$ and $\sum_i p_i = 1$. Clearly, all $p_i \in [0, 1]$.

Now suppose that there is a $j \in \{1, \ldots, n\}$ such that $p_j \in (0, 1)$. Then there must be a $k \neq j$ such that $p_k \in (0, 1)$ as well. Then there is a $\varepsilon > 0$ such that we have $[p_j - \varepsilon, p_j + \varepsilon] \subseteq [0, 1]$ and $[p_k - \varepsilon, p_k + \varepsilon] \subseteq [0, 1]$. Now define

$$r_i = \begin{cases} p_i - \varepsilon : i = j \\ p_i + \varepsilon : i = k \\ p_i \quad\;\; : i \notin \{j, k\} \end{cases}$$

and

$$q_i = \begin{cases} p_i + \varepsilon : i = j \\ p_i - \varepsilon : i = k \\ p_i \quad\;\; : i \notin \{j, k\}. \end{cases}$$

By construction, $\rho_1 := \sum_i r_i \,|v_i\rangle\,\langle v_i|$ and $\rho_2 := \sum_i q_i \,|v_i\rangle\,\langle v_i|$ are density operators too, and $\rho = \frac{1}{2}\rho_1 + \frac{1}{2}\rho_2$. However, $\rho_1 \neq \rho \neq \rho_2$, so ρ is not an extreme point of $\mathscr{D}(M)$. Contradiction, since $\rho \in \partial_e \mathscr{D}(M)$ by assumption. Therefore, all $p_i \in \{0, 1\}$. Combined with $\sum_i p_i = 1$, this gives a unique j such that $p_j = 1$ and $p_k = 0$ for all $k \neq j$. But then, $\rho = \left|v_j\right\rangle\left\langle v_j\right|$, so we see that every extreme point of $\mathscr{D}(M)$ is indeed a one-dimensional projection.

It is clear that every one-dimensional projection is positive and has unit trace, so every one-dimensional projection is clearly a density operator. Now take a one-dimensional projection $\rho = |\psi\rangle\,\langle\psi|$, i.e. a unit vector ψ. Suppose that there are $\rho_1, \rho_2 \in \mathscr{D}(M)$ and a $t \in (0, 1)$ such that $\rho = t\rho_1 + (1 - t)\rho_2$.

Clearly, we have $\langle\psi, \rho\psi\rangle = 1$. Using the spectral decomposition $\sum_i p_i \,|v_i\rangle\,\langle v_i|$ of ρ_1, where the $\{v_i\}$ are orthonormal, all $p_i \geq 0$ and $\sum_i p_i = 1$, we see that:

$$\langle\psi, \rho_1\psi\rangle = \sum_i p_i |\langle\psi, v_i\rangle|^2 \leq \sum_i p_i = 1,$$

by the Cauchy-Schwarz inequality.

By the same token, $\langle\psi, \rho_2\psi\rangle \leq 1$. Therefore,

$$1 = \langle\psi, \rho\psi\rangle = t\langle\psi, \rho_1\psi\rangle + (1 - t)\langle\psi, \rho_2\psi\rangle \leq t + (1 - t) = 1.$$

Therefore, we must have $\langle\psi, \rho_1\psi\rangle = 1$, so for all j such that $p_j \neq 0$, we have $|\langle\psi, v_j\rangle|^2 = 1$. Since ψ is a unit vector and $\{v_i\}$ is an orthonormal set, this means that there is a unique j such that $p_j \neq 0$ and $\psi = zv_i$ with $z \in \mathbb{C}$ such that $|z| = 1$.

But then necessarily $p_j = 1$ and $\rho_1 = \left|v_j\right\rangle\left\langle v_j\right| = |\psi\rangle\,\langle\psi| = \rho$. Likewise, $\rho_2 = \rho$, so indeed, ρ is an extreme point.

So $\partial_e \mathscr{D}(M)$ consists exactly of the one-dimensional projections. Now, under the correspondence of Theorem 2.6,

$$f(a) = \mathrm{Tr}(|\psi\rangle\,\langle\psi|\, a) = \langle\psi, |\psi\rangle\,\langle\psi|\, a\psi\rangle = \langle\psi, a\psi\rangle,$$

where we used an orthonormal basis with ψ as one of the basis vectors for evaluating the trace. □

In the same fashion we can also define (pure) states on D and derive their specific forms. Note that for $a \in D$ the notion of positivity when considering it as an element of M, i.e. $\langle x, ax \rangle \geq 0$ for all $x \in \mathbb{C}^n$, is equivalent to saying that all $a_{ii} \geq 0$.

Definition 2.10 A **state** on D is a linear function $f : D \to \mathbb{C}$ that is positive and unital, meaning that $f(a) \geq 0$ for all $a \geq 0$ and $f(1) = 1$. The set of all states on D is denoted by $S(D)$ and is called the **state space** of D.

In our discussion about the the specific form of states on D, we need (to repeat) the notion of a probability distribution on finite sets.

Definition 2.11 Let X be a finite set. Then a **probability distribution** on X is a function $p : X \to [0, \infty)$ such that $\sum_x p(x) = 1$. The set of all probability distributions on X is denoted by $Pr(X)$.

Note that a probability distribution p on a finite set X is equivalently defined as a map $p : X \to [0, 1]$ such that $\sum_x p(x) = 1$.

Theorem 2.12 *There is a bijective correspondence between states f on D and probability distributions p on $\{1, \ldots, n\}$ such that $f(a) = \sum_i p(i)a_{ii}$ for all $a \in D$.*

Proof We want to show that $S(D) \cong Pr(\{1, \ldots, n\})$ as sets.

Define $\Phi : S(D) \to Pr(\{1, \ldots, n\})$ by

$$\Phi(f)(k) = f(|e_k\rangle \langle e_k|)$$

for all k. Then since f is a state, each $\Phi(f)(k)$ is positive. Furthermore,

$$\sum_i \Phi(f)(i) = \sum_i f(|e_i\rangle \langle e_i|) = f(\sum_i |e_i\rangle \langle e_i|) = f(1) = 1,$$

so $\Phi(f)$ is indeed a probability distribution. Next, define the function $\Psi : Pr(\{1, \ldots, n\}) \to S(D)$ by

$$\Psi(p)(a) = \sum_i p(i)a_{ii}.$$

Since all $p(i)$ are positive, it is clear that $\Psi(p)$ is positive too. Furthermore,

$$\Psi(p)(1) = \sum_i p(i) = 1,$$

so $\Psi(p)$ is indeed a state. Now note that

$$\Psi(\Phi(f))(a) = \sum_i \Phi(f)(i)a_{ii} = \sum_i a_{ii} f(|e_i\rangle \langle e_i|) = f(\sum_i a_{ii} |e_i\rangle \langle e_i|) = f(a),$$

showing that $\Psi \circ \Phi = \mathrm{Id}$.

Furthermore,

$$\Phi(\Psi(p))(k) = \Psi(p)(|e_k\rangle\langle e_k|) = \sum_i p(i)(|e_k\rangle\langle e_k|)_{ii} = p(k),$$

whence $\Phi \circ \Psi = \mathrm{Id}$.

So, indeed, $S(D) \cong Pr(\{1, \ldots, n\})$ as sets and writing $p = \Phi(f)$, the given formula $f(a) = \sum_i p(i)a_{ii}$ holds for every $a \in D$. □

Next, we note that just like in the case of M, the state space $S(D)$ is in fact a convex set, just like $Pr(\{1, \ldots, n\})$. Hence we can again determine the boundary of $S(D)$ and call it the **pure state space** of D. Once again, these pure states have a specific form.

Corollary 2.13 *For every pure state f on D there is an $i \in \{1, \ldots, n\}$ such that $f(a) = a_{ii}$ for all $a \in D$.*

Proof By Theorem 2.12 we know that the states on D correspond to $Pr(\{1, \ldots, n\})$, and by Lemma 2.8 we then know that we only have to determine the boundary of $Pr(\{1, \ldots, n\})$. If we show that these are exactly those probability distributions that have a unique j such that $p(j) = 1$ and $p(k) = 0$ for all $k \neq j$, we are done.

So, suppose that $p \in \partial_e Pr(\{1, \ldots, n\})$. By definition of a probability distribution, we have $p(j) \in [0, 1]$ for all j. Suppose that $p(j) \in (0, 1)$ for some j. Then there must be a $k \neq j$ such that $p(k) \in (0, 1)$ as well. Then there is a $\varepsilon > 0$ such that

$$[p(j) - \varepsilon, p(j) + \varepsilon] \subseteq [0, 1]$$

and

$$[p(k) - \varepsilon, p(k) + \varepsilon] \subseteq [0, 1].$$

By the same reasoning as in the proof of Corollary 2.9, p is not an extreme point. Contradiction. Hence there is no j such that $p(j) \in (0, 1)$, so all $p(j) \in \{0, 1\}$. Therefore, there is a unique j such that $p(j) = 1$ and $p(k) = 0$ for all $k \neq j$.

Now suppose p is a probability distribution such that there is a unique j such that $p(j) = 1$ and $p(k) = 0$ for all $k \neq j$. Then suppose that we have a $t \in (0, 1)$ and $r, q \in Pr(\{1, \ldots, n\})$ such that $p = tr + (1 - t)q$. Suppose that $r(j) \neq 1$. Then $r(j) < 1$, since all $r(k) \geq 0$ and $\sum_k r(k) = 1$. Then $q(j) > 1$, which is a contradiction. Hence $r(j) = 1$. Likewise, $q(j) = 1$. Then, since $r, q \in Pr(\{1, \ldots, n\}$, $r(k) = 0 = q(k)$ for all $k \neq j$. Therefore $p = q = r$ and p is an extreme point. □

2.2 Extensions of Pure States

We have now established the ingredients to get to the main point of this chapter. By definition of the state spaces, it is clear that when restricting a state on M one obtains a state on D. The question we can now ask ourselves is whether this restriction determines the original state completely, i.e. whether we can *uniquely extend* a state on D to a state on M. It turns out that it does when we consider pure states, as formulated in the following theorem.

Theorem 2.14 *For every pure state f on D there is a unique pure state g on M that extends f.*

Proof Let f be a pure state on D. By Corollary 2.13, there is an $i \in \{1, \ldots, n\}$ such that $f(a) = a_{ii}$ for all $a \in D$.

Now simply define the linear function $g : M \to \mathbb{C}$ by

$$g(a) = a_{ii}$$

for all $a \in M$. Then clearly, $g(a) = a_{ii} = \langle e_i, ae_i \rangle \geq 0$ for all $a \geq 0$, so g is positive. Furthermore, g is obviously unital, so g is a state that extends f.

Suppose that g' is another pure state that extends f. Then by Corollary 2.9, there is a $\psi \in \mathbb{C}^n$ such that $g'(a) = \langle \psi, a\psi \rangle$ for all $a \in M$.

Let us write $\psi = \sum_k c_k e_k$. Then, since $|e_k\rangle \langle e_k| \in D$ for all k:

$$|c_k|^2 = g'(|e_k\rangle \langle e_k|) = f(|e_k\rangle \langle e_k|) = \delta_{ik}$$

Therefore, $\psi = c_i e_i$, with $|c_i| = 1$. Then for any $a \in M$,

$$g'(a) = \langle \psi, a\psi \rangle = \overline{c_i} c_i \langle e_i, ae_i \rangle = |c_i|^2 a_{ii} = a_{ii} = g(a),$$

so $g' = g$, and g is the unique pure state extension of f. □

http://www.springer.com/978-3-319-47701-5

The Kadison-Singer Property
Stevens, M.
2016, X, 140 p., Softcover
ISBN: 978-3-319-47701-5