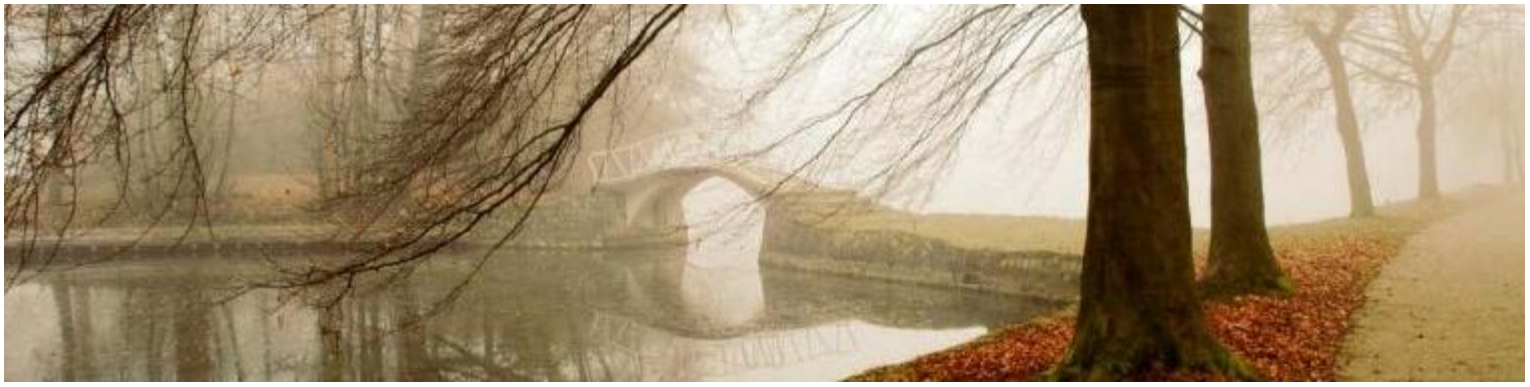
"A good stack of examples, as large as possible, is indispensable for a thorough understanding of any concept, and when I want to learn something new, I make it my first job to build one." – Paul Halmos



« Internal equivalence relations

Notes »

Noncommutative probability

August 18, 2012 by Qiaochu Yuan

The traditional mathematical axiomatization of probability, due to Kolmogorov, begins with a probability space P and constructs random variables as certain functions $P \to \mathbb{R}$. But start doing any probability and it becomes clear that the space P is de-emphasized as much as possible; the real focus of probability theory is on the algebra of random variables. It would be nice to have an approach to probability theory that reflects this.

Moreover, in the traditional approach, random variables necessarily commute. However, in quantum mechanics, the random variables are self-adjoint operators on a Hilbert space H, and these do not commute in general. For the purposes of doing quantum probability, it is therefore also natural to look for an approach to probability theory that begins with an algebra, not necessarily commutative, which encompasses both the classical and quantum cases.

Happily, **noncommutative probability** provides such an approach. Terence Tao's notes on free probability develop a version of noncommutative probability approach geared towards applications to random matrices, but today I would like to take a more leisurely and somewhat scattered route geared towards getting a general feel for what this formalism is capable of talking about.

Classical and quantum probability

(Below, if the reader chooses, she can restrict herself to finite sets and finite-dimensional Hilbert spaces so as to ignore measure-theoretic and analytic difficulties.)

A **classical probability space** $(S, \mathcal{E}, \mathbb{P})$ consists of the following data:

RECENT COMMENTS


Alison Miller on The man who knew elliptic inte…


Qiaochu Yuan on The man who knew elliptic inte…


Alison Miller on The man who knew elliptic inte…


Qiaochu Yuan on The man who knew elliptic inte…


Alison Miller on The man who knew elliptic inte…

BLOGROLL

Academically Interesting
Azimuth
Climbing Mount Bourbaki
Combinatorics and more
Concrete Nonsense
Geometry and the imagination
Gowers's Weblog
John Baez's Stuff
Michael Nielsen
Quomodocumque
Secret Blogging Seminar
Shtetl-Optimized

1. A set S, the **sample space**. Elements of this set describe possible states of some system.
2. A **σ-algebra** $\mathcal{E}$ of subsets of S, the **events**. Events describe properties of states. The pair $(S, \mathcal{E})$ is a **measurable space**
3. A **probability measure** $\mathbb{P}$ on $(S, \mathcal{E})$. This measures the probability that the system has some property.

Example. Let $S = \{H, T\}^n$ be the set of possible outcomes of n coin flips. Letting $\mathcal{E}$ be the set of all subsets of S, we can describe various events like "no heads are flipped" or "at most three tails are flipped," and we can compute their probabilities using the fact that every point in S has probability $\frac{1}{2^n}$.

Example. Let M be a $2n$-dimensional symplectic manifold with symplectic form ω (e.g. the cotangent bundle of some other manifold). The n^{th} exterior power of the symplectic form defines a volume form on M which defines a Borel measure on M called Liouville measure (locally just Lebesgue measure). Since Liouville measure is built from the symplectic form, it is preserved under all symplectomorphisms, and in particular under time evolution with respect to any Hamiltonian.

A **random variable** is a measurable function $X : S \to \mathbb{R}$ (where $\mathbb{R}$ is given the **Borel σ-algebra** generated by the Euclidean topology). In our coin-flipping example, "number of heads flipped" is a random variable. A random variable which only takes the values 0 or 1 encodes the same data as an event (more precisely, it is the indicator function χ_E of a unique event, which takes the value 1 on E and 0 on its complement). More generally, we can construct events from random variables: for any Borel subset $E \subseteq \mathbb{R}$ the preimage $X^{-1}(E)$ is an event (the event that X lies in E), often written $X \in E$, and so we can consider its probability $\mathbb{P}(X \in E)$. (This is the pushforward measure.)

Random variables should be thought of as real-valued observables of our system (and events, as random variables which take the value 0 or 1, are the observables given by answers to yes-no questions). By repeatedly measuring an observable and averaging, we can obtain its **expected value**

$$\mathbb{E}(X) = \int_S X \, d\mu.$$

(if this integral converges). If X only takes the values $0, 1$, then $\mathbb{E}(X)$ reduces to the probability of the corresponding event. In general, if X is a random variable and E is a Borel subset of $\mathbb{R}$, then $\chi_E(X)$ is the indicator function of $X^{-1}(E)$, and $\mathbb{P}(X \in E) = \mathbb{E}(\chi_E(X))$.

If we wanted to define a **quantum probability space** by analogous data, it would consist of the following (not standard):

1. A Hilbert space H, the **space of states**.
2. An abstract σ-algebra $\mathcal{H}$ of closed subspaces of H, the **events**. The intersection of two subspaces is their set-theoretic intersection, the union is the closure of their span, and the complement is the orthogonal complement.
3. A unit vector $\psi \in H$, the **state vector**.

Archives

Example. Every classical probability space $(S, \mathcal{E}, \mathbb{P})$ defines a quantum probability space as follows: H is the Hilbert space $L^2(S, \mathbb{P})$ of (equivalence classes of) square-integrable functions $S \to \mathbb{C}$ under the inner product

$$\langle f, g \rangle = \int_S \overline{f(s)} g(s) \, d\mu,$$

$\mathcal{H}$ consists of the closed subspaces of functions which are (a.e.) equal to zero except on a given event $E \in \mathcal{E}$, and ψ is the function $S \to \mathbb{C}$ which is identically equal to 1.

Example. The quantum probability space describing a qubit comes from applying the above construction to a bit; thus $H = \mathbb{C}^2$ is a 2-dimensional Hilbert space with orthonormal basis $|0\rangle, |1\rangle$, $\mathcal{H}$ consists of four subspaces $0, \text{span}(|0\rangle), \text{span}(|1\rangle), H$, and $\psi = c_0|0\rangle + c_1|1\rangle$ is the state of the qubit.

A quantum probability space does not have points in the classical sense, but we can still talk about the probability of an event E: if P_E denotes the projection onto E, then it is given by

$$\mathbb{P}(E) = \langle \psi, P_E \psi \rangle$$

and writing ψ as the sum of its components parallel and orthogonal to E we see that this is the square of the absolute value of the component of ψ parallel to E. This is a simple form of the **Born rule**, and it describes the probability that ψ, when measured to determine whether or not it lies in E, will in fact lie in E. Applied to a qubit, we conclude that a qubit described by $\psi = c_0|0\rangle + c_1|1\rangle$, when measured, takes the value 0 with probability $|c_0|^2$ and the value 1 with probability $|c_1|^2$.

Note that if E is the entire Hilbert space then the condition that the corresponding probability is 1 is precisely the condition that ψ is a unit vector. Note also that the probability assigned by ψ to an event does not change if ψ is multiplied by a unit complex number; for this reason, state vectors are really points in the projective space over H. Thus the possible states of a qubit are parameterized by the Riemann sphere (called in this context the Bloch sphere).

A (real-valued) **quantum random variable** (probably not standard) is a self-adjoint operator X on H (possibly unbounded and/or densely defined in general). The values taken by X are precisely its **spectral values** (the points in its spectrum $\sigma(X)$). This specializes even to the classical case: the values λ for which a random variable $X : S \to \mathbb{R}$ has the property that $\lambda - X$ fails to be invertible are precisely its values (up to the subtlety that we can ignore the behavior of X on a set of measure zero, but in practice we cannot meaningfully evaluate random variables at points anyway). In particular, for X bounded, X takes only the values $0, 1$ if and only if it is idempotent by Gelfand-Naimark, hence if and only if it is a projection; thus as in the classical case, random variables generalize events.

The expected value of a quantum random variable is

$$\mathbb{E}(X) = \langle \psi, X\psi \rangle$$

(when ψ lies in the domain of X). If X happens to have a countable orthonormal basis ψ_k of eigenvectors with eigenvalues λ_k, then writing $\psi = \sum c_k \psi_k$ we compute that

May 2013
April 2013
March 2013
February 2013
November 2012
October 2012
September 2012
August 2012
July 2012
June 2012
May 2012
April 2012
March 2012
February 2012
January 2012
December 2011
November 2011
October 2011
August 2011
July 2011
June 2011
April 2011
March 2011
February 2011
January 2011
December 2010
November 2010
October 2010
September 2010
August 2010
July 2010
June 2010
May 2010
April 2010
March 2010
February 2010
January 2010
December 2009
November 2009
October 2009
September 2009
August 2009
July 2009
June 2009
May 2009
April 2009
January 2009

Tag Cloud
2-categories

$$\langle \psi, X\psi \rangle = \sum_k \lambda_k |c_k|^2$$

so this really is the expected value of X if we think of it classically as a random variable taking the value λ_k with probability $\langle \psi, P_{\psi_k}\psi \rangle = |c_k|^2$.

As in the classical case, we can make sense of probabilities such as $\mathbb{P}(X \in E)$ where E is a Borel subset of $\mathbb{R}$, but this requires more work. If X has a countable orthonormal basis this is straightforward; in general, we need the **Borel functional calculus** in order to define $\chi_E(X)$ as an operator so that we can compute $\mathbb{E}(\chi_E(X))$ (note that we do not need $\chi_E(X)$ to be a projection whose image lies in $\mathcal{H}$ to compute this expectation, although this would be the appropriate analogue of the random variable X being measurable). Roughly speaking we ought to be able to start from the continuous functional calculus and approximate the indicator function of E by continuous functions, then show that the corresponding limit exists as a self-adjoint operator.

Unlike the classical case, the expected value $\langle \psi, X\psi \rangle$ can be computed independently of any measurability hypotheses on X; in particular, the probability of a particular event occurring (that is, the expected value of an arbitrary projection) is automatically well-defined.

Noncommutative probability

The classical and quantum cases above have several features in common. In both cases we saw that, although we started with a description of events and their probabilities and moved on to a description of random variables and their expected values, we could recover events through their indicator functions as the idempotent random variables and recover probabilities of events as expected values of indicator functions. This suggests that we might fruitfully approach probability in general using algebras of random variables and the expectation.

If the algebra is commutative, we might hope to recover an underlying probability space, but a random-variables-first approach will allow us to work independently of a particular representation of a family of random variables as an algebra of functions on a probability space. If the algebra is noncommutative, we might hope to recover a Hilbert space on which it acts, but again, a random-variables-first approach will allow us to work independently of a particular representation as operators on a Hilbert space. We can also think of the algebra as the algebra of functions on a noncommutative space in the spirit of noncommutative geometry. Although noncommutative spaces don't have a good notion of point, quantum probability spaces suggest that they have a good notion of measure (which we can think of as a "smeared-out" point, the Dirac measures corresponding to ordinary points).

The following definition is morally due to von Neumann and Segal. A **random algebra** (not standard) is a complex $\dagger$-algebra A together with a $\dagger$-linear functional $\mathbb{E} : A \to \mathbb{C}$ such that

$$\mathbb{E}(a^\dagger a) \geq 0, \mathbb{E}(1) = 1.$$

Such a functional is called a **state** on A (as it describes the state of some probabilistic system by describing the expected value of observables). The (real-valued) **random variables** in A are its self-adjoint elements (and an **event** is a projection; that is, a self-adjoint idempotent).

adjoint functors Catalan numbers characteristic classes cobordism cohomology Coxeter groups diagrammatic algebra dualizable objects Dynkin diagrams elliptic curves enriched categories Euler characteristic finite fields fixed point theorems Fourier transforms genera generating functions group actions groupoids Hecke algebras Lawvere theories MaBloWriMo MathOverflow modular forms moments monads monoidal categories Nullstellensatz partition functions pedagogy philosophy of mathematics profinite groups q-analogues quantization quantum groups quaternions self-reference species theory symmetric functions ultrafilters universal properties walks on graphs Young tableaux zeta functions

A morphism of random algebras $(A_1, \mathbb{E}_1) \to (A_2, \mathbb{E}_2)$ is a morphism $\phi : A_1 \to A_2$ of complex $\dagger$-algebras such that $\mathbb{E}_2 \circ \phi = \mathbb{E}_1$. This defines the category Rand of random algebras, and the category of **noncommutative probability spaces** is Rand^{op}. (This is probably the wrong choice of morphisms, but we'll ignore that for now.)

Example. From a classical probability space $(S, \mathcal{E}, \mathbb{P})$ we obtain a random algebra by letting A be the von Neumann algebra $L^\infty(S)$ of essentially bounded measurable functions $S \to \mathbb{C}$ under conjugation and letting $\mathbb{E}$ be the integral.

Example. From a quantum probability space $(H, \mathcal{H}, \psi)$ we obtain a random algebra by letting A be the span of the space of self-adjoint operators $a : H \to H$ such that $\chi_E(a) \in \mathcal{H}$ for all Borel subsets $E \subseteq \mathbb{R}$ and letting $\mathbb{E}$ be the functional $a \mapsto \langle \psi, a\psi \rangle$.

(Because we have not developed the Borel functional calculus, it will be cleaner just to work with an arbitrary $\dagger$-algebra of bounded operators $H \to H$ from which $\mathcal{H}$ can but need not be derived. We can do the same thing in the classical case by starting with a collection of functions $S \to \mathbb{C}$ and taking the preimages under all of them of the Borel subsets of $\mathbb{C}$ to define a σ-algebra on S.)

The above examples require some analysis to define in full generality. However, the reason we do not require any analytic hypotheses on A is to have a formalism flexible enough to discuss more algebraic examples such as the following.

Example. Let G be a group. The group algebra $\mathbb{C}[G]$ is a $\dagger$-algebra in the usual way (with involution extending $g^\dagger = g^{-1}$, so that every element of G is unitary). There is a distinguished state given by $\mathbb{E}(1) = 1$ and $\mathbb{E}(g) = 0$ for every non-identity g.

The axioms we have chosen require some explanation. Working in a complex $\dagger$-algebra is both convenient and has clear ties to quantum mechanics, but I do not have a good explanation of this axiom from first principles. The condition that $\mathbb{E}$ is $\dagger$-linear reflects linearity of expectation, which holds both in the classical and quantum cases, and the fact that we want the expected value of a self-adjoint element to be real. The condition that $\mathbb{E}(a^\dagger a) \geq 0$ (**positivity**) reflects the fact that we want probabilities to be non-negative in the following sense.

In any complex $\dagger$-algebra A, we may define a **positive** (really non-negative) element to be an element of the form $a^\dagger a, a \in A$. A positive element is in particular self-adjoint. In the case of measurable functions on a probability space, the positive elements are precisely the elements which are (a.e.) non-negative, and in the case of operators on a Hilbert space, the positive elements are precisely the self-adjoint elements which have non-negative spectrum (~~by Gelfand representation~~ this is subtle; see the comments below). Hence positivity is a natural analogue in the algebraic setting of the condition that probabilities are non-negative.

Finally, the condition that $\mathbb{E}(1) = 1$ reflects the fact that we want the total probability to be 1.

The semi-inner product

The state allows us to define a bilinear map $\langle a, b \rangle = \mathbb{E}(a^\dagger b)$ on any random algebra A which satisfies all of the axioms of an inner product except that it is not necessarily positive-definite, but only satisfies the weaker axiom that $\langle a, a \rangle \geq 0$.

We call such a gadget a **semi-inner product** (since it is positive-semidefinite).

As for classical random variables we can define the **covariance**

$$\mathrm{Cov}(a,b) = \langle a - \mathbb{E}(a), b - \mathbb{E}(b)\rangle = \mathbb{E}(a^\dagger b) - \mathbb{E}(a^\dagger)\mathbb{E}(b)$$

of two elements, and positive-semidefiniteness implies that the **variance** $\mathrm{Var}(a) = \mathrm{Cov}(a,a)$ is non-negative, hence that $\mathbb{E}(a^\dagger a) \geq \mathbb{E}(a^\dagger)\mathbb{E}(a)$. More generally, the proof of the Cauchy-Schwarz inequality goes through without modification, and we conclude that

$$|\langle a,b\rangle|^2 \leq \langle a,a\rangle\langle b,b\rangle.$$

This is already enough for us to prove the following general version of Heisenberg's uncertainty principle.

Theorem (Robertson uncertainty): Let a, b be self-adjoint elements of a random algebra A. Then

$$\mathrm{Var}(a)\mathrm{Var}(b) \geq \frac{1}{4}|\mathbb{E}([a,b])|^2.$$

Proof. Since both sides are invariant under translation of either a or b by a nonzero constant, we may assume without loss of generality that a, b have mean zero (that is, that $\mathbb{E}(a) = \mathbb{E}(b) = 0$). This gives

$$\mathrm{Var}(a)\mathrm{Var}(b) \geq |\mathbb{E}(ab)|^2$$

by Cauchy-Schwarz. We can write ab as the sum of its real and imaginary parts

$$ab = \frac{ab+ba}{2} + i\frac{ab-ba}{2i}$$

and computing $|\mathbb{E}(ab)|^2$ using the above decomposition gives

$$|\mathbb{E}(ab)|^2 = \left|\mathbb{E}\left(\frac{a\circ b}{2}\right)\right|^2 + \left|\mathbb{E}\left(\frac{[a,b]}{2}\right)\right|^2 \geq \frac{1}{4}|\mathbb{E}([a,b])|^2.$$

where $a \circ b = ab + ba$ and $[a,b] = ab - ba$. The conclusion follows. □

Interpreting Robertson uncertainty will be easier once we do a little more work. By Cauchy-Schwarz, if an element satisfies $\langle n,n\rangle = 0$ then in fact it satisfies $\langle a,n\rangle = 0$ for all a (and the converse is clear). In the classical picture, a function satisfying either of these conditions is equal to zero almost everywhere, which motivates the following definition. An element n of a random algebra A is **null** or **zero almost surely** (abbreviated a.s.) if $\langle a,n\rangle = 0$ for all $a \in A$, which as we have seen is equivalent to $\langle n,n\rangle = 0$. The null elements N form a subspace of A. Two elements a, b are **equal almost surely** if $a - b \in N$, hence equality a.s. is equivalent to equality in the quotient A/N.

An element has variance zero if and only if it is constant almost surely. Robertson uncertainty then says that if two self-adjoint elements $a, b \in A$ have the property that their commutator $[a,b]$ has nonzero expectation, then neither of them can be constant almost surely in a strong sense: the product of their variances is bounded below by a positive constant, so as one increases, the other must decrease. In other words, not only are they uncertain, but a state in which a is less uncertain is a state in which b is more uncertain.

The standard application of Robertson uncertainty is to the case that a, b are the position and momentum operators respectively acting on a quantum particle on $\mathbb{R}$.

This application has the following purely mathematical interpretation: a function in $L^2(\mathbb{R})$ and its Fourier transform cannot simultaneously be too localized.

Independence

A fundamental notion in classical probability theory is the notion of **independence**. It can be generalized to random algebras as follows: two $\dagger$-subalgebras A, B of a random algebra C are independent if

$$\mathbb{E}(ab) = \mathbb{E}(ba) = \mathbb{E}(a)\mathbb{E}(b)$$

for all $a \in A, b \in B$.

Example. Let C be the random algebra $L^\infty(S)$ associated to a classical probability space $(S, \mathcal{E}, \mathbb{P})$ and let A, B be the subalgebras of functions which are measurable with respect to two σ-subalgebras $\mathcal{E}_1, \mathcal{E}_2$ of $\mathcal{E}$. Then A, B are independent in the above sense if and only if $\mathcal{E}_1, \mathcal{E}_2$ are independent in the sense that

$$\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(E_1)\mathbb{P}(E_2)$$

where $E_i \in \mathcal{E}_i$ (by the monotone class theorem). Note that this condition is equivalent to $\mathbb{E}(\chi_{E_1}\chi_{E_2}) = \mathbb{E}(\chi_{E_1})\mathbb{E}(\chi_{E_2})$.

Example. Let A, B be two random algebras with expectations $\mathbb{E}_A, \mathbb{E}_B$. Their tensor product $A \otimes B$ acquires a natural $\dagger$-algebra structure given by $(a \otimes b)^\dagger = a^\dagger \otimes b^\dagger$ on pure tensors (it is the universal $\dagger$-algebra admitting morphisms from A, B whose images commute), and moreover we can define on it a state given by

$$\mathbb{E}_{A \otimes B}(a \otimes b) = \mathbb{E}_A(a)\mathbb{E}_B(b)$$

on pure tensors. Conversely, any state on $A \otimes B$ such that A and B are independent is of this form. This is a noncommutative generalization of product measure; when A, B come from classical probability spaces S_1, S_2, a suitable completion of $A \otimes B$ is the corresponding algebra of functions on the product $S_1 \times S_2$, and the state above comes from integration against the corresponding product measure.

Example. A is independent of itself (in A) if and only if the state $\mathbb{E}$ is actually a homomorphism $A \to \mathbb{C}$ of $\dagger$-algebras. Thinking of the case that A is a C*-algebra in particular, the corresponding states can be thought of as **Dirac measures** supported at points of $\mathrm{MaxSpec}\ A$. In the noncommutative case, A may admit no homomorphisms to $\mathbb{C}$ (for example if A contains the Weyl algebra), hence no Dirac measures, an expression of the general intuition that noncommutative spaces are “smeared out” and not easily expressible in terms of points.

Independence is a formalization of the intuitive idea that knowing the values of the random variables in A doesn’t allow you to deduce anything about the values of the random variables in B and vice versa. One indication of how this works in the setting of random algebras is as follows: if $p \in B$ is a projection with $\mathbb{E}(p) > 0$ (that is, an event that occurs with positive probability) we can define a **conditional expectation**

$$\mathbb{E}(a|p) = \frac{\mathbb{E}(pap)}{\mathbb{E}(p)}.$$

(The first factor of p is necessary in the noncommutative case to ensure that the result is still a state.) This represents the expected value of a given that the event

p occurred. If A, B are independent, it follows that $\mathbb{E}(a|p) = \mathbb{E}(a)$; in other words, knowing that p occurred has no effect on the expected value of any of the elements of A.

Independence is a very strong condition to impose if the subalgebras A, B do not commute. For example, it implies that $\mathbb{E}(ab - ba) = 0$ for all $a \in A, b \in B$, which is the only condition under which Robertson uncertainty cannot relate the variances of a, b. In the particular case of the position and momentum operators, $ab - ba$ is a nonzero scalar, hence always has nonzero expectation; it follows that position and momentum cannot be made independent! (By contrast, in the classical setting any pair of random variables is independent with respect to a Dirac measure.)

In the noncommutative setting, a different notion of independence, free independence (replacing the tensor product with the free product), becomes more natural and useful. We will not discuss this issue further, but see Terence Tao's notes linked above.

The Gelfand-Naimark-Segal construction

If V is any inner product space, A any $\dagger$-algebra of linear operators on V, and $\psi \in V$ is any unit vector, then A is a **concrete random algebra** with expectation $\mathbb{E}(a) = \langle \psi, a\psi \rangle$. This subsumes the examples coming from both classical and quantum probability spaces. The goal of this section is to determine to what extent we can prove a Cayley's theorem for random algebras to the effect that random algebras are concrete.

The above suggests the following definition. If A is a complex $\dagger$-algebra, then a **$\dagger$-representation** of A is a homomorphism $A \to (V \Rightarrow V)$ from A to the endomorphisms of an inner product space V such that

$$\langle v, aw \rangle_V = \langle a^\dagger v, w \rangle_V$$

for all $v, w \in V$. (Note that if V is not a Hilbert space then $\mathrm{End}(V)$ is not necessarily a $\dagger$-algebra because adjoints may not exist in general.) A **Hilbert $\dagger$-representation** is a $\dagger$-representation on a Hilbert space.

The semi-inner product $\langle a, b \rangle = \mathbb{E}(a^\dagger b)$ on a random algebra A descends to the quotient space A/N, where it is becomes an inner product because we have quotiented by the elements of norm zero. Moreover, since

$$\langle a, n \rangle = 0 \forall n \Leftrightarrow \mathbb{E}(a^\dagger n) = 0 \Rightarrow \mathbb{E}(a^\dagger bn) = 0 \Leftrightarrow \langle a, bn \rangle = 0$$

it follows that N is a left ideal, so the quotient map $A \to A/N$ is a quotient of left A-modules; consequently, A acts on A/N by linear operators. Since

$$\langle v, aw \rangle = \mathbb{E}(v^\dagger aw) = \langle a^\dagger v, w \rangle$$

it follows that A/N defines a $\dagger$-representation of A. The procedure we have outlined is essentially the **Gelfand-Naimark-Segal (GNS) construction**: we associate to any state on a $\dagger$-algebra a corresponding $\dagger$-representation such that the state can be recovered from the representation as

$$\mathbb{E}(a) = \langle \psi, a\psi \rangle_{A/N}$$

where $\psi = 1$. This may be regarded as a weak Cayley's theorem: unfortunately, this $\dagger$-representation is not faithful in general. To get a stronger statement about random algebras, we will now assume another condition, namely that the if , then (the state is **faithful**).

$a^\dagger a \neq 0 \qquad \mathbb{E}(a^\dagger a) \neq 0$

The faithfulness axiom is equivalent to requiring $N = 0$ and also equivalent to requiring that A is an inner product space (rather than a semi-inner product space). It implies, but is stronger than, the assumption that the action of A on A/N is faithful. The remarks about the state above then prove the following.

"Cayley's theorem for random algebras": A random algebra with a faithful state is concrete.

This is still not a true analog of Cayley's theorem because the converse is false: the state $\mathbb{E}(a) = \langle \psi, a\psi \rangle$ of a concrete random algebra need not be faithful.

From here we will assume, in addition to faithfulness, another condition, namely that for every $a \in A$ there exists a constant C such that $|\langle v, av \rangle| \leq C \langle v, v \rangle$ (**boundedness**). Boundedness is equivalent to requiring that A acts on itself by bounded linear operators. This action therefore uniquely extends to the completion of A with respect to its inner product, which we'll denote by H, and consequently it follows that in this case A admits a Hilbert $\dagger$-representation $A \mapsto (H \Rightarrow H)$. The closure of the image of A in $(H \Rightarrow H)$ is a C*-algebra $\overline{A}$ of bounded linear operators on H, and moreover since $\mathbb{E}(a) = \langle \psi, a\psi \rangle$ where $\psi = 1$, the expectation uniquely extends to $\overline{A}$.

This motivates the following definition: a **random C*-algebra** is a random algebra which is also a C*-algebra. The above discussion proves the following.

Theorem: Let A be a random algebra with a faithful state satisfying boundedness. Then A canonically embeds as a dense $\dagger$-subalgebra of a random C*-algebra $\overline{A}$ equipped with a Hilbert $\dagger$-representation H via the GNS construction; moreover, there is a canonical vector $\psi \in H$ such that $\mathbb{E}(a) = \langle \psi, a\psi \rangle$ for all $a \in \bar{A}$.

This is a much stronger conclusion than the conclusion that A is concrete, since it allows us to use facts from the theory of C*-algebras.

Corollary: Let A be a commutative random algebra with a faithful state satisfying boundedness. Then A canonically embeds as a dense $\dagger$-subalgebra of the algebra of continuous functions on a compact Hausdorff space X.

Proof. The closure of a commutative $\dagger$-subalgebra of $H \Rightarrow H$ is also commutative, since commutativity is a continuous condition. The conclusion then follows from Gelfand-Naimark. □

Corollary ("Maschke's theorem"): Let A be a finite-dimensional random algebra with a faithful state. Then A is semisimple.

Proof. A finite-dimensional random algebra automatically satisfies boundedness. The GNS construction equips A with a faithful $\dagger$-representation, namely A itself. Let V be a submodule of A. Then for every $a \in A$,

$$\forall v \in V : \langle av, w \rangle = 0 \Leftrightarrow \forall v \in V : \langle v, a^\dagger w \rangle = 0$$

so $V^\perp$ is also a submodule of H. So every submodule of A is a direct summand; consequently, A is semisimple. □

Note that we really do recover Maschke's theorem for complex representations of finite groups as a corollary, since $\mathbb{C}[G]$ is a finite-dimensional random algebra with a faithful state.

Moments

The axioms for a random algebra may not seem strong enough to capture random variables. For example, it does not seem possible to directly access probabilities like $\mathbb{P}(a \in E)$. However, our axioms are enough to define the **moments**

$$\mathbb{E}(a), \mathbb{E}(a^2), \mathbb{E}(a^3), \dots$$

of a random variable, and under suitable hypotheses (discussed under the general heading of the moment problem) it is possible to recover a random variable in the classical sense from its moments. We prove a result of this type for random C*-algebras.

Proposition: Any state $\mathbb{E} : A \to \mathbb{C}$ on a C*-algebra has norm 1 (and in particular is continuous).

Proof. By examining real and imaginary parts, it suffices to show that a self-adjoint element $a \in A$ of norm 1 maps to an element of norm at most 1. Since $1 - a$ has non-negative spectrum, by the continuous functional calculus it has a square root, hence is positive, so

$$\mathbb{E}(1 - a) = 1 - \mathbb{E}(a) \geq 0.$$

Similarly, $1 + a$ has non-negative spectrum, so by the continuous functional calculus it has a square root, hence is positive, so

$$\mathbb{E}(1 + a) = 1 + \mathbb{E}(a) \geq 0.$$

We conclude that $|\mathbb{E}(a)| \leq 1$, with equality if $a = 1$. □

In fact a much stronger statement is true due to the following corollary of the Riesz-Markov theorem, which we will not prove; see Terence Tao's notes.

Theorem: Let X be a compact Hausdorff space and $\mathbb{E} : C(X) \to \mathbb{C}$ be a positive linear functional. Then there is a unique Radon measure μ on X such that

$$\mathbb{E}(f) = \int_X f \, d\mu$$

for all $f \in C(X)$ (and conversely any Radon measure defines a positive linear functional on $C(X)$).

It follows by Gelfand-Naimark that specifying a commutative random C*-algebra is equivalent to specifying a compact Hausdorff space and a Radon measure on it of total measure 1.

Corollary: Let A be a random C*-algebra. If $a \in A$ is normal, then there is a unique Radon measure μ on $X = \text{MaxSpec } \overline{\mathbb{C}[a, a^\dagger]}$ such that

$$\mathbb{E}(f(a)) = \int_X f(a) \, d\mu$$

for all continuous functions $f : \sigma(a) \to \mathbb{C}$ (where $f(a)$ is defined using the continuous functional calculus).

Proof. Since $a, a^\dagger$ is dense in $\overline{\mathbb{C}[a, a^\dagger]} \cong C(X)$ by construction, a morphism $C(X) \to \mathbb{C}$ is uniquely determined by what it does to a, hence a, regarded as a function $X \to \sigma(a) \subset \mathbb{C}$, is injective. Since it is a continuous map between compact Hausdorff spaces, it is also an embedding, so we may regard X as canonically embedded into $\sigma(a)$. (This embedding is actually a homeomorphism but we do not need this.) By Tietze extension, any continuous function $X \to \mathbb{C}$ extends to a continuous function $\sigma(a) \to \mathbb{C}$, so the continuous functions

$f(a) : X \to \mathbb{C}$ given by applying the continuous functional calculus to a include all continuous functions $X \to \mathbb{C}$, and we reduce to the previous result. □

Corollary: With the same hypotheses as above, the Radon measure μ above is uniquely determined by the values

$$\mathbb{E}(p(a, a^\dagger)) = \int_X p \, d\mu$$

where p is a polynomial in z and $\bar{z}$. Consequently, μ is uniquely determined by the **$\dagger$-moments** $\mathbb{E}(a^n (a^\dagger)^m), n, m \geq 0$. If a is self-adjoint, μ is uniquely determined by the values

$$\mathbb{E}(p(a)) = \int_X p \, d\mu$$

where p is a polynomial in one variable. Equivalently, μ is uniquely determined by the moments $\mathbb{E}(a^n), n \geq 0$.

Proof. $\sigma(a)$ is a compact subset of $\mathbb{C}$, so by Stone-Weierstrass the polynomial functions in z and $\bar{z}$ are uniformly dense in the space of continuous functions $\sigma(a) \to \mathbb{C}$. Now recall that the continuous functional calculus and $\mathbb{E}$ both preserve uniform limits. If a is self-adjoint, $\sigma(a)$ is real, so we only need to take polynomial functions in z. □

The proofs above generalize essentially unchanged to the following.

Corollary: Let $a_1, \ldots a_k$ be commuting normal elements of a random C*-algebra A. Then there exists a unique Radon measure μ on $X = \text{MaxSpec } \overline{\mathbb{C}[a_1, a_1^\dagger, \ldots a_k, a_k^\dagger]}$ such that

$$\mathbb{E}(f(a_1, \ldots a_k)) = \int_X f(a_1, \ldots a_k) \, d\mu$$

for all continuous functions $f : \sigma(a_1) \times \ldots \times \sigma(a_k) \to \mathbb{C}$. Furthermore, μ is uniquely determined by the **joint $\dagger$-moments**

$$\mathbb{E}(a_1^{n_1} (a_1^\dagger)^{m_1} a_2^{n_2} (m_2^\dagger)^{e_2} \ldots), n_i, m_i \geq 0$$

of the a_i. If the a_i are self-adjoint, μ is uniquely determined by the **joint moments** $\mathbb{E}(a_1^{n_1} \ldots a_k^{n_k}), n_i \in \mathbb{Z}_{\geq 0}$ of the a_i.

The hypothesis that the a_i commute is crucial in the following sense. We restrict to the self-adjoint case for simplicity.

Proposition: Let a, b be self-adjoint elements of a random C*-algebra A with faithful state such that there exists a measure μ on a measure space X and two measurable functions $x, y : X \to \mathbb{R}$ satisfying

$$\mathbb{E}(a^{n_1} b^{m_1} a^{n_2} b^{m_2} \ldots a^{n_k} b^{m_k}) = \int_X x^{\sum n_i} y^{\sum m_i} \, d\mu$$

for all $n_i, m_i \in \mathbb{Z}_{\geq 0}$. Then $ab = ba$.

Proof. If a, b are self-adjoint then so is $c = \frac{ab - ba}{i}$. The hypothesis above implies that $\mathbb{E}(c^2) = 0$, but since c is self-adjoint c^2 is positive, hence by faithfulness $c = 0$. □

This result may be interpreted as saying that two noncommuting random variables do not in general have a reasonable notion of joint distribution.

Some closing remarks about quantumness

Classical mechanics is in principle deterministic: if the initial state of a system is known deterministically, then classical mechanics can in principle determine all future states. The predictions of quantum mechanics are, however, probabilistic: all that can be determined is a probability distribution on possible outcomes of a given experiment.

The two can be made to seem more similar if classical mechanics is generalized by allowing the state of the system to be probabilistic in the classical sense. Then classical and quantum mechanics can both be subsumed under the heading of random algebras, where in the classical case we do not keep track of the position and momentum of a particle but a probability distribution over all possible positions and momenta. What distinguishes the classical from the quantum cases is the noncommutativity of the random algebras in the latter case, and in particular the fact that the random algebras occurring in quantum mechanics generally do not admit any homomorphisms $A \to \mathbb{C}$, hence admit no Dirac measures, so we are forced to always work probabilistically.

The formal similarity between classical and quantum mechanics described here only applies to states and observables; to get time evolution back into the picture we should endow our random algebras with Poisson brackets, giving us **random Poisson algebras**, and Hamiltonians…

Related

Finite noncommutative

The Heisenberg picture of quantum

Noncommutative probability and

probability, the Born rule, and wave function collapse
In "math.PR"

mechanics
In "physics.quant-ph"

group theory
In "math.GR"

Posted in math.PR, physics.class-ph, physics.quant-ph | Tagged moments | 7 Comments

7 Responses

SNR focus or Quantum / thermodynamics focus ? « Peter's ruminations on December 6, 2012 at 12:41 pm | Reply

[…] https://qchu.wordpress.com/2012/08/18/noncommutative-probability/ […]

home_pw@msn.com on October 11, 2012 at 8:32 am | Reply

Reblogged this on Peter's ruminations and commented:
This nicely explains Turings argument form in his on permutations manuscript. Now I can see the form of the argument is indeed based on sound references to the art of the day.

Finite noncommutative probability, the Born rule, and wave function collapse « Annoying Precision on September 9, 2012 at 11:34 pm | Reply

[…] previous post on noncommutative probability was too long to leave much room for examples of random algebras. In […]

Erik Crevier on August 19, 2012 at 11:23 am | Reply

This is really beautiful stuff.

Maybe I'll comment more later, but for now only one tiny technical quibble: You claimed that an operator on a Hilbert space which is algebraically positive (of the form $a^{\dagger}a$) has non-negative, by Gelfand-Naimark. This is true, but it's not quite that obvious. Gelfand-Naimark can't be applied directly if you're not assuming a to be normal, since then $C*(a,a^{\dagger}$ need not be commutative, so you need to do some extra work. The usual proof (see any $C*$ algebra book) uses at least one unenlightening hat pull computation, and I think this might be unavoidable.

Qiaochu Yuan on August 19, 2012 at 1:24 pm | Reply

Thanks for the comment! You're right that it doesn't directly follow from the Gelfand representation (we don't

a

need the full power of Gelfand-Naimark here) if is not normal. The argument I had in mind is the following: if λ is a positive scalar then we have $\langle v, (a^\dagger a + \lambda)v \rangle \geq \lambda \langle v, v \rangle$ for any $v \in H$, hence $a^\dagger a + \lambda$ is invertible. But I should clarify this.

Erik Crevier on August 19, 2012 at 3:57 pm | Reply

Ahh, that's prettier than what I was thinking.

Also, how do I get LaTeX to display properly in comments?

Qiaochu Yuan on August 19, 2012 at 4:10 pm | Reply

You need to add the word "latex" immediately after the first dollar sign. (This is extra work but it prevents actual dollar signs on WordPress blogs from being unfortunately misinterpreted.)

Comments RSS

Leave a Reply