A categorical view of conditional expectation

Prakash Panangaden

School of Computer Science
McGill University

Applied Category Theory - 8^{th} April 2020

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Joint work with

Chaput, Danos and Plotkin

Philippe Chaput, Vincent Danos, Prakash Panangaden, and Gordon Plotkin. "Approximating Markov processes by averaging." Journal of the ACM (JACM) 61, no. 1 (2014): 1-45.

The idea of functorializing conditional expectation is due to Vincent.

Conditional probability

- Probability triple: (Ω, Σ, p). Random process produces outcome distributed according to p.

Conditional probability

- Probability triple: (Ω, Σ, p). Random process produces outcome distributed according to p.
- If $\omega \in \Omega$ is the outcome and $A \in \Sigma$, then the probability that $\omega \in A$ is $p(A)$.

Conditional probability

- Probability triple: (Ω, Σ, p). Random process produces outcome distributed according to p.
- If $\omega \in \Omega$ is the outcome and $A \in \Sigma$, then the probability that $\omega \in A$ is $p(A)$.
- If we *know* that $\omega \in B$ then we reassess the probability of $\omega \in A$: $P(A|B) = p(A \cap B)/p(B)$.

Conditional probability

- Probability triple: (Ω, Σ, p). Random process produces outcome distributed according to p.
- If $\omega \in \Omega$ is the outcome and $A \in \Sigma$, then the probability that $\omega \in A$ is $p(A)$.
- If we *know* that $\omega \in B$ then we reassess the probability of $\omega \in A$: $P(A|B) = p(A \cap B)/p(B)$.
- If we know whether ω is in B or B^c we can compute $P(A|B)$ and $P(A|B^c)$. Define $f(\omega) = P(A|B)$ if $\omega \in B$ and $f(\omega) = P(A|B^c)$ if $\omega \in B^c$.

Conditional probability

- Probability triple: (Ω, Σ, p). Random process produces outcome distributed according to p.
- If $\omega \in \Omega$ is the outcome and $A \in \Sigma$, then the probability that $\omega \in A$ is $p(A)$.
- If we *know* that $\omega \in B$ then we reassess the probability of $\omega \in A$: $P(A|B) = p(A \cap B)/p(B)$.
- If we know whether ω is in B or B^c we can compute $P(A|B)$ and $P(A|B^c)$. Define $f(\omega) = P(A|B)$ if $\omega \in B$ and $f(\omega) = P(A|B^c)$ if $\omega \in B^c$.
- If we are given a countable partition $\{B_i | B_i \in \Sigma\}$ of Ω we can define a function $f : \Omega \longrightarrow [0, 1]$ such that $f(\omega) = P(A|B_i)$ if $\omega \in B_i$.

Conditional probability

- Probability triple: (Ω, Σ, p). Random process produces outcome distributed according to p.
- If $\omega \in \Omega$ is the outcome and $A \in \Sigma$, then the probability that $\omega \in A$ is $p(A)$.
- If we *know* that $\omega \in B$ then we reassess the probability of $\omega \in A$: $P(A|B) = p(A \cap B)/p(B)$.
- If we know whether ω is in B or B^c we can compute $P(A|B)$ and $P(A|B^c)$. Define $f(\omega) = P(A|B)$ if $\omega \in B$ and $f(\omega) = P(A|B^c)$ if $\omega \in B^c$.
- If we are given a countable partition $\{B_i | B_i \in \Sigma\}$ of Ω we can define a function $f : \Omega \longrightarrow [0, 1]$ such that $f(\omega) = P(A|B_i)$ if $\omega \in B_i$.
- If we are given $\Lambda \subset \Sigma$ and for every $B \in \Lambda$ we know whether $\omega \in B$ we can define the random variable $P[A||\Lambda]$ which is Λ-measurable and
$$\forall B \in \Lambda \int_B P[A||\Lambda] \mathrm{d}p = P(A \cap B).$$

Conditional expectation

- Sample space: (Ω, Σ, p), random variable $f : \Omega \longrightarrow \mathbf{R}$.

Conditional expectation

- Sample space: (Ω, Σ, p), random variable $f : \Omega \longrightarrow \mathbf{R}$.
- Expectation value of f: $\int f \mathrm{d}p$.

Conditional expectation

- Sample space: (Ω, Σ, p), random variable $f : \Omega \longrightarrow \mathbf{R}$.
- Expectation value of f: $\int f \mathrm{d}p$.
- We want to *revise* our expectation based on new information.

Conditional expectation

- Sample space: (Ω, Σ, p), random variable $f : \Omega \longrightarrow \mathbf{R}$.
- Expectation value of f: $\int f \mathrm{d}p$.
- We want to *revise* our expectation based on new information.
- Given a sub-σ-algebra $\Lambda \subset \Sigma$ we define $E[f||\Lambda]$ as a Λ-**measurable** function such that

$$\forall B \in \Lambda, \int_B E[f||\Lambda] \mathrm{d}p = \int_B f \mathrm{d}p.$$

Conditional expectation

- Sample space: (Ω, Σ, p), random variable $f : \Omega \to \mathbf{R}$.
- Expectation value of f: $\int f \mathrm{d}p$.
- We want to *revise* our expectation based on new information.
- Given a sub-σ-algebra $\Lambda \subset \Sigma$ we define $E[f||\Lambda]$ as a Λ-**measurable** function such that

$$\forall B \in \Lambda, \int_B E[f||\Lambda] \mathrm{d}p = \int_B f \mathrm{d}p.$$

- Why isn't $E[f||\Lambda]$ just f?

Conditional expectation

- Sample space: (Ω, Σ, p), random variable $f : \Omega \longrightarrow \mathbf{R}$.
- Expectation value of f: $\int f \mathrm{d}p$.
- We want to *revise* our expectation based on new information.
- Given a sub-σ-algebra $\Lambda \subset \Sigma$ we define $E[f||\Lambda]$ as a Λ-**measurable** function such that

$$\forall B \in \Lambda, \int_B E[f||\Lambda]\mathrm{d}p = \int_B f \mathrm{d}p.$$

- Why isn't $E[f||\Lambda]$ just f?
- Because it is only Λ measurable; so much "smoother."

Banach spaces

- A *norm* on a vector space V is a function $\|\cdot\| : V \longrightarrow \mathbf{R}^{\geq 0}$ such that:

Banach spaces

- A *norm* on a vector space V is a function $\|\cdot\| : V \longrightarrow \mathbf{R}^{\geq 0}$ such that:
  1. $\|v\| = 0$ iff $v = 0$

Banach spaces

- A *norm* on a vector space V is a function $\|\cdot\| : V \longrightarrow \mathbf{R}^{\geq 0}$ such that:
  1. $\|v\| = 0$ iff $v = 0$
  2. $\|r \cdot v\| = \mid r \mid \|v\|$ and

Banach spaces

- A *norm* on a vector space V is a function $\|\cdot\| : V \longrightarrow \mathbf{R}^{\geq 0}$ such that:
  1. $\|v\| = 0$ iff $v = 0$
  2. $\|r \cdot v\| = \mid r \mid \|v\|$ and
  3. $\|x + y\| \leq \|x\| + \|y\|$.

Banach spaces

- A *norm* on a vector space V is a function $\|\cdot\| : V \longrightarrow \mathbf{R}^{\geq 0}$ such that:
  1. $\|v\| = 0$ iff $v = 0$
  2. $\|r \cdot v\| = \mid r \mid \|v\|$ and
  3. $\|x + y\| \leq \|x\| + \|y\|$.
- The norm induces a metric: $d(u, v) = \|u - v\|$ and, hence, a topology. This topology is called the *norm topology*.

Banach spaces

- A *norm* on a vector space V is a function $\|\cdot\| : V \longrightarrow \mathbf{R}^{\geq 0}$ such that:
  1. $\|v\| = 0$ iff $v = 0$
  2. $\|r \cdot v\| = \mid r \mid \|v\|$ and
  3. $\|x + y\| \leq \|x\| + \|y\|$.
- The norm induces a metric: $d(u, v) = \|u - v\|$ and, hence, a topology. This topology is called the *norm topology*.
- If V is complete in this metric it is called a Banach space.

Maps between Banach spaces

- A linear map $T : U \longrightarrow V$ is *bounded* if there exists a positive real number α such that $\forall u \in U, \|Tu\| \leq \alpha \|u\|$.

Maps between Banach spaces

- A linear map $T : U \longrightarrow V$ is *bounded* if there exists a positive real number α such that $\forall u \in U, \|Tu\| \leq \alpha \|u\|$.
- A lineap map between normed spaces is *continuous* iff it is *bounded*.

Maps between Banach spaces

- A linear map $T : U \longrightarrow V$ is *bounded* if there exists a positive real number α such that $\forall u \in U, \|Tu\| \leq \alpha \|u\|$.
- A lineap map between normed spaces is *continuous* iff it is *bounded*.
- Given a bounded linear map between normed spaces $T : U \longrightarrow V$ we define $\|T\| = \sup \{\|Tu\| \mid u \in U, \|u\| \leq 1\}$.

Maps between Banach spaces

- A linear map $T : U \longrightarrow V$ is *bounded* if there exists a positive real number α such that $\forall u \in U, \|Tu\| \leq \alpha \|u\|$.
- A lineap map between normed spaces is *continuous* iff it is *bounded*.
- Given a bounded linear map between normed spaces $T : U \longrightarrow V$ we define $\|T\| = \sup\{\|Tu\| \mid u \in U, \|u\| \leq 1\}$.
- This is a norm on the space of bounded linear maps and is called the *operator norm*.

Maps between Banach spaces

- A linear map $T : U \longrightarrow V$ is *bounded* if there exists a positive real number α such that $\forall u \in U, \|Tu\| \leq \alpha \|u\|$.
- A lineap map between normed spaces is *continuous* iff it is *bounded*.
- Given a bounded linear map between normed spaces $T : U \longrightarrow V$ we define $\|T\| = \sup \{\|Tu\| \mid u \in U, \|u\| \leq 1\}$.
- This is a norm on the space of bounded linear maps and is called the *operator norm*.
- With this norm the space of bounded linear maps between Banach spaces forms a Banach space.

Duality for Banach spaces

- The space of bounded ($=$ continuous) linear maps from V, a Banach space, to $\mathbf{R}$ is itself a Banach space, called the *dual space*, V^*.

Duality for Banach spaces

- The space of bounded ($=$ continuous) linear maps from V, a Banach space, to $\mathbf{R}$ is itself a Banach space, called the *dual space*, V^*.
- For any two vector spaces U, V we say that they are in *algebraic duality* if there is a bilinear form $\langle\cdot,\ \cdot\rangle : U \times V \longrightarrow \mathbf{R}$ such that spaces of functionals $\langle\cdot,\ V\rangle$ and $\langle U,\ \cdot\rangle$ separates points of U and V.

Duality for Banach spaces

- The space of bounded ($=$ continuous) linear maps from V, a Banach space, to $\mathbf{R}$ is itself a Banach space, called the *dual space*, V^*.
- For any two vector spaces U, V we say that they are in *algebraic duality* if there is a bilinear form $\langle\cdot,\ \cdot\rangle : U \times V \longrightarrow \mathbf{R}$ such that spaces of functionals $\langle\cdot,\ V\rangle$ and $\langle U,\ \cdot\rangle$ separates points of U and V.
- We say two Banach spaces are in *duality* if $\langle\cdot,\ V\rangle \subseteq U^*$ and $\langle U,\ \cdot\rangle \subseteq V^*$.

Duality for Banach spaces

- The space of bounded ($=$ continuous) linear maps from V, a Banach space, to $\mathbf{R}$ is itself a Banach space, called the *dual space*, V^*.
- For any two vector spaces U, V we say that they are in *algebraic duality* if there is a bilinear form $\langle \cdot, \cdot \rangle : U \times V \longrightarrow \mathbf{R}$ such that spaces of functionals $\langle \cdot, V \rangle$ and $\langle U, \cdot \rangle$ separates points of U and V.
- We say two Banach spaces are in *duality* if $\langle \cdot, V \rangle \subseteq U^*$ and $\langle U, \cdot \rangle \subseteq V^*$.
- For V a Banach space, the spaces V and V^* are in duality.

Duality for Banach spaces

- The space of bounded ($=$ continuous) linear maps from V, a Banach space, to $\mathbf{R}$ is itself a Banach space, called the *dual space*, V^*.
- For any two vector spaces U, V we say that they are in *algebraic duality* if there is a bilinear form $\langle \cdot, \ \cdot \rangle : U \times V \longrightarrow \mathbf{R}$ such that spaces of functionals $\langle \cdot, \ V \rangle$ and $\langle U, \ \cdot \rangle$ separates points of U and V.
- We say two Banach spaces are in *duality* if $\langle \cdot, \ V \rangle \subseteq U^*$ and $\langle U, \ \cdot \rangle \subseteq V^*$.
- For V a Banach space, the spaces V and V^* are in duality.
- The bilinear form is $\langle v, \ \phi \rangle = \phi(v)$.

Duality for Banach spaces

- The space of bounded ($=$ continuous) linear maps from V, a Banach space, to $\mathbf{R}$ is itself a Banach space, called the *dual space*, V^*.
- For any two vector spaces U, V we say that they are in *algebraic duality* if there is a bilinear form $\langle \cdot, \ \cdot \rangle : U \times V \longrightarrow \mathbf{R}$ such that spaces of functionals $\langle \cdot, \ V \rangle$ and $\langle U, \ \cdot \rangle$ separates points of U and V.
- We say two Banach spaces are in *duality* if $\langle \cdot, \ V \rangle \subseteq U^*$ and $\langle U, \ \cdot \rangle \subseteq V^*$.
- For V a Banach space, the spaces V and V^* are in duality.
- The bilinear form is $\langle v, \ \phi \rangle = \phi(v)$.
- There is a canonical injection $\iota : V \longrightarrow V^{**}$; if this is an isometry we say that the Banach space V is reflexive.

Duality for Banach spaces

- The space of bounded ($=$ continuous) linear maps from V, a Banach space, to $\mathbf{R}$ is itself a Banach space, called the *dual space*, V^*.
- For any two vector spaces U, V we say that they are in *algebraic duality* if there is a bilinear form $\langle\cdot,\ \cdot\rangle : U \times V \longrightarrow \mathbf{R}$ such that spaces of functionals $\langle\cdot,\ V\rangle$ and $\langle U,\ \cdot\rangle$ separates points of U and V.
- We say two Banach spaces are in *duality* if $\langle\cdot,\ V\rangle \subseteq U^*$ and $\langle U,\ \cdot\rangle \subseteq V^*$.
- For V a Banach space, the spaces V and V^* are in duality.
- The bilinear form is $\langle v,\ \phi\rangle = \phi(v)$.
- There is a canonical injection $\iota : V \longrightarrow V^{**}$; if this is an isometry we say that the Banach space V is reflexive.
- Infinite dimensional Banach spaces are **not necessarily** reflexive.

Duality for Banach spaces

- The space of bounded ($=$ continuous) linear maps from V, a Banach space, to $\mathbf{R}$ is itself a Banach space, called the *dual space*, V^*.
- For any two vector spaces U, V we say that they are in *algebraic duality* if there is a bilinear form $\langle \cdot,\ \cdot \rangle : U \times V \longrightarrow \mathbf{R}$ such that spaces of functionals $\langle \cdot,\ V \rangle$ and $\langle U,\ \cdot \rangle$ separates points of U and V.
- We say two Banach spaces are in *duality* if $\langle \cdot,\ V \rangle \subseteq U^*$ and $\langle U,\ \cdot \rangle \subseteq V^*$.
- For V a Banach space, the spaces V and V^* are in duality.
- The bilinear form is $\langle v,\ \phi \rangle = \phi(v)$.
- There is a canonical injection $\iota : V \longrightarrow V^{**}$; if this is an isometry we say that the Banach space V is reflexive.
- Infinite dimensional Banach spaces are **not necessarily** reflexive.
- Finite dimensional Banach spaces are always reflexive.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are μ-almost everywhere the same and we actually work with equivalence classes.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are μ-almost everywhere the same and we actually work with equivalence classes.
- The integral defines a norm on these equivalence classes and gives the Banach space $L_1(X, \Sigma, \mu)$ or just $L_1(\mu)$.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are μ-almost everywhere the same and we actually work with equivalence classes.
- The integral defines a norm on these equivalence classes and gives the Banach space $L_1(X, \Sigma, \mu)$ or just $L_1(\mu)$.
- The space $L_p(\mu)$ is the space obtained by defining the norm $\|f\|_p = (\int |f|^p \, \mathrm{d}\mu)^{\frac{1}{p}}$, where $0 < p < \infty$.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are μ-almost everywhere the same and we actually work with equivalence classes.
- The integral defines a norm on these equivalence classes and gives the Banach space $L_1(X, \Sigma, \mu)$ or just $L_1(\mu)$.
- The space $L_p(\mu)$ is the space obtained by defining the norm $\|f\|_p = (\int | f |^p \mathrm{d}\mu)^{\frac{1}{p}}$, where $0 < p < \infty$.
- The *infinity norm* of a measurable function f is $\|f\|_\infty = \inf \{M > 0 \mid | f(x) | \leq M$ for $\mu -$ almost all $x\}$.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are μ-almost everywhere the same and we actually work with equivalence classes.
- The integral defines a norm on these equivalence classes and gives the Banach space $L_1(X, \Sigma, \mu)$ or just $L_1(\mu)$.
- The space $L_p(\mu)$ is the space obtained by defining the norm $\|f\|_p = (\int | f |^p \mathrm{d}\mu)^{\frac{1}{p}}$, where $0 < p < \infty$.
- The *infinity norm* of a measurable function f is $\|f\|_\infty = \inf \{M > 0 \mid | f(x) | \leq M \text{ for } \mu - \text{almost all } x\}$.
- The collection of all equivalence classes of measurable functions f with $\|f\|_\infty < \infty$ with the norm just defined is the space $L_\infty(\mu)$.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are μ-almost everywhere the same and we actually work with equivalence classes.
- The integral defines a norm on these equivalence classes and gives the Banach space $L_1(X, \Sigma, \mu)$ or just $L_1(\mu)$.
- The space $L_p(\mu)$ is the space obtained by defining the norm $\|f\|_p = (\int |f|^p \mathrm{d}\mu)^{\frac{1}{p}}$, where $0 < p < \infty$.
- The *infinity norm* of a measurable function f is $\|f\|_\infty = \inf \{M > 0 \mid |f(x)| \leq M$ for $\mu -$ almost all $x\}$.
- The collection of all equivalence classes of measurable functions f with $\|f\|_\infty < \infty$ with the norm just defined is the space $L_\infty(\mu)$.
- These are all Banach spaces.

Duality for L_p spaces

- If $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ then L_p and L_q are duals of each other!

Duality for L_p spaces

- If $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ then L_p and L_q are duals of each other!
- However, L_1 and L_∞ are *not duals*.

Duality for L_p spaces

- If $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ then L_p and L_q are duals of each other!
- However, L_1 and L_∞ are *not duals*.
- The dual of L_1 is L_∞ but not the other way around!

Duality for L_p spaces

- If $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ then L_p and L_q are duals of each other!
- However, L_1 and L_∞ are *not duals*.
- The dual of L_1 is L_∞ but not the other way around!
- We will switch to a cone view and the situation will be much improved.

What are cones?

- Want to combine linear structure with order structure.

What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.

What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include 0.

What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include 0.
- Then for any positive $v \in P$ and positive real r, $rv \in P$. For $u, v \in P$ we have $u + v \in P$ and if $v \in P$ and $-v \in P$ then $v = 0$.

What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include 0.
- Then for any positive $v \in P$ and positive real r, $rv \in P$. For $u, v \in P$ we have $u + v \in P$ and if $v \in P$ and $-v \in P$ then $v = 0$.
- We *define* a **cone** C in a vector space V to be a set with exactly these conditions.

What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include 0.
- Then for any positive $v \in P$ and positive real r, $rv \in P$. For $u, v \in P$ we have $u + v \in P$ and if $v \in P$ and $-v \in P$ then $v = 0$.
- We *define* a **cone** C in a vector space V to be a set with exactly these conditions.
- Any cone defines a order by $u \leq v$ if $v - u \in C$.

What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include 0.
- Then for any positive $v \in P$ and positive real r, $rv \in P$. For $u, v \in P$ we have $u + v \in P$ and if $v \in P$ and $-v \in P$ then $v = 0$.
- We *define* a **cone** C in a vector space V to be a set with exactly these conditions.
- Any cone defines a order by $u \leq v$ if $v - u \in C$.
- Unfortunately for us, many of the structures that we want to look at are cones but are not part of any obvious vector space: *e.g.* the measures on a space.

What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include 0.
- Then for any positive $v \in P$ and positive real r, $rv \in P$. For $u, v \in P$ we have $u + v \in P$ and if $v \in P$ and $-v \in P$ then $v = 0$.
- We *define* a **cone** C in a vector space V to be a set with exactly these conditions.
- Any cone defines a order by $u \leq v$ if $v - u \in C$.
- Unfortunately for us, many of the structures that we want to look at are cones but are not part of any obvious vector space: *e.g.* the measures on a space.
- We could artificially embed them in a vector space, for example, by introducing signed measures.

Abstract cones d'après Selinger

Definition of Cones

A **cone** is a commutative monoid $(V, +, 0)$ with an action of $\mathbb{R}^{\geq 0}$. Multiplication by reals distributes over addition and the following cancellation law holds:

$$\forall u, v, w \in V, v + u = w + u \Rightarrow v = w.$$

The following strictness property also holds:

$$v + w = 0 \Rightarrow v = w = 0.$$

Note that every cone comes with a natural order.

An order on a cone

If $u, v \in V$, a cone, one says $u \leq v$ if and only if there is an element $w \in V$ such that $u + w = v$.

Normed cones

Definition of a normed cone

A **normed cone** C is a cone with a function
$||\cdot|| : C \longrightarrow \mathbb{R}^{\geq 0}$ satisfying the usual conditions:
$||v|| = 0$ if and only if $v = 0$
$\forall r \in \mathbf{R}^{\geq 0}, v \in C, ||r \cdot v|| = r||v||$
$||u+v|| \leq ||u|| + ||v||$
$u \leq v \Rightarrow ||u|| \leq ||v||$.

Normally one uses norms to talk about convergence of Cauchy sequences. But without negation how can we talk about Cauchy sequences?

Normed cones

Definition of a normed cone

A **normed cone** C is a cone with a function $||\cdot|| : C \longrightarrow \mathbb{R}^{\geq 0}$ satisfying the usual conditions:
$||v|| = 0$ if and only if $v = 0$
$\forall r \in \mathbf{R}^{\geq 0}, v \in C, ||r \cdot v|| = r||v||$
$||u + v|| \leq ||u|| + ||v||$
$u \leq v \Rightarrow ||u|| \leq ||v||$.

Normally one uses norms to talk about convergence of Cauchy sequences. But without negation how can we talk about Cauchy sequences?

We can write $u_i - u_j$ when we really mean the (unique) w such that $u_j + w = u_i$; needs $u_j \leq u_i$. So, in the case that we have an increasing sequence we can define Cauchy sequence in, more or less, the usual way.

Completeness

However, order-theoretic concepts can be used instead.

Complete normed cones

An ω-**complete normed cone** is a normed cone such that if $\{a_i \mid i \in I\}$ is an increasing sequence with $\{||a_i||\}$ bounded then the lub $\bigvee_{i\in I} a_i$ exists and $\bigvee_{i\in I} ||a_i|| = ||\bigvee_{i\in I} a_i||$.

Convergence in the sense of norm and in the order theory sense match.

Completeness

However, order-theoretic concepts can be used instead.

Complete normed cones

An ω-**complete normed cone** is a normed cone such that if $\{a_i \mid i \in I\}$ is an increasing sequence with $\{||a_i||\}$ bounded then the lub $\bigvee_{i \in I} a_i$ exists and $\bigvee_{i \in I} ||a_i|| = ||\bigvee_{i \in I} a_i||$.

Convergence in the sense of norm and in the order theory sense match.

Selinger's lemma

Suppose that u_i is an ω-chain with a l.u.b. in an ω-complete normed cone and u is an upper bound of the u_i. Suppose furthermore that $\lim_{i \longrightarrow \infty} \|u - u_i\| = 0$. Then $u = \bigvee_i u_i$.

Completeness

However, order-theoretic concepts can be used instead.

Complete normed cones

An ω-**complete normed cone** is a normed cone such that if $\{a_i \mid i \in I\}$ is an increasing sequence with $\{||a_i||\}$ bounded then the lub $\bigvee_{i \in I} a_i$ exists and $\bigvee_{i \in I} ||a_i|| = || \bigvee_{i \in I} a_i||$.

Convergence in the sense of norm and in the order theory sense match.

Selinger's lemma

Suppose that u_i is an ω-chain with a l.u.b. in an ω-complete normed cone and u is an upper bound of the u_i. Suppose furthermore that $\lim_{i \longrightarrow \infty} \|u - u_i\| = 0$. Then $u = \bigvee_i u_i$.

Here we are writing $u - u_i$ informally

We really mean w_i where $u_i + w_i = u$.

Maps between cones

Continuous maps

An ω-**continuous** linear map between two cones is one that preserves least upper bounds of countable chains.

Maps between cones

Continuous maps

An ω-**continuous** linear map between two cones is one that preserves least upper bounds of countable chains.

Bounded maps

A *bounded* linear map of normed cones $f : C \longrightarrow D$ is one such that for all u in C, $||f(u)|| \leq K||u||$ for some real number K. Any linear continuous map of complete normed cones is bounded.

Maps between cones

Continuous maps

An ω-**continuous** linear map between two cones is one that preserves least upper bounds of countable chains.

Bounded maps

A *bounded* linear map of normed cones $f : C \longrightarrow D$ is one such that for all u in C, $||f(u)|| \leq K||u||$ for some real number K. Any linear continuous map of complete normed cones is bounded.

Norm of a bounded map

The norm of a bounded linear map $f : C \longrightarrow D$ is defined as $||f|| = \sup\{||f(u)|| \,:\, u \in C,\ ||u|| \leq 1\}$.

A category of normed cones

The ambient category

The ω-complete normed cones, along with ω-continuous linear maps, form a category which we shall denote ω**CC**.

A category of normed cones

The ambient category

The ω-complete normed cones, along with ω-continuous linear maps, form a category which we shall denote $\omega\mathbf{CC}$.

The subcategory of interest

we define the subcategory $\omega\mathbf{CC}_1$: the norms of the maps are all bounded by 1. Isomorphisms in this category are always isometries.

Dual cones

Dual cone

Given an ω-complete normed cone C, its dual C^* is the set of all ω-continuous linear maps from C to $\mathbf{R}_+$. We define the norm on C^* to be the operator norm.

Dual cones

Dual cone

Given an ω-complete normed cone C, its dual C^* is the set of all ω-continuous linear maps from C to $\mathbf{R}_+$. We define the norm on C^* to be the operator norm.

Basic facts

C^* is an ω-complete normed cone as well, and the cone order corresponds to the point wise order.

The duality functor

In ω**CC**, the dual operation becomes a contravariant functor.
If $f : C \longrightarrow D$ is a map of cones, we define $f^* : D^* \longrightarrow C^*$ as follows: given a map L in D^*, we define a map f^*L in C^* as $f^*L(u) = L(f(u))$.

How does this compare with Banach spaces?

This dual is stronger than the dual in usual Banach spaces, where we only require the maps to be bounded. For instance, it turns out that the dual to $L^+_\infty(X)$ (to be defined later) is isomorphic to $L^+_1(X)$, which is not the case with the Banach space $L_\infty(X)$.

Cones that we use I

- If μ is a measure on X, then one has the well-known Banach spaces L_1 and L_∞.

Cones that we use I

- If μ is a measure on X, then one has the well-known Banach spaces L_1 and L_∞.
- These can be restricted to cones by considering the μ-almost everywhere positive functions.

Cones that we use I

- If μ is a measure on X, then one has the well-known Banach spaces L_1 and L_∞.
- These can be restricted to cones by considering the μ-almost everywhere positive functions.
- We will denote these cones by $L_1^+(X, \Sigma, \mu)$ and $L_\infty^+(X, \Sigma)$.

Cones that we use I

- If μ is a measure on X, then one has the well-known Banach spaces L_1 and L_∞.
- These can be restricted to cones by considering the μ-almost everywhere positive functions.
- We will denote these cones by $L_1^+(X, \Sigma, \mu)$ and $L_\infty^+(X, \Sigma)$.
- These are complete normed cones.

Cones that we use II

- Let (X, Σ, p) be a measure space with finite measure p. We denote by $\mathcal{M}^{\ll p}(X)$, the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p

Cones that we use II

- Let (X, Σ, p) be a measure space with finite measure p. We denote by $\mathcal{M}^{\ll p}(X)$, the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p
- If q is such a measure, we define its norm to be $q(X)$.

Cones that we use II

- Let (X, Σ, p) be a measure space with finite measure p. We denote by $\mathcal{M}^{\ll p}(X)$, the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p
- If q is such a measure, we define its norm to be $q(X)$.
- $\mathcal{M}^{\ll p}(X)$ is also an ω-complete normed cone.

Cones that we use II

- Let (X, Σ, p) be a measure space with finite measure p. We denote by $\mathcal{M}^{\ll p}(X)$, the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p
- If q is such a measure, we define its norm to be $q(X)$.
- $\mathcal{M}^{\ll p}(X)$ is also an ω-complete normed cone.
- The cones $\mathcal{M}^{\ll p}(X)$ and $L_1^+(X, \Sigma, p)$ are isometrically isomorphic in ω**CC**.

Cones that we use II

- Let (X, Σ, p) be a measure space with finite measure p. We denote by $\mathcal{M}^{\ll p}(X)$, the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p
- If q is such a measure, we define its norm to be $q(X)$.
- $\mathcal{M}^{\ll p}(X)$ is also an ω-complete normed cone.
- The cones $\mathcal{M}^{\ll p}(X)$ and $L_1^+(X, \Sigma, p)$ are isometrically isomorphic in ω**CC**.
- We write $\mathcal{M}^p_{\mathsf{UB}}(X)$ for the cone of all measures on (X, Σ) that are uniformly less than a multiple of the measure p: $q \in \mathcal{M}^p_{\mathsf{UB}}$ means that for some real constant $K > 0$ we have $q \le Kp$.

Cones that we use II

- Let (X, Σ, p) be a measure space with finite measure p. We denote by $\mathcal{M}^{\ll p}(X)$, the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p
- If q is such a measure, we define its norm to be $q(X)$.
- $\mathcal{M}^{\ll p}(X)$ is also an ω-complete normed cone.
- The cones $\mathcal{M}^{\ll p}(X)$ and $L_1^+(X, \Sigma, p)$ are isometrically isomorphic in ω**CC**.
- We write $\mathcal{M}^p_{\mathsf{UB}}(X)$ for the cone of all measures on (X, Σ) that are uniformly less than a multiple of the measure p: $q \in \mathcal{M}^p_{\mathsf{UB}}$ means that for some real constant $K > 0$ we have $q \leq Kp$.
- The cones $\mathcal{M}^p_{\mathsf{UB}}(X)$ and $L_\infty^+(X, \Sigma, p)$ are isomorphic.

Duality for cones

A Riesz-like theorem

The dual of the cone $L_\infty^+(X, \Sigma, p)$ is isometrically isomorphic to $\mathcal{M}^{\ll p}(X)$.

Duality for cones

A Riesz-like theorem

The dual of the cone $L_\infty^+(X, \Sigma, p)$ is isometrically isomorphic to $\mathcal{M}^{\ll p}(X)$.

Corollary

Since $\mathcal{M}^{\ll p}(X)$ is isometrically isomorphic to $L_1^+(X)$, an immediate corollary is that $L_\infty^{+,*}(X)$ is isometrically isomorphic to $L_1^+(X)$, which is of course false in general in the context of Banach spaces.

Duality for cones II

Another Riesz-like theorem

The dual of the cone $L_1^+(X, \Sigma, p)$ is isometrically isomorphic to $\mathcal{M}_{\mathsf{UB}}^p(X)$.

Duality for cones II

Another Riesz-like theorem

The dual of the cone $L_1^+(X, \Sigma, p)$ is isometrically isomorphic to $\mathcal{M}^p_{\mathrm{UB}}(X)$.

Corollary

$\mathcal{M}^p_{\mathrm{UB}}(X)$ is isometrically isomorphic to $L_\infty^+(X)$, hence immediate corollary is that $L_1^{+,*}(X)$ is isometrically isomorphic to $L_\infty^+(X)$.

The pairing

Pairing function

There is a map from the product of the cones $L_\infty^+(X,p)$ and $L_1^+(X,p)$ to $\mathbf{R}^+$ defined as follows:

$$\forall f \in L_\infty^+(X,p), g \in L_1^+(X,p) \quad \langle f,\ g\rangle = \int fg\mathrm{d}p.$$

The pairing

Pairing function

There is a map from the product of the cones $L_\infty^+(X,p)$ and $L_1^+(X,p)$ to $\mathbf{R}^+$ defined as follows:

$$\forall f \in L_\infty^+(X,p), g \in L_1^+(X,p) \quad \langle f,\ g\rangle = \int fg\mathrm{d}p.$$

This map is bilinear and is continuous and ω-continuous in both arguments; we refer to it as the pairing.

Duality expressed via pairing

This pairing allows one to express the dualities in a very convenient way. For example, the isomorphism between $L^+_\infty(X,p)$ and $(L^+_1(X,p))^*$ sends $f \in L^+_\infty(X,p)$ to $\lambda g.\langle f,\ g\rangle = \lambda g. \int fg\mathrm{d}p$.

Summary of cones

We fix a probability triple (X, Σ, p) and focus on six spaces of cones that are based on them. They break into two natural groups of three isomorphic spaces. The first three spaces are:

A1 $\mathcal{M}^{\ll p}(X)$ - the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p,

Summary of cones

We fix a probability triple (X, Σ, p) and focus on six spaces of cones that are based on them. They break into two natural groups of three isomorphic spaces. The first three spaces are:

A1 $\mathcal{M}^{\ll p}(X)$ - the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p,

A2 $L_1^+(X, p)$ - the cone of integrable almost-everywhere positive functions,

Summary of cones

We fix a probability triple (X, Σ, p) and focus on six spaces of cones that are based on them. They break into two natural groups of three isomorphic spaces. The first three spaces are:

- A1 $\mathcal{M}^{\ll p}(X)$ - the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p,
- A2 $L_1^+(X, p)$ - the cone of integrable almost-everywhere positive functions,
- A3 $L_\infty^{+,*}(X, p)$ - the dual cone of the the cone of almost-everywhere positive bounded measurable functions.

Summary of cones II

The next group of three isomorphic spaces are:

B1 $\mathcal{M}^{p}_{\mathsf{UB}}(X)$ - the cone of all measures that are uniformly less than a multiple of the measure p,

Summary of cones II

The next group of three isomorphic spaces are:

- B1 $\mathcal{M}^{p}_{\mathsf{UB}}(X)$ - the cone of all measures that are uniformly less than a multiple of the measure p,
- B2 $L^{+}_{\infty}(X,p)$ - the cone of almost-everywhere positive functions in the normed vector space $L_{\infty}(X,p)$,

Summary of cones II

The next group of three isomorphic spaces are:

- B1 $\mathcal{M}^{p}_{\mathsf{UB}}(X)$ - the cone of all measures that are uniformly less than a multiple of the measure p,
- B2 $L^{+}_{\infty}(X,p)$ - the cone of almost-everywhere positive functions in the normed vector space $L_{\infty}(X,p)$,
- B3 $L^{+,*}_{1}(X,p)$ - the dual of the cone of almost-everywhere positive functions in the normed vector space $L_{1}(X,p)$.

Summary of dualities and isos

The spaces defined in A1, A2 and A3 are dual to the spaces defined in B1, B2 and B3 respectively. The situation may be depicted in the diagram

$$\begin{tikzcd}
\mathcal{M}^{\ll p}(X) \arrow[r, leftrightarrow, "\sim"] \arrow[d, leftrightarrow, dotted] & L_1^+(X,p) \arrow[r, leftrightarrow, "\sim"] \arrow[d, leftrightarrow, dotted] & L_\infty^{+,*}(X,p) \arrow[d, leftrightarrow, dotted] \\
\mathcal{M}^p_{\mathsf{UB}} \arrow[r, leftrightarrow, "\sim"] & L_\infty^+(X,p) \arrow[r, leftrightarrow, "\sim"] & L_1^{+,*}(X,p)
\end{tikzcd} \tag{1}$$

where the vertical arrows represent dualities and the horizontal arrows represent isomorphisms.

Some measure theory

1. Given (X, Σ, p) and (Y, Λ) and a measurable function $f : X \longrightarrow Y$ we obtain a measure q on Y by $q(B) = p(f^{-1}(B))$. This is written $M_f(p)$ and is called the *image measure* of p under f.

Some measure theory

1. Given (X, Σ, p) and (Y, Λ) and a measurable function $f : X \longrightarrow Y$ we obtain a measure q on Y by $q(B) = p(f^{-1}(B))$. This is written $M_f(p)$ and is called the *image measure* of p under f.
2. We say that a measure ν is **absolutely continuous** with respect to another measure μ if for any measurable set A, $\mu(A) = 0$ implies that $\nu(A) = 0$. We write $\nu \ll \mu$.

Some measure theory

1. Given (X, Σ, p) and (Y, Λ) and a measurable function $f : X \longrightarrow Y$ we obtain a measure q on Y by $q(B) = p(f^{-1}(B))$. This is written $M_f(p)$ and is called the *image measure* of p under f.
2. We say that a measure ν is **absolutely continuous** with respect to another measure μ if for any measurable set A, $\mu(A) = 0$ implies that $\nu(A) = 0$. We write $\nu \ll \mu$.
3. For *finite* measures ν, $\nu \ll \mu$ is equivalent to:

$$\forall \varepsilon > 0, \exists \delta > 0, \ s.t. \ \forall A \text{ with } \mu(A) \leq \delta, \nu(A) \leq \varepsilon.$$

The Radon-Nikodym Theorem

The Radon-Nikodym theorem is a central result in measure theory allowing one to define a “derivative” of a measure with respect to another measure.

Radon-Nikodym

If $\nu \ll \mu$, where ν, μ are finite measures on a measurable space (X, Σ) there is a positive measurable function h on X such that for every measurable set B

$$\nu(B) = \int_B h \,\mathrm{d}\mu.$$

The function h is defined uniquely up to a set of μ-measure 0. The function h is called the Radon-Nikodym derivative of ν with respect to μ; we denote it by $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$. Since ν is finite, $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} \in L_1^+(X, \mu)$.

Notation for Radon-Nikodym

1. Given an (almost-everywhere) positive function $f \in L_1(X, p)$, we let $f \cdot p$ be the measure which has density f with respect to p.

Notation for Radon-Nikodym

1. Given an (almost-everywhere) positive function $f \in L_1(X, p)$, we let $f \cdot p$ be the measure which has density f with respect to p.
2. Two identities that we get from the Radon-Nikodym theorem are:

Notation for Radon-Nikodym

1. Given an (almost-everywhere) positive function $f \in L_1(X, p)$, we let $f \cdot p$ be the measure which has density f with respect to p.
2. Two identities that we get from the Radon-Nikodym theorem are:
   - given $q \ll p$, we have $\frac{\mathrm{d}q}{\mathrm{d}p} \cdot p = q$.

Notation for Radon-Nikodym

1. Given an (almost-everywhere) positive function $f \in L_1(X,p)$, we let $f \cdot p$ be the measure which has density f with respect to p.
2. Two identities that we get from the Radon-Nikodym theorem are:
   - given $q \ll p$, we have $\frac{\mathrm{d}q}{\mathrm{d}p} \cdot p = q$.
   - given $f \in L_1^+(X,p)$, $\frac{\mathrm{d}f \cdot p}{\mathrm{d}p} = f$

Notation for Radon-Nikodym

1. Given an (almost-everywhere) positive function $f \in L_1(X, p)$, we let $f \cdot p$ be the measure which has density f with respect to p.
2. Two identities that we get from the Radon-Nikodym theorem are:
   - given $q \ll p$, we have $\frac{\mathrm{d}q}{\mathrm{d}p} \cdot p = q$.
   - given $f \in L_1^+(X, p)$, $\frac{\mathrm{d}f \cdot p}{\mathrm{d}p} = f$
3. These two identities just say that the operations $(-) \cdot p$ and $\frac{\mathrm{d}(-)}{\mathrm{d}p}$ are inverses of each other as maps between $L_1^+(X, p)$ and $\mathcal{M}^{\ll p}(X)$ the space of finite measures on X that are absolutely continuous with respect to p.

Expectation and conditional expectation

1. The expectation $\mathbb{E}_p(f)$ of a measurable function f is the average computed by $\int f \mathrm{d}p$ and therefore it is just a number.

Expectation and conditional expectation

1. The expectation $\mathbb{E}_p(f)$ of a measurable function f is the average computed by $\int f \mathrm{d}p$ and therefore it is just a number.
2. The *conditional* expectation is not a mere number but a random variable.

Expectation and conditional expectation

1. The expectation $\mathbb{E}_p(f)$ of a measurable function f is the average computed by $\int f \mathrm{d}p$ and therefore it is just a number.
2. The *conditional* expectation is not a mere number but a random variable.
3. It is meant to measure the expected value in the presence of additional information.

Expectation and conditional expectation

1. The expectation $\mathbb{E}_p(f)$ of a measurable function f is the average computed by $\int f \mathrm{d}p$ and therefore it is just a number.
2. The *conditional* expectation is not a mere number but a random variable.
3. It is meant to measure the expected value in the presence of additional information.
4. The additional information takes the form of a sub-σ algebra, say Λ, of Σ. The experimenter knows, for every $B \in \Lambda$, whether the outcome is in B or not.

Expectation and conditional expectation

1. The expectation $\mathbb{E}_p(f)$ of a measurable function f is the average computed by $\int f \mathrm{d}p$ and therefore it is just a number.
2. The *conditional* expectation is not a mere number but a random variable.
3. It is meant to measure the expected value in the presence of additional information.
4. The additional information takes the form of a sub-σ algebra, say Λ, of Σ. The experimenter knows, for every $B \in \Lambda$, whether the outcome is in B or not.
5. Now she can recompute the expectation values given this information.

Formalizing conditional expectation

- It is an immediate consequence of the Radon-Nikodym theorem that such conditional expectations exist.

Formalizing conditional expectation

- It is an immediate consequence of the Radon-Nikodym theorem that such conditional expectations exist.

Kolmogorov

Let (X, Σ, p) be a measure space with p a finite measure, f be in $L_1(X, \Sigma, p)$ and Λ be a sub-σ-algebra of Σ, then there exists a $g \in L_1(X, \Lambda, p)$ such that for all $B \in \Lambda$

$$\int_B f \mathrm{d}p = \int_B g \mathrm{d}p.$$

Formalizing conditional expectation

- It is an immediate consequence of the Radon-Nikodym theorem that such conditional expectations exist.

Kolmogorov

Let (X, Σ, p) be a measure space with p a finite measure, f be in $L_1(X, \Sigma, p)$ and Λ be a sub-σ-algebra of Σ, then there exists a $g \in L_1(X, \Lambda, p)$ such that for all $B \in \Lambda$

$$\int_B f \mathrm{d}p = \int_B g \mathrm{d}p.$$

- This function g is usually denoted by $\mathbb{E}(f|\Lambda)$.

Formalizing conditional expectation

- It is an immediate consequence of the Radon-Nikodym theorem that such conditional expectations exist.

Kolmogorov

Let (X, Σ, p) be a measure space with p a finite measure, f be in $L_1(X, \Sigma, p)$ and Λ be a sub-σ-algebra of Σ, then there exists a $g \in L_1(X, \Lambda, p)$ such that for all $B \in \Lambda$

$$\int_B f \mathrm{d}p = \int_B g \mathrm{d}p.$$

- This function g is usually denoted by $\mathbb{E}(f|\Lambda)$.
- We clearly have $f \cdot p \ll p$ so the required g is simply $\frac{\mathrm{d}f \cdot p}{\mathrm{d}p|_\Lambda}$, where $p\mid_\Lambda$ is the restriction of p to the sub-σ-algebra Λ.

Properties of conditional expectation

1. The point of requiring Λ-measurability is that it "smooths out" variations that are too rapid to show up in Λ.

Properties of conditional expectation

1. The point of requiring Λ-measurability is that it "smooths out" variations that are too rapid to show up in Λ.
2. The conditional expectation is *linear*, *increasing* with respect to the pointwise order.

Properties of conditional expectation

1. The point of requiring Λ-measurability is that it “smooths out” variations that are too rapid to show up in Λ.
2. The conditional expectation is *linear*, *increasing* with respect to the pointwise order.
3. It is defined uniquely p-almost everywhere.

Where the action happens

- We define two categories $\mathbf{Rad}_\infty$ and $\mathbf{Rad}_1$ that will be needed for the functorial definition of conditional expectation.

Where the action happens

- We define two categories $\mathbf{Rad}_\infty$ and $\mathbf{Rad}_1$ that will be needed for the functorial definition of conditional expectation.
- This will allow for L_∞ and L_1 versions of the theory.

Where the action happens

- We define two categories $\mathbf{Rad}_\infty$ and $\mathbf{Rad}_1$ that will be needed for the functorial definition of conditional expectation.
- This will allow for L_∞ and L_1 versions of the theory.
- Going between these versions by duality will be very useful.

The "infinity" category

$\mathbf{Rad}_\infty$

The category $\mathbf{Rad}_\infty$ has as objects probability spaces, and as arrows $\alpha : (X, p) \longrightarrow (Y, q)$, measurable maps such that $M_\alpha(p) \leq Kq$ for some real number K.

The reason for choosing the name $\mathbf{Rad}_\infty$ is that $\alpha \in \mathbf{Rad}_\infty$ maps to $d/dqM_\alpha(p) \in L^+_\infty(Y, q)$.

The "one" category

$\mathbf{Rad}_1$

The category $\mathbf{Rad}_1$ has as objects probability spaces and as arrows $\alpha : (X, p) \longrightarrow (Y, q)$, measurable maps such that $M_\alpha(p) \ll q$.

The "one" category

$\mathbf{Rad}_1$

The category $\mathbf{Rad}_1$ has as objects probability spaces and as arrows $\alpha : (X, p) \longrightarrow (Y, q)$, measurable maps such that $M_\alpha(p) \ll q$.

1. The reason for choosing the name $\mathbf{Rad}_1$ is that $\alpha \in \mathbf{Rad}_1$ maps to $d/dqM_\alpha(p) \in L_1^+(Y, q)$.

The "one" category

$\mathbf{Rad}_1$

The category $\mathbf{Rad}_1$ has as objects probability spaces and as arrows $\alpha : (X, p) \longrightarrow (Y, q)$, measurable maps such that $M_\alpha(p) \ll q$.

1. The reason for choosing the name $\mathbf{Rad}_1$ is that $\alpha \in \mathbf{Rad}_1$ maps to $d/dqM_\alpha(p) \in L_1^+(Y, q)$.
2. The fact that the category $\mathbf{Rad}_\infty$ embeds in $\mathbf{Rad}_1$ reflects the fact that L_∞^+ embeds in L_1^+.

Pairing function revisited

Recall the isomorphism between $L_\infty^+(X,p)$ and $L_1^{+,*}(X,p)$ mediated by the pairing function:

$$f \in L_\infty^+(X,p) \mapsto \lambda g : L_1^+(X,p).\langle f,\ g\rangle = \int fg\mathrm{d}p.$$

Precomposition

1. Now, precomposition with α in $\mathbf{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y,q)$ to $L_1^+(X,p)$.

Precomposition

1. Now, precomposition with α in $\mathbf{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y,q)$ to $L_1^+(X,p)$.
2. Dually, given $\alpha \in \mathbf{Rad}_1 : (X,p) \longrightarrow (Y,q)$ and $g \in L_\infty^+(Y,q)$ we have that $P_\infty(\alpha)(g) \in L_\infty^+(X,p)$.

Precomposition

1. Now, precomposition with α in $\mathbf{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y,q)$ to $L_1^+(X,p)$.
2. Dually, given $\alpha \in \mathbf{Rad}_1 : (X,p) \longrightarrow (Y,q)$ and $g \in L_\infty^+(Y,q)$ we have that $P_\infty(\alpha)(g) \in L_\infty^+(X,p)$.
3. Thus the subscripts on the two precomposition functors describe the *target* categories.

Precomposition

1. Now, precomposition with α in $\mathbf{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y,q)$ to $L_1^+(X,p)$.
2. Dually, given $\alpha \in \mathbf{Rad}_1 : (X,p) \longrightarrow (Y,q)$ and $g \in L_\infty^+(Y,q)$ we have that $P_\infty(\alpha)(g) \in L_\infty^+(X,p)$.
3. Thus the subscripts on the two precomposition functors describe the *target* categories.
4. Using the $*$-functor we get a map $(P_1(\alpha))^*$ from $L_1^{+,*}(X,p)$ to $L_1^{+,*}(Y,q)$ in the first case and

Precomposition

1. Now, precomposition with α in $\mathbf{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y,q)$ to $L_1^+(X,p)$.
2. Dually, given $\alpha \in \mathbf{Rad}_1 : (X,p) \longrightarrow (Y,q)$ and $g \in L_\infty^+(Y,q)$ we have that $P_\infty(\alpha)(g) \in L_\infty^+(X,p)$.
3. Thus the subscripts on the two precomposition functors describe the *target* categories.
4. Using the $*$-functor we get a map $(P_1(\alpha))^*$ from $L_1^{+,*}(X,p)$ to $L_1^{+,*}(Y,q)$ in the first case and
5. dually we get $(P_\infty(\alpha))^*$ from $L_\infty^{+,*}(X,p)$ to $L_\infty^{+,*}(Y,q)$.

Expectation value functor

- The **functor** $\mathbb{E}_\infty(\cdot)$ is a functor from $\mathbf{Rad}_\infty$ to ω**CC** which, on objects, maps (X, p) to $L^+_\infty(X, p)$ and on maps is given as follows:

Expectation value functor

- The **functor** $\mathbb{E}_\infty(\cdot)$ is a functor from $\mathbf{Rad}_\infty$ to ω**CC** which, on objects, maps (X, p) to $L_\infty^+(X, p)$ and on maps is given as follows:
- Given $\alpha : (X, p) \longrightarrow (Y, q)$ in $\mathbf{Rad}_\infty$ the action of the functor is to produce the map $\mathbb{E}_\infty(\alpha) : L_\infty^+(X, p) \longrightarrow L_\infty^+(Y, q)$ obtained by composing $(P_1(\alpha))^*$ with the isomorphisms between $L_1^{+,*}$ and L_∞^+

Expectation value functor

- The **functor** $\mathbb{E}_\infty(\cdot)$ is a functor from $\mathbf{Rad}_\infty$ to ω**CC** which, on objects, maps (X,p) to $L^+_\infty(X,p)$ and on maps is given as follows:
- Given $\alpha : (X,p) \longrightarrow (Y,q)$ in $\mathbf{Rad}_\infty$ the action of the functor is to produce the map $\mathbb{E}_\infty(\alpha) : L^+_\infty(X,p) \longrightarrow L^+_\infty(Y,q)$ obtained by composing $(P_1(\alpha))^*$ with the isomorphisms between $L_1^{+,*}$ and L^+_∞

$$\begin{tikzcd}
L_1^{+,*}(X,p) \arrow[d, "(P_1(\alpha))^*"'] & L^+_\infty(X,p) \arrow[l, dotted] \arrow[d, "\mathbb{E}_\infty(\alpha)"] \\
L_1^{+,*}(Y,q) \arrow[r, dotted] & L^+_\infty(Y,q)
\end{tikzcd}$$

Consequences

1. It is an immediate consequence of the definitions that for any $f \in L_\infty^+(X,p)$ and $g \in L_1(Y,q)$

$$\langle \mathbb{E}_\infty(\alpha)(f), g\rangle_Y = \langle f, P_1(\alpha)(g)\rangle_X.$$

Consequences

1. It is an immediate consequence of the definitions that for any $f \in L_\infty^+(X, p)$ and $g \in L_1(Y, q)$

$$\langle \mathbb{E}_\infty(\alpha)(f), g \rangle_Y = \langle f, P_1(\alpha)(g) \rangle_X.$$

$$\begin{tikzcd}
\lambda h : L_1^+(X,p).\langle f,\ h \rangle \arrow[d, mapsto] & f \arrow[l, mapsto] \arrow[d, mapsto, dotted] \\
\lambda g : L_1^+(Y,q).\langle f,\ g \circ \alpha \rangle_X \arrow[r, mapsto] & \langle \mathbb{E}_\infty(\alpha)(f),\ g \rangle_Y
\end{tikzcd}$$

Consequences

1. It is an immediate consequence of the definitions that for any $f \in L_\infty^+(X, p)$ and $g \in L_1(Y, q)$

$$\langle \mathbb{E}_\infty(\alpha)(f), g \rangle_Y = \langle f, P_1(\alpha)(g) \rangle_X.$$

$$\begin{tikzcd}
\lambda h : L_1^+(X,p).\langle f,\ h \rangle \arrow[d, mapsto] & & f \arrow[ll, mapsto'] \arrow[d, mapsto, dotted] \\
\lambda g : L_1^+(Y,q).\langle f,\ g \circ \alpha \rangle_X \arrow[rr, mapsto] & & \langle \mathbb{E}_\infty(\alpha)(f),\ g \rangle_Y
\end{tikzcd}$$

2. One can informally view this functor as a “left adjoint” in view of this proposition.

Consequences

1. It is an immediate consequence of the definitions that for any $f \in L_\infty^+(X,p)$ and $g \in L_1(Y,q)$
$$\langle \mathbb{E}_\infty(\alpha)(f), g\rangle_Y = \langle f, P_1(\alpha)(g)\rangle_X.$$

$$\begin{tikzcd}
\lambda h : L_1^+(X,p).\langle f,\ h\rangle \arrow[d, mapsto] & f \arrow[l, mapsto] \arrow[d, mapsto, dotted] \\
\lambda g : L_1^+(Y,q).\langle f,\ g \circ \alpha\rangle_X \arrow[r, mapsto] & \langle \mathbb{E}_\infty(\alpha)(f),\ g\rangle_Y
\end{tikzcd}$$

2. One can informally view this functor as a “left adjoint” in view of this proposition.
3. Note that since we started with α in $\mathbf{Rad}_\infty$ we get the expectation value as a map between the L_∞^+ cones.

The other expectation value functor

The **functor** $\mathbb{E}_{\mathbf{1}}(\cdot)$ is a functor from $\mathbf{Rad}_1$ to ω**CC** which maps the object (X,p) to $L_1^+(X,p)$ and on maps is given as follows: Given $\alpha : (X,p) \longrightarrow (Y,q)$ in $\mathbf{Rad}_1$ the action of the functor is to produce the map $\mathbb{E}_1(\alpha) : L_1^+(X,p) \longrightarrow L_1^+(Y,q)$ obtained by composing $(P_\infty(\alpha))^*$ with the isomorphisms between $L_\infty^{+,*}$ and L_1^+ as shown in the diagram below

$$\begin{tikzcd}
L_\infty^{+,*}(X,p) \arrow[d, "(P_\infty(\alpha))^*"'] & L_1^+(X,p) \arrow[l, dotted] \arrow[d, "\mathbb{E}_1(\alpha)"] \\
L_\infty^{+,*}(Y,q) \arrow[r, dotted] & L_1^+(Y,q)
\end{tikzcd}$$

Another “adjoint”

Once again we have an “adjointness” statement; this time it is a right adjoint.

Right adjoint

Given $f \in L_\infty^+(Y, q)$ and $g \in L_1^+(X, p)$ we have

$$\langle f, \mathbb{E}_1(\alpha)(g)\rangle_Y = \langle P_\infty(\alpha)(f), g\rangle_X.$$

Relating the two expectation value functors

Given $\alpha \in \mathbf{Rad}_\infty[(X,p),(Y,q)]$ we have

$$
\begin{aligned}
&(a)\ \ \mathbb{E}_1(\alpha)(f\circ\alpha) = \mathbb{E}_\infty(\alpha)(\mathbf{1}_X)f, && \text{for } f \in L_1^+(Y,q) \text{ and}\\
&(b)\ \mathbb{E}_\infty(\alpha)(f\circ\alpha) = \mathbb{E}_1(\alpha)(\mathbf{1}_X)f, && \text{for } f \in L_\infty^+(Y,q).
\end{aligned}
$$

Why?

- This is a piece pulled out of a larger work on approximating Markov processes.

Why?

- This is a piece pulled out of a larger work on approximating Markov processes.
- Instead of compressing the state space we compressed the σ-algebra and used the conditional expectation to define approximate transition kernels.

Why?

- This is a piece pulled out of a larger work on approximating Markov processes.
- Instead of compressing the state space we compressed the σ-algebra and used the conditional expectation to define approximate transition kernels.
- But that is the subject of a different talk.