From Moore-Penrose to Markov via Gauss




JS PL (he/him), joint work with Cole Comfort

Email: `js.lemay@mq.edu.au`

Website: `https://sites.google.com/view/jspl-personal-webpage`

Today's Story

- **Categorical probability theory** (CPT) is a relatively "new[1]" active research program (2020?) that's been making lots of exciting buzz thanks to the wonderful work of T. Fritz, P. Perrone, E. Di Lavore, M. Román, and many others.
- The main framework for CPT are **Markov categories** and one of the fundamental concepts is the notion of **conditionals**.
- A key example of a Markov category captures **Gaussian probability theory**.
- Recently myself and others have been interested in **generalized inverses** including the Moore-Penrose inverse. In particular we've been studying **Moore-Penrose dagger categories** (QPL2023).

TODAY'S STORY: From a Moore-Penrose dagger additive category, we construct a Markov category (with conditionals) by generalizing the Gaussian probability theory example.

[1] Synthetic/categorical approaches to probability theory seem to have been independently rediscovered multiple times over the years, such as by Golubtsov and Gadducci.

References for today

- For CPT and Markov Categories:

A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. T. Fritz

Markov Categories and Entropy. P. Perrone

Partial Markov Categories. E. Di Lavore, M. Román, and P. Sobociński

Lots more reference here (including Youtube video links to tutorials):

`https://ncatlab.org/nlab/show/Markov+category`

- For Moore-Penrose dagger categories:

Moore-Penrose Dagger Categories. R. Cockett and J.-S. P. Lemay

On Traces in Categories of Contractions. A. D. Fairbanks and P. Selinger

Disclaimer

- I am still learning about CPT and Markov categories! My knowledge of probability theory is very basic and my intuition is terrible[2]...
- Luckily Markov categories are very easy to understand categorically!

[2] See my struggles with probability theory when I was in undergrad at the University of Ottawa

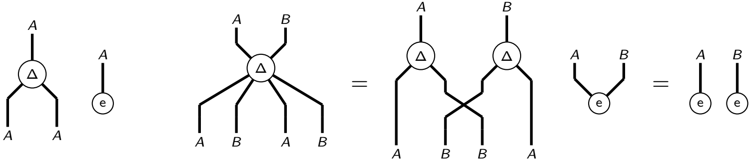
Markov Categories slogan

A **Markov category** is a symmetric monoidal category where every object is a cocommutative comonoid and every morphism preserves the counit.

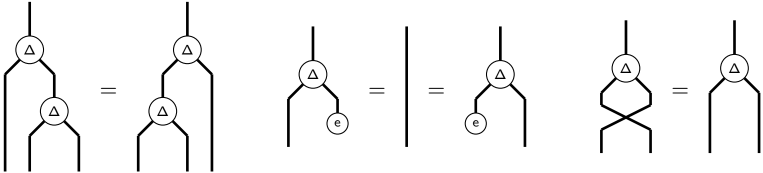
Markov Categories

A **Markov category** is a symmetric monoidal category $\mathbb{X}$, with tensor $\otimes$ and unit I, such that:

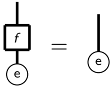
- Equipped with families of maps $\Delta_A : A \to A \otimes A$ and $e_A : A \to I$, which are coherent in the sense that $\Delta_{A \otimes B}$ and $e_{A \otimes B}$ satisfy:



- which make A into a cocommutative comonoid:

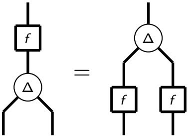


- such that e is a natural transformation (note that this makes I into a terminal object):



Markov Categories: Determinism

In a Markov category, a map $f : A \to B$ is **deterministic** if it preserves the comultiplication:



This means that a deterministic morphism is a comonoid morphism.

Let $\mathbb{X}_{det}$ be the subcategory of deterministic morphisms. Then $\mathbb{X}_{det}$ has finite products.

Markov Categories: Idea

- We usually interpret morphisms of a Markov category as **Markov kernels** (i.e "a function with random outcomes"): basically that it assigns to every input value a probability distribution over output values.
- The comultiplication (or copy) Δ takes an input value and outputs two copies of the input, without introducing any randomness.
- The counit (or discard) e simply discards the input and produces no output.
- Deterministic morphisms are those whose random outputs are independent of themselves, so we can copy them without any issue.

Markov Categories: Examples

Many examples of Markov categories which capture various models of probability theory.

- The category of finite sets and stochastic matrices;
- The category of measurable spaces and Markov kernels;
- Kleisli categories of "probability monads" (symmetric comonoidal monads which preserves the unit), this includes the Giry monad;
- Our favourite example for today is **Gaussian probability theory**, which we will review in a few slides.

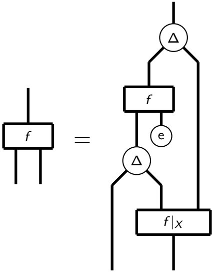
Markov Categories: Things we can do with them

We can formalize many important concepts in probability theory in a Markov category, such as:

- Probability measures;
- Joint states and marginals;
- Almost sure equality;
- Independence;
- Positivity;
- Causality;
- Kolmogorov products and Kolmogorov extension theorem;
- Basu's Theorem;
- **CONDITIONALS**

Markov Categories: Conditionals

A Markov category is said to have **conditionals**[3] if for every morphism of type $f : A \to X \otimes Y$ there exists a morphism of type $f|_X : X \otimes A \to Y$ such that:



A morphism of type $X \otimes A \to Y$ is called **a conditional** of f.

WARNING: Conditionals are **not** usually unique! In fact, conditionals are unique if and only if the Markov category is a poset (i.e. parallel maps are equal).

Instead, conditionals are unique up to **almost sure equality** (which we won't talk about today), which is usually a more suitable notion of equality/equivalence in Markov categories (in the sense that many notions in a Markov category are better behaved with a.s.e).

[3] Some authors take having conditionals as part of the definition of a Markov category.

Conditionals: Idea

Conditionals capture the notion of conditional probability distributions.

- $f : A \to X \otimes Y$ says that A is a known outcome, while X and Y are joint random variables/outcomes. We write:

  $$f(x, y|a)$$

  To say the probability of outcome (x, y) if a is true.

- The conditional $f|_X : X \otimes A \to Y$ is the probability distribution of Y when the outcome of X is known, hence why it's in the domain.

  $$f|_X(y|x, a)$$

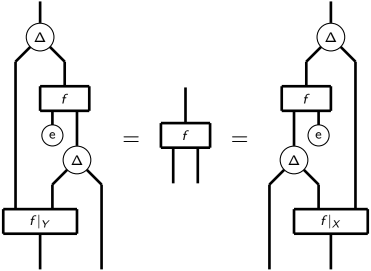
- Then the idea of the necessary equality says that:

  $$f(x, y|a) = f(x|a)f|_X(y|x, a)$$

  so the probability of outcome (x, y) if a is equal to the probability of outcome x if a (forgetting about y) multiplied by the probability outcome of y if a knowing x.

Bayes' Theorem

If $f : A \to X \otimes Y$, then we get its conditional $f|_X : X \otimes A \to Y$, but symmetrically we can also get its conditional $f|_Y : A \otimes Y \to X$, and so we have the equality:



which tells us that:

$$f|_Y(x|y,a)f(y|a) = f(x|a)f|_X(y|x,a)$$

Setting $I = A$ recaptures Bayes' theorem:

$$f(x|y) = \frac{f(y|x)f(x)}{f(y)}$$

the probability outcome of y if x is true is equal to the probability outcome of y if x is true, times the probability outcome of x, divided by the probability outcome of y.

The Gauss Example

Let's go over over the Markov category with conditionals we are most interested in: the example which captures Gaussian probability theory.

The Gauss Example

Let Gauss be the Markov category where:

- The objects are the natural numbers $n \in \mathbb{N}$;
- A morphism is a triple $(M, C, s) : n \to m$ where:
  - M is an $m \times n$ $\mathbb{R}$-matrix,
  - C is a positive semidefinite square $m \times m$ $\mathbb{R}$-matrix,
  - $s \in \mathbb{R}^m$, which we view as a column vector (so an $m \times 1$ $\mathbb{R}$-matrix);

We think of this data as defining the conditional distribution of a random variable $Y \in \mathbb{R}^m$ in terms of $X \in \mathbb{R}^n$ as:

$$Y = MX + \xi$$

where $\xi \in \mathbb{R}^m$ is Gaussian noise independent from X, with expected value/mean:

$$\mathrm{E}[\xi] = s$$

and (co)variance matrix:

$$\mathrm{VAR}[\xi] = \mathrm{E}\left[(\xi - \mathrm{E}[\xi])(\xi - \mathrm{E}[\xi])^{\mathsf{T}}\right] = C$$

The Gauss Example

Let Gauss be the Markov category where:

- The objects are the natural numbers $n \in \mathbb{N}$;
- A morphism is a triple $(M, C, s) : n \to m$ where:
  - M is an $m \times n$ $\mathbb{R}$-matrix,
  - C is a positive semidefinite square $m \times m$ $\mathbb{R}$-matrix,
  - $s \in \mathbb{R}^m$, which we view as a column vector (so an $m \times 1$ $\mathbb{R}$-matrix);
- The identity is the triple $(I_n, 0, 0) : n \to n$ where I_n is the $n \times n$ identity matrix;
- Composition is defined as follows:

$$(N, D, t) \circ (M, C, s) = (NM, NCN^{\mathsf{T}} + D, Ns + t)$$

If $Y = MX + \xi$ with $\mathrm{E}[\xi] = s$ and $\mathrm{VAR}[\xi] = C$, and $Z = NY + \eta$ with $\mathrm{E}[\eta] = t$ and $\mathrm{VAR}[\eta] = D$, then:

$$Z = NMX + (N\xi + \eta)$$

and the noises are independent of each other:

$$\mathrm{E}[N\xi + \eta] = N\mathrm{E}[\xi] + \mathrm{E}[\eta] = Ns + t$$

$$\mathrm{VAR}[N\xi + \eta] = N\mathrm{VAR}[\xi]N^{\mathsf{T}} + \mathrm{VAR}[\eta] = NCN^{\mathsf{T}} + D$$

The Gauss Example

Let Gauss be the Markov category where:

- The objects are the natural numbers $n \in \mathbb{N}$;
- A morphism is a triple $(M, C, s) : n \to m$ where:
  - M is an $m \times n$ $\mathbb{R}$-matrix,
  - C is a positive semidefinite square $m \times m$ $\mathbb{R}$-matrix,
  - $s \in \mathbb{R}^m$, which we view as a column vector (so an $m \times 1$ $\mathbb{R}$-matrix);
- The identity is the triple $(I_n, 0, 0) : n \to n$ where I_n is the $n \times n$ identity matrix;
- Composition is defined as follows: $(N, D, t) \circ (M, C, s) = (NM, NCN^{\mathsf{T}} + D, Ns + t)$
- The monoidal product is defined on objects as $n \otimes m = n + m$ and on maps as:

$$(M, C, s) \otimes (N, D, t) = \left(\begin{bmatrix} N & 0 \\ 0 & M \end{bmatrix}, \begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix}, \begin{bmatrix} s \\ t \end{bmatrix} \right)$$

- The monoidal unit is 0
- The counit is the triple $e_n := (0, 0, 0) : n \to 0$;
- The comultiplication is the triple $\Delta_n := \left(\begin{bmatrix} I_n \\ I_n \end{bmatrix}, 0, 0 \right) : n \to n + n$
- The deterministic maps are those of the form $(M, 0, s)$ (so no randomness involved)

Conditionals for Gauss

Gauss also has conditionals

Consider a map of type $(M, C, s) : n \to m + k$, so a triple, where recall:

- M is an $m + k \times n$ $\mathbb{R}$-matrix,
- C is a positive semidefinite square $m + k \times m + k$ $\mathbb{R}$-matrix,
- $s \in \mathbb{R}^{m+k}$.

which for our purposes we write in block notation:

$$\left(\begin{bmatrix} M_1 \\ M_2 \end{bmatrix}, \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix}, \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \right)$$

Then a conditional for it is map $m + n \to k$ given by the triple:

$$\left(\begin{bmatrix} C_3 C_1^\circ & M_2 - C_3 C_1^\circ M_1 \end{bmatrix}, C_4 - C_3 C_1^\circ C_2, s_2 - C_3 C_1^\circ s_1 \right)$$

where C_1° is the **Moore-Penrose inverse** of C_1.

Objective

The objective is to generalize the Gauss example.

To do so, we need the following ingredients:

- For objects, we need objects.
- For maps, we will need triples of maps:
  - A map
  - Some notion of positive semidefinite maps
  - Some notion of point.
- For identities, we need:
  - Identity maps
  - Zeroes
- For composition, we need:
  - Composition
  - Sums
  - Transpose
- For the monoidal structure, we need biproducts.
- For conditionals, we need:
  - Negatives
  - Moore-Penrose inverses

We have all these ingredients in a **Moore-Penrose dagger additive category**!

Moore-Penrose Dagger Additive Categories

A Moore-Penrose dagger additive category is:

- A dagger category;
- which is dagger-enriched over Abelian groups
- with dagger biproducts
- such that every map has a Moore-Penrose inverse.

Moore-Penrose Dagger Additive Categories

A Moore-Penrose dagger additive category is:

- A **dagger category**;
- which is dagger-enriched over Abelian groups
- with dagger biproducts
- such that every map has a Moore-Penrose inverse.

Dagger Categories

A **dagger** on a category $\mathbb{X}$ is a contravariant functor $(\_)^\dagger : \mathbb{X} \to \mathbb{X}$ which is the identity on objects and involutive. We refer to $f^\dagger$ as the **adjoint** of f.

A †-**category** is a category $\mathbb{X}$ equipped with a chosen dagger †.

Concretely, a †-category can be described as a category $\mathbb{X}$ where for each map $f : A \to B$, there is a chosen map of dual type $f^\dagger : B \to A$ such that:

$$1_A^\dagger = 1_A \qquad (g \circ f)^\dagger = f^\dagger \circ g^\dagger \qquad (f^\dagger)^\dagger = f$$

Lots of examples of dagger categories including (finite-dimensional) Hilbert spaces, sets and relations, set and partial injections, and **categories of matrices**.

Dagger Categories - Matrices Examples

Example

For any commutative ring R, let be MAT(R) whose objects are natural numbers $n \in \mathbb{N}$, and where a map $A : n \to m$ is an $n \times m$ R -matrix A. Then MAT(R) is a dagger category whose dagger is the transpose operator T, $A^{\mathsf{T}}(i,j) = A(j,i)$.

We are particular interested in $R = \mathbb{R}$ and MAT($\mathbb{R}$).

Moore-Penrose Dagger Additive Categories

A Moore-Penrose dagger additive category is:

- A dagger category;
- which is **dagger-enriched over Abelian groups**
- with dagger biproducts
- such that every map has a Moore-Penrose inverse.

Dagger Enriched over Abelian Groups

A **dagger pre-additive category** is a dagger category that is **dagger enriched over commutative monoids**, that is, if each homset is an abelian group, so we can:

- We can sum parallel maps together $f + g : A \to B$
- We have zero maps $0 : A \to B$
- We can have negatives $-f : A \to B$, and so can subtract parallel maps $f - g : A \to B$

which is preserved by composition:

$$f \circ (g + h) \circ k = (f \circ g \circ k) + (f \circ h \circ k) \quad f \circ 0 \circ k = 0 \quad f \circ (-g) = -(f \circ g) = (-f) \circ g$$

but also by the dagger:

$$(f + g)^\dagger = f^\dagger + g^\dagger \qquad 0^\dagger = 0 \qquad (-f)^\dagger = -f^\dagger$$

Moore-Penrose Dagger Additive Categories

A Moore-Penrose dagger additive category is:

- A dagger category;
- which is dagger-enriched over Abelian groups
- with **dagger biproducts**
- such that every map has a Moore-Penrose inverse.

Dagger Biproducts

A **dagger additive category** is a dagger pre-additive category with:

- A terminal/intial/zero object 0
- Dagger biproducts $A_1 \oplus \ldots \oplus A_n$ with projections $\pi_j : A_1 \oplus \ldots \oplus A_n \to A_j$ and injections $\iota_j : A_j \to A_1 \oplus \ldots \oplus A_n$ such that:

$$\pi_j \circ \iota_i = \begin{cases} 1_{A_i} & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \qquad \sum_{j=1}^{n} \iota_j \circ \pi_j = 1_{A_1 \oplus \ldots \oplus A_n} \qquad \pi_j^\dagger = \iota_j$$

Matrix Representation

In a dagger additive category, maps can be represented by matrices!

A map $f : A_1 \oplus \ldots \oplus A_n \to B_1 \oplus \ldots \oplus B_m$ is uniquely determined by a family of maps $f(i,j) : A_j \to B_i$. Therefore f can be represented as a $m \times n$ matrix where the term in the i-th row and j-th column is $f(i,j)$, so:

$$f = \left[f(i,j)\right] = \begin{bmatrix} f(1,1) & f(1,2) & \ldots & f(1,n) \\ f(2,1) & f(2,2) & \ldots & f(2,n) \\ \vdots & \vdots & \ddots & \vdots \\ f(m,1) & f(m,2) & \ldots & f(m,n) \end{bmatrix}$$

The dagger is given by the adjoint transpose of the matrix, so:

$$f^\dagger = \begin{bmatrix} f(1,1)^\dagger & f(2,1)^\dagger & \ldots & f(n,1)^\dagger \\ f(1,2)^\dagger & f(2,2)^\dagger & \ldots & f(n,2)^\dagger \\ \vdots & \vdots & \ddots & \vdots \\ f(1,m)^\dagger & f(2,m)^\dagger & \ldots & f(n,m)^\dagger \end{bmatrix}$$

while composition is given by matrix multiplication:

$$g \circ f = \left[g(k,l)\right] \circ \left[f(i,j)\right] = \left[\sum_{j=1}^{n} g(j,k) \circ f(i,j)\right]$$

Dagger Additive Categories - Matrices Examples

Example

For any commutative ring R, $\mathrm{MAT}(R)$ is a dagger additive category (with dagger given by transpose) where $n \oplus m = n + m$ and the abelain group enrichment is given by the usual sum of matrices $(A + B)(i, j) = A(i, j) + B(i, j)$.

From Dagger Additive Categories to Markov Categories

We can already build a Markov category from a dagger additive category!

To do so, we need the following ingredients:

- For objects, we need objects.
- For maps, we will need triples of maps:
  - A map
  - Some notion of positive semidefinite maps
  - Some notion of point.
- For identities, we need:
  - Identity maps
  - Zeroes
- For composition, we need:
  - Composition
  - Sums
  - Transpose
- For the monoidal structure, we need biproducts.

From Dagger Additive Categories to Markov Categories

We can already build a Markov category from a dagger additive category!

To do so, we need the following ingredients:

- For objects, we need objects ✓.
- For maps, we will need triples of maps:
  - A map ✓
  - Some notion of positive semidefinite maps
  - Some notion of point.
- For identities, we need:
  - Identity maps ✓
  - Zeroes ✓
- For composition, we need:
  - Composition ✓
  - Sums ✓
  - Transpose ✓
- For the monoidal structure, we need biproducts ✓.

Positive Maps

In a dagger category a **positive map** is an endomorphism $p : A \to A$ such that there exists a map $f : A \to B$ such that:

$$p = f^\dagger \circ f$$

Note that this f need not be unique!

Example

In MAT(R), T-positive maps are precisely the positive semidefinite matrices.

From Dagger Additive Categories to Markov Categories

We can already build a Markov category from a dagger additive category!

To do so, we need the following ingredients:

- For objects, we need objects ✓.
- For maps, we will need triples of maps:
  - A map ✓
  - Some notion of positive semidefinite maps ✓
  - Some notion of point.
- For identities, we need:
  - Identity maps ✓
  - Zeroes ✓
- For composition, we need:
  - Composition ✓
  - Sums ✓
  - Transpose ✓
- For the monoidal structure, we need biproducts ✓.

Points

Recall that in Gauss, part of our morphisms where vectors $s \in \mathbb{R}^m$, or rather $m \times 1$ matrices.

In $\mathrm{MAT}(\mathbb{R})$ these correspond to maps of type $1 \to m$.

So every map in Gauss has a map of type $1 \to$ codomain from $\mathrm{MAT}(\mathbb{R})$.

In our construction we will do the same: we will fix an object X and take maps $s : X \to A$, which some call a generalized point.

From Dagger Additive Categories to Markov Categories

We can already build a Markov category from a dagger additive category!

To do so, we need the following ingredients:

- For objects, we need objects ✓.
- For maps, we will need triples of maps:
  - A map ✓
  - Some notion of positive semidefinite maps ✓
  - Some notion of point. ✓
- For identities, we need:
  - Identity maps ✓
  - Zeroes ✓
- For composition, we need:
  - Composition ✓
  - Sums ✓
  - Transpose ✓
- For the monoidal structure, we need biproducts ✓.

Gauss Construction

For an dagger category $(\mathbb{X}, \dagger)$, for each object $X \in \mathbb{X}$, define the category $\mathfrak{G}[(\mathbb{X}, \dagger)]_X$ as follows:

- The objects of $\mathfrak{G}[(\mathbb{X}, \dagger)]_X$ are the same as the object of $\mathbb{X}$;
- A map $F : A \to B$ in $\mathfrak{G}[(\mathbb{X}, \dagger)]_X$ is a triple $F = (f, p, x) : A \to B$ consisting of:
  - A map $f : A \to B$
  - A $\dagger$-positive map $p : B \to B$
  - A generalized point $x : X \to B$.
- The identity of A in $\mathfrak{G}[(\mathbb{X}, \dagger)]_X$ is the triple $I_A = (1_A, 0, 0) : A \to A$;
- Composition in $\mathfrak{G}[(\mathbb{X}, \dagger)]_X$ of $F = (f, p, x) : A \to B$ and $G = (g, q, y) : B \to C$ is defined as

$$G \circ F = (g, q, y) \circ (f, p, x) = (g \circ f, q + g \circ p \circ g^\dagger, y + g \circ x) : A \to C$$

Hold on... why did we need biproducts to construct this category? Yes we will need them for the monoidal structure, but why do I need them here?

Composition might not be well-defined! There's a problem with the positive map component...

Problem with Positive Maps

In a dagger pre-additive category:

- Zero $0 : A \to A$ is always positive, $0 = 0^\dagger \circ 0$, so identities in $\mathfrak{G}\left[(\mathbb{X}, \dagger)\right]_X$ are well-defined.

- However, the sum of positive maps is not necessarily positive!

However in a **dagger additive category** we can use the dagger biproducts to solve the issue:

- If $\phi : A \to A$ and $\psi : A \to A$ are positive, where $\phi = \alpha^\dagger \circ \alpha$ for some $\alpha : A \to B$ and $\psi = \beta^\dagger \circ \beta$ for some $\beta : A \to C$. Consider the map $\begin{bmatrix} \alpha \\ \beta \end{bmatrix} : A \to B \oplus C$, and we compute that:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}^\dagger \circ \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \alpha^\dagger & \beta^\dagger \end{bmatrix} \circ \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \alpha^\dagger \circ \alpha + \beta^\dagger \circ \beta = \phi + \psi$$

  So $\phi + \psi$ is positive as well.

  So in the presence of dagger biproducts, the sum of dagger positive maps is positive. Thus composition in $\mathfrak{G}\left[(\mathbb{X}, \dagger)\right]_X$ is well-defined.

Gauss Construction - Monoidal Structure

$\mathfrak{G}\left[(\mathbb{X}, \dagger)\right]_X$ is a symmetric monoidal category as follows:

- On objects, the monoidal product is defined as the †-biproduct, $A \otimes B = A \oplus B$;
- On maps, the monoidal product of $F = (f, p, x) : A \to B$ and $G = (g, q, y) : C \to D$ is defined as follows:

$$F \otimes G = (f, p, x) \otimes (g, q, y) =$$
$$\left(f \oplus g, p \oplus q, x \oplus \begin{bmatrix} x \\ y \end{bmatrix}\right) = \left(\begin{bmatrix} f & 0 \\ 0 & g \end{bmatrix}, \begin{bmatrix} p & 0 \\ 0 & q \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix}\right) : A \oplus B \to C \oplus D$$

- The monoidal unit is the †-zero object 0

The monoidal product is well-defined since the biproduct of positive maps is again positive.

Gauss Construction - Markov

$\mathfrak{G}\left[(\mathbb{X}, \dagger)\right]_X$ is a Markov category where:

- The comultiplication $\Delta_A : A \to A \otimes A$ in $\mathfrak{G}\left[(\mathbb{X}, \dagger)\right]_X$ is the triple defined as follows:

$$\Delta_A = \left(\begin{bmatrix} 1_A \\ 1_A \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right)$$

- The monoidal unit 0 is a terminal object in $\mathfrak{G}\left[(\mathbb{X}, \dagger)\right]_X$, where the unique map $e_A : A \to 0$ is the triple defined as follows:

$$e_A = (0, 0, 0)$$

This comonoid structure also comes from the canonical comonoid structure given by biproducts of the base category.

Lemma

For every dagger additive category $(\mathbb{X}, \dagger)$ and every object $X \in \mathbb{X}$, $\mathfrak{G}\left[(\mathbb{X}, \dagger)\right]_X$ is a Markov category.

Gauss Construction - Deterministic

In $\mathfrak{G}\left[(\mathbb{X}, \dagger)\right]_X$ a map $F : A \to B$ is deterministic if and only if $F = (f, 0, s)$.

Applications of the Gauss Construction

Example

Consider $(\mathsf{MAT}(\mathbb{R}), \mathsf{T})$. Then $\mathfrak{G}\left[(\mathsf{MAT}(\mathbb{R}), \mathsf{T})\right]_1 = \mathsf{Gauss}$.

If we want to take $n \neq 1$, this results in the category Gauss_n whose objects are the natural numbers and where a map is an n-tuple $(M, C, s_1, \ldots, s_n)$ which we interpret as Gaussian distributions of random variables $Y_1 = MX_1 + \xi_1$, ..., $Y_n = MX_n + \xi_n$, where each $\xi_j \in \mathbb{R}^m$ is Gaussian noise independent from X_j with $\mathsf{E}[\xi_j] = s_j$ but which all have the same covariance matrix $\mathsf{VAR}[\xi_1] = \ldots = \mathsf{VAR}[\xi_k] = C$.

Example

We can instead consider the dagger category of complex matrices $\mathsf{MAT}(\mathbb{C})$ with dagger $\dagger$ given by conjugate transpose. What does $\mathfrak{G}\left[(\mathsf{MAT}(\mathbb{C}), \mathsf{T})\right]_1$ represent as a categorical framework for probability?

The answer seems to be this notion of *complex normal distributions* `https://en.wikipedia.org/wiki/Complex_normal_distribution`.

Thanks to Dario Stein for pointing this out to me.

Adding Conditionals

To add conditionals we need the following ingredients:

- Negatives
- Moore-Penrose inverses

Adding Conditionals

To add conditionals we need the following ingredients:

- Negatives ✓
- Moore-Penrose inverses

Moore-Penrose Dagger Additive Categories

A Moore-Penrose dagger additive category is:

- A dagger category;
- which is dagger-enriched over Abelian groups
- with dagger biproducts
- such that **every map has a Moore-Penrose inverse**.

Moore-Penrose Inverses for Dagger Categories

In a dagger category $(\mathbb{X}, \dagger)$, a **Moore-Penrose inverse** (M-P inverse) of a map $f : A \to B$ is a map $f^\circ : B \to A$ such that the following equalities hold:

- **[MP.1]** $f \circ f^\circ \circ f = f$
- **[MP.2]** $f^\circ \circ f \circ f^\circ = f^\circ$
- **[MP.3]** $(f \circ f^\circ)^\dagger = f \circ f^\circ$
- **[MP.4]** $(f^\circ \circ f)^\dagger = f^\circ \circ f$

While a Moore-Penrose inverse may not always exists, **if a Moore-Penrose inverse exists, then it is unique**.

A **Moore-Penrose dagger category** is a dagger category such that every map is M-P invertible.

Moore-Penrose Inverses Examples

Example

For a commutative ring R, for an $n \times m$ R-matrix A, its **Moore-Penrose inverse** (if it exists) is the unique $m \times n$ $\mathbb{R}$-matrix A° satisfying the following equalities:

$$AA^\circ A = A \qquad A^\circ AA^\circ = A^\circ \qquad (AA^\circ)^\mathsf{T} = AA^\circ \qquad (A^\circ A)^\mathsf{T} = A^\circ A$$

In general $(\mathsf{MAT}(R), \mathsf{T})$ is not a Moore-Penrose dagger category.

However, $(\mathsf{MAT}(\mathbb{R}), \mathsf{T})$ is a Moore-Penrose dagger additive category. In fact, for any subfield of $\mathbb{K} \subseteq \mathbb{R}$ (ex. $\mathbb{Q}$), $(\mathsf{MAT}(\mathbb{K}), \mathsf{T})$ is a Moore-Penrose dagger additive category. We don't need to know the construction of the Moore-Penrose, it is enough to know that it exists.

References for More Examples

The category of complex matrices or the category of quaternion matrices (where importantly the dagger is conjugate transpose) are Moore-Penrose dagger additive categories.

Lots more on Moore-Penrose dagger additive categories can be found in:

On Traces in Categories of Contractions. A. D. Fairbanks and P. Selinger

I recently also came across a new example here (new paper from this year!):

Pseudoinversos de morfismos entre variedades abelianas. R. Auffarth

Adding Conditionals

To add conditionals to our Gauss construction we need the following ingredients:

- Negatives ✓
- Moore-Penrose inverses ✓

Gauss Construction: Conditionals

A map of type $F : A \to Z \otimes Y$ in $\mathfrak{G}\left[(\mathbb{X}, \dagger)\right]_X$ is a triple:

$$F = \left(\begin{bmatrix} f \\ g \end{bmatrix}, \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}, \begin{bmatrix} s \\ t \end{bmatrix} \right)$$

where:

$$\begin{array}{cccc} & f : A \to Z & g : A \to Y & \\ \alpha : Z \to Z & \beta : Y \to Z & \gamma : Z \to Y & \delta : Y \to Y \\ & s : X \to Z & t : X \to Y & \end{array}$$

Moreover, the second component block matrix is dagger positive, so there is a map type $\begin{bmatrix} \phi & \psi \end{bmatrix} : Z \oplus Y \to W$, with $\phi : Z \to W$ and $\psi : Y \to W$, such that:

$$\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} = \begin{bmatrix} \phi & \psi \end{bmatrix}^\dagger \circ \begin{bmatrix} \phi & \psi \end{bmatrix} = \begin{bmatrix} \phi^\dagger \\ \psi^\dagger \end{bmatrix} \circ \begin{bmatrix} \phi & \psi \end{bmatrix} = \begin{bmatrix} \phi^\dagger \circ \phi & \phi^\dagger \circ \psi \\ \psi^\dagger \circ \phi & \psi^\dagger \circ \psi \end{bmatrix}$$

So in particular, α and δ are positive maps, and also $\beta = \gamma^\dagger$. So our matrix is of the form:

$$\begin{bmatrix} \alpha & \gamma^\dagger \\ \gamma & \delta \end{bmatrix} \qquad\qquad F = \left(\begin{bmatrix} f \\ g \end{bmatrix}, \begin{bmatrix} \alpha & \gamma^\dagger \\ \gamma & \delta \end{bmatrix}, \begin{bmatrix} s \\ t \end{bmatrix} \right)$$

Gauss Construction: Conditionals

Now suppose that $(\mathbb{X}, \dagger)$ is also a Moore-Penrose dagger additive category.

So for a map of type $F : A \to Z \otimes Y$, a possible conditional for it will be of type $F|_Z : Z \otimes A \to Y$ which is the triple:

$$F|_Z = \left(\begin{bmatrix} \gamma \circ \alpha^\circ & g - \gamma \circ \alpha^\circ \circ f \end{bmatrix}, \delta - \gamma \circ \alpha^\circ \circ \gamma^\dagger, t - \gamma \circ \alpha^\circ \circ s \right)$$

Showing that this is a conditional is a long calculation... but I did it!

Theorem

For every Moore-Penrose dagger additive category $(\mathbb{X}, \dagger)$ and every object $X \in \mathbb{X}$, $\mathfrak{G}\left[(\mathbb{X}, \dagger)\right]_X$ is a Markov category with conditionals.

Comments about conditionals in the Gauss construction

Recall that conditionals are not unique! The Moore-Penrose inverse is just one way of constructing a conditional. But there could be other possible conditionals.

However it turns out that in the Gauss construction, even if you don't have Moore-Penrose inverses, we can give a description of all possible conditionals.

Proposition

In $\mathfrak{G}\left[(\mathbb{X},\dagger)\right]_X$, $F = \left(\begin{bmatrix} f \\ g \end{bmatrix}, \begin{bmatrix} \alpha & \gamma^\dagger \\ \gamma & \delta \end{bmatrix}, \begin{bmatrix} s \\ t \end{bmatrix}\right) : A \to B \otimes C$ *has a conditional if and only if there is a map* $m : B \to C$ *such that:*

1. $m \circ \alpha = \gamma$
2. $\delta - m \circ \alpha \circ m^\dagger$ *is* $\dagger$*-positive;*

Moreover, every conditional of F *is of the form:*

$$G_m = \left(\begin{bmatrix} m & g - m \circ f \end{bmatrix}, \delta - m \circ \alpha \circ m^\dagger, t - m \circ s\right)$$

If we have Moore-Penrose inverses, we can take $m = \gamma \circ \alpha^\circ$ to get the conditionals from the previous slide.

Thanks for listenining! Merci!

Paper in progress.

I'm always excited about more applications and constructions involving Moore-Penrose dagger categories. So if you have ideas: let me know!