A Category for Unifying Gaussian Probability and Nondeterminism

Dario Stein[1], Richard Samuelson[2]

[1]Radboud University Nijmegen, [2]Humming Inc.

CALCO 2023
Bloomington, Indiana – 19 June 2023

Overview

Question

In which ways can we express *not knowing something*?

Overview

Question

In which ways can we express *not knowing something*?

Consider the outcome of a coin flip in $X = \{H, T\}$:

Overview

Question

In which ways can we express *not knowing something*?

Consider the outcome of a coin flip in $X = \{H, T\}$:

- probabilistic: $\frac{1}{2}|\,H\,\rangle + \frac{1}{2}|\,T\,\rangle \in \mathcal{D}(X)$

Overview

Question

In which ways can we express *not knowing something*?

Consider the outcome of a coin flip in $X = \{H, T\}$:

- probabilistic: $\frac{1}{2}|\,H\,\rangle + \frac{1}{2}|\,T\,\rangle \in \mathcal{D}(X)$
- nondeterministic: $\{H, T\} \in \mathcal{P}(X)$

Overview

Question

In which ways can we express *not knowing something*?

Consider the outcome of a coin flip in $X = \{H, T\}$:

- probabilistic: $\frac{1}{2}|H\rangle + \frac{1}{2}|T\rangle \in \mathcal{D}(X)$
- nondeterministic: $\{H, T\} \in \mathcal{P}(X)$
- quotient: $1 \to X/X$

Overview

Question

In which ways can we express *not knowing something*?

Consider the outcome of a coin flip in $X = \{H, T\}$:

- probabilistic: $\frac{1}{2}|H\rangle + \frac{1}{2}|T\rangle \in \mathcal{D}(X)$
- nondeterministic: $\{H, T\} \in \mathcal{P}(X)$
- quotient: $1 \to X/X$
- indiscrete: $1 \to \mathcal{G}(X, \mathcal{E}_X)$ where $\mathcal{E}_X = \{\emptyset, X\}$

Overview

Question

In which ways can we express *not knowing something*?

Consider the outcome of a coin flip in $X = \{H, T\}$:

- probabilistic: $\frac{1}{2}|\,H\,\rangle + \frac{1}{2}|\,T\,\rangle \in \mathcal{D}(X)$
- nondeterministic: $\{H, T\} \in \mathcal{P}(X)$
- quotient: $1 \to X/X$
- indiscrete: $1 \to \mathcal{G}(X, \mathcal{E}_X)$ where $\mathcal{E}_X = \{\emptyset, X\}$
- **How can we combine these notions?**

Open Stochastic Systems

Case study [Jan Willems, 2012–]

n-dimensional stochastic system: probability space $(\mathbb{R}^n, \mathcal{E}, P)$

- 'closed' if $\mathcal{E} = \mathcal{B}(\mathbb{R}^n)$. $\leftarrow$ fully resolved random vector
- 'open' if $\mathcal{E} \subset \mathcal{B}(\mathbb{R}^n)$. $\leftarrow$ 'openness' $=$ lack of information

Open Stochastic Systems

Example: Noisy Resistor

Ohm's law constrains pairs $(V, I) \in \mathbb{R}^2$ to lie in the subspace

$$D = \{(V, I) : V = RI\}$$

In a noisy resistor, $V = RI + \epsilon$ with $\epsilon \sim \mathcal{N}(0, \sigma^2)$.

According to Willems, we should model this as the open system

$$(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2/D), P_{VI})$$

V, I are not considered random variables in isolation.

From [Willems'12]:

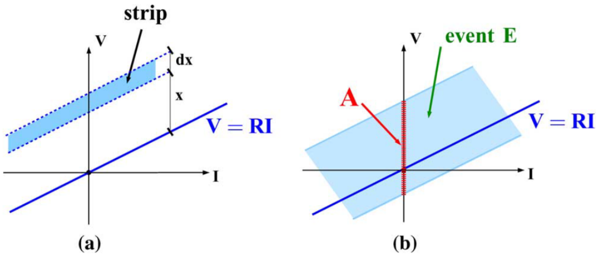


Fig. 2. Events for the noisy resistor.

Open Stochastic Systems

How to find a cleaner picture? Make σ-algebras part of the morphisms, not objects?

$$
\begin{tikzcd}
 & & \mathbb{R} \\
I \arrow[r, "P_{VI}"] \arrow[urr, dashed, bend left, "P_V"] \arrow[drr, dashed, bend right, "P_I"'] & \mathbb{R}^2 \arrow[ur, "\pi_1"] \arrow[dr, "\pi_2"] & \\
 & & \mathbb{R}
\end{tikzcd}
$$

Goal:

Define a Markov category

$$\mathsf{Gauss} \to \textbf{GaussEx} \leftarrow \mathsf{LinRel}^+$$

which unifies probabilistic and nondeterministic contributions. [Caveat: no-go theorems]

- corresponds to ‘linear open stochastic systems’ [Willems]
- ‘open‘ (morphism) $\neq$ ‘open‘ (lack of information)

Extended Gaussian distributions

Definition

An *extended Gaussian distribution* on $\mathbb{R}^n$ is pair (D, ψ) where

$$D \subseteq \mathbb{R}^n \text{ subspace}, \quad \psi \in \mathbf{Gauss}(\mathbb{R}^n/D)$$

Call D *nondeterministic fibre*, and write $\psi + D$ $\leftarrow$ just like coset notation $3 + 2\mathbb{Z}$

- $D = 0 \leftrightarrow$ purely probabilistic ('closed')
- $\psi = 0 \leftrightarrow$ purely nondeterministic
- $D = \mathbb{R}^n \leftrightarrow$ ideal uniform distribution

Example: One-dimensional distributions

On $\mathbb{R}$, either $D = 0$ and $\psi \in \mathbf{Gauss}(\mathbb{R})$, or $D = \mathbb{R}$, and $\mathbf{Gauss}(\mathbb{R}/\mathbb{R}) \cong \{0\}$, so

$$\mathbf{GaussEx}(\mathbb{R}) = \mathbf{Gauss}(\mathbb{R}) + \{\mathbb{R}\}$$

The uniform distribution $\mathbb{R}$ is translation invariant, e.g. $\mathcal{N}(0,1) + \mathbb{R} = \mathbb{R}$.

Noisy Resistor Revisited

The extended Gaussian model

$$\begin{aligned}\epsilon &\sim \mathcal{N}(0, \sigma^2)\\ I &\sim \mathbb{R}\\ V &= R \cdot I + \epsilon\end{aligned}$$

has joint distribution

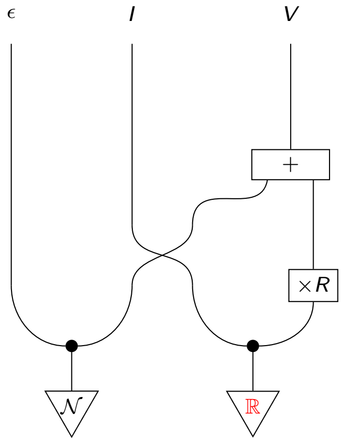
$$P_{\epsilon IV} = \mathcal{N}\left(\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}\sigma^2 & 0 & \sigma^2\\ 0 & 0 & 0\\ \sigma^2 & 0 & \sigma^2\end{pmatrix}\right) + \left\{\begin{pmatrix}0\\I\\V\end{pmatrix} : V = R \cdot I\right\}$$

with marginals [assuming $R \neq 0$]

$$P_{IV} = \mathcal{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & \sigma^2\end{pmatrix}\right) + \left\{\begin{pmatrix}I\\V\end{pmatrix} : V = R \cdot I\right\}, \quad P_I = P_V = \mathbb{R}$$

String Diagrams

How to compute these compositionally?



Categorical Structure

Extended Gaussian distributions can be added, tensored, pushed forward etc. just like ordinary Gaussians. E.g. for $A \in \mathbb{R}^{m \times n}$,

$$A_*(\mathcal{N}(\mu, \Sigma) + D) = A_*\mathcal{N}(\mu, \Sigma) + A[D] = \mathcal{N}(A\mu, A\Sigma A^T) + A[D]$$

Gaussian maps [Fritz'20] $\to$ Extended Gaussian maps

$\mathbf{Gauss}(\mathbb{R}^m, \mathbb{R}^n)$ = linear maps + Gaussian noise

$$x \mapsto f(x) + \psi \text{ where } \psi \in \mathbf{Gauss}(\mathbb{R}^n)$$

$\mathbf{GaussEx}(\mathbb{R}^m, \mathbb{R}^n)$ = linear maps + extended Gaussian noise

$$x \mapsto f(x) + \psi + D \text{ where } \psi \in \mathbf{Gauss}(\mathbb{R}^n), D \subseteq \mathbb{R}^n$$

but: Representatives are not unique; we really want $f : \mathbb{R}^m \to \mathbb{R}^n/D$, $\psi \in \mathbf{Gauss}(\mathbb{R}^n/D)$. How to compose these?

Decorated Cospans

Recall that a *partial function* $X \rightharpoonup Y$ is a span[1] $X \xleftarrow{m} A \xrightarrow{f} Y$

Definition: a *copartial function* $X \rightharpoondown Y$ is a cospan[1] $X \xrightarrow{f} P \xtwoheadleftarrow{p} Y$

Copartial functions compose via pushout

$$
\begin{tikzcd}
 & & W \arrow[dd, phantom, "\vee", very near start] & & \\
 & P \arrow[ur] & & Q \arrow[ul, two heads] & \\
X \arrow[ur, "f"] & & Y \arrow[ul, two heads, "p"'] \arrow[ur, "g"] & & Z \arrow[ul, two heads, "q"']
\end{tikzcd}
$$

Linear Relations and cospans

Theorem [Fong]: To give a copartial function $X \rightharpoondown Y$ in Vec is to give a total linear relation $R \subseteq X \times Y$. That is every such relation can be written as

$$R(x) = f(x) + D \text{ for a unique } f : X \to Y/D$$

[1] up to equivalence

Decorated Cospans

Given a monoidal functor $F : \mathbb{C} \to \mathrm{Set}$, a decorated cospan is a pair $(X \xrightarrow{f} P \xleftarrow{p} Y; \psi)$ with $\psi \in F(P)$ [Fong'15]. Decorations are composed by taking their coproduct and projecting them to the pushout.

Theorem/Definition

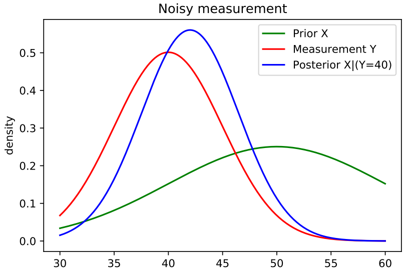
The extended Gaussian map
$$x \mapsto f(x) + \psi + D$$
formally corresponds to the following **Gauss**-decorated copartial function on Vec
$$(\mathbb{R}^n \xrightarrow{f} \mathbb{R}^m/D \xleftarrow{\pi} \mathbb{R}^m; \psi)$$
and composition is pushout.

Applications

Extended Gaussians have been considered at different times, for different purposes.

Applications I – Improper priors

Consider Bayesian inference with Gaussians (e.g. linear regression, Gaussian Processes, Kalman filters)



There is no uniform prior over $\mathbb{R}$. In which sense is

$$\lim_{\sigma^2 \to \infty} \mathcal{N}(\mu, \sigma^2) = \mathbb{R}?$$

Applications II – Duality

Covariance-Precision duality

For a Gaussian distribution $\mathcal{N}(\mu, \Sigma)$, the *precision matrix* is defined as

$$\Omega = \Sigma^{-1}$$

if Σ is regular.

Precision forms generalize logprobability densities:

- covariance is additive for convolution
- precision is additive for conditioning!

Fix the asymmetry [e.g. James'73]:

- that's the precision for singular Σ? $\leftarrow$ a partial quadratic form
- which distributions correspond to singular Ω? $\leftarrow$ extended Gaussians!

$$\ker(\Omega) = D = \mathrm{dom}(\Sigma)^{\perp}$$

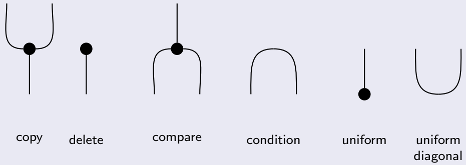
Applications III – Gaussian Relations

Definition

Define the category of *Gaussian relations* as

$$\mathbf{GaussRel}(X, Y) = \mathbf{GaussEx}(X \otimes Y) + \{\perp\}$$

This is a hypergraph category:



Alternative characterizations:

1. equivalence classes of decorated cospans $X \xrightarrow{f} P \xleftarrow{g} Y$.
2. convex functions (partial quadratic forms)

Take Home Message

Summary:

- **Extended Gaussian = probability + nondeterminism**
- Compositional manipulation using decorated cospans & categorical probability
- Categorical account of Willems' open systems [extending Zanasi & al]

Take Home Message

Summary:

- **Extended Gaussian = probability + nondeterminism**
- Compositional manipulation using decorated cospans & categorical probability
- Categorical account of Willems' open systems [extending Zanasi & al]

Outlook:

- presentations, signal flow diagrams
- convex analysis, nonlinear systems

Fig. 10. **Price/demand/supply event**

Thank you!