A Hypergraph Category for Exact Gaussian Inference

Dario Stein, Radboud University Nijmegen
Sam Staton, University of Oxford

Applied Category Theory, Strathclyde – July 22, 2022

Outline

Part I – What is Exact Inference? [LICS'21, Thesis§IV]

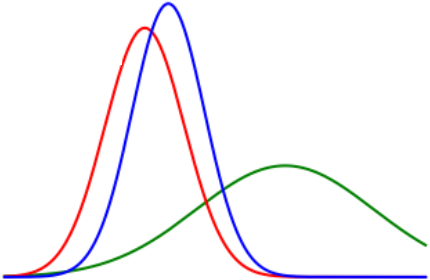
- Key Example: Gaussian probability, Markov categories
- Cond-construction for compositional inference

Part II – Duality: How to add *uninformative priors* to Gaussians? [arxiv]

- What is a uniform distribution over $\mathbb{R}$?
- Construct Extended Gaussians via decorated linear relations

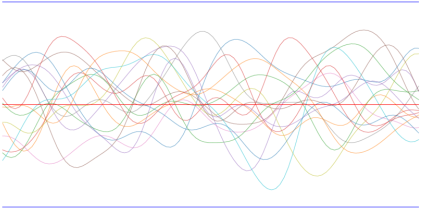
Example: Noisy measurement

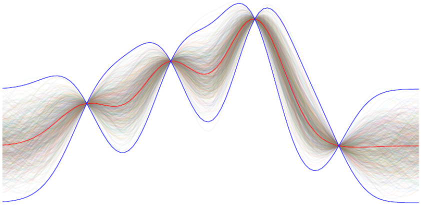
```
X = normal(μ=50, σ=10) # prior
Y = normal(μ=X, σ=5)   # noisy measurement
Y =:= 40               # exact observation (careful! prob = 0)

# X = normal(μ=42,σ=4.47)
# Y = 40
```



Example: Gaussian Process Regression

```
f = gaussian_process(kernel=...)
for (x, y) in obs:
  f(x) =:= y
```



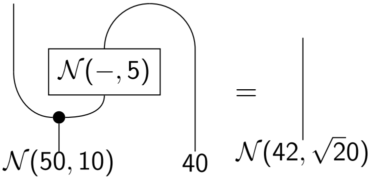


Example: Noisy measurement

GaussianInfer: Probabilistic language for Gaussian probability

- first-class exact inference operator (=:=)
- implementation in Python&F# (github/damast93/GaussianInfer)
- Kalman filters, ridge regression, Gaussian processes, ...

Categorical semantics in CD category Cond(Gauss):



Semantics of Exact Conditioning

Towards denotational semantics

- Define Markov category $\mathbb{C} = \text{Gauss}$ for purely probabilistic computation
- Understand conditioning in Markov categories
- Build a category $\text{Cond}(\mathbb{C})$ which internalizes (=:=)

Define $\mathbb{C}$ = Gauss

The Markov category Gauss is extremely simple:

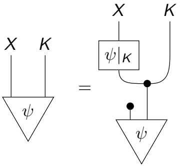
Gauss

- objects are vector spaces X, monoidal structure $(\oplus, 0)$
- morphisms are linear maps + noise: $f(x) = Ax + \mathcal{N}(\mu, \Sigma)$

Gauss is closed under conditionals, but conditioning is not a morphism in Gauss!

Inference problems

- A *conditional* for $\psi : I \to X \otimes K$ is



- An *inference problem* is a pair (ψ, k_0) of $\psi : I \to X \otimes K$, $k_0 : I \to K$ deterministic.
  - ψ is the model
  - k_0 is an observation
- A *solution of the inference problem* is
  - the composite $\psi|_K \circ k_0 : I \to X$ if $k_0 \in \mathrm{supp}(\psi_K)$,
  - otherwise $\perp$.

By definition of supports [Fritz], the solution is uniquely defined!

Cond-construction

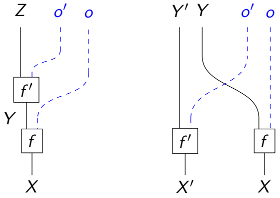
A morphism in $\mathrm{Cond}(\mathbb{C})$ is an open inference problem

$$\mathrm{Cond}(\mathbb{C})(X, Y) = \sum_K \mathbb{C}(X, Y \otimes K) \times \mathbb{C}_{\mathrm{det}}(I, K)/\sim$$

quotiented by contextual equivalence $\sim$

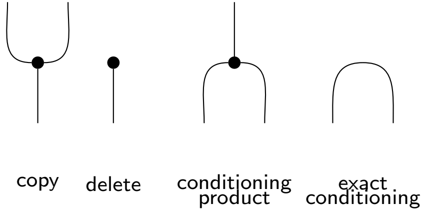
- roughly "same solutions in all contexts"

Theorem: This is a well-defined CD category.

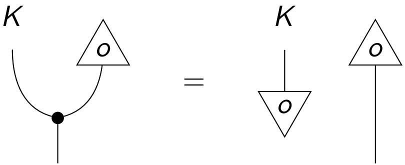


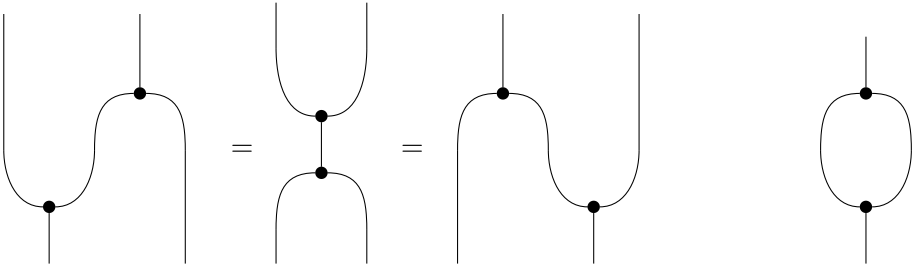
Cond-construction

Synthetic dictionary – Part I



Diagrammatic reasoning

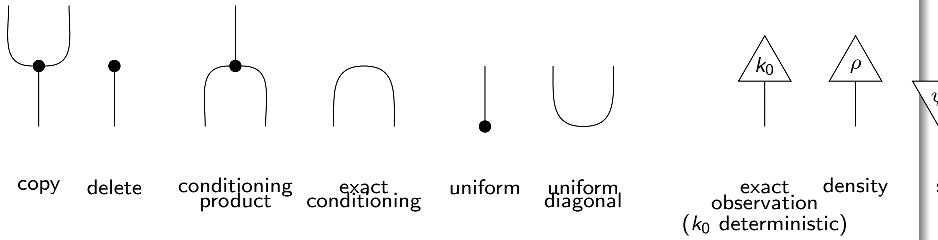




Cond-construction

- Conditioning product satisfies the Frobenius law!
- If it had a unit, we would obtain a hypergraph category: compact closed, self-dual

Synthetic dictionary – Part II



The quest for units

Unit for conditioning should be perfectly uniform on $\mathbb{R}$! Desired properties

- translation invariant
- invariant under rescaling

Problem: Such a distribution does not exist

- $\mathcal{N}(\mu, \sigma) \to 0$ for $\sigma \to \infty$
- the Lebesgue measure is not a probability measure

Extended Gaussians

Hope: Relations can do it

Consider $\mathbb{R} \subseteq \mathbb{R}$ as a linear relation

- Goal: somehow combine "GaussEx = Gauss + LinRel"
- Caution, there is no distributive law between powerset and Giry monad

Re-analying linear relations

A relation $R \subseteq X \times Y$ is

- left-total if $R(x) = \{y : (x, y) \in R\} \neq \emptyset$
- linear (affine) if R is a vector (affine) subspace of $X \times Y$

Lemma: Decomposing linear relations

To give a left-total linear relation $R \subseteq X \times Y$ is to give a 'subspace of nondeterminism' $D \subseteq Y$ and a linear map $f : X \to Y/D$

$$R(x) = f(x) + D$$

Defining GaussEx

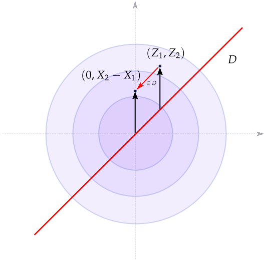
From Gaussians to Extended Gaussians

- Gaussian map $f(x) = Ax + \psi$ with ψ Gaussian
- Total linear relation $f(x) = Ax + D$ with $D \subseteq Y$ subspace
- Extended Gaussian map $f(x) = Ax + \psi + D$
  Note that $A : X \to Y/D$ and ψ is a Gaussian on Y/D.

Example: There is a unique Extended Gaussian on $\mathbb{R}$ with $D = \mathbb{R}$.

Two-variable example: $D = \{(x, y) : x = y\}$

$$\mathcal{N}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + D = \mathcal{N}\begin{pmatrix}0 & 0\\ 0 & 2\end{pmatrix} + D$$



A general construction

Decorated maps and relations

For *any* 'decoration' functor $S : \mathsf{Vec} \to \mathsf{CMon}$, define

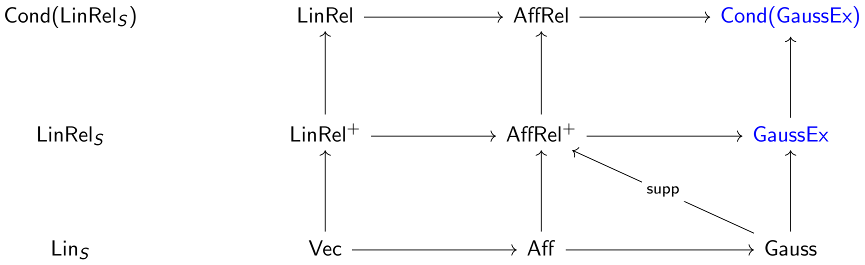
$$\mathrm{Lin}_S(X, Y) = \mathsf{Vec}(X, Y) \times S(Y)$$
$$\mathrm{LinRel}_S(X, Y) = \sum_{D \subseteq Y} \mathsf{Vec}(X, Y/D) \times S(Y/D)$$

Lin_S is precisely the op-Grothendieck construction for

$$S : \mathsf{Vec} \to \mathsf{CMon} \to \mathsf{Cat}$$

LinRel_S extends the construction of LinRel^+ with decoration S.

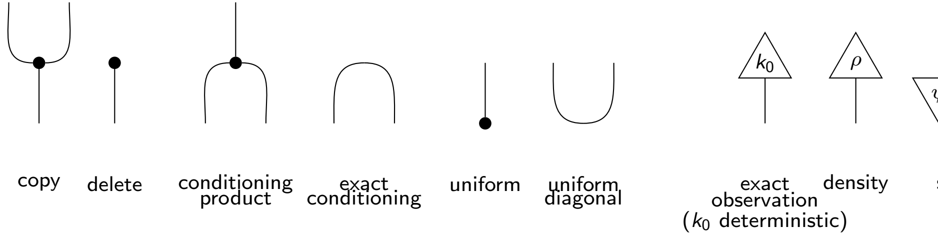
Functoriality



$S(X) = 0$ $\qquad$ $S(X) = X$ $\qquad$ $S(X) = \{(\mu, \Sigma)\}$

Summary

- GaussEx extends Gaussian distributions with idealized uniform distributions
- Uniform distributions act as units for exact conditioning
- Cond(GaussEx) is a hypergraph category (compact closed, self-dual), which is an ideal setting for inference [Morton'14]



Talking points

- Compare with [Coecke&Spekkens'12]
- Cond is optic-like, cf. CoPara($\mathbb{C}$), Bayesian lenses etc.
- PROPs, presentations & graphical calculi
  - graphical linear algebra, ZX-calculus, electrical circuits
- how to combine LinRel and Gauss more directly?