INTRODUCTION

Objectives:
• Design flexible univariate margins in characterizing the posterior skewness, high kurtosis and multi-modality
• Preserve interpretable posterior dependencies among hidden variables
Main idea: Decoupled proposal construction

Sklar’s theorem: Any multivariate distribution can be completely described by its copula and univariate margins

F(x_1, \ldots, x_p) = C(F_1(x_1), \ldots, F_p(x_p)).

VARIATIONAL COPULA INFERENCE

Copula-based variational proposal:
q_c(x) = \int F(x_1, \ldots, x_p) \prod_{j=1}^p f_j(x_j), \quad c \in C.

KL additive decomposition:
\text{KL}(q_c(x)||p(x|y)) = \text{KL}(F(x)||p(F(x)|y)) + \sum_j \text{KL}(f_j(x_j)||p_j(x_j)).

The Unified Variational Copula Inference Framework

FLEXIBLE MARGINAL PROPOSALS

Fixed-form margins:
Marginal densities ⇔ Monotone transformations
f_j(x_j) = q_j(h_j^{-1}(x_j); \mu_j, \sigma_j^2) \frac{d}{dx_j} h_j^{-1}(x_j) = q_j(h_j^{-1}(x_j); \mu_j, \sigma_j^2) \frac{1}{h_j'(h_j^{-1}(x_j))}.

Limitations:
(1) Difficult to specify the correct parametric form
(2) Tractable inverse CDF with optimizable parameters

Bernstein polynomials (BPs)
The BPs have a uniform convergence property for continuous functions on unit interval [0, 1].

\begin{align*}
B(u; k, \omega) &= \frac{k}{\omega^k} \sum \binom{k}{j} u^j \bigg(1 - u \bigg)^{k-j}, \\
&= \frac{k}{\omega^k} \sum \binom{k}{j} \left(\frac{u}{\omega}\right)^j \left(1 - \frac{u}{\omega}\right)^{k-j}, \\
&= \frac{k}{\omega^k} \sum \binom{k}{j} \left(\frac{u}{\omega}\right)^j.
\end{align*}

Desirable properties:
• Bijective
• Monotonic non-decreasing
• Having constrained range
• Differentiable w.r.t. both its argument and parameters
• Sufficiently flexible

STOCHASTIC OPTIMIZATION

Limitations of deterministic VGC:
(1) Highly model-dependent
(2) Cross-terms involved in the log densities
(3) Non-convex optimization

Stochastic Variational Gaussian Copula (VGC)

ELBO under Jacobian representation:
\mathcal{L} = \mathbb{E}[\ln p(x) | z] - \ln q_c(z),
\ell_i(z, h) = \ln p(y_i, h(z)) + z_i \ln q_i(z).

Gradients of the ELBO:
\nabla \mathcal{L} = \mathbb{E}[\nabla \ln p(y_i, h(z)) | z] - \mathbb{E}[\ln q_i(z)],
\nabla C_{\ell} = \mathbb{E}[\nabla \ell_i(z, h) | z] - \mathbb{E}[\ln q_i(z)]e^T,
\n\mathbb{E}[C_{\ell}] = \mathbb{E}[\nabla \ell_i(z, h, y)].

Stochastic gradient terms:
\ell_i(z, h) = \frac{\partial \ln p(y_i, h(z))}{\partial z_j} h_j(z_i) + \frac{\partial \ln p(y_i, h(z))}{\partial z_j} \ln h_j(z_i),
\mathbb{E}[\nabla \ell_i(z, h, y)] = \frac{\partial \ln p(y_i, h(z))}{\partial z_j} \ln h_j(z_i) + \frac{\partial \ln p(y_i, h(z))}{\partial z_j} \ln h_j(z_i).

The only two model-dependent terms are \ln p(y_i, h(z)) and \partial \ln p(y_i, h(z)) / \partial z_j.

ACKNOWLEDGEMENTS

The research was supported in part by grants from ARO, DARPA, DOE, NGA, ONR and NSF.

IV. Poisson log-linear regression:

Stochastic Variational Gaussian Copula (VGC)

ELBO under Jacobian representation:
\mathcal{L} = \mathbb{E}[\ln p(x) | z] - \ln q_c(z),
\ell_i(z, h) = \ln p(y_i, h(z)) + z_i \ln q_i(z).

Gradients of the ELBO:
\nabla \mathcal{L} = \mathbb{E}[\nabla \ln p(y_i, h(z)) | z] - \mathbb{E}[\ln q_i(z)],
\nabla C_{\ell} = \mathbb{E}[\nabla \ell_i(z, h) | z] - \mathbb{E}[\ln q_i(z)]e^T,
\n\mathbb{E}[C_{\ell}] = \mathbb{E}[\nabla \ell_i(z, h, y)].

Stochastic gradient terms:
\ell_i(z, h) = \frac{\partial \ln p(y_i, h(z))}{\partial z_j} h_j(z_i) + \frac{\partial \ln p(y_i, h(z))}{\partial z_j} \ln h_j(z_i),
\mathbb{E}[\nabla \ell_i(z, h, y)] = \frac{\partial \ln p(y_i, h(z))}{\partial z_j} \ln h_j(z_i) + \frac{\partial \ln p(y_i, h(z))}{\partial z_j} \ln h_j(z_i).

The only two model-dependent terms are \ln p(y_i, h(z)) and \partial \ln p(y_i, h(z)) / \partial z_j.

ACKNOWLEDGEMENTS

The research was supported in part by grants from ARO, DARPA, DOE, NGA, ONR and NSF.

SUMMARY

Highlights:
• A unified variational copula inference framework
• Semi-parametric (flexibility + tractability)
• Automated with minimal model-specific derivations

Generalizability:
• The t-copula
• Mixture of Gaussian copulas
• Nonparametric copulas

Reproducibility: Code is available from GitHub: https://github.com/shabobhan/VariationalGaussianCopula