On Variational Bounds of Mutual Information

A number of machine learning problems make use of Mutual Information (MI) as a metric for estimating the relationship between variables. However, bounding MI is a challenging task. Variational bounds allow tractabe objectives which are scalable to high dimensional tasks. The work presents a unified framework for variational bounds which consists of existing lower and upper bounds on MI. The bounds exhibit high bias or variance when MI is large. To that end, the framework is improved by providing a continuum on lower bounds which trades off bias and variance. Suitability of bounds is assessed on high dimensional control problems and representation learning.

Variational bounds provide tractability of objectives and improved approximations in the case f parameterized distributions. The work reviews existing bounds in literature by accumulating variational objectives in a unified framework. The framework consists of bounds which maximize or minimize MI pertaining to the auxilary objective. In order to improve the performance of bounds with large values of MI, a continuum of multi-sampple bounds is proposed which reduces variance by sample dependence. The multi-sample unnormalized setup consists of samples from $p(x_1)p(y|x_1)$ and access to $K-1$ additional samples which are used to estimate MI $I(X;Y)$. The multi-sample InfoNCE bound is obtained as a lower bound the MI $I(X;Y)$. Validation of the bound is carried out on