[Peters, Bühlmann, and Meinshausen, "Causal Inference by using invariant prediction," JRSSB, 2016] and its various extensions/applications introduce the invariant causal prediction (ICP) framework that tries to avoid spurious correlation in learning. The basic idea is to explore the invariance in causal relationships through different environments to make prediction models more robust. For example data from different hospitals or linguistic communities. The aim is, in part, to identify features that are predictive across environments (and therefore not spurious).

We have a set of environments $\mathcal{E}$.

For linear models, the underlying assumption is the existence of invariant (w.r.t. environments) coefficients and noise variables.

There exists a vector of coefficients $\gamma^* = [\gamma_1^*, \ldots, \gamma_m^*]^T \in \mathbb{R}^m$ with support $S^* = \{i : \gamma_i^* \neq 0\}$ that satisfy:

For all $e \in \mathcal{E}$: $Y^e = X^e \gamma^* + N^e$, where $N^e \sim \mathcal{N}(0, \sigma^2 I)$.

Here σ^2 is unknown but in the range $[\sigma_{min}^2, \sigma_{max}^2]$, and $X^e = [x_1^e, \ldots, x_m^e] \in \mathbb{R}^{n_e \times m}$ is a deterministic design matrix with n_e samples from environment e. (In the original ICP framework, X^e are the input variables, but here a little different)

[Goddard, Xiang, Soloveychik, "Error Probability Bounds for Invariant Causal Prediction via Multiple Access Channels," 2023]

Try to connect this version of the ICP framework to the Gaussian MAC setting in order to bound error probabilities.

→ Bounds on the probability of error for Gaussian MAC with stored codebook exist for case of positive rate, where m is exponential in n. Here we want to focus on when m does not grow with n.

Consider a setting in which a vector of coefficients $\gamma^* \in \mathbb{R}^m$ is generated such that its support $S^* \subseteq \{1, \dots, m\}$ is uniformly chosen from all subsets of $\{1, \dots, m\}$. Let $w \in \mathbb{R}^m$ be a fixed vector where each element w_i is non-zero. The generation of $\gamma^* = [\gamma_1^*, \dots, \gamma_m^*]^T$ then follows from $\gamma_i^* = w_i$ if $i \in S^*$ and $\gamma_i^* = 0$ if $i \notin S^*$, for every $i \in \{1, \dots, m\}$.

The number of non-zero coefficients in γ^* is $k = |S^*|$.

Let X_S^e be matrix with only the columns of X^e indexed by set S. Those columns are causal predictors.

Upon receiving Y^e for each $e \in \mathcal{E}$, one wants to recover the support S^* (or equivalently γ^*). Let $\hat{S}_e^*$ be the estimate of S^* for some $e \in \mathcal{E}$. We correctly recover the support S^* if $\hat{S}_e^* = S^*$ for all e.

Error probability of error in recovering S^* is

$$P_{err} = \Pr[\hat{S}_e^* \neq S^* \text{ for any } e \in \mathcal{E}],$$ where probability over random support S^* and noise N^e

Again to emphasize: in the original ICP framework, X^e are modeled as random, but here X^e are considered deterministic due to focus on lower bounds.

(The other extreme is compressed sensing, and its information-theoretic perspective, where X^e has iid entries)

Relation between ICP and Gaussian MAC.

In a MAC, each sender $i \in \{1, \dots, k\}$ has access to codebook $C^i = \{c_1^i, c_2^i, \dots c_m^i\}$ where $c_j^i \in \mathbb{R}^n$ and m is # codewords in C^i.

To transmit information, the ith sender first chooses a codeword and then sends the tth element of the chosen codeword at time t as input symbol $x_{i,t}$, so receiver gets

$$y_t = h_1 x_{1,t} + h_2 x_{2,t} + \quad + h_k x_{k,t} + N_t$$

where h_i is channel gain for sender i and $N_t \sim \mathcal{N}(0, \sigma^2)$, for all $t \in \{1, \dots, n\}$

Usually the number of senders k is known.

Differences between this formulation of Gaussian MAC and model of ICP

1. In ICP, there is a single shared "codebook" X^e whereas in Gaussian MAC, each sender has own codebook
2. In ICP, k is arbitrary whereas in Gaussian MAC, k is fixed.
3. No notion of environment in Gaussian MAC.

To bridge these:

(a) Consider a codebook shared by all senders

(b) assume number of senders k is arbitrary.

Then recovery S^* for an individual environment is equivalent

to codeword recovery in a Gaussian MAC