Structural Causal Model Distance (SCMD) for Comparing Structural Causal Models

Théotime Le Goff [1] Émilie Devijver [1]

[1]Univ Grenoble Alpes, CNRS, Grenoble INP, LIG

Too Long; Didn't Read

We propose a pseudo-distance combining graphical and distributional criteria to compare structural causal models.

Motivation

A structural causal model over endogenous variables $\mathbf{V} = \{V_1, \dots, V_d\}$ and its corresponding graph are defined as

$$\mathcal{M}^i := \langle \boldsymbol{U}, \mathbf{V}, \mathcal{F}^i, P^i(\boldsymbol{U}) \rangle, \qquad \mathcal{G}^i := \langle \mathbf{V}, \mathcal{E}^i \rangle.$$

Graphical criterion (SID)

The structural intervention distance (SID, [1]; extended in [2,3]) provides a metric for comparing causal graphs $\mathcal{G}^1$ and $\mathcal{G}^2$:

$$\mathrm{SID}(\mathcal{G}^1, \mathcal{G}^2) = \#\Big\{(i,j) \in \mathcal{E}^1 \ \Big|\ \text{the interventional distribution from } i \text{ to } j \text{ is falsely estimated by } \mathcal{G}^2 \text{ wrt } \mathcal{G}^1\Big\}$$

Distribution criterion (MMD)

Joint distributions can also be compared using the maximum mean discrepancy (MMD, [4]). Let $\mathcal{H}$ be a reproducing kernel Hilbert space (RKHS) parameterized by a kernel k, for two distributions P^1 and P^2 over $\mathbf{V}$:

$$\mathrm{MMD}(P^1(\mathbf{V}), P^2(\mathbf{V})) := \|\mu_{P^1(\mathbf{V})}(\cdot) - \mu_{P^2(\mathbf{V})}(\cdot)\|_{\mathcal{H}}$$

where the kernel mean embedding of $P^1(\mathbf{V})$ is defined by

$$\mu_{P^1(\mathbf{V})}(\cdot) = \mathbb{E}_{\mathbf{V}\sim P^1(\mathbf{V})}[k(\mathbf{V}, \cdot)].$$

For a sample $(\mathbf{v}_1^1, \dots, \mathbf{v}_N^1)$, it is estimated by $\hat{\mu}^N_{P^1(\mathbf{V})}(\cdot) = \frac{1}{N}\sum_{n=1}^N k(\mathbf{v}_n^1, \cdot)$.

Example

$\mathcal{M}^1$

$\mathcal{G}^1$: $V_1 \rightarrow V_2$

$$\begin{cases} V_1 \sim \mathcal{N}(0,\,1) \\ V_2 \mid V_1 = v_1 \sim \mathcal{N}(av_1,\,1) \end{cases}$$

$\mathcal{M}^2$

$\mathcal{G}^2$: $V_1 \leftarrow V_2$

$$\begin{cases} V_2 \sim \mathcal{N}\left(0,\,1+a^2\right) \\ V_1 \mid V_2 = v_2 \sim \mathcal{N}\left(\frac{a}{1+a^2}v_2,\, \frac{1}{1+a^2}\right) \end{cases}$$

$$P_a^1 = P_a^2 = \mathcal{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & a \\ a & 1+a^2\end{pmatrix}\right)$$

$$\boxed{\begin{array}{c}\mathrm{MMD}(P_a^1, P_a^2; k) = 0 \\ \mathrm{SID}(\mathcal{G}^1, \mathcal{G}^2) \neq 0\end{array}} \qquad \boxed{\begin{array}{c}\mathrm{MMD}(P_a^1, P_b^1; k) \neq 0 \\ \mathrm{SID}(\mathcal{G}^1, \mathcal{G}^1) = 0\end{array}}$$

References

1. Peters et al. (2015). *Structural intervention distance for evaluating causal graphs.* Neural computation.
2. Henckel et al. (2024). *Adjustment identification distance: A gadjid for causal structure learning.* UAI
3. Wahl et al. (2025). *Separation-Based Distance Measures for Causal Graphs.* AISTATS
4. Gretton et al. (2012). *A kernel two-sample test.* JMLR
5. Park et al. (2020). *A measure-theoretic approach to kernel conditional mean embeddings.* NeurIPS
6. Muandet et al. (2017). *Kernel Mean Embedding of Distributions: A Review and Beyond.* Found. and Trends in ML.
7. Dhanakshirur et al. (2024). *A continuous structural intervention distance to compare causal graphs.* PCIC.
8. Bongers et al. (2021). *Foundations of structural causal models with cycles and latent variables.* AoS
9. Wildberger et al. (2023). *On the interventional kullback-leibler divergence.* CLeaR

Proposed distance

Tools: existing extensions of MMD

For conditional distributions [5,6]

$$\mathrm{MCMD}(P^1(V_j \mid V_i = v_i), P^2(V_j \mid V_i = v_i)) = \|\mu_{P^1(V_j|V_i=v_i)}(\cdot) - \mu_{P^2(V_j|V_i=v_i)}(\cdot)\|_{\mathcal{H}_j}$$

For interventional distributions [7]

$$\mathrm{MIMD}(P^1_{\mathrm{do}(V_i=v_i)}(V_j), P^2_{\mathrm{do}(V_i=v_i)}(V_j)) = \|\mu_{P^1_{\mathrm{do}(V_i=v_i)}(V_j)}(\cdot) - \mu_{P^2_{\mathrm{do}(V_i=v_i)}(V_j)}(\cdot)\|_{\mathcal{H}_j}$$

SCMD

For two SCMs $\mathcal{M}^1, \mathcal{M}^2$, for a vector of arbitrary interventions $\boldsymbol{v}$:

$$\mathrm{SCMD}(\mathcal{M}^1, \mathcal{M}^2; \boldsymbol{v}) := \sum_{1 \le i \neq j \le d} \mathrm{MIMD}\left(P^1_{\mathrm{do}(V_i=v_i)}(V_j), P^2_{\mathrm{do}(V_i=v_i)}(V_j)\right)$$

Properties: SCMD is non-negative, symmetric, satisfies the triangular inequality. But it is not separable in general:

$$\mathrm{SCMD}(\mathcal{M}^1, \mathcal{M}^2; \boldsymbol{v}) = 0$$
$$\not\Rightarrow \mathcal{M}_1 \text{ and } \mathcal{M}_2 \text{ are interventionally equivalent [8] wrt } \boldsymbol{v}.$$

Estimation

For two samples $\mathcal{D}^1 = (\mathbf{v}_1^1, \dots, \mathbf{v}_N^1)$ and $\mathcal{D}^2 = (\mathbf{v}_1^2, \dots, \mathbf{v}_N^2)$, assuming known the causal graphs $\mathcal{G}^1$ and $\mathcal{G}^2$,

- Conditional embeddings are estimated by solving a regularized vector-valued regression problem [5].
- Interventional embeddings are derived by integrating the conditional embedding over the adjustment variables [7] determined from the causal graphs.

$$\widehat{\mathrm{SCMD}}(\mathcal{D}^1, \mathcal{D}^2, \mathcal{G}^1, \mathcal{G}^2; \boldsymbol{v}) = \sum_{1 \le i \neq j \le d} \|\widehat{\mu}^{\lambda,N}_{P^1_{\mathrm{do}(V_i=v_i)}(V_j)}(\cdot) - \widehat{\mu}^{\lambda,N}_{P^2_{\mathrm{do}(V_i=v_i)}(V_j)}(\cdot)\|_{\mathcal{H}_j}$$

Illustration

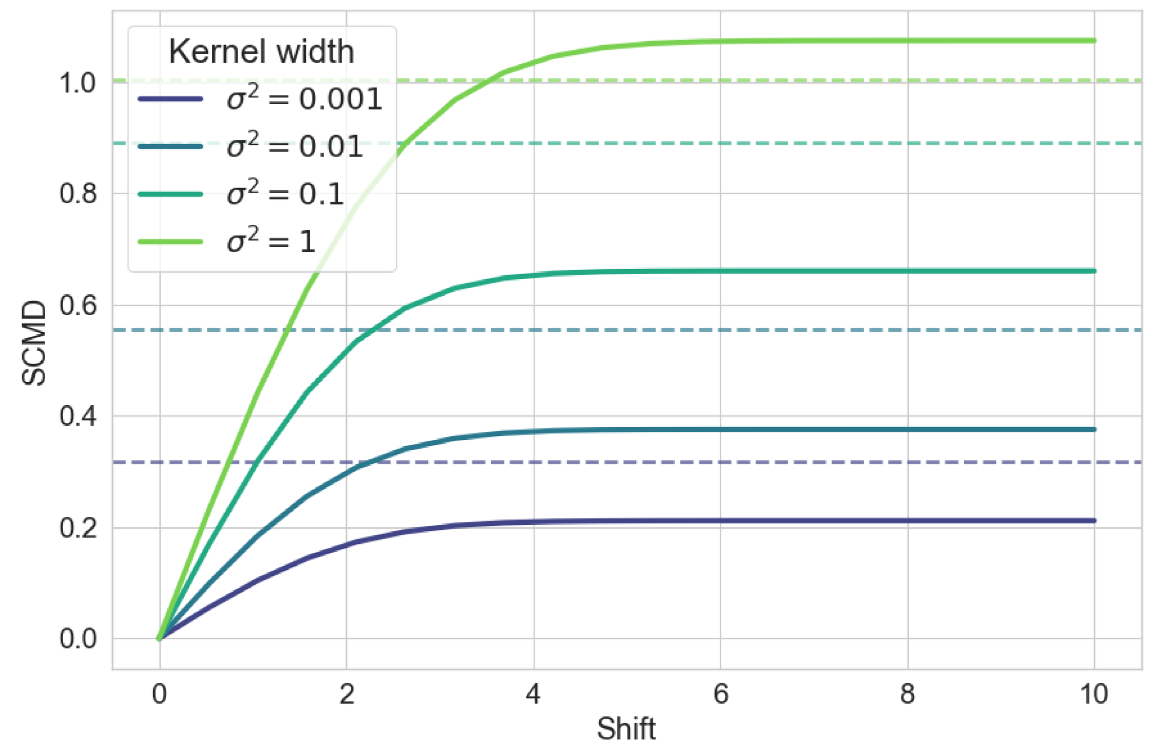
For a Gaussian kernel with bandwidth $\sigma^2 = 0.1$ and $v = (1, 1)$,

$$\mathrm{SCMD}(\mathcal{M}_a^1, \mathcal{M}_a^2; v)$$
$$= \|\mu_{\mathcal{N}(0,\,1)} - \mu_{\mathcal{N}\left(\frac{av_2}{1+a^2},\, \frac{1}{1+a^2}\right)}\|_{\mathcal{H}} + \|\mu_{\mathcal{N}(av_1,\,1)} - \mu_{\mathcal{N}(0,\,1+a^2)}\|_{\mathcal{H}}$$

$\mathrm{SCMD} = 0.8921$ and $\widehat{\mathrm{SCMD}} = 1.0046(\pm 0.02)$

$$\mathrm{SCMD}(\mathcal{M}_a^1, \mathcal{M}_b^1; v)$$
$$= \|\mu_{\mathcal{N}(0,\,1)} - \mu_{\mathcal{N}(0,\,1)}\|_{\mathcal{H}} + \|\mu_{\mathcal{N}(av_1,\,1)} - \mu_{\mathcal{N}(bv_1,\,1)}\|_{\mathcal{H}}$$

$\mathrm{SCMD} = 0.5177$ and $\widehat{\mathrm{SCMD}} = 0.5360(\pm 0.02)$



Dotted lines $\mathrm{SCMD}(\mathcal{M}_a^1, \mathcal{M}_a^2)$
Solid lines $\mathrm{SCMD}(\mathcal{M}_a^1, \mathcal{M}_{a+Shift}^1)$

theotime.legoff@univ-grenoble-alpes.fr