Poincaré inequalities for dimension reduction and efficient sampling

Olivier Zahm
olivier.zahm@inria.fr

INRIA – Université Grenoble-Alpes
https://team.inria.fr/airsea/en/

ETICS
October 2025

Overview

- **Today:** the (very) basics of ETICS scientific background
  - Part 0: **Sensitivity Analysis** for newbies (HSIC & Sobol)
  - Part I: The **curse of dimensionality** illustrated with Kernel Ridge Regression
  - Part II: **Gradient-based dimension reduction** to identify and exploit low-dimensional structure

- **Tomorrow:** some topics related to sampling
  - Part III: Dimension reduction for **Bayesian inverse problems**
  - Part IV: Optimal preconditionner for **Langevin-type sampler**

Overview

- **Today:** the (very) basics of ETICS scientific background
  - Part 0: **Sensitivity Analysis** for newbies (HSIC & Sobol)
  - Part I: The **curse of dimensionality** illustrated with Kernel Ridge Regression
  - Part II: **Gradient-based dimension reduction** to identify and exploit low-dimensional structure
- **Tomorrow:** some topics related to sampling
  - Part III: Dimension reduction for **Bayesian inverse problems**
  - Part IV: Optimal preconditionner for **Langevin-type sampler**

Poincaré inequality:

$$\mathrm{Var}(u(X)) \leq \mathbb{C}\ \mathbb{E}[\|\nabla u(X)\|^2]$$

Part 0: **Sensitivity Analysis** for newbies

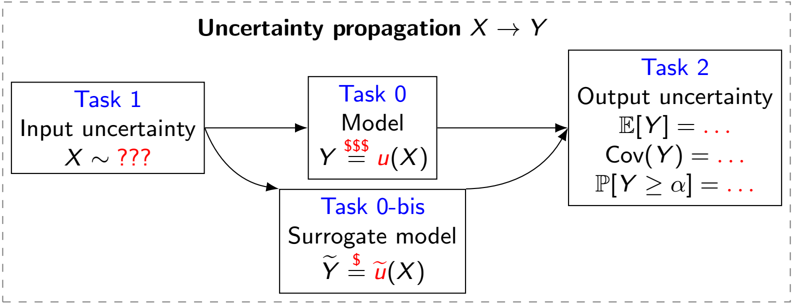
Context

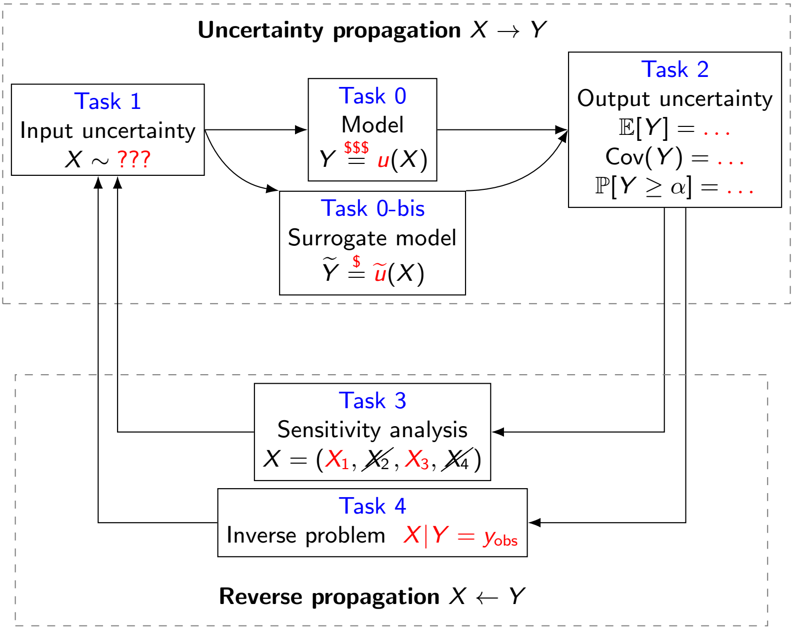
$$Y = u(X_1, \ldots, X_d)$$

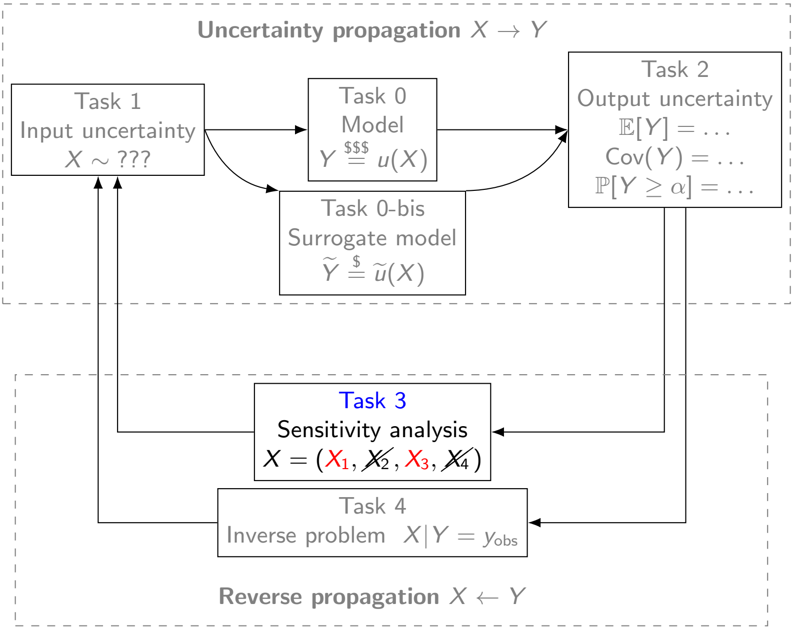
- $X \sim \pi_X$: input parameter, typically with product density $\pi_X(x)$
- Y: output of interest, generally scalar $Y \in \mathbb{R}$
- u: expensive numerical model

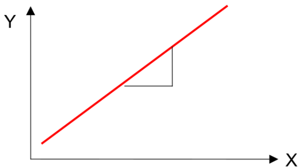
Example: $u(X) = \int v \mathrm{d}\Omega$, where v solved the parametrized PDE

$$-\mathrm{div}(\kappa \nabla v) = f \quad \text{on } \Omega, \quad \text{where} \quad \kappa = \kappa(X),$$

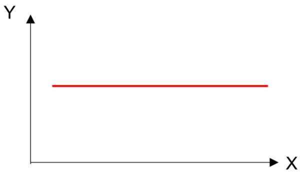

Uncertainty propagation $X \to Y$
Task 1
Input uncertainty
$X \sim ???$
Task 0
Model
$Y \overset{\$\$\$}{=} u(X)$
Task 0-bis
Surrogate model
$\widetilde{Y} \overset{\$}{=} \widetilde{u}(X)$
Task 2
Output uncertainty
$\mathbb{E}[Y] = \dots$
$\mathrm{Cov}(Y) = \dots$
$\mathbb{P}[Y \geq \alpha] = \dots$


Uncertainty propagation $X \to Y$
Task 1
Input uncertainty
$X \sim ???$
Task 0
Model
$Y \overset{\$\$\$}{=} u(X)$
Task 0-bis
Surrogate model
$\widetilde{Y} \overset{\$}{=} \widetilde{u}(X)$
Task 2
Output uncertainty
$\mathbb{E}[Y] = \ldots$
$\mathrm{Cov}(Y) = \ldots$
$\mathbb{P}[Y \geq \alpha] = \ldots$
Task 3
Sensitivity analysis
$X = (X_1, \cancel{X_2}, X_3, \cancel{X_4})$
Task 4
Inverse problem $X|Y = y_{\mathrm{obs}}$
Reverse propagation $X \leftarrow Y$

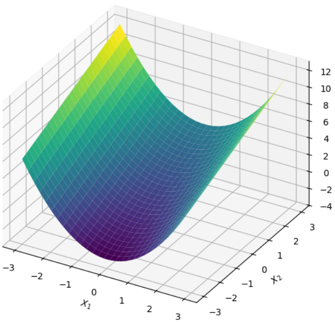

Uncertainty propagation $X \to Y$
Task 1
Input uncertainty
$X \sim ???$
Task 0
Model
$Y \overset{\$\$\$}{=} u(X)$
Task 0-bis
Surrogate model
$\widetilde{Y} \overset{\$}{=} \widetilde{u}(X)$
Task 2
Output uncertainty
$\mathbb{E}[Y] = \dots$
$\mathrm{Cov}(Y) = \dots$
$\mathbb{P}[Y \geq \alpha] = \dots$
Task 3
Sensitivity analysis
$X = (X_1, \cancel{X_2}, X_3, \cancel{X_4})$
Task 4
Inverse problem $X|Y = y_{\mathrm{obs}}$
Reverse propagation $X \leftarrow Y$



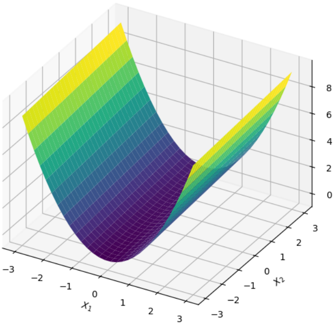
Y is sensitive to X



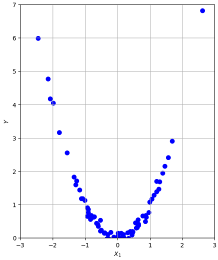
Y is not sensitive to X



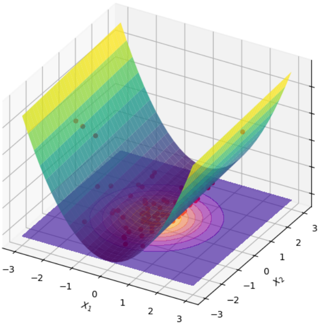
Y is sensitive to both X_1 and X_2

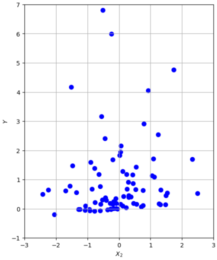


Y is not sensitive to X_2



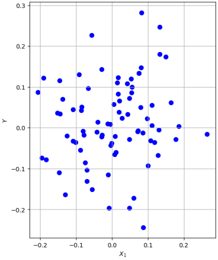
Y V.S. X_1



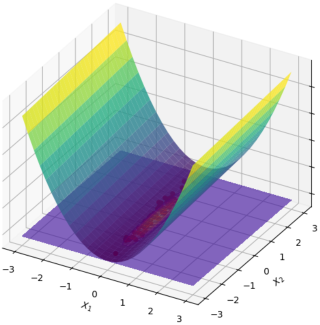


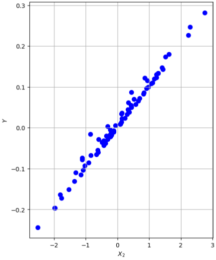
Y V.S. X_2

Y can be explained by X_1, whereas Y seems independent on X_2



Y V.S. X_1





Y V.S. X_2

~~Y can be explained by X_1, whereas Y seems independent on X_2~~

If we change the law π_X of X, the conclusion can be reversed.

Today, we focus on two standard estimators

1/ HSIC

2/ Sobol' indices

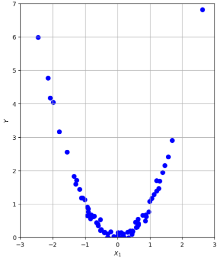
Today, we focus on two standard estimators

1/ HSIC

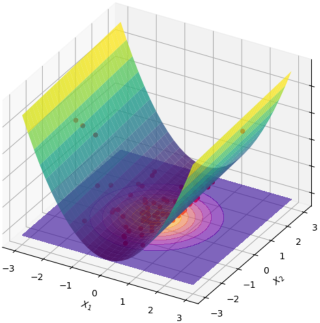
2/ Sobol' indices

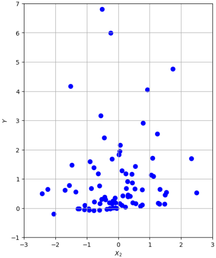
[Gretton et.al. 2005: Measuring statistical dependence with Hilbert-Schmidt norms]

[Da Veiga 2015: Global sensitivity analysis with dependence measures]



Y V.S. X_1





Y V.S. X_2

Consider $X_\tau = (X_i)_{i \in \tau}$ for some $\tau \subset \{1, \dots, d\}$.

$$Y \perp X_\tau \qquad \Rightarrow \qquad X_\tau \text{ is not influential on } Y$$

How independent Y and X_τ are? We need to measure some distance between

$$\pi_{Y,X_\tau} \overset{???}{\approx} \pi_Y \otimes \pi_{X_\tau}$$

Integral probability metric

$$D_{\mathcal{H}}(\pi, \rho) = \sup_{\substack{f \in \mathcal{H} \\ \|f\|_{\mathcal{H}} \leq 1}} \left| \mathbb{E}_{X \sim \pi}[f(X)] - \mathbb{E}_{Y \sim \rho}[f(Y)] \right|$$

- $\mathcal{H}$ = bounded functions $\Rightarrow$ total variation distance
- $\mathcal{H}$ = lipschitz functions $\Rightarrow$ Wasserstein-1 distance
- $\mathcal{H}$ = **Reproducing Kernel Hilbert Space** (RKHS) $\Rightarrow$ Maximum Mean Discrepancy

$$\mathcal{H}_\kappa = \left\{ f(x) = \sum_{i=1}^{\infty} \alpha_i \kappa(x, x_i) \;\middle|\; \begin{array}{l} x_i \in \mathbb{R}^d \text{ and } \alpha_1, \alpha_2, \ldots \in \mathbb{R} \text{ s.t.} \\ \|f\|_{\mathcal{H}}^2 := \sum_{i,j=1}^{\infty} \alpha_i \alpha_j \underbrace{\kappa(x_i, x_j)}_{\succeq 0} < \infty \end{array} \right\}$$

for some symmetric positive definite (SPD) function $\kappa(\cdot, \cdot)$.

Integral probability metric

$$D_{\mathcal{H}}(\pi, \rho) = \sup_{\substack{f \in \mathcal{H} \\ \|f\|_{\mathcal{H}} \le 1}} \left| \mathbb{E}_{X \sim \pi}[f(X)] - \mathbb{E}_{Y \sim \rho}[f(Y)] \right|$$

- $\mathcal{H}$ = bounded functions $\Rightarrow$ total variation distance
- $\mathcal{H}$ = lipschitz functions $\Rightarrow$ Wasserstein-1 distance
- $\mathcal{H}$ = **Reproducing Kernel Hilbert Space** (RKHS) $\Rightarrow$ Maximum Mean Discrepancy

$$\mathcal{H}_\kappa = \left\{ f(x) = \sum_{i=1}^{\infty} \alpha_i \kappa(x, x_i) \;\middle|\; \begin{array}{l} x_i \in \mathbb{R}^d \text{ and } \alpha_1, \alpha_2, \ldots \in \mathbb{R} \text{ s.t.} \\ \|f\|_{\mathcal{H}}^2 := \sum_{i,j=1}^{\infty} \alpha_i \alpha_j \underbrace{\kappa(x_i, x_j)}_{\succeq 0} < \infty \end{array} \right\}$$

for some symmetric positive definite (SPD) function $\kappa(\cdot, \cdot)$.

  - **Reproducing** property $\boxed{f(x) = \langle f, \kappa(x, \cdot) \rangle_{\mathcal{H}}}$
  - κ is **characteristic** when $\boxed{D_{\mathcal{H}_\kappa}(\pi, \rho) = 0 \quad \Leftrightarrow \quad \pi = \rho}$
  - Classic choice of kernel:

$$\kappa(x, x') = \exp\left(- \frac{\|x - x'\|^2}{2\ell^2} \right)$$

for some length scale ℓ, classically $\ell = \mathrm{median}(\|X^{(i)} - X^{(j)}\|)$

Maximum Mean Discrepancy: $\mathrm{MMD}(\pi, \rho) = D_{\mathcal{H}_\kappa}(\pi, \rho)$

$$\mathrm{MMD}(\pi, \rho) = \sup_{\substack{f \in \mathcal{H}_\kappa \\ \|f\|_{\mathcal{H}} \leq 1}} \left| \mathbb{E}_{X\sim\pi}[f(X)] - \mathbb{E}_{Y\sim\rho}[f(Y)] \right|$$

$$\text{(reproducing property)} \quad = \sup_{\substack{f \in \mathcal{H} \\ \|f\|_{\mathcal{H}} \leq 1}} \left| \mathbb{E}_{X\sim\pi}[\langle f, \kappa(X, \cdot)\rangle_{\mathcal{H}}] - \mathbb{E}_{Y\sim\rho}[\langle f, \kappa(Y, \cdot)\rangle_{\mathcal{H}}] \right|$$

$$\text{(mean embedding)} \quad = \sup_{\substack{f \in \mathcal{H} \\ \|f\|_{\mathcal{H}} \leq 1}} \left| \left\langle f \ , \ \underbrace{\mathbb{E}_{X\sim\pi}[\kappa(X, \cdot)]}_{\mu_\pi} - \underbrace{\mathbb{E}_{Y\sim\rho}[\kappa(Y, \cdot)]}_{\mu_\rho} \right\rangle_{\mathcal{H}} \right|$$

$$\left(f = \frac{\mu_\pi - \mu_\rho}{\|\mu_\pi - \mu_\rho\|_{\mathcal{H}}}\right) \quad = \|\mu_\pi - \mu_\rho\|_{\mathcal{H}}$$

Then

$$\mathrm{MMD}(\pi, \rho)^2 = \langle \mu_\pi, \mu_\pi \rangle_{\mathcal{H}} - 2\langle \mu_\pi, \mu_\rho \rangle_{\mathcal{H}} + \langle \mu_\rho, \mu_\rho \rangle_{\mathcal{H}}$$

$$= \mathbb{E}_{\substack{X\sim\pi \\ X'\sim\pi}}[\kappa(X, X')] - 2\mathbb{E}_{\substack{X\sim\pi \\ Y\sim\rho}}[\kappa(X, Y)] + \mathbb{E}_{\substack{Y\sim\rho \\ Y'\sim\rho}}[\kappa(Y, Y')]$$

$\rightarrow$ Monte Carlo!

Hilbert-Schmidt Independence Criterion (HSIC)

$$\boxed{\mathrm{HSIC}_\kappa(Y, X_\tau) = \mathrm{MMD}(\pi_{Y,X_\tau}, \pi_Y \otimes \pi_{X_\tau})^2}$$

where κ is a kernel in the $X_\tau \times Y$ space, typically

$$\kappa\Big((y, x_\tau), (y', x'_\tau)\Big) = \underbrace{\exp\left(-\frac{\|x_\tau - x'_\tau\|^2}{2\sigma_x^2}\right)}_{=\kappa_{\sigma_x}(x_\tau, x'_\tau)} \underbrace{\exp\left(-\frac{\|y - y'\|^2}{2\sigma_y^2}\right)}_{=\kappa_{\sigma_y}(y, y')}$$

Hilbert-Schmidt Independence Criterion (HSIC)

$$\boxed{\mathrm{HSIC}_\kappa(Y, X_\tau) = \mathrm{MMD}(\pi_{Y,X_\tau}, \pi_Y \otimes \pi_{X_\tau})^2}$$

where κ is a kernel in the $X_\tau \times Y$ space, typically

$$\kappa\Big((y, x_\tau), (y', x'_\tau)\Big) = \underbrace{\exp\left(-\frac{\|x_\tau - x'_\tau\|^2}{2\sigma_x^2}\right)}_{=\kappa_{\sigma_x}(x_\tau, x'_\tau)} \underbrace{\exp\left(-\frac{\|y - y'\|^2}{2\sigma_y^2}\right)}_{=\kappa_{\sigma_y}(y, y')}$$

Given a sample $\{(Y^{(i)}, X_\tau^{(i)})\}_{i=1}^N$, assemble the matrices

$$\boxed{K = \left[\kappa_{\sigma_x}(X_\tau^{(i)}, X_\tau^{(j)})\right]_{i,j=1}^N \qquad L = \left[\kappa_{\sigma_y}(Y^{(i)}, Y^{(j)})\right]_{i,j=1}^N}$$

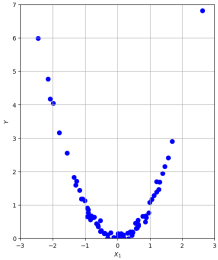
- Monte-Carlo (U-statistics): Unbiased, but no guarantee of positivity

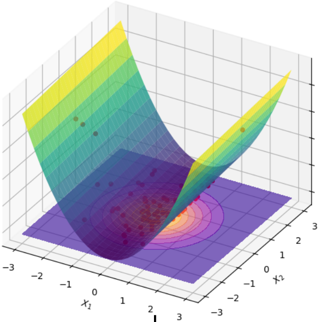
$$\mathrm{HSIC}_\kappa^{U,N}(Y, X_\tau) = \frac{1}{N(N-3)}\left(\mathrm{trace}(KL) - \frac{(\mathbf{1}^\top K\mathbf{1})(\mathbf{1}^\top L\mathbf{1})}{(N-1)(N-2)} - \frac{2}{N-2}\mathbf{1}^\top KL\mathbf{1}\right)$$

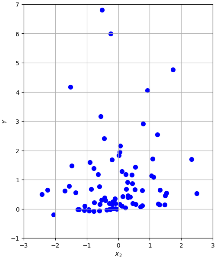
- Monte-Carlo (V-statistics): Positive and simple, but biased

$$\boxed{\mathrm{HSIC}_\kappa^{V,N}(Y, X_\tau) = \frac{1}{N^2}\mathrm{trace}(KHLH)} \quad \text{where } H = I_d - \frac{1}{N}\mathbf{1}\mathbf{1}^\top$$

$$\boxed{\mathrm{HSIC}_{\kappa}^{V,N}(Y, X_{\tau}) = \frac{1}{N^2}\,\mathrm{trace}(KHLH) \;\approx\; \langle K, L\rangle_{Frobenius}}$$



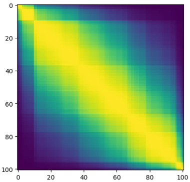




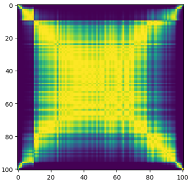
Y V.S. X_1

Y V.S. X_2

K

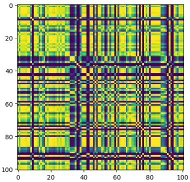


L

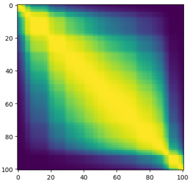


sort the samples $X_1^{(1)} \leq \ldots \leq X_1^{(N)}$

L



K



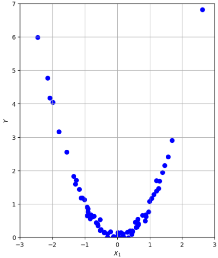
sort the samples $X_2^{(1)} \leq \ldots \leq X_2^{(N)}$

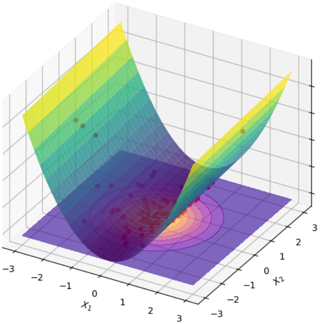
hsic_sensitivity_demo.ipynb

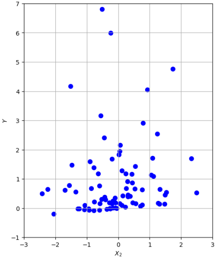
Today, we focus on two standard estimators

1/ HSIC

2/ Sobol' indices

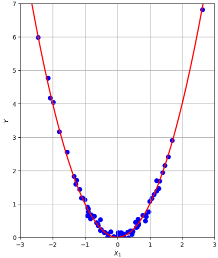




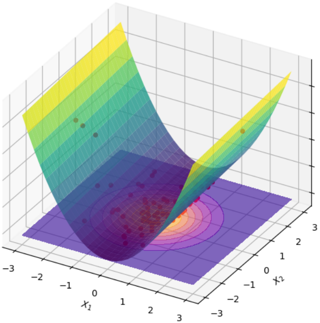


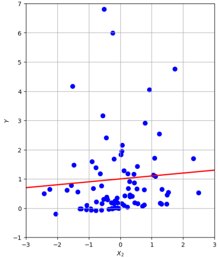
Y V.S. X_1

Y V.S. X_2



Y V.S. X_1



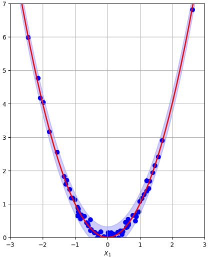


Y V.S. X_2

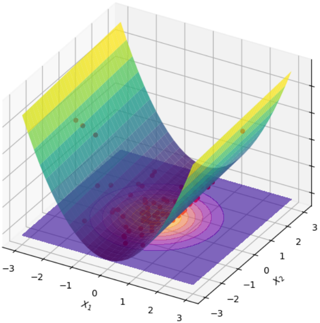
For $\tau \subset \{1, \dots, d\}$, let $X_\tau = (X_i)_{i \in \tau}$ and $X_{-\tau} = (X_i)_{i \notin \tau}$

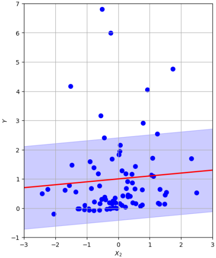
- Conditional expectation: if the variance of $\mathbb{E}[u(X)|X_\tau]$ is close to the variance of Y, then X_τ is an explanatory factor. **Closed Sobol' index:**

$$S_\tau = \frac{\mathrm{Var}(\mathbb{E}[u(X)|X_\tau])}{\mathrm{Var}(u(X))} \approx 1 \quad \Rightarrow \quad X_\tau = \text{explanatory factor}$$



Y V.S. X_1





Y V.S. X_2

For $\tau \subset \{1, \ldots, d\}$, let $X_\tau = (X_i)_{i \in \tau}$ and $X_{-\tau} = (X_i)_{i \notin \tau}$

- Conditional expectation: if the variance of $\mathbb{E}[u(X)|X_\tau]$ is close to the variance of Y, then X_τ is an explanatory factor. **Closed Sobol' index:**

$$S_\tau = \frac{\mathrm{Var}(\mathbb{E}[u(X)|X_\tau])}{\mathrm{Var}(u(X))} \approx 1 \quad \Rightarrow \quad X_\tau = \text{explanatory factor}$$

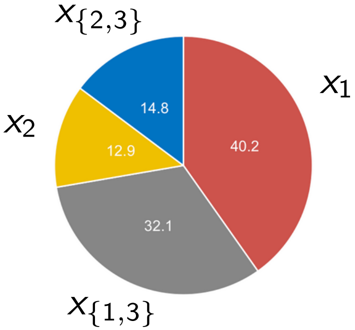
- Conditional variance: if $\mathrm{Var}(u(X)|X_1) \ll \mathrm{Var}(Y)$, then $u(X_1^{\text{freeze}}, X_2)$ doesn't recover the variance of $Y = u(X)$. **Total Sobol' index:**

$$T_\tau = \frac{\mathbb{E}[\mathrm{Var}(u(X)|X_{-\tau})]}{\mathrm{Var}(u(X))} \approx 0 \quad \Rightarrow \quad \Rightarrow X_\tau = \text{irrelevant factor}$$

ANOVA: the **An**alysis **o**f **Va**riance

What percentage of the variance of $Y = u(X_1, \ldots, X_d)$ is attributed to $X_\tau = (X_{\tau_1}, \ldots, X_{\tau_m})$ for some $\tau \subset \{1, \ldots, d\}$?

$$\mathrm{Var}(Y) =$$



[Sobol' 1993: Sensitivity estimates for nonlinear mathematical models]

ANOVA: the Analysis of Variance

Assuming π_X is a product density[1], the **Hoeffding decomposition** of u is

$$u(X) = \mathbb{E}[u(X)] + \sum_{i=1}^{d} u_i(X_i) + \sum_{i\neq j}^{d,d} u_{i,j}(X_i, X_j) + \sum_{i\neq j\neq k}^{d,d,d} u_{i,j,k}(X_i, X_j, X_k) + \dots$$

where $u_\alpha(X_\alpha)$ are the **decorrelated**. We deduce

$$\boxed{\mathrm{Var}(u(X)) = \sum_{\alpha\subset\{1,\dots,d\}} \mathrm{Var}(u_\alpha(X_\alpha))}$$

[1] to ensure $\mathbb{E}\left[\mathbb{E}[u(X)|X_\tau]\middle|X_{\tau'}\right] = \mathbb{E}\left[\mathbb{E}[u(X)|X_{\tau'}]\middle|X_\tau\right] = \mathbb{E}[u(X)|X_{\tau\cap\tau'}]$

ANOVA: the **An**alysis **of** **Va**riance

Assuming π_X is a product density[1], the **Hoeffding decomposition** of u is

$$u(X) = \mathbb{E}[u(X)] + \sum_{i=1}^{d} u_i(X_i) + \sum_{i \neq j}^{d,d} u_{i,j}(X_i, X_j) + \sum_{i \neq j \neq k}^{d,d,d} u_{i,j,k}(X_i, X_j, X_k) + \dots$$

where $u_\alpha(X_\alpha)$ are the **decorrelated**. We deduce

$$\boxed{\mathrm{Var}(u(X)) = \sum_{\alpha \subset \{1,\dots,d\}} \mathrm{Var}(u_\alpha(X_\alpha))}$$

Assuming $\mathrm{Var}(u) = 1$:

- **Closed Sobol' index** $\quad S_\tau(u) = \sum_{\alpha \subset \tau} \mathrm{Var}(u_\alpha)$
- **Total Sobol' index** $\quad T_\tau(u) = \sum_{\alpha \cap \tau \neq \emptyset} \mathrm{Var}(u_\alpha)$
- Superset importance $\quad \Upsilon_\tau^2(u) = \sum_{\alpha \supset \tau} \mathrm{Var}(u_\alpha)$
- Shapley-Owen value $\quad \phi_\tau(u) = \sum_{\alpha \supset \tau} \frac{\mathrm{Var}(u_\alpha)}{|\alpha| - |\tau| + 1}$
- ...

[1] to ensure $\mathbb{E}\left[\mathbb{E}[u(X)|X_\tau]\middle|X_{\tau'}\right] = \mathbb{E}\left[\mathbb{E}[u(X)|X_{\tau'}]\middle|X_\tau\right] = \mathbb{E}[u(X)|X_{\tau\cap\tau'}]$

The **function approximation** perspective

Let $L^2_{\pi_X}$ be the space of square-integrable functions endowed with the norm

$$\|u\|^2 = \int u(x)^2 \mathrm{d}\pi_X(x)$$

Expectations and conditional expectations are orthogonal projections in $L^2_{\pi_X}$:

The **function approximation** perspective

Let $L^2_{\pi_X}$ be the space of square-integrable functions endowed with the norm

$$\|u\|^2 = \int u(x)^2 \mathrm{d}\pi_X(x)$$

Expectations and conditional expectations are orthogonal projections in $L^2_{\pi_X}$:

- The constant $c \in \mathbb{R}$ which best approximates u in $L^2_{\pi_X}$ is the **expectation** $c = \mathbb{E}[u(X)]$

$$\min_{c \in \mathbb{R}} \|u - c\|^2 =: \mathrm{Var}(u(X))$$

The **function approximation** perspective

Let $L^2_{\pi_X}$ be the space of square-integrable functions endowed with the norm

$$\|u\|^2 = \int u(x)^2 \mathrm{d}\pi_X(x)$$

Expectations and conditional expectations are orthogonal projections in $L^2_{\pi_X}$:

- The constant $c \in \mathbb{R}$ which best approximates u in $L^2_{\pi_X}$ is the **expectation** $c = \mathbb{E}[u(X)]$

$$\min_{c\in\mathbb{R}} \|u - c\|^2 =: \mathrm{Var}(u(X))$$

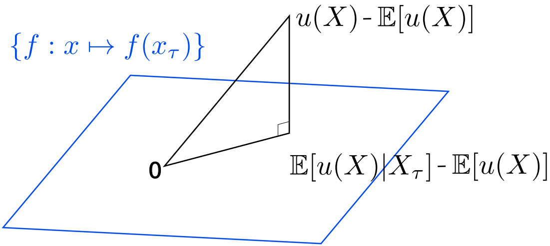
- For any $\tau \subset \{1, \ldots, d\}$, the function $f : x \mapsto f(x_\tau)$ which best approximates u as

$$\min_{f : x \mapsto f(x_\tau)} \|u - f\|^2 =: \mathbb{E}[\mathrm{Var}(u(X)|X_\tau)]$$

is the **conditional expectation** $f(x_\tau) = \mathbb{E}[u(X)|X_\tau = x_\tau]$.

Total variance formula = Pythagorean theorem in $L^2_{\pi_X}$

$$\|u - \mathbb{E}[u(X)]\|^2 = \|u - \mathbb{E}[u(X)|X_\tau]\|^2 + \|\mathbb{E}[u(X)|X_\tau] - \mathbb{E}[u(X)]\|^2$$



Put in statistical language:

$$\boxed{\mathrm{Var}(u(X)) = \underbrace{\mathbb{E}[\mathrm{Var}(u(X)|X_\tau)]}_{\min\limits_{f:x\mapsto f(x_\tau)} \|u-f\|^2} + \mathrm{Var}(\mathbb{E}[u(X)|X_\tau]) \qquad (\star)}$$

Interpretation of **Sobol' indices**

The **closed Sobol' indices** writes

$$S_\tau(u) := \frac{\mathrm{Var}(\mathbb{E}[u(X)|X_\tau])}{\mathrm{Var}(u(X))} \overset{(\star)}{=} 1 - \frac{\min\limits_{f : x \mapsto f(x_\tau)} \|u - f\|^2}{\mathrm{Var}(u(X))}$$

$$\boxed{\begin{aligned} S_\tau(u) \approx 1 \quad &\Leftrightarrow \quad u(X) \approx f(X_\tau) \\ &\Leftrightarrow \quad X_\tau \text{ ``explains'' well } Y = u(X) \\ &\quad\; \text{usefull for ranking the } X_\tau \end{aligned}}$$

Interpretation of **Sobol' indices**

The **closed Sobol' indices** writes

$$S_\tau(u) := \frac{\mathrm{Var}(\mathbb{E}[u(X)|X_\tau])}{\mathrm{Var}(u(X))} \overset{(\star)}{=} 1 - \frac{\min_{f:x\mapsto f(x_\tau)} \|u-f\|^2}{\mathrm{Var}(u(X))}$$

$$\boxed{\begin{array}{rcl} S_\tau(u)\approx 1 & \Leftrightarrow & u(X)\approx f(X_\tau) \\ & \Leftrightarrow & X_\tau \text{ ``explains'' well } Y=u(X) \\ & & \text{usefull for ranking the } X_\tau \end{array}}$$

Similarly, the **total Sobol' indices** writes

$$T_\tau(u) := 1 - \frac{\mathrm{Var}(\mathbb{E}[u(X)|X_{-\tau}])}{\mathrm{Var}(u(X))} \overset{(\star)}{=} \frac{\min_{f:x\mapsto f(x_{-\tau})} \|u-f\|^2}{\mathrm{Var}(u(X))}$$

$$\boxed{\begin{array}{rcl} T_\tau(u)\approx 0 & \Leftrightarrow & u(X)\approx f(X_{-\tau}) \\ & \Leftrightarrow & X_\tau \text{ is useless to ``explain'' } Y=u(X) \\ & & \text{usefull for screening the } X_\tau \end{array}}$$

Pick-freeze estimator of **total Sobol' indices**

Assume π_X is a product density. For $\tau \subset \{1, \dots, d\}$, let $X' = (X'_\tau, X'_{-\tau})$ be an iid copie of $X = (X_\tau, X_{-\tau})$. Then

$$\mathbb{E}[\mathrm{Var}(u(X)|X_{-\tau})] = \dots \text{ (exercise) } \dots = \frac{1}{2}\mathbb{E}[(\underbrace{u(X)}_{Y} - \underbrace{u(X'_\tau, X_{-\tau})}_{Y_{-\tau}})^2]$$

Then

$$\boxed{T_\tau(u) = \frac{\frac{1}{2}\mathbb{E}[(Y - Y_{-\tau})^2]}{\mathbb{E}[Y^2] - \mathbb{E}[Y]^2}}$$

Pick-freeze estimator of **total Sobol' indices**

Assume π_X is a product density. For $\tau \subset \{1, \ldots, d\}$, let $X' = (X'_\tau, X'_{-\tau})$ be an iid copie of $X = (X_\tau, X_{-\tau})$. Then

$$\mathbb{E}[\mathrm{Var}(u(X)|X_{-\tau})] = \ldots \text{ (exercise) } \ldots = \frac{1}{2}\mathbb{E}[(\underbrace{u(X)}_{Y} - \underbrace{u(X'_\tau, X_{-\tau})}_{Y_{-\tau}})^2]$$

Then

$$\boxed{T_\tau(u) = \frac{\frac{1}{2}\mathbb{E}[(Y - Y_{-\tau})^2]}{\mathbb{E}[Y^2] - \mathbb{E}[Y]^2}}$$

Given N iid copies of $(Y, Y_{-\tau})$, a Pick-Freeze estimator of $T_\tau(u)$ reads

$$T_\tau^{N,\mathrm{PF}}(u) = \frac{\frac{1}{2N}\sum_{i=1}^N (Y^{(i)} - Y_{-\tau}^{(i)})^2}{\frac{1}{N}\sum_{i=1}^N Y^{(i)2} - \left(\frac{1}{N}\sum_{i=1}^N Y^{(i)}\right)^2}$$

Pick-freeze estimator of **closed Sobol' indices**

Assume π_X is a product density. Then

$$\mathrm{Var}(\mathbb{E}[u(X)|X_\tau]) = \ldots \text{ (exercise) } \ldots = \mathrm{Cov}(\underbrace{u(X)}_{Y}, \underbrace{u(X_\tau, X'_{-\tau})}_{Y_\tau})$$

so that

$$\boxed{S_\tau(u) = \frac{\mathrm{Cov}(Y, Y_\tau)}{\mathrm{Var}(Y)} = \frac{\mathbb{E}[Y Y_\tau] - \mathbb{E}[Y]^2}{\mathbb{E}[Y^2] - \mathbb{E}[Y]^2}}$$

Pick-freeze estimator of **closed Sobol' indices**

Assume π_X is a product density. Then

$$\mathrm{Var}(\mathbb{E}[u(X)|X_\tau]) = \dots \text{ (exercise) } \dots = \mathrm{Cov}(\underbrace{u(X)}_{Y}, \underbrace{u(X_\tau, X'_{-\tau})}_{Y_\tau})$$

so that

$$\boxed{S_\tau(u) = \frac{\mathrm{Cov}(Y, Y_\tau)}{\mathrm{Var}(Y)} = \frac{\mathbb{E}[Y Y_\tau] - \mathbb{E}[Y]^2}{\mathbb{E}[Y^2] - \mathbb{E}[Y]^2}}$$

Given N iid copies of (Y, Y_τ), a Pick-Freeze estimator of $S_\tau(u)$ reads

$$S_\tau^{N,\mathrm{PF}}(u) = \frac{\frac{1}{N}\sum_{i=1}^N Y^{(i)} Y_\tau^{(i)} - \left(\frac{1}{N}\sum_{i=1}^N Y^{(i)}\right)^2}{\frac{1}{N}\sum_{i=1}^N Y^{(i)2} - \left(\frac{1}{N}\sum_{i=1}^N Y^{(i)}\right)^2}$$

Pick-freeze estimators

Are Pick-Freeze estimators **“good estimator”** ?

Pick-freeze estimators

Are Pick-Freeze estimators **“good estimator”** ?

Yes: efficient estimators, converge in $\mathcal{O}(1/\sqrt{N})$, etc etc etc

[Janon et.al. 2013: symptotic normality and efficiency of two Sobol index estimators]

[Owen 2013: Better estimation of small Sobol’ indices]

[Da Veiga, Gamboa 2013: Efficient estimation of sensitivity indices]

[Gamboa et.al. 2015: Statistical Inference for Sobol’ Pick Freeze Monte Carlo method]

Pick-freeze estimators

Are Pick-Freeze estimators **"good estimator"** ?

Yes: efficient estimators, converge in $\mathcal{O}(1/\sqrt{N})$, etc etc etc

[Janon et.al. 2013: symptotic normality and efficiency of two Sobol index estimators]

[Owen 2013: Better estimation of small Sobol' indices]

[Da Veiga, Gamboa 2013: Efficient estimation of sensitivity indices]

[Gamboa et.al. 2015: Statistical Inference for Sobol' Pick Freeze Monte Carlo method]

but....

- Computationally expensive: estimating **all first order indices** would requires $(d+1)N$ model evaluations.
- This approach relies on a **specific experimental design**, namely $(u(X_\tau, X_{-\tau}), u(X_\tau, X'_{-\tau}))$, which may not be feasible in practice. For instance in a **given data** situation where we can't query the model...

Pick-freeze estimators

Are Pick-Freeze estimators **"good estimator"** ?

Yes: efficient estimators, converge in $\mathcal{O}(1/\sqrt{N})$, etc etc etc

[Janon et.al. 2013: symptotic normality and efficiency of two Sobol index estimators]

[Owen 2013: Better estimation of small Sobol' indices]

[Da Veiga, Gamboa 2013: Efficient estimation of sensitivity indices]

[Gamboa et.al. 2015: Statistical Inference for Sobol' Pick Freeze Monte Carlo method]

but....

- Computationally expensive: estimating **all first order indices** would requires $(d+1)N$ model evaluations.
- This approach relies on a **specific experimental design**, namely $(u(X_\tau, X_{-\tau}), u(X_\tau, X'_{-\tau}))$, which may not be feasible in practice. For instance in a **given data** situation where we can't query the model...

Other estimators:

[Da Veiga, Gamboa, Lagnoux, Klein 2023: New estimation of Sobol'indices using kernels]

[Gamboa, Gremaud, Klein, Lagnoux 2022: Global Sensitivity Analysis : a new generation of mighty estimators based on rank statistics]

Rank-based estimator of order-1 closed Sobol' indices

Assume both Y and X_k are **real-valued, continuous** randon variables. Given iid samples $\{(Y^{(i)}, X_k^{(i)})\}_{i=1}^N$, renumber

$$(Y^{[1]}, X_k^{[1]}), \ldots, (Y^{[N]}, X_k^{[N]})$$

by **ranking** the $X_k \in \mathbb{R}$ so that

$$X_k^{[1]} \leq X_k^{[2]} \leq \ldots \leq X_k^{[N-1]} \leq X_k^{[N]}$$

[Gamboa, Gremaud, Klein, Lagnoux 2022: Global Sensitivity Analysis : a new generation of mighty estimators based on rank statistics]

Rank-based estimator of order-1 closed Sobol' indices

Assume both Y and X_k are **real-valued, continuous** randon variables. Given iid samples $\{(Y^{(i)}, X_k^{(i)})\}_{i=1}^N$, renumber

$$(Y^{[1]}, X_k^{[1]}), \ldots, (Y^{[N]}, X_k^{[N]})$$

by **ranking** the $X_k \in \mathbb{R}$ so that

$$X_k^{[1]} \leq X_k^{[2]} \leq \ldots \leq X_k^{[N-1]} \leq X_k^{[N]}$$

(we cannot rank $X_\tau = (X_{\tau_1}, X_{\tau_2}, ...)$ when $\#\tau \geq 2$!!)

[Gamboa, Gremaud, Klein, Lagnoux 2022: Global Sensitivity Analysis : a new generation of mighty estimators based on rank statistics]

Rank-based estimator of order-1 closed Sobol' indices

Assume both Y and X_k are **real-valued, continuous** randon variables. Given iid samples $\{(Y^{(i)}, X_k^{(i)})\}_{i=1}^N$, renumber

$$(Y^{[1]}, X_k^{[1]}), \ldots, (Y^{[N]}, X_k^{[N]})$$

by **ranking** the $X_k \in \mathbb{R}$ so that

$$X_k^{[1]} \leq X_k^{[2]} \leq \ldots \leq X_k^{[N-1]} \leq X_k^{[N]}$$

(we cannot rank $X_\tau = (X_{\tau_1}, X_{\tau_2}, \ldots)$ when $\#\tau \geq 2$!!)

The rank-estimator for the closed Sobol' index reads

$$S_{\{k\}}^{N,\mathrm{Rank}}(u) = \frac{\boxed{\frac{1}{N}\sum_{i=1}^{N-1} Y^{[i]} Y^{[i+1]}} - \left(\frac{1}{N}\sum_{i=1}^N Y^{[i]}\right)^2}{\frac{1}{N}\sum_{i=1}^N Y^{[i]2} - \left(\frac{1}{N}\sum_{i=1}^N Y^{[i]}\right)^2}$$

- Based on **[“A new coefficient of correlation” by Chatterjee 2020]**
- Consistent, converges in $\mathcal{O}(1/\sqrt{N})$
- N **model evaluations suffice to compute all order-1 closed indices**

[Gamboa, Gremaud, Klein, Lagnoux 2022: Global Sensitivity Analysis : a new generation of mighty estimators based on rank statistics]

Intuition behind the rank-estimator

$$S_{\{k\}}^{N,\mathrm{Rank}}(u) = \frac{\boxed{\dfrac{1}{N}\displaystyle\sum_{i=1}^{N-1} Y^{[i]} Y^{[i+1]}} - \left(\frac{1}{N}\sum_{i=1}^{N} Y^{[i]}\right)^2}{\frac{1}{N}\sum_{i=1}^{N} Y^{[i]2} - \left(\frac{1}{N}\sum_{i=1}^{N} Y^{[i]}\right)^2}$$

Intuition behind the rank-estimator

$$S_{\{k\}}^{N,\mathrm{Rank}}(u) = \frac{\boxed{\frac{1}{N}\sum_{i=1}^{N-1} Y^{[i]} Y^{[i+1]}} - \left(\frac{1}{N}\sum_{i=1}^{N} Y^{[i]}\right)^2}{\frac{1}{N}\sum_{i=1}^{N} Y^{[i]2} - \left(\frac{1}{N}\sum_{i=1}^{N} Y^{[i]}\right)^2}$$

- If Y and X_k are **independent**, then $Y^{[i]}$ and $Y^{[i+1]}$ are also independent so that

$$Y^{[i]} \perp Y^{[i+1]} \quad \Rightarrow \quad \frac{1}{N}\sum_{i=1}^{N-1} Y^{[i]} Y^{[i+1]} \approx \mathbb{E}[YY'] \overset{Y \perp Y'}{=} \mathbb{E}[Y]^2 \quad \Rightarrow \quad \boxed{S_{\{k\}}^{N,\mathrm{Rank}}(u) \approx 0}$$

Intuition behind the rank-estimator

$$S^{N,\mathrm{Rank}}_{\{k\}}(u) = \frac{\boxed{\frac{1}{N}\sum_{i=1}^{N-1} Y^{[i]}Y^{[i+1]}} - \left(\frac{1}{N}\sum_{i=1}^{N} Y^{[i]}\right)^2}{\frac{1}{N}\sum_{i=1}^{N} Y^{[i]2} - \left(\frac{1}{N}\sum_{i=1}^{N} Y^{[i]}\right)^2}$$

- If Y and X_k are **independent**, then $Y^{[i]}$ and $Y^{[i+1]}$ are also independent so that

$$Y^{[i]} \perp Y^{[i+1]} \quad \Rightarrow \quad \frac{1}{N}\sum_{i=1}^{N-1} Y^{[i]}Y^{[i+1]} \approx \mathbb{E}[YY'] \overset{Y \perp Y'}{=} \mathbb{E}[Y]^2 \quad \Rightarrow \quad \boxed{S^{N,\mathrm{Rank}}_{\{k\}}(u) \approx 0}$$

- If $Y = u(X_k)$ then a Taylor expansion $Y^{[i+1]} = u(X_k^{[i]}) + \mathcal{O}(\Delta X_k^{[i]})$ yields

$$\frac{1}{N}\sum_{i=1}^{N-1} Y^{[i]}Y^{[i+1]} = \frac{1}{N}\sum_{i=1}^{N-1} Y^{[i]}\left(Y^{[i]} + \mathcal{O}(\Delta X_k^{[i]})\right) = \frac{1}{N}\sum_{i=1}^{N-1} Y^{[i]2} + \mathcal{O}(1)$$

so that $\boxed{S^{N,\mathrm{Rank}}_{\{k\}}(u) \approx 1}$

(disclaimer) HSIC $\longleftrightarrow$ Sobol rank-estimator

$$K = \frac{1}{2}\begin{pmatrix} 0 & 1 & & & 0 \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 1 \\ 0 & & & 1 & 0 \end{pmatrix} \qquad \text{and} \qquad L = \begin{pmatrix} Y^{[1]}Y^{[1]} & \dots & Y^{[1]}Y^{[N]} \\ \vdots & \ddots & \vdots \\ Y^{[N]}Y^{[1]} & \dots & Y^{[N]}Y^{[N]} \end{pmatrix}$$

Because $K\mathbf{1} = \mathbf{1}$ (well, almost...), we have (remember $H = I_N - \frac{1}{N}\mathbf{1}\mathbf{1}^\top$)

$$\begin{aligned} \frac{1}{N^2}\operatorname{trace}(KHLH) &= \frac{1}{N^2}\operatorname{trace}(KL) - \frac{\mathbf{1}L\mathbf{1}^\top}{N^3} \\ &= \frac{1}{N}\left(\frac{1}{N}\sum_{i=1}^{N-1} Y^{[i]}Y^{[i+1]} - (\frac{1}{N}\sum_{i=1}^{N} Y^{[i]})^2\right) \end{aligned}$$

This is the numerator of the rank-estimator of the closed Sobol index!!

(disclaimer) HSIC $\longleftrightarrow$ Sobol rank-estimator

$$K = \frac{1}{2}\begin{pmatrix} 0 & 1 & & & 0 \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 1 \\ 0 & & & 1 & 0 \end{pmatrix} \quad \text{and} \quad L = \begin{pmatrix} Y^{[1]}Y^{[1]} & \dots & Y^{[1]}Y^{[N]} \\ \vdots & \ddots & \vdots \\ Y^{[N]}Y^{[1]} & \dots & Y^{[N]}Y^{[N]} \end{pmatrix}$$

Because $K\mathbf{1} = \mathbf{1}$ (well, almost...), we have (remember $H = I_N - \frac{1}{N}\mathbf{1}\mathbf{1}^\top$)

$$\begin{aligned} \frac{1}{N^2}\operatorname{trace}(KHLH) &= \frac{1}{N^2}\operatorname{trace}(KL) - \frac{\mathbf{1}L\mathbf{1}^\top}{N^3} \\ &= \frac{1}{N}\left(\frac{1}{N}\sum_{i=1}^{N-1} Y^{[i]}Y^{[i+1]} - (\frac{1}{N}\sum_{i=1}^{N} Y^{[i]})^2\right) \end{aligned}$$

This is the numerator of the rank-estimator of the closed Sobol index!!

Question: what kernels on X and on Y permit to recover the above K and L?

$$\kappa_Y(y, y') \overset{\text{easy!}}{=} yy' \qquad \kappa_X(x, x') \overset{????}{=} \lim_{\ell \to \varepsilon > 0} \exp\left(-\frac{\|x - x'\|^2}{2\ell^2}\right)$$

`Sobol_demo.ipynb`

(Notebook by Agnès Lagnoux)